Publications of the Astronomy Department of the Eötvös University

Volume 14

British-Hungarian N+N Workshop for Young Researchers

On Computer processing and use of satellite data in astronomy and astrophysics

and

3rd Workshop of Young Researchers in Astronomy & Astrophysics

Edited by
E. Forgács-Dajka
K. Petrovay
R. Erdélyi

BRITISH COUNCIL

Budapest 2004

Publications of the Astronomy Department of
the Eötvös Loránd University
Vol. 14.

British - Hungarian N+N Workshop for Young Researchers

on Computer processing and use of satellite data in astronomy and astrophysics

and

3rd Workshop of Young Researchers in Astronomy & Astrophysics

Budapest, Hungary, 3-7 February 2004

EDITED BY

E. Forgács-Dajka
K. Petrovay
R. Erdélyi

Published by the Department of Astronomy of the Eötvös University
Budapest 2004

Publication of this volume has been aided by the British Council.

ISBN 963 463 557
Responsible publisher: Dr. Bálint Érdi

Contents

Solar Physics

Optical studies and distant space

Celestial mechanics

Preface

The N+N Young Researchers' Workshop scheme, initiated and funded by the British Council, aims to provide an opportunity for young researchers to exchange ideas, knowledge and information by coming together in the form of N+N workshops and meetings. Here the term N+N workshop refers to a workshop involving a number of researchers from the UK and an equal number of local researchers. The workshops are followed by real and virtual networking to sustain the contacts made with a view to produce a proposal for longer term externally funded collaboration or applications for further funding.

As at the Department of Astronomy of the Eötvös University we have had a tradition of national workshops of young researchers in astronomy and astrophysics with similar characteristics, the call for meeting proposals by the British Council seemed an ideal way to extend the geographical scope of our meeting, while still maintaining its general format and spirit. Thus, this year's Hungarian Young Researchers' Workshop in Astronomy and Astrophysics was held in tandem with the British–Hungarian N+N workshop, and the contributions are also presented together in this volume.

For the topic of this year's meeting we chose "Computer processing and use of satellite data in astronomy and astrophysics". The reason for this is that, thanks to a high number of space probes, in the past decades a vast amount data has been collected from the extraterrestrial world, from the magnetosphere to the most distant galaxies and beyond. We are now in a situation where the amount data grows much faster than the speed by which they can be processed and duly analyzed. The workshop was devoted to methods aimed at improving on this situation, as well as to scientific results born out of the use of space data. A web page was set up for the meeting at the address `http://astro.elte.hu/nandn`

The workshop was open to post-doctoral scientists and engineers and those tenured for five years or less. PhD students in an advanced phase of their project were also admitted. The number of participants from the U.K. was nine, while a total of 13 applications were accepted from Hungary for the N+N workshop (with many more participants at the adjoining national workshop). It is to be noted that from the nine UK participants, only four were British citizens, the rest hailing from Armenia, Germany, Hungary, and Italy. Nevertheless, all

have been residing in the UK for years, and this composition reflects well the international character and openness of current British academic research.

The workshop took place from the 5th to the 7th of February at the Lágymányos campus of Eötvös University (Northern Block, Lecture Hall 1.71). The technical background was provided by the host institution, the Department of Astronomy of Eötvös University. The event was opened by J. McGrath from the British Council Hungary.

In view of the the British Council's original purpose with these workshops, special attention was paid to providing ample opportunity for informal discussions and get-togethers, not only to discuss scientific problems but also to forge a lasting link between participants. Besides the formal, and sometimes quite heated, discussions after each talk, there was plenty of opportunity to talk science in the coffee breaks and in lunchtime. The opening and closing receptions and the conference banquet gave further opportunities for discussions.

A half-day excursion was also an integral part of the programme. The trip consisted of a professional part (a visit to the Institute of Satellite Geodesy at Penc, near Budapest) and of a more general cultural part (visit to Grassalkovich Palace in neearby Gödöllő). The excursion was very successful in getting the participants together as a group in a relaxed atmosphere.

Acknowledgment

Finally, the co-directors would like to express their thanks to The British Council for the initiation of the project and for its generous support. In particular, we are grateful to the staff of British Council Hungary, Éva Salamon, Paul Clementson and Jim McGrath for their enthusiastic moral support and for their helpfulness in the management of the event.

Budapest, 11 May 2004

Kristóf Petrovay
Co-director of the meeting

Part One

Space physics and particle detection

PAST AND FUTURE OF SCIENTIFIC SPACE MISSIONS: AN OVERVIEW

K. Petrovay
Eötvös University, Department of Astronomy
H-1518 Budapest, P.O.Box 32., Hungary
E-mail: K.Petrovay@astro.elte.hu

Abstract

After briefly mentioning a few highlights of recent, current and future scientific space missions, an extensive list of all the most important such missions, together with web links, is given.
KEYWORDS: *Sun, MHD, plasma physics*

1. Introduction

The opening of the space age in 1957 led to the emergence of several new disciplines of science: planetology, extraterrestrial astronomy and space physics. Solar physics, while also existing before, has also entered a qualitatively new phase, thanks to its strong links with space science. The hundreds or dozens of space missions with some science payload in the past decades have produced an enormous and exponentially increasing quantity of data. We have now reached a point where data are being collected faster than they can be analyzed. Therefore, efficient methods of data processing, storage and access are of crucial importance for the further development of space research.

The sheer number of missions has made it very hard for most of us to gain a general outlook on the actual situation and recent developments in areas other than our own narrow field of research. The aim of this paper is to try to remedy this situation somewhat by providing such an overview. The most important part of the paper is an extensive appendix, listing the names and web pages of all the most important scientific space missions. An HTML version of this list can be found at the web page of the present meeting.

In the following sections I will briefly mention a few highlights of the missions listed in the Appendix, according to my own subjective choice.

2. Planetary missions

Planetary exploration started with the Soviet Luna-1 probe in 1959. The dozens of missions launched since then can be crudely divided in two "generations",

the dividing line being the year 1990 or so. The SDI ("space war") initiative of the US government arguably played a role in the development of the new technologies employed in second-generation missions: one of the first such probe, the Clementine lunar probe was in fact a Pentagon project, using revolutionary technologies.

The largest number of missions in the past decade were, of course, aimed at the exploration of Mars. Unfortunately, their failure rate has been around 50 %, two especially lamentable losses being Mars-96 with many Hungarian-built instruments on board, and recently the British-built Beagle 2. But the spectacular success of Mars Express and the Mars Exploration Rovers has given us some consolation.

The two "hottest" topics in planetary research nowadays seem to be sample retrieval and ion propulsion. An especially ambitious mission combining these two novelties is the Japan's Hayabusa, launched in 2003, set to land on asteroid Itokawa in 2005 and then return a sample to Earth by 2007.

3. Astronomy from space

By getting above the atmosphere, astonomers were able to open several new windows on the Universe at different wavelength domains, such as infrared, ultraviolet, X-ray and gamma ray. An extensive list of space missions operating in these domains is given in the Appendix. Equally important has been, however, the possibility of making observations unimpeded by atmospheric seeing in the optical domain. The HIPPARCOS mission and the Hubble Space Telescope are the two most celebrated missions of this kind.

In addition to avoiding the problems caused by the atmosphere, space astronomy also offers another advantage: the possibility of performing interferometric measurements on unprecedented, cosmic scales. The space VLBI projects described by S. Frey in this volume are one example of this application. Another exciting new prospect is the ultimate detection of gravitational waves. The LISA project, essentially consisting of a cosmic-scale Michelson interferometer and to be launched in 2012, aims at this.

4. Space physics and solar science

Space science was born in 1958 when the American satellite Explorer-1 discovered van Allen's radiation belts around the Earth. In the same year Eugene Parker predicted the existence of the solar wind. The prediction was confirmed

by the first space vehicle to leave the Erth's magnetosphere, Mariner-2, in 1961. In 2003 Parker was awarded the Kyoto Price of Physics for his discovery.

The interplanetary mission that has reached the farthest parts of the solar system, Voyager-1, created some excitement in the past year by detecting slower than usual solar wind in late 2002. This raised the possibility that the probe, now at a distance of 90 astronomical units from the Sun, may have crossed the termination shock of the solar wind where it slows to subsonic speeds. However, six months later the detected wind conditions were back to normal, indicating that the event was a local anomaly. The consensus now is that no crossing has occurred yet, but it is expected soon.

In the field of solar physics, beside the celebrated SOHO and TRACE missions, the most exciting novelty is RHESSI, the first hard X-ray and gamma imager of the Sun, operative since early 2002. It has already made the most convincing case yet for magnetic reconnection taking place on the Sun, but it also led to some surprises. One of these is the apparent displacement of gamma-ray emitting regions in flares compared to sof X-ray emitting ones.

Among new and upcoming missions of interest one may mention Genesis, currently collecting solar wind particles at the L1 point of the Sun–Earth system, to return them to Earth in September 2004. NASA's Solar Probe, with a preliminary launch date of 2007, aims at sampling the solar corona from a much closer distance of only four solar radii.

5. Appendix: List of missions

5.1. Planetary missions

Links to mission homepages:

NASA JPL
`http://www.jpl.nasa.gov/missions/`
`http://astrogeology.usgs.gov/Missions/`
ESA
`http://www.sci.esa.int/science-e/www/area/index.cfm?fareaid=1`
IKI
`http://www.iki.rssi.ru/eng/`
`http://www.russianspaceweb.com/spacecraft.html`
ISAS
`http://www.isas.ac.jp/e/enterp/missions/catalogue.shtml`

First generation:

Moon:
Ranger 1-9 (1961-65, NASA): impact
Cosmos-159 (1967, IKI): orbiter
Lunar Orbiter 1-5 (1966-67, NASA), Surveyor 1-7 (1966-68, NASA)
Luna 1-24 (1959-76, IKI): soft landings, orbiters, rovers
Apollo 7-17 (1967-73, NASA): human
First interplanetary missions; outer planets:
Pioneer 3-11 (1962-79, NASA); Zond 4-8 (1968-70, IKI)
Inner planets:
Mariner 1-10 (1962-75, NASA), Venera 2-16 (1965-1983, IKI)
Mars 2-7 (1971-73, IKI), Viking 1-2 (1975-77)
Outer planets
Voyager 1-2 (1977- , NASA)

Second generation:

Moon:
Clementine (1994, US Defense)

Venus:
Magellan (1989-94, NASA)

Mars:
Mars Observer (1992-93, NASA), lost
Mars-96 (1996, IKI+KFKI), lost
Mars Global Surveyor + Pathfinder (1996- , NASA)
Mars Climate Orbiter (1998, NASA), lost
Nozomi (1998, ISAS), lost
Mars Polar Lander (1999, NASA), lost
2001 Mars Odyssey (2001- , NASA)

"Mars Attack":
ESA: Mars Express + Beagle 2
`http://www.rssd.esa.int/MarsExpress/`
NASA: Mars Exploration Rovers Mission:
`http://marsrovers.jpl.nasa.gov/newsroom/`

Outer SS:
Galileo (1989-2003, NASA): Venus, asteroids (Gaspra, Ida), Jupiter
Cassini+Huygens (1997- NASA+ESA+KFKI): Saturn, Titan
Hayabusa (2003-, ISAS): to retrieve asteroid sample.

Comet flybys:
ICE (1978-85, NASA), Giotto (1985-92, ESA),
Sakigake+Suisei (1985-86, ISAS)
VEGA-1,2 (1984-86, Intercosmos incl. KFKI),
Stardust (1999- , NASA): to retrieve sample in 2006!
Rosetta (Feb 2004-, ESA+KFKI): to 67P/Churyumov-Gerasimenko

5.2. Astronomy from space

Advantages:

No atmosphere ⇒ new spectral domains

Gamma:
Vela (1969-79, US Defense)
http://heasarc.gsfc.nasa.gov/docs/vela5b/vela5b.html
COS-B (1975, ESA+NASA+UK)
http://www.rssd.esa.int/index.php?project=COSB&page=index
GRANAT (1989-98, IKI)
http://heasarc.gsfc.nasa.gov/docs/granat/granat.html
Compton GRO (1991-2000, NASA)
http://cossc.gsfc.nasa.gov/
INTEGRAL (2002-, ESA)
http://astro.estec.esa.nl/SA-general/Projects/Integral/

X-ray:
SAS 1-3 (incl. Uhuru) (1970-79, NASA)
http://heasarc.gsfc.nasa.gov/docs/uhuru/uhuru.html
ANS (1974-77, Holland)
http://heasarc.gsfc.nasa.gov/docs/ans/ans.html
Ariel-V (1974-80, UK+NASA)
http://heasarc.gsfc.nasa.gov/docs/ariel5/ariel5.html
HEAO 1-3 (incl. Einstein) (1977-81, NASA)
http://heasarc.gsfc.nasa.gov/docs/einstein/heao2.html
EXOSAT (1983-86, ESA)
http://www.rssd.esa.int/SA-general/Projects/Exosat/
ROSAT (1990-99, Germany+UK+NASA)
http://wave.xray.mpe.mpg.de/rosat
Hakucho (Corsa-B, ISAS)
http://heasarc.gsfc.nasa.gov/docs/hakucho/hakucho.html
Astro B-D (incl. GINGA and ASCA) (1983-2000, ISAS+NASA)
http://heasarc.gsfc.nasa.gov/docs/asca/ascagof.html

RXTE (1995- , NASA)
http://heasarc.gsfc.nasa.gov/docs/xte/xte_1st.html
BeppoSAX (1996-2002, Italy+Holland)
http://heasarc.gsfc.nasa.gov/docs/sax/saxgof.html
XMM-Newton (1999- , ESA)
http://xmm.vilspa.esa.es/
Chandra (1999- , NASA)
http://chandra.harvard.edu/

Ultraviolet:
OAO-2 (1968-71, NASA)
http://www.sal.wisc.edu/~meade/OAO/
Copernicus =OAO-3 (1972-81, NASA+UK)
http://heasarc.gsfc.nasa.gov/docs/copernicus/copernicus.html
IUE (1978-96, ESA)
http://www.sci.esa.int/science-e/www/
area/index.cfm?fareaid=22
GALEX (2003- , NASA)
http://www.galex.caltech.edu/

Infrared:
IRAS (1983, NASA+Holland+UK)
http://irsa.ipac.caltech.edu/IRASdocs/iras.html
ISO (1995-98, ESA)
http://www.iso.vilspa.esa.es/
Spitzer (SIRTF) (2003- , NASA)
http://sirtf.caltech.edu/

Microwaves:
COBE (1989-93, NASA)
http://aether.lbl.gov/www/projects/cobe/
WMAP (2001-, NASA)
http://map.gsfc.nasa.gov/

No atmosphere ⇒ no seeing in optical:

Hipparcos (1989-93, ESA)
http://sci.esa.int/science-e/www/area/index.cfm?fareaid=20
HST (1990-, NASA)
http://hubble.nasa.gov/

Large scales ⇒ interferometry

⇒ Radio: Space VLBI: HALCA (1997- , ISAS)
http://www.vsop.isas.ac.jp/
⇒ Gravity waves:
LISA Laser Interferometer Space Antenna (2012?, ESA)
2 satellites in solar orbit. Gravitational wave detection
http://www.sci.esa.int/science-e/www/area/index.cfm?fareaid=27

5.3. Space physics

Ulysses (Oct 1990 - , NASA+ESA): Fast solar wind at high ecl. latitudes
http://ulysses.jpl.nasa.gov/

Geotail (Jul 1992 - , ISAS + NASA)
http://www.isas.ac.jp/e/enterp/missions/geotail/index.shtml

GGS (Global Geospace Science) Program by NASA:
WIND (Nov 1994 - , NASA) Around L1 point.
http://www-spof.gsfc.nasa.gov/istp/wind/
POLAR (Feb 1996 -)
http://www-spof.gsfc.nasa.gov/istp/polar/
http://science.nasa.gov/uvi/LatestImage.htm

INTERBALL (1995-2000, IKI+...,Hungary,UK)
2 probes, Auroral and Tail
http://www.iki.rssi.ru/interball/

ACE Advanced Composition Explorer (Aug 1997 - , NASA) In L1 point.
`http://www.srl.caltech.edu/ACE/`

IMAGE (March 2000 - , NASA): Duskside plasma tail discovered.
`http://image.gsfc.nasa.gov/`

CLUSTER (Jul 2000 - , ESA)
`http://sci.esa.int/science-e/www/area/index.cfm?fareaid=8`

Genesis (Aug 2001- , NASA)
`http://genesismission.jpl.nasa.gov/`

Towards the heliopause:
Pioneer 10 (March 1972 - Jan 2003!)
Will reach Aldebaran in 2 million yrs.
Pioneer 11 (April 1973 - Sep 1995)
Will reach Aquila in 4 million yrs.
`http://spaceprojects.arc.nasa.gov/Space_Projects/pioneer/PNhome.html`
Voyager 1 (Sep 1977 -) 90 AU from Sun; most distant.
Will reach Camelopardalis in 40 000 yrs.
Voyager 2 (Aug 1977 -) Now at 72 AU.
Will reach Sirius in 300 000 yrs.
`http://voyager.jpl.nasa.gov/`

Plans:

Double Star (Dec 2003 / Jul 2004, China + ESA)
Equatorial and polar satellites. CLUSTER followup.
`http://sci.esa.int/science-e/www/area/index.cfm?fareaid=70`

THEMIS (Jul 2006?, NASA): Study of magnetic storms
`http://sprg.ssl.berkeley.edu/themis/`

TWINS (NASA + ...): Magnetosphere in stereo
`http://nis-www.lanl.gov/nis-projects/twins/`

5.4. Solar science

Yohkoh (Aug 1991 - Dec 2001)
http://isass1.solar.isas.ac.jp/

SOHO (Dec 1995 - , ESA)
http://sohowww.nascom.nasa.gov/

TRACE (1998 -)
http://vestige.lmsal.com/TRACE/

AcrimSat (1999- , NASA) to measure solar constant
ACRIM I aboard SMM, 1980-89 ;
ACRIM II aboard UARS Upper Atm. Research Sat. (1991-)
http://acrim.jpl.nasa.gov/

RHESSI (Feb 2002 -)
http://hesperia.gsfc.nasa.gov/hessi/

Plans:

Solar-B (Sep 2005?, ISAS + PPARC + NASA)
British Instrument: EIS (EUV Imaging Spectrograph).
Also: Solar Optical Telescope (SOT, 50 cm)
Solar X-Ray Telescope (XRT)
http://science.nasa.gov/ssl/pad/solar/solar-b.stm

STEREO (Nov 2005?, NASA SEC): Solar physics in stereo
http://stp.gsfc.nasa.gov/missions/stereo/stereo.htm

Solar Probe (Feb 2007?, NASA): In situ solar physics!
http://umbra.nascom.nasa.gov/spd/solar_probe.html

SDO (April 2008?, NASA): Solar Dynamics Observatory
http://sdo.gsfc.nasa.gov/

Solar Orbiter (Jan 2011, ESA): The Sun from 45 solar radii
http://sci.esa.int/science-e/www/area/index.cfm?fareaid=45

SPACECRAFT MEASUREMENTS OF SOLAR ENERGETIC PARTICLES

S. Dalla
Physics Dept., University of Manchester Institute of Science and Technology
P.O. Box 88, Manchester M60 1QD
E-mail: s.dalla@umist.ac.uk

Abstract

Solar energetic particles (SEPs) are accelerated in the solar atmosphere during flares and coronal mass ejections. They can escape from the Sun and propagate through interplanetary space, where they can be measured by particle detectors on board spacecraft. Key to understanding the acceleration and propagation of these particles are: 1) combined analysis of SEP data and solar flare/CME observations and 2) easy comparison of SEP data from different spacecraft/instruments. This paper will demonstrate the importance of the multi-spacecraft and multi-instrument approach by showing new results from the analysis of SEP data from Ulysses and Wind.

KEYWORDS: *Solar Energetic Particles, flares, Coronal Mass Ejections*

1. Introduction

Sudden enhancements in the fluxes of energetic ions and electrons in interplanetary space have been observed by spacecraft for decades. The short time delay between many of these events and flares - short bursts of electromagnetic radiation released from the solar atmosphere - was identified very early on and they were termed Solar Energetic Particle (SEP) events. To clarify the meaning of 'energetic' in the above definition, proton energies typically fall in the range ~50 KeV – ~30 MeV.

During the late seventies, the development of white light coronographs led to the discovery of a second type of solar energy release event, in addition to flares: Coronal Mass Ejections (CMEs). These are ejections of coronal material seen in sequences of white light images. CME material can also be detected in situ by spacecraft in interplanetary space, and in some cases a collisionless shock in front of the ejecta is detected.

Prior to the discovery of CMEs, solar energetic particles were thought to be accelerated during the flare process. In the 1980s and 1990s, however, a new paradigm of interpretation of SEP measurements emerged, according to which

two distinct classes of particle events exist: so-called *impulsive* events, of flare origin, and *gradual* events, thought to result from particle acceleration at the shock driven by CMEs (Reames, 1999).

The new paradigm of interpretation emerged from a large body of data, mostly collected by spacecraft located close to the ecliptic plane. Between 2000 and 2002 however, the Ulysses spacecraft was able to detect solar particles at an entirely different location in space: at large heliographic latitudes, i.e. over the solar poles. In this paper the Ulysses SEP observations will be reviewed and the two-class paradigm for interpretation of SEP measurements will be tested against these new data.

2. Geometry of SEP detection

SEPs accelerated in the solar atmosphere can travel through space and be detected by instrumentation on board spacecraft, typically located near the Earth, i.e. at 1 Astronomical Unit (AU) from the accceleration region. Figure 1 describes the average structure of the interplanetary magnetic field, through which the particles propagate from the Sun to the detecting spacecraft. Here the Sun is in the center of the spiral plot. The solar wind is a continous flow of material out of the corona, and it carries magnetic field into interplanetary space. Because of the Sun's rotation, the magnetic field lines are twisted into a spiral shape, called Parker spiral. In Figure 1 the Parker spiral magnetic field lines are plotted at heliolatitudes 0°, 40° and 80°, up to a distance from the Sun of 1.3 AU. For a spacecraft located at 1 AU from the Sun and in the ecliptic, the direction of the magnetic field is at ~45° to the spacecraft-Sun line. The smooth magnetic field lines of Figure 1 are a representation of the average magnetic field. Many years of spacecraft observations have shown that the interplanetary magnetic field is in fact turbulent, and this influences energetic particle propagation.

The time profile of particle flux measured by a spacecraft in interplanetary space will be the result of (1) the injection profile of energetic particles at the Sun and (2) their propagation between the Sun and spacecraft. The key question in the interpretation of SEP data is how to disentangle the contribution of these two processes to the measured energetic particle profiles.

In the current paradigm of interpretation, particles in gradual events are thought to be accelerated at the shock driven by a CME through the corona and interplanetary space (Reames, 1999). As the shock travels outwards from the Sun, particles continue to be accelerated. Once they escape from the shock region, they travel through interplanetary space and can reach the detecting

Figure 1: Schematic of Parker spiral in 3-D. The field lines are drawn for heliolatitudes 0°, 40° and 80°, up to a distance from the Sun of 1.3 AU.

spacecraft. In this model, particle propagation is assumed to take place mostly in the direction parallel to the magnetic field, with negligible scatter of guiding centers in the direction perpendicular to the magnetic field. As a result, it is thought that a spacecraft will be able to detect SEPs only when it is directly connected to the acceleration region (a travelling shock front) via the interplanetary magnetic field.

3. Ulysses high heliolatitude observations

The Ulysses spacecraft, launched in 1990, is in a ~6 year orbit around the Sun, at an inclination of 80° with respect to the ecliptic plane, i.e. it passes over the solar poles (Figure 2). The COSPIN suite of particle instruments detected several solar energetic particle events during Ulysses' second orbit passage over the poles, at a time of high solar activity.

Large SEP fluxes were detected at Ulysses even when it was located at the highest latitudes (McKibben *et al*, 2003). This was a surprising finding, because solar flares and CMEs typically take place at low to mid heliolatitudes: how did the energetic particles reach high heliolatitudes?

To address this question, studies of several properties of the time profiles at Ulysses were carried out for 9 large SEP events (Dalla *et al*, 2003a; Dalla *et al*, 2003b). Features of high latitude data from the COSPIN instruments were compared with those of data recorded at low latitudes during the same SEP events, by the Wind/3DP and the SOHO/COSTEP instruments.

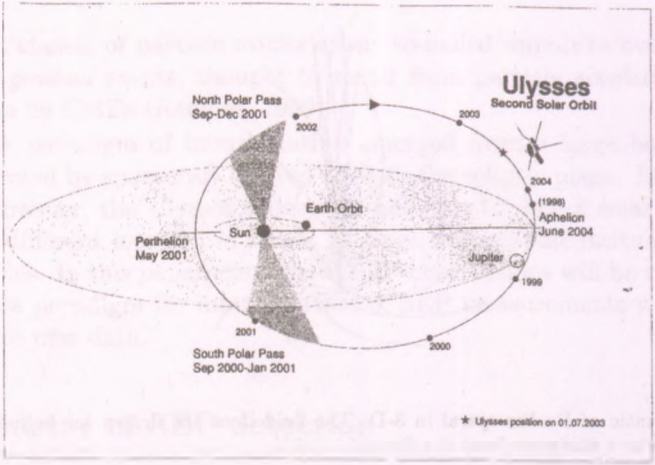

Figure 2: Plot of the second solar orbit of the Ulysses spacecraft (Source: Ulysses ESA site http://helio.esa.int/ulysses). Ulysses is in a highly inclined orbit with respect to the ecliptic plane, and passed over the solar poles at a time of high solar activity during 2000–2002.

3.1. Onset times

Following a flare or CME on the Sun, faster SEPs are expected to reach a spacecraft in interplanetary space before slower ones, assuming that both acceleration and injection into interplanetary space occur simultaneously at all energies. A way to verify this is to measure the time of onset of a flux enhancement in several energy channels, and plot this time versus the inverse speed associated to the channel.

Studies of SEP electron onsets measured by the Earth's satellite Wind have shown that in most cases the onset times versus c/v data points lie on a straight line, with the faster electrons arriving first (Krucker *et al*, 1999). Here c is the speed of light and v the particle speed. If no scattering took place between the Sun and spacecraft, a fit to the plots would allow one to obtain the time t_{Sun} at which particles were released at the Sun, and the distance they travelled prior to reaching the spacecraft (path length, L). In the presence of scattering, if the data points can still be fitted by a straight line, the fitting parameters t_{Sun} and L do not have such a straightforward interpretation. It is expected that strong scattering would result in a non linear plot of onset times versus c/v.

A very important consideration when studying particle onset times, is the choice of the reference time t_0 of the solar event, with respect to which the onset

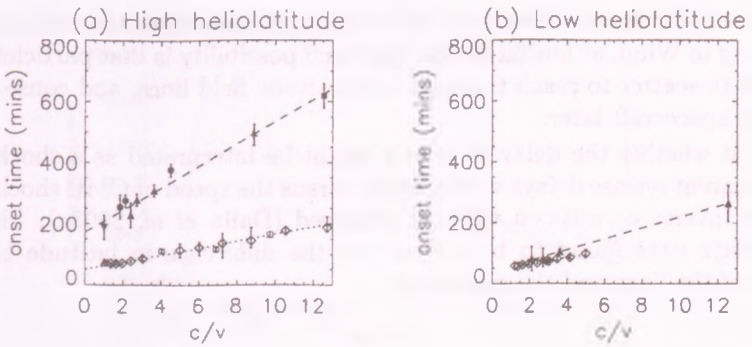

Figure 3: Plots of onset times versus c/v at Ulysses (filled symbols) and at Wind and SOHO (empty symbols). Plot (a) is for an event taking place when Ulysses was at high heliolatitudes, and (b) for one when Ulysses was close to the ecliptic. Wind and SOHO are always located near the ecliptic plane.

time is calculated. Electromagnetic emission associated to flares can start at different times in different wavelength ranges. It is generally agreed that the best time indicator to be used for energetic particle analysis is the time of a radio emission called 'type III', thought to be generated by electrons streaming out of the corona into interplanetary space.

Using type III onsets as the reference, it has been shown that the release times t_{Sun} obtained from fitting of onset times plots are in some cases coincident with the type III onset times, and in other cases show a delay up to about 30 minutes (Krucker *et al*, 1999).

Figure 3a shows a plot of onset times at Ulysses at high heliolatitudes (filled symbols) and at Wind in the ecliptic (open symbols) versus c/v. One can see that onsets at Ulysses are considerably delayed compared to in-ecliptic ones. This is the case for all large SEP events taking place at times when Ulysses was at high heliolatitudes (Dalla *et al*, 2003a). Instead, for an SEP event taking place at a time when Ulysses was at low heliolatitudes, no large delay is observed, as shown in Figure 3b. It should be noted that for the event of Figure 3b Ulysses and Wind were separated in longitude by 77°.

What is causing the delay in onset in the high heliolatitude events? One possibility is that particles might have been released onto the high heliolatitude interplanatary field lines later than onto the near ecliptic ones. Within the CME acceleration model this could result from the accelerating shock taking

longer to reach magnetic field lines connecting to Ulysses than to reach field lines connecting to Wind, at low latitudes. A second possibility is that particles might have had to scatter to reach the high heliolatitude field lines, and consequently reach the spacecraft later.

To test whether the delay in release might be interpreted as a shock travel time, apparent release delays were plotted versus the speed of CME shocks. The expected inverse correlation was not observed (Dalla *et al*, 2003a). However onset delays were found to be ordered by the difference in latitude between location of the flare and the spacecraft.

3.2. Times to maximum

A second feature of SEP time profiles which was analysed in Ulysses data is the time to peak intensity, defined as the interval between the time at which intensities peak in a given channel, and the type III emission onset time. What was found is that in general intensities peaked at Ulysses at high latitudes much later than in the ecliptic (Dalla *et al*, 2003b).

The interpretation given to times to maximum by the CME shock acceleration model is as follows. A shock front is assumed to have different acceleration efficiencies in different regions: the nose of the shock, i.e. its central region, is typically the one with highest compression ratio and is therefore assumed to be the most efficient particle accelerator. On the contrary, the flanks of a shock are assumed to accelerate particles less efficiently. As a consequence it is thought that the time of maximum intensity is an an indication of the time at which the detecting spacecraft is magnetically connected to the portion of the shock which most efficiently accelerates particles.

From the model and the Ulysses observations, one would conclude that Ulysses must have established a magnetic connection with an acceleration efficient portion of the CME driven shock very late on, after a day or more. This is unlikely for the following reason: a spacecraft at high latitudes would most likely be connected to the flanks of a CME shock, as the source regions of CMEs and flares are typically in the mid to low heliolatitude band. As the shock propagated outward the connection would remain to the flanks as the shock propagates radially outward. This is unlike in the ecliptic, where often a spacecraft's connection point moves along the shock from the flanks to the nose or viceversa.

Hence the large delays in time to maximum observed at high heliolatitudes are not well explained by the CME shock acceleration model.

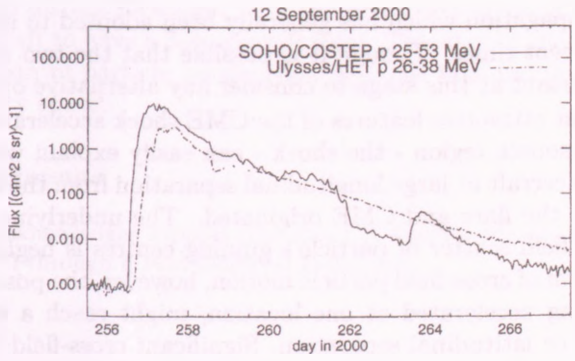

Figure 4: Time profile of an SEP event detected at Ulysses at heliolatitude 71°S and near the ecliptic by SOHO/COSTEP.

3.3. Decay phases

After intensities peak, SEP fluxes typically display a long duration quasi-exponential decay. Comparison of SEP profiles measured at high latitudes and in the ecliptic show that the decay time constant is similar at the two locations, despite the large separation between them. This can be seen in Figure 4, showing SEP flux profiles at Ulysses and near the ecliptic.

At present, long duration decays are interpreted as resulting from particle trapping by the CME shock, once it has swept past the detecting spacecraft. However, in 6 out of the 9 SEP events analysed, no interplanetary shock was observed at Ulysses. One such event is the one shown in Figure 4.

4. Discussion

In the above sections, a comparison of various features of SEP time profiles at spacecraft located far apart was presented. Data from the COSPIN experiment on Ulysses, taken at times when Ulysses was at high heliographic latitudes, were compared with those from the Wind/3DP and SOHO/COSTEP instruments, located near the ecliptic plane.

Solar energetic particles were found to reach Ulysses at high heliolatitudes with a delay of ~100 minutes compared with their arrival time to the near-ecliptic spacecraft. Times to maximum were also delayed at high heliolatitudes.

These observations pose a challenge to the model of energetic particle ac-

celeration and propagation which has generally been adopted to interpret SEP observations in recent times. Though it is possible that the two might be reconciled, it is important at this stage to consider any alternative options.

One of the most attractive features of the CME shock acceleration model, is that an extended source region - the shock - can easily explain why SEPs can be detected by spacecraft at large longitudinal separation from the active region on the Sun where the flare and CME originated. The underlying assumption here is that cross-field scatter of particle's guiding centres is negligible. If one allows for some form of cross-field particle motion, however, the possibility exists that particles being accelerated at one location, might reach a spacecraft at large longitudinal or latitudinal separation. Significant cross-field diffusion has recently been invoked to explain features of the type III burst emission (Cane *et al*, 2003), and galactic cosmic ray (McKibben *et al*, 2003) and jovian electron data (Heber *et al*, 2002).

If cross-field diffusion were confirmed as an important feature of energetic particle propagation in the heliosphere, the two-class model of acceleration at the Sun might have to be revised. In fact, cross-field diffusion would allow for particles of so called gradual events to be of flare origin, rather than be shock accelerated. This would have profound consequences on the interpretation of SEP data and on our understanding of acceleration processes at the Sun.

5. Conclusions

The analysis presented in this paper demonstrates the importance of multi-spacecraft, multi-instrument studies in the interpretation of solar energetic particle data. At present this can be complicated, due to incomplete data sets being made publicly available, a variety of data formats being used, and generally a need to understand the operation of instrumentation in great detail in order to make sensible use of the provided data products.

Another desirable development in future years will be the creation of a better interface between solar flare and coronal mass ejection observations, and energetic particle data.

Several current and future space missions will provide improved observations of the energy release events during which particles are accelerated at the Sun. The RHESSI mission is currently operating and taking high resolution images and spectra of flares in the photon range from 3 KeV and 17 MeV. The STEREO mission, scheduled to be launched at the end of 2005, will provide 3-dimensional imaging of CMEs. Its particle detectors will measure SEPs over a wide energy

range, at two distant locations, with the possibility of using detectors from the ACE spacecraft to have a 3-spacecraft view. This should allow new insights into the mechanism of particle acceleration near the Sun and propagation through interplanetary space.

Acknowledgments

The author acknowledges support from the UK Particle Physics and Astronomy Research Council through a Post-Doctoral Fellowship.

References

Cane H.V., et al., 2003, *J. Geophys. Res.* 108, 1203, doi: 10.1029/2002JA009488

Dalla S., et al., 2003a, *Annales Geophys.* 21, 1367

Dalla S., et al., 2003b, *Geophys. Res. Lett.*, 30, 8035, doi:10.1029/2003GL017139

Heber B., et al., 2002, *Astrophys. J.* 579, 888

Krucker S., et al., 1999, *Astrophys. J.* 519, 864

McKibben R.B., et al., 2003, *Annales Geophys.* 21, 1217

Reames D.V., 1999, *Space Science Rev.* 90, 413

Low-Energy Particle Studies Near the Earth Using Data of Cluster RAPID Measurements

G. Facskó
KFKI Research Institute for Particle and Nuclear Physics
H-1525 Budapest, P.O.Box 49
E-mail: gfacsko@rmki.kfki.hu

Abstract

The Cluster mission has been designed to perform simultaneous 3D plasma, magnetic field and energetic particles measurements in the Earth's plasma environment. The advanced energetic charged particle spectrometer RAPID on board the four spacecraft provide a complete coverage of energy, mass, and angular distribution of ions and electrons, over the energy range from 30 keV to 1.5 MeV. By combining with other Cluster measurements, primarily FGM and CIS, small-scale features can be observed and examined with unprecedented detail in the magnetosphere of the Earth and near the bow shock. In this paper I present a data processing method together with the observations of a hot flow anomaly (HFA) event as an illustrative example.

Keywords: *Cluster, RAPID, magnetosphere, bow shock, solar wind, hot flow anomaly*

1. Introduction

ESA's cornerstone mission Cluster provides excellent conditions to study the plasma environment of the Earth. The very interesting phenomena include diamagnetic cavities, so called hot flow anomalies, which develop when an interplanetary tangential discontinuity intersects the bow shock of our planet. We have to solve special problems when analyzing the measurements of Cluster because the satellites and their observations are made in a frames moving relative to the frame of the plasma. First I describe the Cluster spacecraft and the RAPID instrument, and then a simple transformation method is presented. Finally I show the benefits of the method in analyzing a HFA event.

2. The Cluster Mission to the Magnetosphere

The four Cluster satellites were launched on July 16, 2000 by two Russian booster rockets. The main purpose of the mission is investigating of the small-scale structures of the Earth's plasma environment in three dimensions. Their

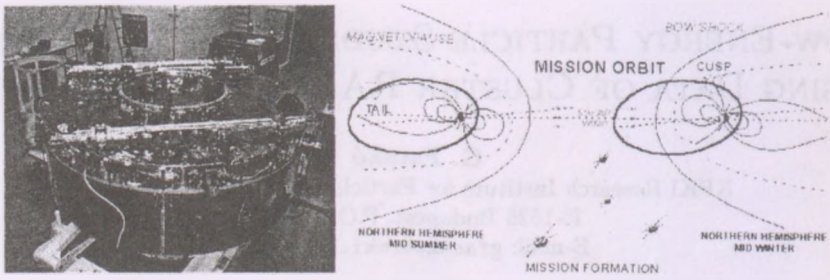

Figure 1: The CLUSTER Mission. *Left:* One of the four Cluster satellites. *Right:* The orbit of the fleet.

elongated elliptical orbits were chosen accordingly: the orbit has high inclination (close to a polar orbit) with a perigee of 19000 km and an apogee of 119000 km (*Fig. 1*). These orbits covers very interesting regions of the Earth's environment including the bow shock, the magnetosheath, the polar cusp regions, the magnetosphere, and the magnetotail as well (*Fig. 1*). The observations have seasons depending on the position of the Sun, the Earth and the satellites: the best season for a bow shock, solar wind, magnetopause survey is Northern Hemisphere Winter, on the other hand we can best study the magnetotail on Northern Hemisphere Summer because the orbit of the fleet leads across these regions that time.

The four nearly identical satellites form a tetrahedron in space which has a size of about 100 km to 10000 km, varying over the mission. The satellites are cylinder-shaped with two booms and are equipped with identical payloads of eleven instruments. The instruments are located on the circle of the satellites as well as at the ends of the booms. Including the FGM, CIS and RAPID, the detectors were developed and built in collaboration of many universities and institutes worldwide (Table 1). The KFKI RMKI took part developing, testing and building the FGM and RAPID. The satellites are stabilized by rotation, their rotational period is 4 sec (Escoubet et al., 1997).

3. The RAPID Instrument

One RAPID instrument consists of two sub-instruments and a common data processing unit (*Fig. 2*). One of them detects electrons whereas the other observes ions including protons, helium nuclei, and heavier particles. The two parts have different apertures, closed until the satellites reached the orbit. The

data processing unit provides two modes: the first mode has 32^s time resolution (but excellent angular resolution), the second mode has 4^s time resolution. Both instruments consist of three detector heads; each head built from three (electrons) or four (ions) detectors dividing the polar angle range into 9/12 sectors. The processing unit reads the state of detectors every 0.25 seconds, saves it and after turning around adds the newly read state. After eight cycles the DPU records the detected particle fluxes in all energy channels. This method provides 192 "pixels" (ions) or 14 (electrons). Each detector is made of two thin flat silicon pieces. Particles, which have sufficient energy, flight across the first and reach the second. The DPU calculates the impulse, the spread and the charge of each particle from the current pulses triggered (*Fig. 2*). The second mode of the DPU collect only the last mentioned data in eight energy channels.

3.1. A Bow Shock Crossing Event on February 10, 2001

On *Fig. 3* a bow shock crossing event is illustrated. The absolute value of magnetic field is about 25 nT before *line A* on the second and the fourth panels. Between lines *A* and *B* the value decreases suddenly to 5 nT. The first value is close to the typical value of B in the magnetosheath, the latter is typical

FGM	**F**luxgate **M**agnetometer
RAPID	**R**esearch with **A**daptive **P**article **I**maging **D**etectors
CIS	**C**luster **I**on **S**pectrometry experiment
EDI	**E**lectron **D**rift **I**nstrument
ASPOC	**A**ctive **S**pacecraft **P**otential **C**ontrol experiment
STAFF	**S**patio-**T**emporal **A**nalysis of **F**ield **F**luctuation experiment
EFW	**E**lectric **F**ield and **W**ave experiment
DWP	**D**igital **W**ave **P**rocessing experiment
WHISPER	**W**aves of **H**igh frequency and **S**ounder for **P**robing of **E**lectron density by **R**elaxation experiment
WBD	**W**ide **B**and **D**ata instrument
PEACE	**P**lasma **E**lectron **A**nd **C**urrent **E**xperiment

Table 1: Instruments of Cluster.

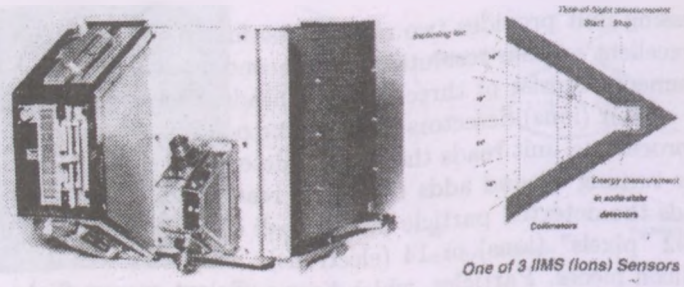

Figure 2: RAPID instrument.

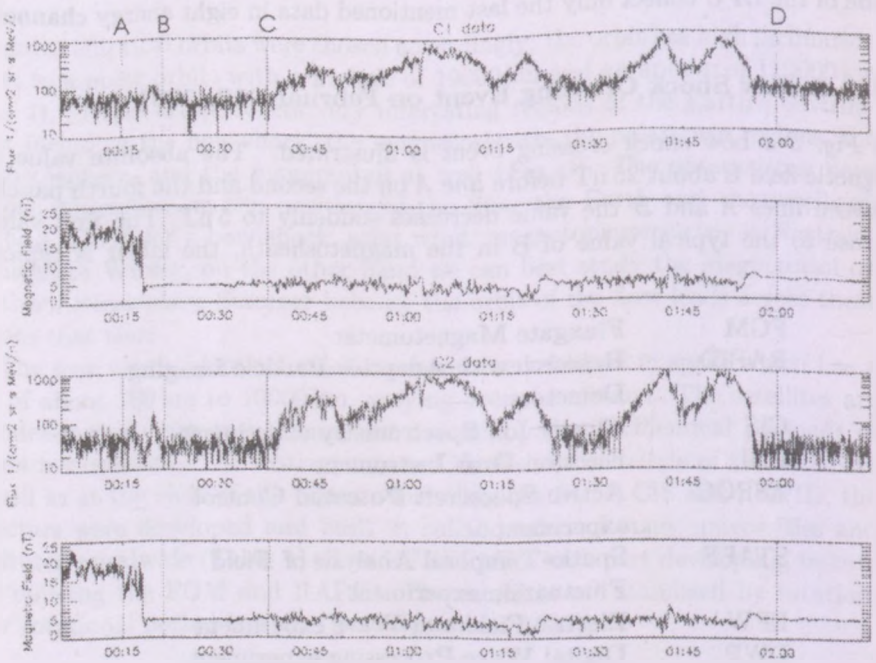

Figure 3: Bow shock crossing event on 2001 February 10. *Top panel:* The flux of 28.2–68.9 keV ions aboard Cluster-1. *Second panel:* The absolute value of magnetic field as measured by FGM. *Third and fourth panels:* Measurements of Cluster-2, see the last two panels.

upstream of the Earth's bow shock: the fleet (C1 and C2) crossed the bow

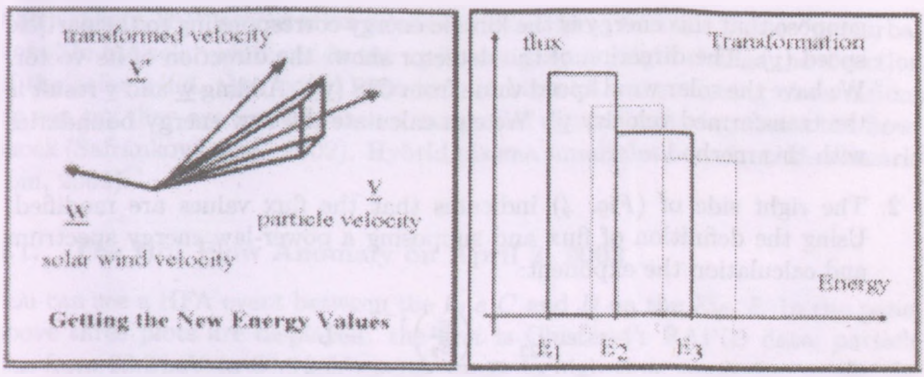

Figure 4: Transformation into the frame of the solar wind. *Left:* Calculation of the new energy channel limits. *Right:* Calculation of the transformed fluxes.

shock. There is a small peak between the *A* and *B lines* on the first and third panels after the bow shock crossing. This is the sign of particles bounced back from bow shock into the upstream region. There is an interesting thing, a small flux drop on the first and third panels between *C* and *D lines*. Disturbances are observable at same time on the second and fourth panels related to a *Solar Energetics Particle event* (SEP). The curves are similar on panels, indicating that a SEP event took place and the observed fluctuations are not due to the motion of satellites.

4. The Compton-Getting effect: Transformation into the Frame of the Solar Wind

When we observe plasma features, charges, and currents in a moving reference frame, these are different from what would be seen in the plasma frame. This is the Compton-Getting effect (Compton & Getting, 1935). Many sophisticated and complex methods have been developed to obtain the real values in the frame of reference of the plasma. Here a simple algorithm is used as introduced below.

The method has three main steps (*Fig. 4*):

1. Calculation of the new energy limit values using the solar wind velocity data as measured by CIS. Let us see the left panel of (*Fig. 4*). The RAPID detector has excellent angular distribution in its first mode. The instrument detects fluxes in the frame of satellite with energy *E*. Let us

suppose that this energy is the kinetic energy corresponding to the particle speed ($\underline{v}$). The direction of the detector shows the direction of its vector. We have the solar wind speed value from CIS ($\underline{w}$). Adding $\underline{w}$ and $\underline{v}$ result is the transformed velocity $\underline{v}$'. We can calculate the new energy boundaries with this method.

2. The right side of (*Fig. 4*) indicates that the flux values are modified. Using the definition of flux and supposing a power-law energy spectrum and calculation the exponent:

$$\frac{f_{23}}{f_{12}} = \frac{\left(\frac{E_3}{E_2}\right)^{1-\gamma} - 1}{1 - \left(\frac{E_1}{E_2}\right)^{1-\gamma}}$$

Here f_{12} and f_{23} are the fluxes of the first and second energy channels, E_1, E_2 and E_3 the boundaries of energy channels in the s/c frame, γ is the exponent what we find.

3. If we have the exponent, γ we can calculate the transformed fluxes:

$$f_{12}^{'} = f_{12} \cdot \frac{\left(E_2^{'}\right)^{1-\gamma} - \left(E_1^{'}\right)^{1-\gamma}}{(E_2)^{1-\gamma} - (E_1)^{1-\gamma}}$$

Here I use as tokens as above. The energy and flux values with comma are the transformed values.

The procedure has to be repeated for each pixel.

5. Hot Flow Anomalies

Hot flow anomalies are very interesting phenomena to demonstrate the advantages of this method. They were first observed in 1985 by AMPTE and in 1986 by ISEE satellites (Schwartz et al., 1985; see also Thomsen et al., 1986). They wanted to study the bow shock and flight on parallel orbit with the surface of the Earth's bow shock. Suddenly strong temperature increases, disturbances in the magnetic field and energized particle beams have been detected related to the passage of an interplanetary discontinuity. These hot, diamagnetic cavities so called hot flow anomalies (HFAs) have temperatures about 300000 K. They are transient perturbations of the terrestrial bow-shock with significant magnetospheric responses and ground signature. Energized particles are detectable

which means that acceleration processes must form beams. These "perturbations" in the solar wind and in the upstream plasma might change the direction of the solar wind, they could even turn it back. After analyzing observations we can say they are formed when a tangential discontinuity intersects the bow-shock (Šafránková et al., 2002). Hybrid plasma simulations confirm this scenario (Lin, 2002).

5.1. The Hot Flow Anomaly on April 2, 2002

You can see a HFA event between the *line C* and *D* on the *Fig. 5.* In the panel above three plots are displayed: the first is Cluster-4's RAPID data, particle flux from 28.2 keV to 68.9 keV energy range in high time resolution mode, the second is Cluster4's FGM magnetic field strength, and the third is the solar wind speed (CIS). The most important x (antisunward) component is assigned with V_x. The absolute value of magnetic field decreases between the *line A* and *B* on the second plot from 25 nT to 5 nT, when the satellite crossed the bow shock and leaved magnetosheath, entered the upwind region. As the position of the bow shock is determined by pressure and magnetic field of the solar wind, which change very quickly, the bow shock moves with a speed of about 10 km/s. In turn it overtakes the satellites and they intersect the bow shock again and again. One can notice this process on the second plot of *Fig. 5*'s upper panel between *line B* and *C*. Suddenly, the solar wind speed decreases close to 0 km/s at *4:30 UT* (see third plot) and the direction of magnetic field changes strongly (second plot). The particle flux as measured by RAPID increases and a peak appears.

However, there is another aspect of studying the HFA event: the excellent angular resolution mode of the RAPID solves new informations (See *Fig. 5* lower left panel). The left panel does not show correct values as I mentioned in the *Sec. 4.*, (Compton & Getting, 1935). The lower right panel shows the transformed values. It indicates that the HFA has beam like behavior and shows the direction of the beam. The empty belt in the middle is the consequence of a failure of the middle head of RAPID.

6. Summarize

After an introduction to the Cluster project, the features of the RAPID instrument is presented, then HFA events and the Compton-Getting effect are briefly discussed. The advantages of the parallel measurements are demonstrated by a bow shock crossing event and I gave an example to the transformation: the

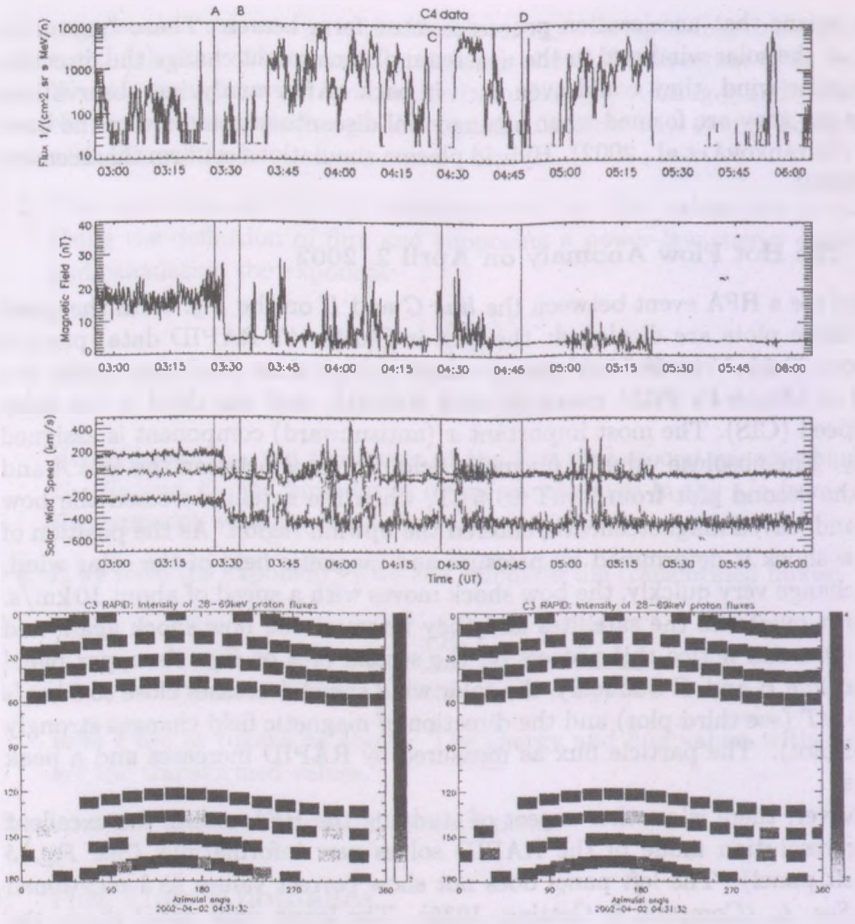

Figure 5: Hot flow anomaly on 2002 April 2 measured by Cluster-3 and Cluster-4 satellites. *Upper panels:* The event measured with Cluster-4 RAPID (28.2 keV-68.9 keV channel, top), FGM (magnetic field magnitude, second) and CIS (solar wind speed, third). The X component is marked on V_x. *Bottom left:* RAPID measurement aboard Cluster-3 in high angular resolution mode at 4:31:32 UT. *Bottom right:* The distribution transformed to the frame of solar wind.

HFA event on April 2, 2002. Henceforward I will search other HFA events in Cluster's data series using aforementioned method.

Acknowledgments

The Author thanks Dr. Károly Kecskeméty for his help and advices and the KFKI RMKI, Department of Space Physics for using its infrastructure.

References

Compton, H., Getting, I. 1935, Physical Review, 47(11), 817

Escoubet, C. P., Russell, C. T., Schmidt, R. (eds.) 1997, The Cluster and Phoenix Missions, Kluwer Academic Publishers

Lin, Y. 2002, Planet. Space Sci., 50, 577

Schwartz, S. J., Chaloner, C. P., Christiansen, P. J., Coates, A. J., Hall, D. S., Johnstone, A. D., Gough, M. P., Norris, A. J., Rijnbeek, R. J., Southwood, D. J., Woolliscroft., L. J. C. 1985, Nature, (318), 269

Thomsen, M. F., Gosling, J. T., Fuselier, S. A., Bame, S. J., Russell, C. T. 1986, JGR, 93, 11311

Šafránková, J., Přech, L., Nemeček, Z., Sibeck, D. 2002, Advanced Space Researches, 30(12), 2737

STEREO MISSION: OVERVIEW, THE PLASMA INSTRUMENT, CALIBRATIONS AND DATA

A. Opitz, R. Karrer, P. Bochsler, L. Blush, J. Fischer, J. Jost, M. Sigrist and P. Wurz
Physikalisches Institut, University of Bern
CH-3012 Sidlerstrasse 5, Bern, Switzerland
http://www.phim.unibe.ch/stereo
E-mail: opitz@stereo.unibe.ch

Abstract

This is an overview of the next generation solar space mission called STEREO. We focus on the plasma instrument and its calibrations at the University of Bern. The management of the plasma data is also discussed briefly.

KEYWORDS: *solar-terrestrial mission, space physics*

1. Mission overview

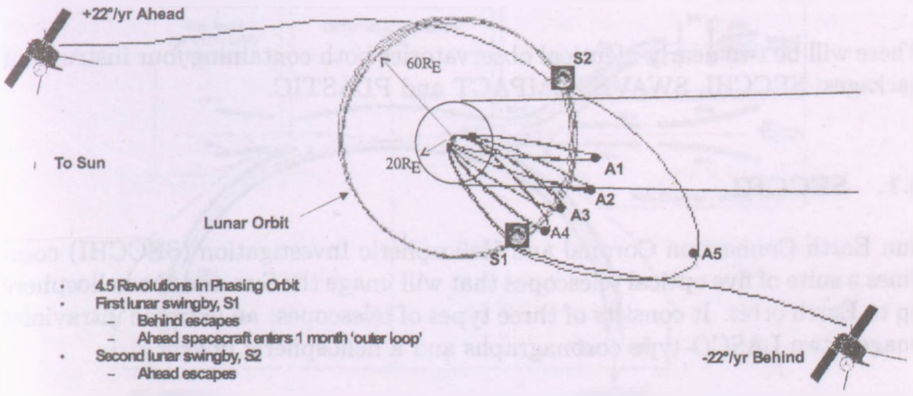

Figure 1: STEREO launch trajectory (figure provided by JHU/APL under NASA contract NAS5-97271 Task Order 28).

Solar Terrestrial Relations Observatory (STEREO) is the third mission in NASA's Solar Terrestrial Probes (STP) program. Its primary goal is to advance

the understanding of the three dimensional structure of the Sun's corona and its temporal evolution, especially regarding coronal mass ejections (CME). This mission will provide a totally new perspective on solar eruptions by imaging from two nearly identical space-based observatories simultaneously. One observatory will be placed ahead of the Earth in its orbit and the other behind using a series of lunar swingbys (see Figure 1). The launch is currently scheduled for February 2006.

The overall objective of the STEREO mission is to obtain the necessary measurements and observations to develop an understanding of the fundamental nature and origin of CMEs, which are the most energetic eruptions on the Sun and the primary cause of major geomagnetic storms.

The specific science objectives of STEREO are to understand the causes and mechanisms triggering CMEs, to characterize their propagation through the heliosphere, to discover the mechanisms and sites of energetic particle acceleration in the low corona and the interplanetary medium, and to develop a 3-dimensional, time-dependent model of the magnetic topology, temperature, density and velocity structure of the ambient solar wind.

2. Instrumentation

There will be two nearly identical observatories both containing four instrument packages: SECCHI, SWAVES, IMPACT and PLASTIC.

2.1. SECCHI

Sun Earth Connection Coronal and Heliospheric Investigation (SECCHI) combines a suite of five optical telescopes that will image the Sun and the heliosphere up to Earth orbit. It consists of three types of telescopes: an extreme ultraviolet imager, two LASCO-type coronagraphs and a heliospheric imager.

2.2. SWAVES

STEREO/WAVES (SWAVES) consists of three radio receivers (fixed, high and low frequency receivers) and a time domain sampler. It will track electromagnetic disturbances through the heliosphere.

2.3. IMPACT

In-situ Measurements of Particles and CME Transients (IMPACT) consists of two packages. The solar wind electron analyzer, the suprathermal electron telescope and the magnetometer are located on a mast anti-sunward. The solar energetic particle experiment consists of a solar electron proton telescope, a suprathermal ion telescope, a low and a high energy telescope. This experiment will focus on *in situ* observations of energetic particles.

2.4. PLASTIC

Plasma and Suprathermal Ion Composition (PLASTIC) consists of an entrance system containing an electrostatic analyzer (ESA) connected to a time-of-flight (TOF) chamber (see Figure 2). It is the primary *in situ* solar wind instrument.

3. PLASTIC instrument

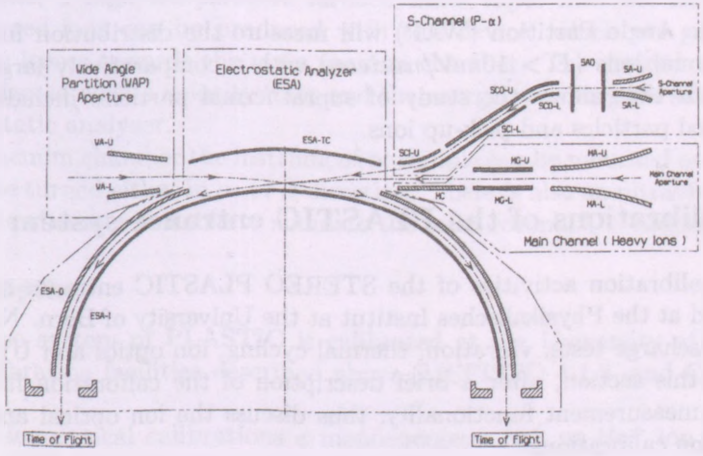

Figure 2: Drawing of the PLASTIC entrance system.

The instrument will measure *in situ* various solar wind species in the energy-

per-charge range of 0.2 to 100 keV/e. There are three science measurements combined into one instrument: S-channel, main channel and WAP (Allegrini, 2002).

3.1. S-channel

The Solar Wind Sector S-channel will measure the distribution functions of solar wind protons and alpha particles providing density, velocity, kinetic temperature and its anisotropy. The time resolution will be about one minute.

3.2. Main channel

The Solar Wind Sector main channel will measure the elemental composition, charge state distribution, kinetic temperature and velocity of the more abundant solar wind heavy ions (e.g. C, O, Ne, Mg, Si, and Fe) on at least five minute resolution.

3.3. WAP

The Wide Angle Partition (WAP) will measure the distribution functions of suprathermal ions ($E > 10keV/nucleon$) with a comparatively large geometrical factor that allows the study of suprathermal particles including shock-accelerated particles and pick-up ions.

4. Calibrations of the PLASTIC entrance system

Several calibration activities of the STEREO PLASTIC entrance system are performed at the Physikalisches Institut at the University of Bern. Namely the partial discharge tests, vibration, thermal cycling, ion optics and UV suppression. In this section, after a brief description of the calibration facilities we focus on measurement functionality, thus discuss the ion optical and the UV suppression calibrations.

4.1. Calibration facilities at the University of Bern

There are two main calibration facilities for solar wind instrumentations at the University of Bern: CASYMS and MEFISTO.

4.1.1. CASYMS

CASYMS stands for CAlibration SYstem for Mass Spectrometers. This facility provides a large-area ($\sim$ 250 cm^2), parallel beam of ions over the energy per charge (E/q) range from 5 eV/e up to 100 keV/e (Ghielmetti et al., 1983). The gas atoms (e.g. argon, helium, hydrogen) are ionized by an electron filament and then redirected and accelerated by electrostatic lenses.

The beam is guided into a vacuum chamber, where an instrument can be mounted on a x-y-z-table, which can also be rotated either in α- or the perpendicular β-direction. All table manipulation are controlled by computer as well as the beam energy and monitoring.

4.1.2. MEFISTO

MEFISTO stands for MEsskammer für FlugzeitInStrumente und Time-Of-Flight (calibration facility for solar wind instrumentations). This system provides a well focused ion beam at a beam energy per charge up to 100 keV/e (Marti et al., 2001). The ions are produced in an electron-cyclotron-resonance ion source (ECRIS) operating at 2.45 GHz from elements in gaseous or solid phase. For solid elements, a high temperature furnace emits vapor into the ion source. Highly charged ions can be produced with this device, which are guided by electrostatic lenses towards the main vacuum chamber. The desired mass per charge is selected with a velocity filter and the energy per charge is selected by the electrostatic analyzer.

In the vacuum chamber the instrument or sample can be mounted on a table, which can be turned either in α- or β-direction. There is also an ultraviolet light source in the chamber suitable to simulate the solar coronal UV radiation.

4.2. Ion optics

The entrance system of PLASTIC is calibrated at the University of Bern in the two calibration facilities described above (MEFISTO 4.1.2. and CASYMS 4.1.1.).

For the ion optical calibrations a monoenergetic Ar^+ or He^+ ion beam is used to simulate incident solar wind. During the calibration tests the voltage of the entrance system electrodes (ESA, SCO-L, SCI-U, etc.; see Figure 2) were tuned to particular voltages so that the ions were deflected through the entrance apertures and the ESA-hemisphere. The ratio between beam energy per charge E/q and the tuned voltage difference over the ESA-hemisphere voltages is called *analyzer constant*: $k = \frac{E/q}{V_{electrode}}$ (Allegrini, 2002 and Blush et al., 2003).

The ion optical calibrations are in progress. Preliminary results show good agreement with the data obtained when calibrating the engineering qualification model and the simulation results.

4.3. UV suppression

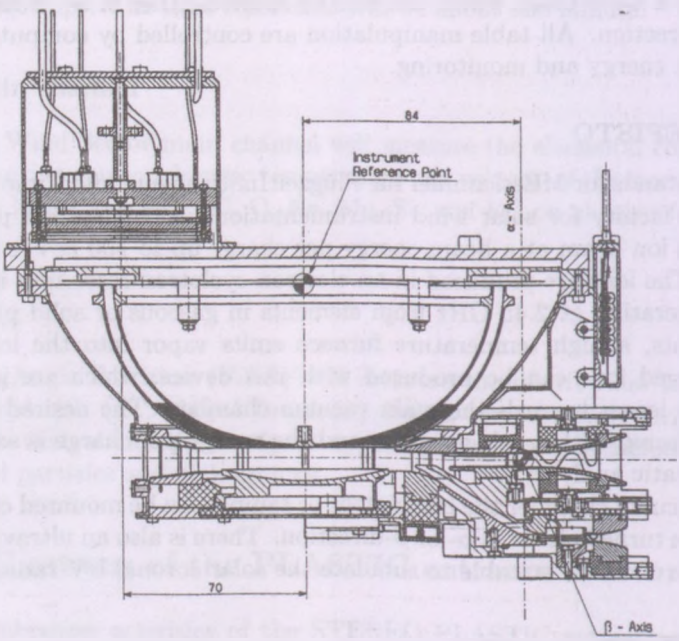

Figure 3: UV suppression calibration setup of the PLASTIC entrance system.

Ion measurements might be contaminated by energetic ($> 5eV$) photons. Solar UV can generate spurious detector counts on the carbon foil and the multi channel plates (MCP). Hence the ultraviolet radiation arriving onto the detectors must be suppressed as much as possible. We define the suppression factor as the ratio of the flux on the entrance slit and the flux on the exit slit.

The solar UV flux is approximately $2.5 \cdot 10^{14} \frac{photons}{cm^2 \cdot s}$ from the photosphere,

and subtends an angle of 0.5° at 1 AU. The entrance system is specially designed that photons entering at an angle lower than 1.24° cannot get into the instrument, hence photons from the photosphere are excluded. Photons with a coronal origin of 1.24° to 4.1° may enter the hemisphere, thus they might affect the measurements. The coronal UV flux is lower than the photospheric flux by a factor of 10^6. As a result, a suppression factor of 10^8-10^9 is necessary if we want $1\frac{photon}{cm^2 \cdot s}$ to avoid a high background caused by energetic photons.

That's why the inner surfaces of the PLASTIC entrance system were blackened by CuS. It is a rough surface, which isotropically scatters and absorbs photons. Due to the blackening only a very small portion of the incoming photons can get through the instrument and affect the measurement count rate, which was confirmed by the calibrations done in the MEFISTO laboratory. The suppression factor is of the order of 10^9.

5. PLASTIC data

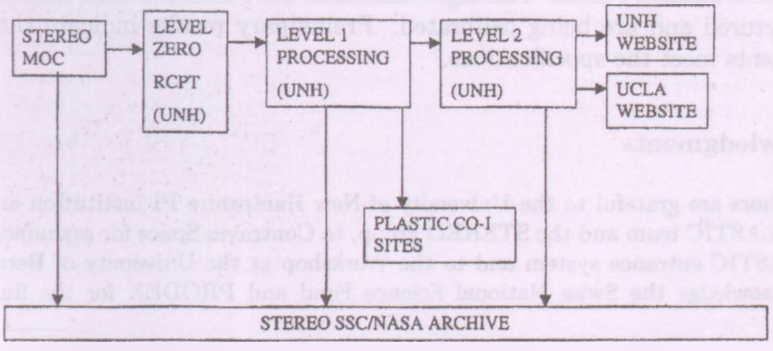

Figure 4: PLASTIC data flow from receipt from the STEREO Mission Operation Center through to the data archive at the STEREO Science Center.

The data products from PLASTIC are the composition, distribution functions, and time history of the solar wind and suprathermal positive ions. Solar wind measurements and identification of structures will be possible under nearly all types of solar wind conditions, including high-speed streams, interplanetary CMEs, and slow solar wind. The PLASTIC Data Center at the University

of New Hampshire will collect the Level-0 (binary) data and a software will automatically convert it into Level-1 (ASCII) data. Level-2 data consists of one-minute summary data sets containing solar wind proton density and speed, and proton temperatures, proton anisotropy referenced to the magnetic field, major ion species' densities, and suprathermal ion fluxes and anisotropies in several energy ranges, and beacon data for comparisons with the calibrated and cleaned final values of the same quantities. The Level-2 data will be available on the website of the University of California at Los Angeles (UCLA). Level-3 products result from basic scientific analyses of the PLASTIC Level-2 data with cooperation with the IMPACT, SWAVES and SECCHI teams. The planned dataflow is shown in Figure 4 (Kistler, personal communication).

6. Summary

A new solar mission has been discussed, which will give a three dimensional view on the solar corona. The STEREO mission is in progress and will be launched in February 2006. The flight models of the PLASTIC instrument were manufactured and are being calibrated. Preliminary results indicate that the instruments meet the specifications.

Acknowledgments

The authors are grateful to the University of New Hampshire PI institution and the whole PLASTIC team and the STEREO group, to Contraves Space for manufacturing the PLASTIC entrance system and to the Workshop at the University of Bern. We also acknowledge the Swiss National Science Fond and PRODEX for the financial support.

References

F. Allegrini **2002** *The PLASTIC sensor on STEREO: Design of the entrance system, energy analyzer and numerical simulations of solar wind measurements* PhD thesis at University of Bern

L.M. Blush, F. Allegrini, P. Bochsler, A. Galvin, M. Hohl, R. Karrer, L. Kistler, B. Klecker, E. Möbius, M. Popecki, B. Thompson, X. Wang, R.F. Wimmer-Schweingruber, P. Wurz **2003** *Test and Calibrations of the PLASTIC Entrance System: Design Verifications for Flight Models on the STEREO spacecraft* EPS Conference proceedings

A.G. Ghielmetti, H. Balsiger, R. Bänninger, P. Eberhardt, J. Geiss and D.T. Young **1983** *Calibration system for satellite and rocket-borne ion mass spectrometers in the energy range from 5 eV/charge to 100 keV/charge* Review of Scientific Instruments 54

A. Marti, R. Schletti, P. Wurz and P. Bochsler **2001** *Calibration facility for solar wind plasma instrumentation* Review of Scientific Instruments 72

The Plasma Environment of Jupiter as observed by the onboard Plasma Instruments of Cassini

Zs. Bebesi
KFKI Research Institute for Particle and Nuclear Physics
H-1525 Budapest, P.O.Box 49
E-mail: bzsofi@rmki.kfki.hu

Abstract

The Cassini spacecraft - on its way to Saturn - explored the plasma environment of Jupiter during the millenium flyby near the ecliptic plane along the dusk side of the planet. In this paper I give a general overview of the Cassini-Huygens mission focusing on the plasma instruments onboard of Cassini, and summarize the main scientific concepts. The orbit of Cassini at Jupiter was oriented along the Jovian Bow Shock (BS), so there were several occasions for the spacecraft to intercept the BS and to detect other various types of plasma events too. Therefore I present some results concerning the BS crossings, and some of the events that are suggested to be the consequence of the interaction of the BS with the Solar Wind (SW) flow. The more precise analysis of these "unusual" events will be the subject of future studies.

Keywords: *Jupiter, solar wind, magnetosphere, shock waves, plasma physics*

1. Introduction

The Cassini-Huygens spacecraft was built as a joint project of NASA and the European Space Agency with the main scientific objective of the exploration of Saturn, its system of rings and moons during a four year orbiting period. The spacecraft consists of the Cassini Orbiter - to discover the direct planetary environment of the Saturnian system - and the Huygens Probe - which is going to descend below the clouds of Saturn's moon, Titan. Cassini carries 12 different scientific instruments onboard, and Huygens has an additional 6. The mission was launched on October 15, 1997, and it takes almost 7 years for the spacecraft to arrive at Saturn - which is about to happen on July 1st this year. The long journey in the Solar System provided several opportunities for scientific measurements, especially the occasions of the four planetary flybys (Table 1.), of which the present paper is concerned with the Jupiter encounter. Among the

Date of Flyby	Planet
April 26, 1998	Venus
June 24, 1999	Venus
August 18, 1999	Earth
December 30, 2000	Jupiter

Table 1: *Former planetary flybys of Cassini.*

multiple scientific goals (multispectral analysis, plasma-particle measurements, magnetospherical studies, imaging, observation of dust, rings and satellites, atmospherical analysis) of the mission we will focus on the magnetic field observations and the data provided by the onboard plasma instruments of Cassini. First we will have an overview of the Cassini Plasma Spectrometer.

2. Cassini Plasma Spectrometer

The Cassini Plasma Spectrometer (CAPS) (Young et al., to be published; Szego et al., 2003) is one of the 12 onboard hardware investigations of Cassini Orbiter. The instrument consists of 3 different plasma sensors (Figure 1.), which are designed to fullfill the objective of covering the broadest possible range of plasma energy, composition and temporal variation. These 3 sensors of CAPS are the Electron Spectrometer (ELS), the Ion Mass Spectrometer (IMS) and the Ion Beam Spectrometer (IBS). The ELS instrument is designed for the study of electron plasma dynamics measuring the velocity distribution of electrons between 0.7eV-30keV. The IMS sensor is for ion plasma dynamics and composition studies concerning diffuse magnetospheric plasmas and low concentration ion species. ELS is mounted IMS so their fields-of-view are coplanar and coaxial as well. The IBS instrument is specialised for narrow beam-like structures expected in the interplanetary region and near the planets, detecting ion velocity velocity distributions with high angular and energy resolution between 1eV-50keV. The 3 sensors are positioned upon the Data Processing Unit (DPU), which acts as a mechanical base. Furthermore the DPU is set on a motor driven Actuator (ACT), and that rotates the entire CAPS instrument around a rotation axis parallel to the z-axis of the spacecraft. Therefore it is possible for the sensors to provide a sweeping motion in the plasma flow during measurements. For the present paper the most important one of the 3 sensors is the Ion Beam Spectrometer (Figure 2.) (Vilppola et al. 2001; Young et al., to be published), which is a high resolution hemispherical section electrostatic analyser, that has

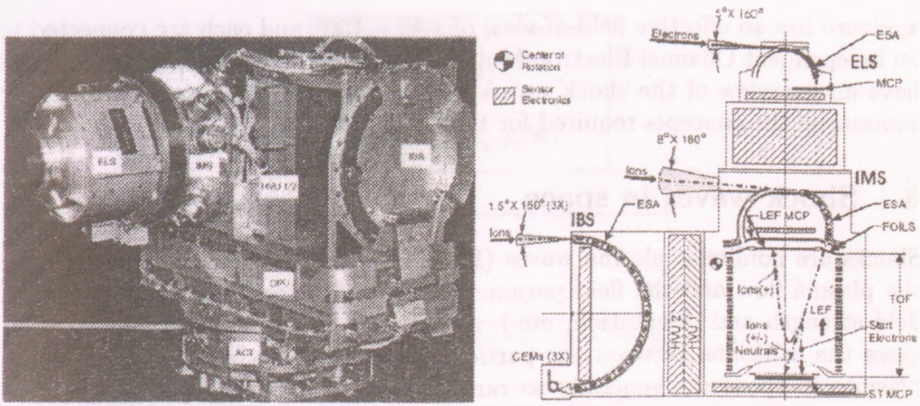

Figure 1: *The CAPS intrument (left) with the optical layout (right), fields-of-view and key sensor elements.*

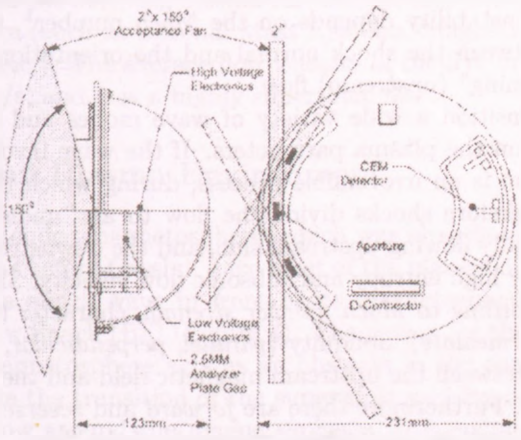

Figure 2: *The scheme of the IBS sensor of CAPS.*

been optimized to study the 3 dimensional characteristics of ion beams mainly in the neighbourhood of Jupiter and Saturn and in the solar wind. The instrument has 3 entrance apertures (of which the center one is coplanar with ELS and IMS apertures) offset by 30° to create a triple "crossed-fan" array in order to achieve a 3 dimensional mapping of the incoming target ion distribution. Each entrance

aperture has an effective field-of-view of 1.5° × 150° and each are connected to an independent Channel Electron Multiplier (CEM). In the next section we will have an overview of the shock waves observed in space in general in order to summarise the concepts required for the further discussion.

3. Shock waves in space

Shocks are nonlinear plasma waves (Burgess, 1995) occuring at places, where the plasma flow and the field parameters (flow velocity, temperature, density, field strength and orientation, etc.) go through abrupt changes. In ordinary gases the collisions between the particles transfer energy and momentum, yet plasmas in space are generally so rare, that this coupling practically doesn't appear - we call that medium a *collisionless plasma*. However according to spacecraft observations shocks do exist in space, but the mechanism is somewhat different - the dissipation comes from microscopic instabilities (Cravens, 1997) which lead to the increase of electric and magnetic field perturbations. The type of the instability depends on the Mach number[1], the plasma beta[2] and the angle between the shock normal and the orientation of the magnetic field in the "incoming" (upstream) flow.

At the shock transition a wide variety of wave modes and shock types may occur depending on the plasma parameters. If the wave involves compression, then the transition is an irreversible process, during which the entropy of the gas changes. Therefore shocks divide the flow to an "undisturbed" (low entropy) supersonically flowing upstream side, and the downstream region, which is characterized by high entropy and subsonic flow velocity. Shock types can be distinguished according to *Mach number strength*, character (whether they are *fast*, *slow*, or *intermediate*), obliquity (*parallel*, *perpendicular*, *oblique*) depending on the angle between the upstream magnetic field and the shock normal, as mentioned earlier. Furthermore there are *forward* and *reverse* shocks according to their sense of travel in the frame of reference of the flow.

In the Solar System there is a continuous flow of charged particles emerging from the Solar Corona, filling the entire interplanetary space - that is the Solar wind (SW) plasma. It mainly consists of protons (96%) and He α particles (4%) and a small fraction of heavier ion species, but the composition and parameters evolve as the flow propagates towards the outer regions of the Solar System.

[1]Definition of Mach number: $M = u_{flow}/c_S$, where c_S is the speed of wave propagation in the medium - sound waves, Alfven waves, etc., therefore we can define several types of Mach numbers.

[2]Ratios of thermal to magnetic pressures

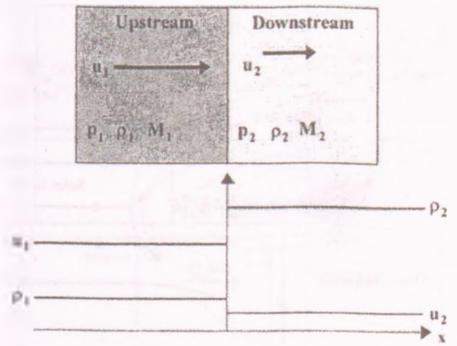

Figure 3: *Variation of plasma parameters during the shock transition.*

There are several types of plasma structures in the SW due to the frosen in magnetic field, and the occuring plasma turbulences and due to the cause of the shock's appearence (such as Coronal Mass Ejection, Magnetic Cloud, interacting SW streams, etc.). The average flow velocity of the SW in the ecliptic plane is around 400 km/s, and it is a highly supersonic flow.

4. The Jovian Plasma Environment

Jupiter has an extended magnetosphere, which was already visited by six other spacecraft missions before Cassini. Like most of the planets of the Solar System Jupiter also has a shock wave in front of its magnetosphere in the sunward direction - the Bow Shock (BS) (Figure 4). The reason for the formation of the BS is that the magnetosphere acts as an obstacle to the supersonic SW flow. The BS is to make the transition of the supersonic velocity to subsonic values, so the SW can flow around the magnetosphere. The region behind the BS - generally characterized by subsonic flow (depending on the parallel component of the bulk velocity) - is called the magnetosheath. The position of the BS is not steady - it moves dynamically according to the outside plasma pressure - depending on the solar activity. The plasma flow at near Jupiter is characterized by high Mach number (≥ 10) and low plasma beta (≤ 0.7) values.
For missions studying the magnetosphere the Jovian rotation is also important - together with the pressure increasing continouos mass injections from the moon Io -, which adds a ten hour periodicity to the events observed by spacecraft instruments.

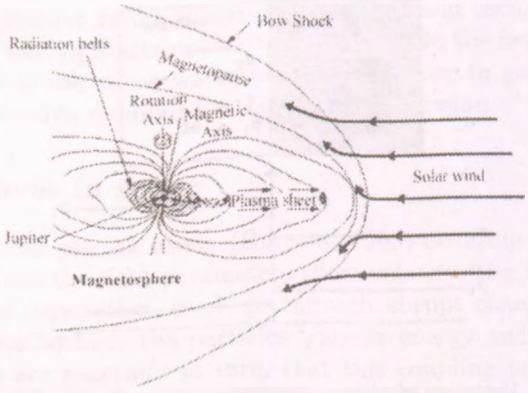

Figure 4: *The Jovian plasma environment.*

5. Cassini's plasma measurements at Jupiter

Cassini arrived at Jupiter in the end of December 2000, and during the following few month the onboard instruments almost continuously monitored the Jovian environment. The orbit of Cassini is plotted on Figure 5. with an embedded figure in the bottom left corner (from Simpson et al., 1992) showing the former spacecraft passing by the system of Jupiter (Galileo is not plotted, since it spent several years orbiting). During the flyby Cassini flew just along the BS (Joy et al., 2002), so there were more than 40 shock observations, of which the first on accured on December 30, 2000. In the following sections of this study we will present some illustrous plasma discontinuity events, of which some are ordinary BS crossings, but we have found several rather peculiar plasma structures as well. Since the interaction of the SW flow with the with the Jovian plasma environment gives rise to many different types of plasma waves, we have investigated and classified some of the observed phenomenae. The variation of the plasma parameters during the shock transition is a kind of "signature" of the shock itself as illustrated on Figure 6.

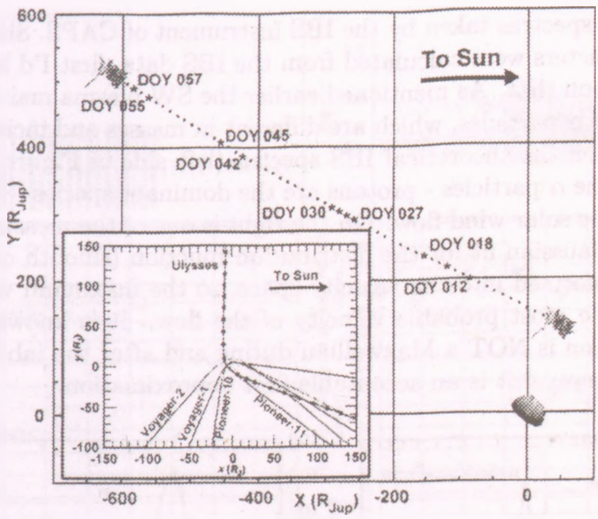

Figure 5: *The orbit of Cassini during the Jupiter flyby. The embedded figure in the bottom left corner (Simpson et al., 1992) shows the orbits of the spacecrafts that had ever visited Jupiter in history (in fact Galileo is one of them too, but it spent several years orbiting Jupiter).*

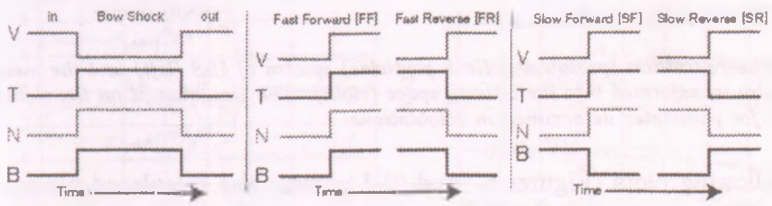

Figure 6: *Variation of plasma parameters (where V is the flow velocity, T is the Temperature, N is density and B is the magnetic field strength) during the different types of shock transitions.*

5.1. Shock waves in the neighbourhood of Jupiter

In this section I present some of the plasma events that Cassini encountered during the Jupiter flyby between the end of December 2000 and the beginning of March 2001. The plasma paramereters were calculated from the measurements of the CAPS-IBS sensor, and the Magnetometer data of Cassini was used as well. The presented plots show the variation the parameters in accordance with the notations of Figure 6. except for the top panels (Figures 8. and 9.),

which are the ion spectras taken by the IBS instrument of CAPS. Since most of the plasma parameters were calculated from the IBS data, first I'd like to have a little discussion on that. As mentioned earlier the SW plasma mainly consists of protons and He α particles, which are different in masses and incidence rates too. It is shown on the theoretical IBS spectra (left side of Figure 7.), that - comparing with the α particles - protons are the dominant species - with higher count rates - in the solar wind flow. On the right is one of the measured proton spectras with a Gaussian fit for the distribution function (smooth curve). The spectra was transformed into the velocity space, so the maximum value of the Gaussian gives the most probable velocity of the flow. It is known, that the velocity distribution is NOT a Maxwellian during and after the inbound shock transition - yet however it is an acceptable first approximation.

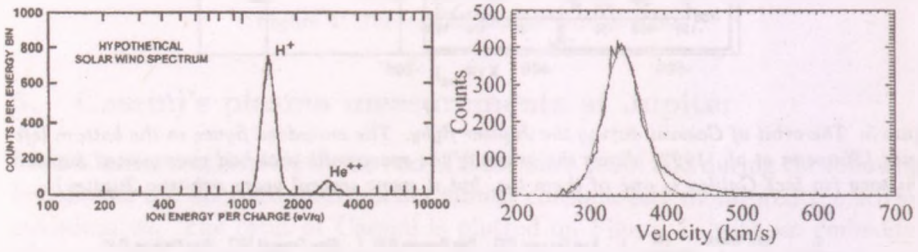

Figure 7: *Theoretical ion (proton and He α particles) spectra of IBS (left) and the measured proton spectra transformed into the velocity space (right). The Gaussian fit on the right hand side plot is for parameter determination calculations.*

On the following plots (Figures 8. and 9) I present the calculated plasma parameters. The top panel is the IBS spectra - as mentioned above -, the second one is the bulk velocity of the flow in km/s units, the third is thermal velocity in km/s, the fourth is density in relative units, and on the bottom is the magnetic field data. The total magnetic field variation (the upper curve) is plotted with the 3 components (B_x is not shifted, but B_y, B_z and B_{total} are shifted up by 2, 4 and 5 nT respectively.) using the 24 s resolution Magnetometer data, which is measured in the RTN[3] frame of reference.

The first well observed bow shock crossing event during the Cassini flyby took place on DOY[4] 012 (January 12) 2001 (left side of Figure 8.) (Szego et al.,

[3]The RTN reference system: the R unit vector is oriented from the Sun towards the spacecraft direction, the T unit vector points to the vector product of R and the Sun's rotation axis, and the N vector completes the right-hand system.

[4]Conventional abbreviation of "Day Of Year"

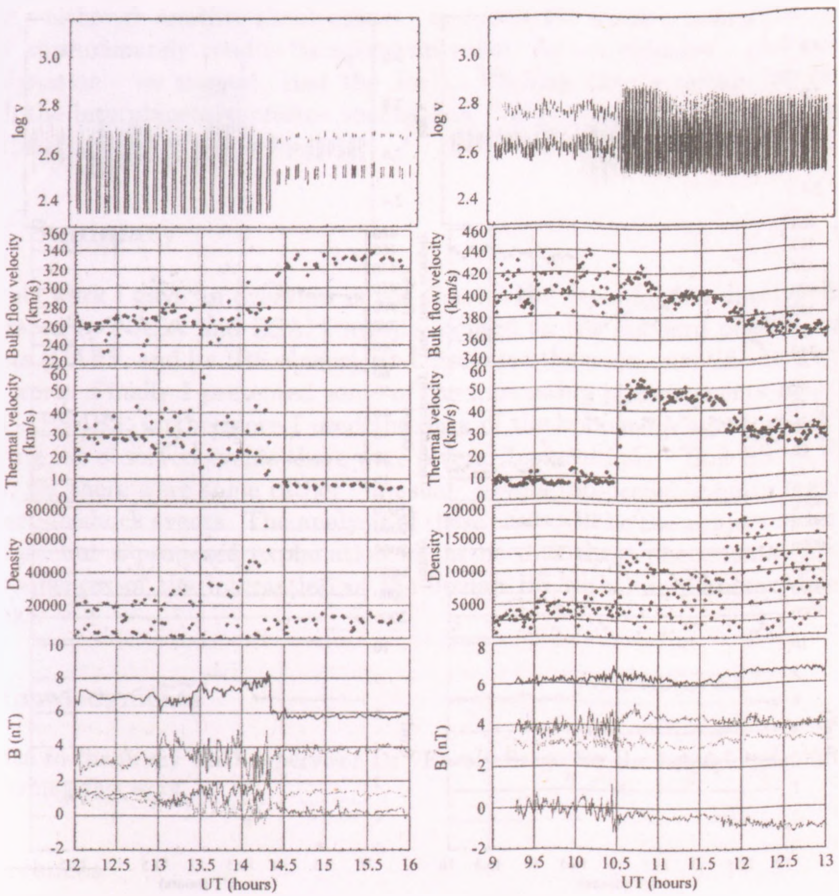

Figure 8: *Bow shock crossing on DOY 012 (left) and on DOY 030 (right).*

2003) at 14:20 UT. It was a quasi-perpendicular BS crossing with $\vartheta_{Bn} > 80°$, where ϑ_{Bn} is the angle between the shock normal and the upstream magnetic field (B), and the spacecraft's velocity was almost perpendicular to the shock normal. The BS crossing on DOY 030 (right hand side of Figure 8.) was a quasi-parallel (B_t is almost 0) BS transition (Szego et al., to be published). There is an interesting furher aspect of this event - the serial acceleration and deceleration of the flow at the foot region might be the effect of firehose in-

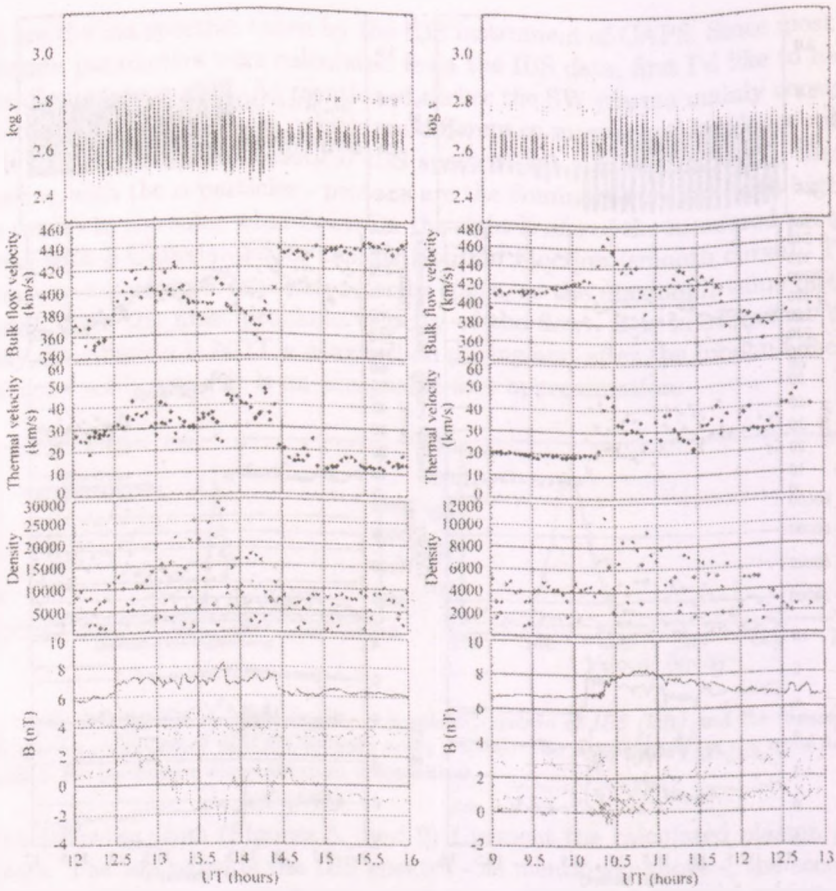

Figure 9: *Secondary discontinuities accompany the BS transitions.*

stability. Finally I present two of the rather "unusual" cases mentioned earlier (Figure 9). Comparing these shock transitions with the BS crossings introduced above we can see, that there are multiple events near the shocks, so the identification of the exact instants of the transitions themselves is a little ambiguous too. On DOY 027 the upstream and downstream regions can be separated, but they were divided by another shock between 12:30 and 14:30 UT. The DOY 028 transition at 10:20 UT is confirmed by the thermal velocity and magnetic field

data - although another shock occurs - and after the shock crossing the plasma flow approximately retains its upstream value. As a conclusion - and possible explanation - we suggest, that the Jovian BS may have a certain interaction with the interplanetary plasma shocks, but this problem will be the subject of further studies.

6. Summary

In this work I gave an overview of the Cassini-Huygens mission developed and operated by NASA and ESA. I mainly focused on the onboard plasma instruments (CAPS, and its IBS sensor), and discussed the main scientific concepts of our work. Finally I presented some of the interesting plasma events measured by CAPS-IBS, furthermore I used the data of the onboard Magnetometer too. During the observed events there were several "conventional" Bow Shock crossings, yet there were some rather "unusual" events too occuring quite regularly at certain shock events. The analysis of these cases will be the subject of further studies, but a proposed explanation might be that these phenomenae are the consequences of the interaction of the Jovian BS with interplanetary plasma shocks.

Acknowledgments

I'd like to thank my PhD supervisor Dr. Karoly Szego for the helpful consultations concerning this work.

References

Burgess, D., Collisionless shocks, in Introduction to Space Physics, edited by M. G. Kivelson and C.T. Russell, Cambridge University Press, New York, 1995

Cravens, T. E., Physics of Solar System Plasmas, edited by A. J. Dressler, J. T. Houghton, M. J. Rycroft, Cambridge University Press, New York, 1997

Joy, S. P., Kivelson, M. G., Walker, R. J., Khurana, K. K., Russell, C. T., Ogino, T., Probabilistic models of the Jovian magnetopause and bow shock locations, JGR Vol. 107, No. A10, 1309, 2002

Simpson, J.A., Anglin, J. D., Balogh, A., Burrows, J. R., Cowley, S. W. H., Ferrando, P., Heber, B., Hynds, R. J., Kunow, H., Marsden, R. G., McKibben, R. B., Miller-Mellin, R., Page, D. E., Raviart, A., Sanderson, T. R., Staines, K., Wenzel, K. P., Wilson, M. D., Zhang, M., Energetic Charged-Particle Phenomena in the Jovian

Magnetosphere: First Results from the Ulysses COSPIN Collaboration, Science, 257, 1543, 1992

Szego, K., Young, D. T., Barraclough, B. L., Berthelier, J. J., Coates, A. J., McComas, D. J., Crary, F. J., Dougherty, M. K., Erdos, G., Gurnett, D. A., Kurth, W. S., Thomsen, M. F., "Cassini plasma spectrometer measurements of Jovian bow shock structure", JGR, Vol. 108, No. A7, 1287, 2003

Szego, K., Young, D.T., Bagdonat, T., Barraclough, B., Berthelier, J.-J., Coates, A.J., Crary, F.J., Dougherty, M.K., Erdos, G., Gurnett, D.A., Kurth, W.S., Opitz, A., Rymer, A., Thomsen, M.F., A pre-shock event at Jupiter on 30 January 2001, submitted to J.G.R.

Vilppola, J. H., Tanskanen, P. J., Barraclough, B. L., McComas D. J., Comparison between simulations and calibrations of a high resolution electrostatic analyzer, Rev. Sci. Instrum. Vol. 72, No. 9, 2001

Young, D. T., Berthelier, J. J., Blanc, M., Burch, J. L., Coates, A. J., Goldstein, R., Grande, M., Hill, T. W., Johnson, R. E., Kelha, V., McComas, D. J., Sittler, E. C., Svenes, K. R., Szeg, K., Tanskanen, P., Ahola, K., Anderson, D., Bakshi, S., Baragiola, R. A., Barraclough, B. L., Black, R. K., Bolton, S., Booker, T., Bowman, R., Casey, P., Crary, F. J., Delapp, D., Dirks, G., Eaker, N., Funsten, H., Furman, J. D., Gosling, J. T., Hannula, H., Holmlund, C., Huomo, H., Illiano, J. M., Jensen, P., Johnson, M. A., Linder, D. R., Luntama, T., Maurice, S., McCabe, K. P., Mursula, K., Narheim, B. T., Nordholt, J. E., Preece, A., Rudzki, J., Ruitberg, A., Smith, K., Szalai, S., Thomsen, M. F., Viherkanto, K., Vilppola, J., Vollmer, T., Wahl, T. E., West, M., Ylikorpi, T., Zinsmeyer, C., Cassini Plasma Spectrometer Investigation, submitted to Space Science Reviews.

A Study of Solar Energetic Particle Events and Coronal Mass Ejections Using SOHO Data

G. Facskó
KFKI Research Institute for Particle and Nuclear Physics
H-1525 Budapest, P.O.Box 49
E-mail: gfacsko@rmki.kfki.hu

Abstract

The relation of coronal mass ejections (CMEs) and solar energetic particle events (SEPs) is a field where our understanding of the underlying physical processes has been subject to rapid changes recently. Two paradigms are still competing: the earlier one is that solar energetic particle events (SEPs) are triggered by flares. The recently more popular viewpoint is that from the two classes of SEPs, gradual events and impulsive events, only the latter are caused by flares, large spectacular gradual events are related to CMEs, and entirely unrelated to flares. Most recent observations seem to indicate that many events do not fit into this simplified scheme, and many combinations termed as hybrid events may exist. Here I present two typical SEPs, and then two new "hybrid" class SEPs are reviewed.
KEYWORDS: *Sun, solar energetic particle events, coronal mass ejections, SOHO*

1. Introduction

The Sun is a permanent source of electromagnetic radiation and particles. The particles and the invisible part of this radiation were discovered in the middle of the last century by space vehicles. The flux of the particle radiation often decreases with orders of magnitude, this phenomena is called solar energetic particle event. These phenomena relate the active regions of the Sun and their processes: flares and coronal mass ejections (CMEs). Solar energetic particle events (SEPs) are divided into two classes: impulsive and gradual events. Nowadays, this theory seems to be reassessing because the new spacecraft detected events, which cannot be classified in the simple scheme. In this paper I will describe two typical solar energetic particle events then present two hybrid SEPs observed by SOHO and the WIND spacecraft simultaneously.

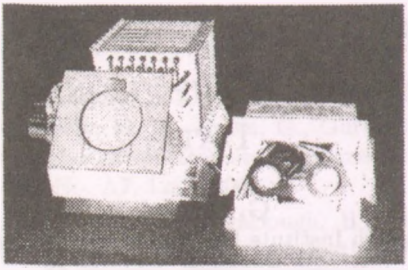

Figure 1: The EPHIN and LION instruments. *Left:* EPHIN *Right:* LION

1.1. The Solar and Heliospheric Observatory

The SOHO[1] is mission of ESA and NASA collaboration. Launched on 2th of December, 1995 by an Atlas-IIAS booster, it has reached a final halo orbit around the L_1 Langrange point 2 months later. There the direction of the Sun changes no more than 10" with 1/15" angular velocity. SOHO's location is very suitable to study the corona, the solar interior, as well as the solar wind (NASA, 2000). The spacecraft is 3-axis stabilized. In this paper I use data of four instruments only: LASCO, EIT, LION and EPHIN. The LASCO consists of three coronographs having different viewing fields, allowing easy detection and observation of CMEs. EIT (extreme ultraviolet telescope) images help understanding processes before and after eruptions. The latter two instruments form a group named COSTEP[2] (*Fig. 1*) and our group has full access to their data.

1.1.1. The LION Instrument

LION employs semiconductor detectors to detect energetic particles. It consists of two tubes, both of them containing a double telescope recording ions and electrons (44 keV to 6 MeV). In front of one of the telescopes a deflecting magnet filters < 300 keV electrons out. The energy deposited can be determined from the impulse caused by the impact of the particles. A Lexan foil protects the sensors from the direct (UV) radiation of the Sun and the scattered light. The two tubes form an angle of 7° and tilted at 45° of the Sun-Eart direction so they look along the mean direction of the interplanetary magnetic field (R.

[1]**SO**lar and **H**eliospheric **O**bservatory
[2]**C**omprehensive **S**uprathermal and **E**nergetic **P**article **A**nalyzer

Müller-Mellin et al., 1995). The time resolution is 15 s.

1.1.2. The EPHIN Instrument

The sensors of EPHIN are built up from solid-state detectors as well. It measures the energy spectrum of electrons from 250 keV to 8.7 MeV, protons and helium isotopes from 4 MeV/nucleon to 53 MeV/nucleon. The instrument looks into the direction of the interplanetary magnetic field like LION. The heart of the EPHIN, the sensor head contains five detectors. Both the A and B detectors are composed of six pieces; they can detect protons and helium isotopes. The C, D and E detectors can observe electrons. The purpose of the G detector is filtering the background radiation. The aperture of the instrument is covered with a foil, which protects the detectors from the radiation of the Sun (R. Müller-Mellin et al., 1995).

2. Solar Energetic Particle Events

2.0.3. An Impulsive Event

An impulsive SEP event took place on July 9, 1996 recorded by both LION and EPHIN (*Fig. 2*). The event was also detected by SOHO EIT (on FeXV, 28 nm wavelength) and the accompanying CME was observed by LASCO. A typical impulsive development can be seen: a short, strong increase followed by a smooth decay of several days on the EPHIN plot, nicely complemented by the measurements of LION. Electrons and protons are plotted separately. The shapes of the profiles are similar on all plots. This event demonstrates the advantage of continuous measurements SOHO is capable of simultaneously observing over wide ranges of electromagnetic and particle radiation. This allowed identifying the active region situated on the west side of the Sun from which the CME and the flare were developed.

2.0.4. A Gradual Event

A gradual SEP event took place between 7 and 12th of April, 1997, observed by EIT, LASCO and COSTEP as well (*Fig. 3*). We took EIT images at 19.5 nm wavelength, LASCO C2 coronograph data, the 0.25-0.73 MeV electron and 4.3-7.8 MeV energy channels of EPHIN and the 80-125 keV electron channels of LION. One can see a slow increase and a longer decay phase. The Moreton wave in the low corona compared with this event is visible in the image of the

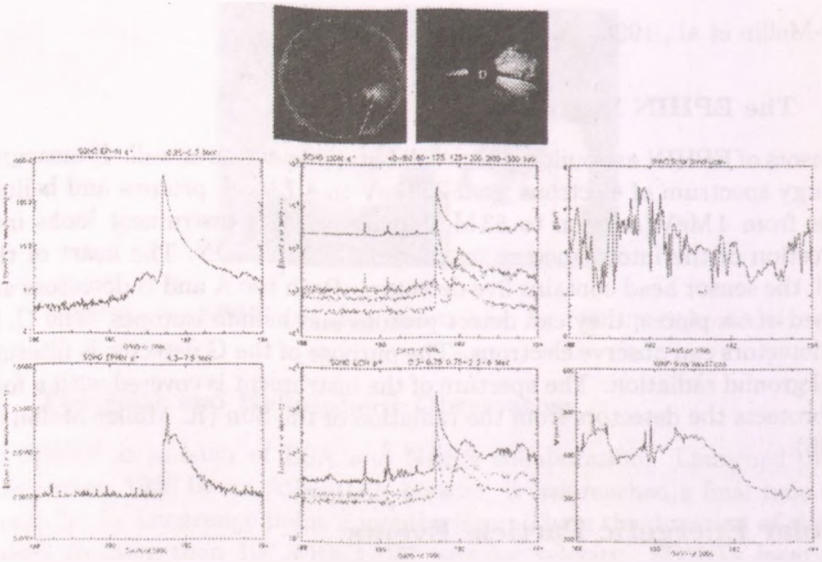

Figure 2: The 9th of July, 1996 impulsive event. *Upper panels:* EIT image, LASCO C3 coronograph image taken at 15^h38^m (UT) *Bottom 2 rows:* COSTEP EPHIN, LION fluxes and WIND magnetic field and solar wind speed (time interval: 6 days).

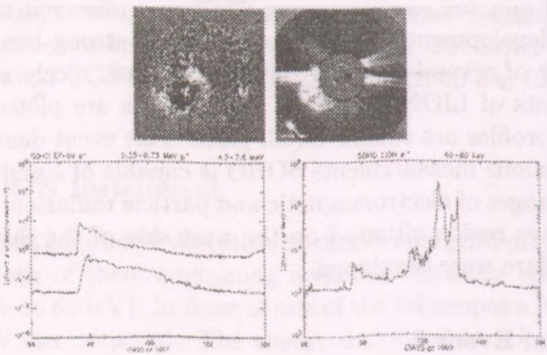

Figure 3: The gradual event on 7th of April, 1997. *Top:* EIT image and LASCO images *Bottom:* EPHIN and LION ion fluxes (time interval: 8 days).

EIT. There is no sign of a flare, confirming that gradual events are generated by CMEs.

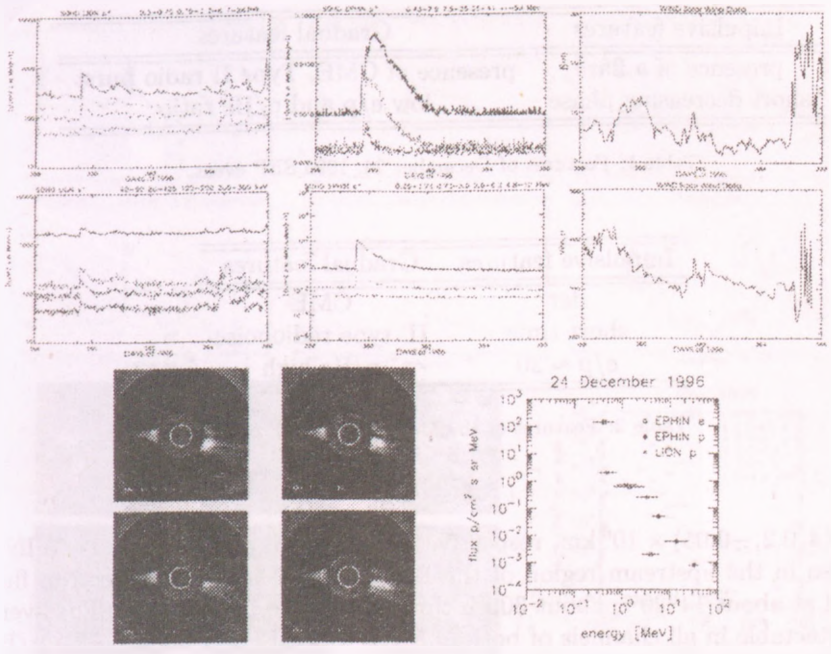

Figure 4: A hybrid SEP event between 23th and 31th of December, 1996. *Upper panels:* LION ions, EPHIN protons, and WIND magnetic field. *Middle:* EPHIN and LION electrons, WIND solar wind speed (time interval: 8 days). *Lower panels:* CME on LASCO images and the energy spectrum of the SEP.

2.1. Hybrid Events

Two special events will be analyzed in the next two sections. Their peculiarity is that we are not able to classify them, using classical the impulsive/gradual categories, as they exhibit both impulsive and gradual features.

2.1.1. SEP on 24th of December, 1996

After a long quiet period a small flux increase was detected on December 24, 1996. At 13^h23^m UT the intensity of 300 keV electrons suddenly increased (*Fig. 4*). At 13^h11^m a C2.1 class flare was observed by GOES-8 and a type II radio burst appeared in the radio range at 13:09 (UT) (*Table 1*).

The middle panel of (*Fig. 4*) shows the flux of 40-300 keV protons of LION for 8 days. The position of SOHO and WIND were $(1.59, -0.37, -0.05) \times 10^6$ km

Impulsive features	Gradual features
presence of a flare short decreasing phase	presence of CME, Type II radio burst low e/p and p/He ratio

Table 1: Features of December 24, 1996 SEP event.

Impulsive features	Gradual features
fler short time $e/p \sim 20$	CME II. type radionoise p/He high

Table 2: Features of the May 6, 1998 SEP event.

and $(0.4, 0.2, -0.05) \times 10^6$ km, respectively in the GSE system, that is, WIND was also in the upstream region of the Earth's bow shock. The electron flux peaked at about 14^h20^m, about 300th times above the background. The event was detectable in all channels of both of LION and EPHIN chanels. The p/He ratio did not increase significantly.

A closer inspection of the magnetic field shows that the WIND was near to the border of a magnetic sector: at 22-23th of December the azimuth suddenly changed, first pointed to the Sun (315°), then pointed to outward (135°). On December 25 magnetic field pointed inwards. Magnetic field data indicate that SOHO intersected two magnetic flux tubes, SEP appearing in the middle of a CME which left the Sun earlier. Another CME followed the event from the west side of the Sun, at 264° position angle. The view-direction speed of the CME was 290 km/s and it accelerated slightly only. The WIND magnetometer detected an shock December 29, at 23^h accompanying this CME, resulting in a transition CME speed of 320 km/s.

Fig. 4 displays the energy spectrum of the SEP, based on maximum fluxes in each energy channel, providing approximately the energy spectra of the source region. LION ion fluxes together with EPHIN proton data are visible on the energy spectrum in (*Fig. 4*). They fit well, however at lower place the spectra are not so steep. The proton spectrum is hard with an exponent of -2.55, whereas the electrons spectrum is softer, the exponent is -4.8 (S. McKenna-Lawlor et al., 1999).

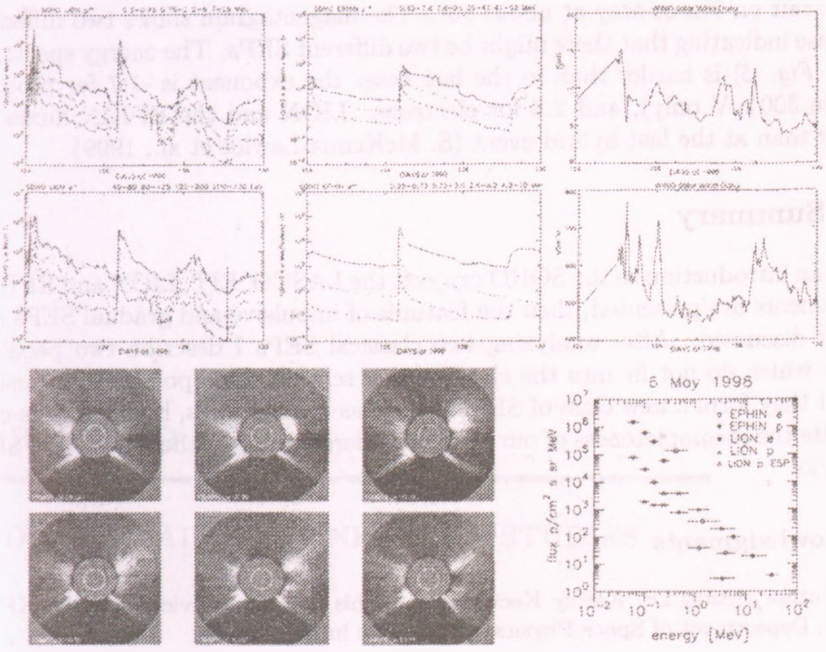

Figure 5: The hybrid SEP event between 6th and 9th of May, 1998. *Upper panels:* LION and EPHIN protons. WIND magnetic field. *Middle:* EPHIN and LION electrons, WIND solar wind speed (time interval: 6 days). *Lower panels:* CME on LASCO images; *Lower right:* Energy spectrum of the SEP.

2.1.2. SEP on 6th of May, 1998

This event was followed a stronger and longer brightening in the X-ray range between 6^h06^m to 7^h45^m. The source of the radiation was an X2.7 class flare. The H_α class of the flare 1N, its position is S11, W65 (*Table 2*).

The LASCO C2 telescope observed a CME at 8^h04^m at 311° position angle (Northwest direction) with a 105s km/s calculated speed. LASCO and EIT worked as particle detector: this process is called proton snow-storm. A piece of the CME arrived at the Earth (*Fig. 5*) what proved by the jump in the magnetic field. The average positions of SOHO and WIND were $(1.62, 0.40, 0.06) \times 10^6$ km and $(1.35, 0.6, 0.18) \times 10^6$ km in GSE. The SEP begun at 8^h11^m with an increase of the electrons flux, peaking at about 8^h35^m. All EPHIN and LION flux above 100 keV exhibited strong increases. The shock crossed the position of the

spacecraft on 7th of May at about 16^h. The magnetogram shows two different increase indicating that there might be two different SEPs. The energy spectrum (See: *Fig. 5*) is harder than in the last case: the exponent is -1.7 for protons (above 300 keV only), and 2.8 for electrons. LION and the EPHIN fluxes fit better than at the last hybrid event (S. McKenna-Lawlor et al., 1999).

3. Summary

After an introduction to the SOHO project, the LASCO, EIT, LION and EPHIN instruments are presented, then the features of impulsive and gradual SEPs are briefly discussed. After analyzing two classical SEPs I describe two peculiar SEPs, which do not fit into the classification scheme. One possible conclusion is that they form a new class of SEPs termed as hybrid SEPs, however, this can indicate the incompleteness of our present understanding of the flare-CME-SEP scenario.

Acknowledgments

The Author thanks Dr. Károly Kecskeméty for his help and advices and the KFKI RMKI, Department of Space Physics for using its infrastructure.

References

NASA 2000, http://sohowww.nascom.nasa.gov/

R. Müller-Mellin et al. 1995, Solar Physics, (162), 483

S. McKenna-Lawlor, S., Kecskeméty, K., Bothmer, V., Rodriguez-Pacheco, J., Facskó, G., Cyr, C. S. 1999, in Proc. 26th Int. Cosmic Ray Conf., Vol. 6

Part Two

Radio domain and infrared studies

Very Long Baseline Interferometry from Space: Past, Present and Future

S. Frey
FÖMI Satellite Geodetic Observatory
P.O. Box 585, H-1592 Budapest, Hungary
E-mail: frey@sgo.fomi.hu

Abstract

Space VLBI (SVLBI) is a relatively new radio astronomical technique used for high angular resolution imaging of bright, compact active galactic nuclei. SVLBI is an extension of ground-based VLBI networks by placing another interferometer element into Earth orbit. The first dedicated SVLBI satellite, the Japanese HALCA was launched in 1997. The technical background and highlights of HALCA scientific observations are reviewed, with special emphasis on the Hungarian contribution. The current perspectives for next-generation SVLBI satellites are outlined.
KEYWORDS: *VLBI, space VLBI, interferometry, radio continuum, quasars*

1. The principle of Space VLBI

Very Long Baseline Interferometry (VLBI) is a radio astronomical observing techique that involves an array of radio telescopes distributed over large – often inter-continental – distances. These telescopes observe the same radio sources in a coordinated way. The data are recorded at each telescope site. The interference is achieved later by playing back the data at a central data processing facility called correlator. This way a large synthesized telescope is created with a "diameter" equivalent with the longest distance between any pair of the individual telecopes (i.e. the length of the longest baseline). The diffraction-limited angular resolution of a telescope at a given observing frequency is inversely proportional to its diameter. VLBI at present is superior to any other means of astronomical observing in terms of angular resolution. Since its invention in 1967, VLBI has become a very successful tool for high-resolution imaging of compact celestial radio sources and for extremely precise astrometry. VLBI is also a geodetic technique that is used for determining antenna positions, Earth rotation and orientation parameters and crustal movements.

Ground-based VLBI baseline lenghths are naturally limited to $\sim$10 000 km set by the Earth diameter. Consequently, this limits the angular resolution

achievable at a given observing frequency. At 5 GHz, this is about 1 milliarcsecond (mas). However, many radio sources are known to be unresolved or only slightly resolved with ground-based VLBI. The baseline lenghts can in principle be increased by putting an interferometer element into Earth orbit as was proposed already in the early days of VLBI. The feasibility of SVLBI was successfully demonstrated first using a TDRSS communications satellite at 2.3 and 15 GHz in 1986–1988 (Linfield et al. (1990) and references therein). The practical realization of SVLBI, however, poses several technical difficulties. First of all, there is a substantial difference between SVLBI and other space astronomy missions. The orbiting antenna is just an element of a VLBI network, thus an SVLBI mission requires an extensive support of globally distributed and inhomogeneous ground-based facilities (radio telescopes, tracking stations, orbit determination systems and VLBI correlators). An important restriction is that the data taken by the orbiting element have to be down-linked to the Earth in real time, using a wide-band communications link to one of the dedicated satellite tracking stations. Also, the on-board electronics has to be synchronized from the ground to provide the frequency and timing accuracy necessary for later correlation of data. Unlike ground-based VLBI antennas, the position of the orbiting antenna is not known to high accuracy, and the position changes rapidly. This necessitates a special orbit determination system and poses serious additional requirements on the SVLBI correlators. At present, it is technologically possible to launch a 10-m class parabolic antenna with sufficient accuracy of its reflecting surface into Earth orbit. However, such an aperture is much smaller than that of a typical ground-based radio telescope (25–100 m in diameter). This fact, together with the long baselines, limits the sensitivity of the space–ground interferometer. The image fidelity is somewhat degraded by the uneven spatial sampling, i.e. the lack of intermediate baseline lengths between ground-only and space–ground baselines. In many cases, the "holes" in the (u, v) coverage limit the practically achievable image dynamic range to $\sim 100 : 1$ (Lister et al., 2001a).

More detailed reviews of the realization of space–ground radio interferometry can be found in e.g. Hirabayashi (1999, 2001) and Ulvestad (1999).

2. The VLBI Space Observatory Programme

The first satellite dedicated to SVLBI was launched in Japan by the Institute of Space and Astronautical Science (ISAS) on 12 February 1997 (Hirabayashi et al., 1998). After the successful launch, the MUSES-B satellite developed as

an orbiting element of the VLBI Space Observatory Programme (VSOP)[1] was renamed to Highly Advanced Laboratory for Communications and Astronomy (HALCA). The HALCA satellite carries an on-board 8-m diameter antenna. Its eccentric orbit has an apogee height of 21 400 km and perigee height of 560 km above the Earth's surface. In the standard observing mode, SVLBI data from two adjacent 16-MHz-wide base-band channels are digitized with 2-bit sampling and formatted. The digital data stream is then transmitted to a ground tracking station at the aggregate bit rate of 128 Mbit/s, in real time. The observing frequency bands of VSOP are 1.60–1.73 GHz, 4.7–5.0 GHz and 21.9–22.3 GHz. The performance of the latter receiving system did not meet the specifications, due to too high attenuation between the main antenna and the low-noise amplifier (Kobayashi et al., 2000a). Although the high system temperature made routine observing impossible at 22 GHz, the bursting Orion-KL water maser with flux density of $\sim 10^5$ Jy allowed the mission to detect fringes on baselines to HALCA in February 1998 (Kobayashi et al., 2000b).

More than 40 radio telescopes from all over the world (USA, Europe, Australia, China, South Africa, Japan in particular) took occasionally part in the VSOP observations, forming a truly global interferometric network. Five dedicated tracking stations (at Goldstone and Green Bank in the USA, Robledo in Spain, Tidbinbilla in Australia, and Usuda in Japan) were used to establish the two-way communications link to HALCA to transmit a stable reference signal to the satellite and to receive and record the digitized astronomical data. Three VLBI data correlating facilities (at Mitaka in Japan, Penticton in Canada, and Socorro in the USA) participated in the program. More information on the tracking and orbit determination of HALCA, the co-observing ground radio telescopes and the correlators can be found in e.g. Hirabayashi et al. (2000c) and references therein. The nominal lifetime of HALCA was originally set to 3 years. The HALCA spacecraft is still operational after more than 7 years from the launch, although most of the ground support oputside Japan has been terminated in 2002 due to financial reasons and time allocation difficulties. The mission at present is conducting Survey observations, and HALCA serves as an invaluable testbed for desingning its successors.

3. VSOP scientific observations in brief

A significant fraction – nominally 50% – of the in-orbit time was devoted to observations based on peer-reviewed proposals. Six calls for proposals were

[1]The VSOP web site is at `www.vsop.isas.jaxa.jp`

issued between 1995 and 2001. These opportunities were open for the entire scientific community. A brief and somewhat subjective census of some of the mission's scientific highlights is given here. A more detailed compilation is given in Frey (2004). Large collections of VSOP-related papers have been published in the proceedings of the COSPAR symposium on VSOP (*Advances in Space Research*, Vol. 26, No. 4, 2000), the VSOP Symposium (Hirabayashi et al., 2000a, *APRSV*), held in 1998 and 2000, respectively, and in the special issue of the *Publications of the Astronomical Society of Japan* (Vol. 52, No. 6, 2000).

One of the key science programs was to image relatively nearby, bright active galactic nuclei (AGNs). At low redshift, high angular resolution corresponds to high linear resolution. At 5 GHz, the close vicinity of the central engine in the elliptical galaxy M87 (the host for the bright radio source Virgo A) could be studied with a resolution corresponding to ~ 300 Schwarzschild radii of the central super-massive black hole (Junor et al., 2000). The effect of improved angular resolution is spectacularly demonstrated in the case of the quasar 3C 273. The transverse structure of the compact jet at 5 GHz is resolved by VSOP observations. Emission profiles taken accross the jet show regular oscillating patterns, which can be understood in terms of Kelvin–Helmholtz instabilities developing and propagating in the relativistic plasma along the jet (Lobanov & Zensus, 2001). A series of eight VSOP observations of the nearby quasar 1928+738 revealed a very complex and rapidly changing core–jet structure. The observed kinematics support the model of a ballistic precessing jet produced by a massive binary black hole (Murphy et al., 2003).

Polarization SVLBI observations offer uniquely high resolution imaging of compact polarization structures in AGNs. HALCA receives only left-circularly polarized radiation. However, with dual-polarization observations in the ground network, VSOP can provide polarization images for sources with sufficiently high correlated polarized flux density (Kemball et al., 2000). The parsec-scale linear polarization structures in a number of BL Lac objects have been studied (e.g. Gabuzda, 1999; Gabuzda & Gómez, 2001; Gabuzda et al., 2001).

VSOP provides a unique tool for high resolution radio spectral index mapping of AGNs. The angular resolution at 1.6 and 5 GHz can be matched with ground-only VLBI observations at ~ 3 times higher observing frequencies. The spectral index map of the γ-ray blazar 3C 279 shows highly inverted spectrum of the core between 1.6 and 5 GHz (Piner et al., 2000). In the case of the X-ray and TeV γ-ray source Mkn 501, the spectral index map also reveals the flat-spectrum core and the steep-spectrum inner jet. There is also significant spectral variation across the complex jet. Surrounding the inner jet, there are lower surface brightness regions with unexpectedly flatter spectra, possibly due

to processes in a slower outer shear layer (Edwards et al., 2000). The central 10-pc region of NGC 6251, a nearby radio galaxy was studied with VSOP at 5 GHz and with the VLBA at 15 GHz, at two close epochs. A sub-parsec-scale counterjet was discovered based on its optically thin spectrum derived from the spectral index image (Sudou et al., 2000).

Equipartition and inverse Compton theories limit the brightness temperature (T_B) of incoherent synchrotron sources to $\sim 10^{11} - 10^{12}$ K (Readhead, 1994; Kellermann & Pauliny-Toth, 1969). Curiously, the ability of ground-based VLBI for diretcly measuring radio source brightness temperatures is also constrained to this limit. The improved resolution of SVLBI makes this technique a unique tool to directly measure T_B values apparently violating the theoretical limits. Brightness temperatures in excess of 10^{12} K have often been measured with VSOP for individual AGNs, e.g. NRAO 530 (Bower & Backer, 1998) and 1921−293 (Shen et al., 1999). This indicates that Doppler boosting of the radiation, which enhances the apparent T_B, is indeed a common phenomenon in the cores of radio-loud AGNs.

A sub-sample of the well-studied Pearson-Readhead survey (Pearson & Readhead, 1988) sources was imaged with high angular resolution and high dynamic range using VSOP (Lister et al., 2001a). New information on the parsec-scale radio stucture of this statistically complete core-selected sample of AGNs, together with extensive data taken at various wavelengths allowed studies of the relativistic beaming effects in compact radio sources. Many previously known correlations between different source properties have been confirmed. Several new trends that support the beaming model have been discovered (Lister et al., 2001b). A significant fraction of the sources studied have brightness temperature in excess of 10^{12} K. A relationship between the brightness tempearuture and the intra-day variability type has also been found (Tingay et al., 2001).

4. Hungarian participation in VSOP

The potential geodetic and geodynamic application of the new SVLBI technique was pointed out first in the Satellite Geodetic Observatory (SGO) the of the Institute of Geodesy, Cartography and Remote Sensing (FÖMI) in the 1980's (Fejes et al., 1986). We got subsequently interested in many aspects of SVLBI, and – partly due to the unique openness of the VSOP – participated in many science programs, in most cases in international collaboration. I review here our main activities and results.

4.1. Application of SVLBI for geodesy

Ground-based VLBI is a key observing technique in global geodesy and geodynamics. Through the best realization of a quasi-inertial reference system defined by the highly accurate astrometry of distant radio-loud AGNs, it provides external reference to measure Earth rotation, orientation and VLBI antenna positions on the Earth's surface (Sovers et al., 1998). In SVLBI, one of the interferometer elements is orbiting around the geocenter. It is therefore possible in principle to directly connect different terrestrial and celestial reference frames, as well as to use SVLBI observables for impoving the satellite orbit determination (Fejes et al., 1996). The concept could be tested with VSOP data and proved to be technically feasible, although a number of practical restrictions prevented us from obtaining accurate results (Frey et al., 2002a).

4.2. SVLBI user assistance software development

SVLBI is a rather complicated observing technique involving a satellite, a tracking network and an array of ground radio telescopes, each with a series of specific operational and observing constraints. Successful planning of a SVLBI experiment requires a software tool to visualise the geometry, to check the technical constraints, to estimate the sensitivity and the imaging capability of the network, etc. One of our early activities was to develop a user assistance software in order to help the community to simulate observing programs. The first version of the Space VLBI Assistance Software (SPAS) was released in 1995. The PC-based software with user-friendly graphical interface was accepted by the VSOP as one of the three packages recommended for proposal preparation (Frey et al., 1998).

4.3. The VSOP Survey Program

About a quarter of the HALCA observing time is used by the mission-led VSOP Survey Program (Hirabayashi et al., 2000b). All-sky surveys are always of high importance, especially if done by a new instrument with unprecedented capabilities – the angular resolution in the case of VSOP. The major goal of the VSOP Survey is to collect data on sub-mas scale structural properties of a complete flux density-limited sample of 402 extragalactic radio sources at 5 GHz. The program was designed to use typically 3–4 ground-based radio telescopes during $\sim$ 1 orbital period of HALCA, and thus to make effective use of the in-orbit time when less ground resources are available. The Survey observations and the data reduction are still in progress. The data and results are being made

publicly available. One of the most important preliminary results is that the brightness temperature distribution of the sources shows a clear tail well above 10^{12} K (Hirabayashi et al., 2000b). As a part of a large collaborative effort, we participate in the data reduction of VSOP Survey observations (Lovell et al., 2004; Scott et al., 2004).

Prior to the launch of HALCA, 374 sources from the VSOP Survey list were imaged with the US Very Long Baseline Array (VLBA) at 5 GHz. The aim of this Pre-Launch Survey (VLBApls) was to select the most compact objects that are suitable for SVLBI observations (Fomalont et al., 2000). The VLBApls data base is also valuable in its own right, containing first-time VLBI images and improved positions for a number of radio sources.

4.4. Astrophysical studies

Our earliest VSOP project was the imaging of the core of the optically violently variable quasar 3C 446 in 1997. We found a core position offset of a few tenth of a mas, giving reasonable estimate of the equipartition magnetic field (~ 1 G) at 1 pc from the core (Paragi et al., 2000a). After investigating the compact core–jet region of the superluminal quasar 3C 216 at 5 GHz, a lower limit of 3 for the Doppler-boosting factor was estimated. The jet is oriented closer than 19° to our line of sight (Paragi et al., 2000b). The compact quasar 1546+027 is also a candidate for detecting apparent superluminal motion if we associate a new jet component seen in the 5 GHz VSOP image with an earlier total flux density outburst (Mosoni et al., 2002). Space VLBI observations at 5 GHz indicate misalignment between the mas- and sub-mas-scale jet of the quasar 0458−020 (Paragi et al., 2000c).

The highest rest-frame brightness temperature measured to date with VSOP is $T_{\rm B} > 5.8 \times 10^{13}$ K for the compact, violently variable BL Lac object AO 0235+164 (Frey et al., 2000a). This value was obtained at 5 GHz in 1999. Subsequent polarization-sensitive VSOP monitoring observations in 2001 and 2002 indicated that the source became larger – clearly resolved on grond–space baselines. At the same time, its total flux density decreased. AO 0235+164 was only weakly ($< 1\%$) polarized (Frey et al., 2002b).

Among the most distant ($z > 3$) radio-loud quasars, about 20 are bright enough to be imaged with VSOP. Due to the extremely high redhifts, mas-scale radio structures at emitted frequencies up to ~ 30 GHz can be studied. For steep-spectrum jet components, the high resolution of SVLBI is unique and cannot be supplemented with ground-based VLBI observations at higher frequencies (Gurvits, 2000). The quasar 2215+020 ($z = 3.57$) shows an unex-

pectedly prominent jet structute at 1.6 GHz, out to ~ 80 mas angular distance from the core. One of the jet components seen in the ground-only VLBI image is resolved out in the VSOP image, leading to an estimate of the transverse dimension of the jet, and the mass of the central black hole ($4 \times 10^9 M_\odot$; Lobanov et al., 2001). Based on the VSOP image taken at 5 GHz observing frequency, the jet in the quasar 1351−018 ($z = 3.71$) bends sharply ($\sim 130°$) between the sub-mas and 10-mas scales. Such jet misalignments are found in other high-redshift sources as well. Another 5-GHz VSOP image of 1351−018 made less than 3 years later with sub-mas resolution allowed us to identify a jet component within 1 mas from the core showing apparent superluminal motion of $\sim 9c$ (Fig. 1; Frey et al., 2002c). Other VSOP results on our sample of extremely high redshift quasars are published in Gurvits et al. (2000) and Frey et al. (2000b).

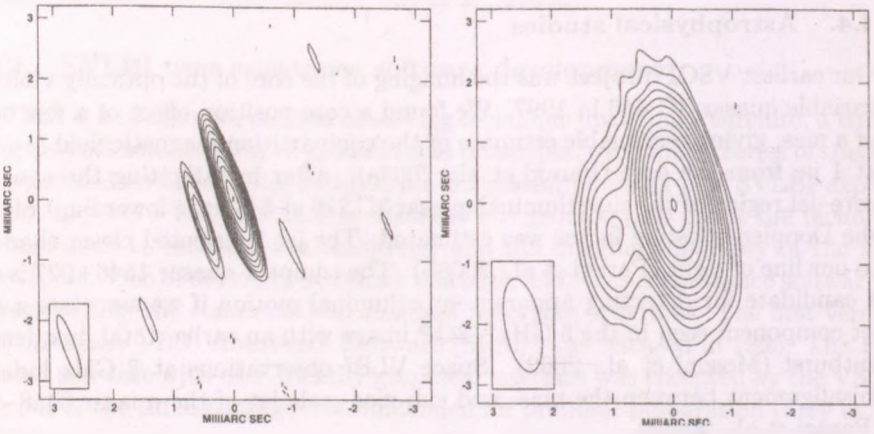

Figure 1: 5-GHz VSOP images of 1351−018 on 1998 Apr 12 (*left*) and 2001 Jan 23 (*right*). The pc-scale jet component to the south-east of the core has changed its position considerably. The restoring beam size is indicated in the bottom-left corners (Frey et al., 2002c).

5. Next-generation SVLBI satellites

The success of VSOP, in both technical and scientific senses, gives a good perspective for next-generation SVLBI missions. One of the key issues is how to increase the sensitivity in order to enable studying many more sources and astrophysical phenomena. In principle, this can be achieved in number of ways, by e.g. increasing the antenna collecting area, the down-link data rate (i.e. the

recording bandwidth), or reducing the receiver system temperature (cooled receivers). The increase in sensitivity with respect to VSOP is a factor of ~ 10 for the VSOP-2 mission being considered at JAXA in Japan (Hirabayashi, 2003). If successful, the earliest possible launch date for VSOP-2 is 2010. The VSOP-2 spacecraft would carry a 9-m antenna in an orbit characterised by 25 000 and 1000 km apogee and perigee heights, respectively. The planned frequency bands are 8, 22 and 43 GHz, in dual circular polarization. One promising possibility is to make VSOP-2 able to conduct phase-reference observations routinely. This technique is based on rapidly alternating (or simultaneous) observations of two sources close in direction. The phase fluctuations of mainly atmospheric origin are nearly the same for both sources. If one of the sources is bright and compact enough, the corrections can be determined. This method is useful for imaging the other source which may otherwise be too weak for self-calibration. With HALCA, which cannot switch rapidly between different sources, phase-reference observations were only possible with two objects close enough to lie within the primary beam of the 8-m antenna on board (Porcas et al., 2000).

The RadioAstron mission has been considered in Russia (formerly in the Soviet Union) for a long time (Kardashev, 1998). Due to the unfavourable economic circumstances, the construction and launch of the satellite was delayed several times. Now the projected launch time is early 2006. The spacecraft would observe at four frequency bands (0.3, 1.6, 5 and 22 GHz) with its 10-m diameter on-board radio telescope. Because of its extremely high and elongated orbit (9.5 days period, 350 000 km apogee height, 29 000 km perigee height), RadioAstron would primarily be a "discovery" mission to explore the most compact celestial radio sources.

Acknowledgments

The VSOP Project is led by the Japanese Aerospace Exploration Agency (JAXA) – formerly the Institute of Space and Astronautical Science (ISAS) – in cooperation with many organizations and radio telescopes around the world. During the last ten years, VSOP-related work at the FÖMI SGO was led by István Fejes with the participation of Gábor Heitler, László Mosoni, István Noszticzius, Zsolt Paragi and the author. I acknowledge the Bolyai Research Scholarship received from the Hungarian Academy of Sciences.

References

Bower, G.C., Backer, D.C. 1998, *ApJ* , 507, L117

Edwards, P.G., Giovannini, G., Cotton, W.D., et al. 2000, *PASJ* , 52, 1015

Fejes, I., Almár, I., Ádám, J., Mihály, Sz. 1986, Adv. Space Res., 6, 205

Fejes, I., Kawaguchi, N., Mihály, Sz. 1996, *Ap&SS* , 239, 275

Fomalont, E.B., Frey, S., Paragi, Z., et al. 2000, *ApJS* , 131, 95

Frey, S. 2004, in: *The Role of VLBI in Astrophysics, Astrometry and Geodesy*, ed. Mantovani, F., Kluwer, Dordrecht, in press

Frey, S., Fejes, I., Heitler, G., Noszticzius, I., Paragi, Z. 1998, in: *Radio Emission from Galactic and Extragalactic Compact Sources, IAU Coll. 164*, eds. Zensus, J.A., Taylor, G.B., Wrobel, J.M., ASP Conference Series 144, 399

Frey, S., Gurvits, L.I., Altschuler, D.R., et al. 2000a, *PASJ* , 52, 975

Frey, S., Gurvits, L.I., Schilizzi, R.T., et al. 2000b, in: *Proc. 5th European VLBI Network Symposium*, eds. Conway, J.E., Polatidis, A.G., Booth, R.S., Pihlström, Y., Onsala Space Observatory, 41

Frey, S., Meyer, U., Fejes, I., et al. 2002a, Adv. Space Res., 30, 357

Frey, S., Gurvits, L.I., Gabuzda, D.C., et al. 2002b, in: *Proc. 6th European VLBI Network Symposium*, eds. Ros, E., Porcas, R.W., Lobanov, A.P., Zensus, J.A., Max-Planck-Institute für Radioastronomie, Bonn, 91

Frey, S., Gurvits, L.I., Lobanov, A.P., et al. 2002c, in: *Proc. 6th European VLBI Network Symposium*, eds. Ros, E., Porcas, R.W., Lobanov, A.P., Zensus, J.A., Max-Planck-Institute für Radioastronomie, Bonn, 89

Gabuzda, D.C. 1999, New Astron. Rev., 43, 691

Gabuzda, D.C., Gómez, J.L. 2001, *MNRAS* , 320, L49

Gabuzda, D.C., Gómez, J.L., Agudo, I. 2001, *MNRAS* , 328, 719

Gurvits, L.I. 2000, *APRSV*, 151

Gurvits, L.I., Frey, S., Schilizzi, R.T., et al. 2000, Adv. Space Res., 26, 719

Hirabayashi, H. 1999, in: *Modern Radio Science 1999*, ed. Stuchly, M.A., Oxford University Press, 51

Hirabayashi, H. 2001, in: *Galaxies and Their Constituents at the Highest Angular Resolutions, IAU Symp. 205*, eds. Schilizzi, R.T., Vogel, S., Paresce, F., Elvis, M., Kluwer, Dordrecht, 422

Hirabayashi, H. 2003, Communications Research Laboratory Technology Development Center News, No. 23, CRL, Tokyo, 51

Hirabayashi, H., Hirosawa, H., Kobayashi, H., et al. 1998, Sci, 281, 1825

Hirabayashi, H., Edwards, P.G., Murphy, D.W. (eds.) 2000a, *Astrophysical Phenomena Revealed by Space VLBI*, Institute of Space and Astronautical Science, Sagamihara, Japan [*APRSV*]

Hirabayashi, H., Fomalont, E.B., Horiuchi, et al. 2000b, *PASJ* , 52, 997

Hirabayashi, H., Hirosawa, H., Kobayashi, H., et al. 2000c, *PASJ* , 52, 955

Junor, W., Biretta, J.A., Owen, F.N., Begelmann, M.C. 2000, *APRSV*, 13

Kardashev, N.S. 1998, Exp. Astron., 7, 329

Kellermann, K.I., Pauliny-Toth, I.I.K. 1969, *ApJ* , 155, L71

Kemball, A., Flatters, C., Gabuzda, D., et al. 2000, *PASJ* , 52, 1055

Kobayashi, H., Wajima, K., Hirabayashi, H., et al. 2000a, *PASJ* , 52, 967

Kobayashi, H., Shimoikura, T., Omodaka, T., Diamond, P.J. 2000b, *APRSV*, 109

Linfield, R.P., Levy, G.S., Edwards, C.D., et al. 1990, *ApJ* , 358, 350

Lister, M.L., Tingay, S.J., Murphy, D.W., et al. 2001a, *ApJ* , 554, 948

Lister, M.L., Tingay, S.J., Preston, A.R. 2001b, *ApJ* , 554, 964

Lobanov, A.P., Gurvits, L.I., Frey, S., et al. 2001, *ApJ* , 547, 714

Lobanov, A.P., Zensus, J.A. 2001, Sci, 294, 128

Lovell, J.E.J., Moellenbrock, G.A., Horiuchi, S., et al. 2004, *ApJ* , submitted

Mosoni, L., Frey, S., Paragi, Z., et al. 2002, in: *Proc. 6th European VLBI Network Symposium*, eds. Ros, E., Porcas, R.W., Lobanov, A.P., Zensus, J.A., Max-Planck-Institute für Radioastronomie, Bonn, 97

Murphy, D.W., Preston, R.A., Hirabayashi H. 2003, New Astron. Rev., 47, 633

Paragi, Z., Frey, S., Fejes, I., et al. 2000a, Adv. Space Res., 26, 697

Paragi, Z., Frey, S., Fejes, I., et al. 2000b, *PASJ* , 52, 983

Paragi, Z., Frey, S., Fejes, I., Murphy, D.W. 2000c, in: *Proc. 5th European VLBI Network Symposium*, eds. Conway, J.E., Polatidis, A.G., Booth, R.S., Pihlström, Y., Onsala Space Observatory, 43

Pearson, T.J., Readhead, A.C.S. 1988, *ApJ* , 328, 114

Piner, B.G., Edwards, P.G., Wehrle, A.E., et al. 2000, *ApJ* , 537, 91

Porcas, R.W., Rioja, M.J., Machalski, J., Hirabayashi, H. 2000, *APRSV*, 245

Readhead, A.C.S. 1994, *ApJ* , 426, 51

Scott, W.K., Fomalont, E.B., Horiuchi, S., et al. 2004, *ApJS* , submitted

Shen, Z.-Q., Edwards, P.G., Lovell, J.E.J., et al. 1999, *PASJ* , 51, 513

Sovers, O.J., Fanselow, J.L., Jacobs C.S. 1998, Rev. Mod. Phys., 70, 1393

Sudou, H., Taniguchi, Y., Ohyama, Y., et al. 2000, *PASJ* , 52, 989

Tingay, S.J., Preston, R.A., Lister, M.L., et al. 2001, *ApJ* , 549, L55

Ulvestad, J.S. 1999, in: *Synthesis Imaging in Radio Astronomy II*, eds. Taylor, G.B., Carilli, C.L., Perley R.A., ASP Conference Series 180, 513

A Close Look on Thousands of Faint Radio Sources: Pilot Results of the DEVOS Survey

L. Mosoni[1,2] and S. Frey[2]
[1]Konkoly Observatory of the Hungarian Academy of Sciences
P.O. Box 67, H-1525 Budapest, Hungary
[2]FÖMI Satellite Geodetic Observatory
P.O. Box 585, H-1592 Budapest, Hungary
E-mail: [1]mosoni@konkoly.hu, [2]frey@sgo.fomi.hu

Abstract

We briefly demonstrate the pilot results of the Deep Extragalactic VLBI-Optical Survey (DEVOS). Our ultimate aim is to eventually construct a large sample of compact sources ($\sim 10^4$ objects) up to two orders of magnitude fainter than studied in VLBI surveys until now. This would lead to an unprecedented data base for various astrophysical and cosmological studies. The first global VLBI observations were successfully conducted in May 2002. We selected sources from the VLA FIRST catalogue, detected with MERLIN in the areas expected to be surveyed by the Sloan Digital Sky Survey with spectroscopic redshift measured. By means of observations of mJy-level sources in this pilot project we want to estimate the outcome of and the resources needed for a full-scale DEVOS project and obtain results valuable in their own right.
KEYWORDS: *techniques: interferometric – radio continuum: galaxies – galaxies: active – quasars: general – cosmology – astrometry*

1. Introduction

Very Long Baseline Interferometry (VLBI) imaging surveys consist of data – milli-arcsecond (mas) resolution images – of thousands of extragalactic radio sources, providing invaluable information on the nature of physical phenomena in the close vicinity of active galactic nuclei (AGNs). Completed and ongoing VLBI surveys are reviewed by Gurvits (2003).

At present, VLBI surveys – although conducted using different VLBI arrays and different sample selection criteria – are in fact flux-density limited. The lower limit of the flux density is ~ 100 mJy (see Fig. 1). In order to reach conclusive results on the intrinsic properties of sources as well as possible imprints of cosmological models on the source structures, one has to match in luminosity sources detected and imaged with VLBI at low redshift (e.g. $z < 0.1$) with those at high redshift ($z > 1$). This requires VLBI study of $z > 1$ sources

of luminosities as low as $10^{23} - 10^{25}$ W/Hz. It corresponds to mJy-level flux densities. This work can be carried out with the modern VLBI arrays using innovative observing techniques. It seems now feasible to bring the total number of VLBI-imaged extragalactic sources to 10–20 000.

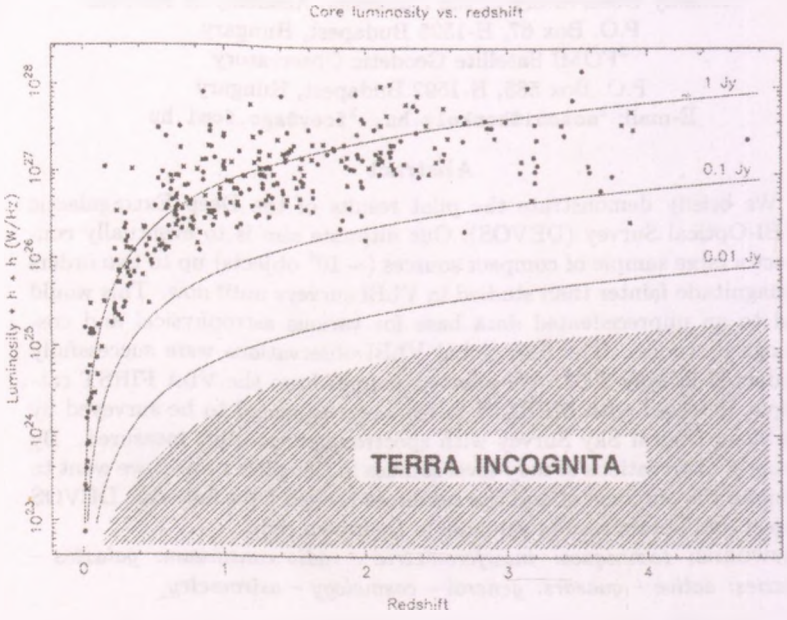

Figure 1: Luminosities (Lh^2) of 330 sources observed at $\nu = 5$ GHz as a function of redshift (Gurvits et al., 1999). Solid lines show luminosities of sources with particular flux densities ($S = 1, 0.1$ and 0.01 Jy), calculated assuming spectral index $\alpha = 0$ where $S \sim \nu^{\alpha}$.

The applications of the data from VLBI surveys include astrophysical and cosmological studies of powerful compact radio sources. With the advent of large optical surveys (e.g. the SDSS[1], e.g. Abazajian et al., 2003), the problem of missing optical identifications (morphological type, redshift) will be practically eliminated within the next several years. We can select sources for VLBI observations from celestial areas covered by the optical surveys.

[1]Sloan Digital Sky Survey, `http://www.sdss.org/`

1.1. Potential applications of the DEVOS data base

It is hard to predict in details what kind of scientific insight will become possible with such a dramatic increase in the number of sources. One can briefly mention a few possible applications.

- The *cosmological evolution* of radio-loud active galaxy population could be studied with the data base in the light of orientation-dependent unification models of powerful radio sources (cf. Wall & Jackson, 1997). Note that we do not introduce initial sample selection criteria based on source spectral index.

- *Morphological classification* of low-luminosity sources could be compared with that of their high-luminosity counterparts.

- The dependence of the source *compactness* on the luminosity could be verified on mas angular scales.

- The data base would be useful to study *gravitational lensing* since the survey of weak compact radio sources probes the parent population of lensed objects being found in e.g. the CLASS[2] of $\sim 10^4$ weak flat-spectrum sources (e.g. Myers et al., 2003).

- Several thousand mas-scale images of optically identified extragalactic sources with known redshifts are needed to bring about conclusive estimates of fundamental *cosmological parameters*, such as the density parameters Ω_m and Ω_Λ (Dabrowsky et al., 1995). As shown by Gurvits (1994), such the cosmological test could be conducted with a very limited VLBI (u, v)-coverage using visibility data, and does not require high dynamic range imaging.

- Getting VLBI images of $\sim 10^4$ extragalactic sources will be an essential supplement to and basis for future development of the *astrometric* VLBI data bases (e.g. Johnston et al., 1995) and the next generation space-borne optical astrometry missions (e.g. Gaia[3]).

[2] Cosmic Lens All-Sky Survey, `http://www.nrao.edu/~smyers/class.html`

[3] `http://astro.estec.esa.nl/GAIA/`

2. Observations

2.1. The technique: phase-referencing

Radio sources can be observed with mas-scale angular resolution at GHz frequencies with VLBI. A brief introduction to VLBI can be found in Frey (2004, these proceedings) and references therein. The "standard" VLBI observations are not sensitive enough to detect mJy-sources which we are interested in. The signal-to-noise ratio achievable is limited by the coherent integration time spent on the target source. Phase-referencing is a tool to image faint radio sources with VLBI. The sensitivity is increased by extending the signal coherence time using regularly interleaving observations of a nearby bright and compact calibrator source. The delay, delay-rate and phase solutions obtained for the reference source are interpolated and applied for the faint target source, thus removing most of the phase errors introduced by geometric, ionospheric, atmospheric and instrumental effects (e.g. Beasley & Conway, 1995).

The phase-referencing technique can also be applied for surveying a number of faint objects. It can be done in the vicinity of an adequately chosen reference source. In this case we reverse the usual logic of first selecting the target object and then looking for a suitable reference source. If a calibrator is selected first, and the potential targets populate its close vicinity densely enough, then the single calibrator can serve as phase-reference source for all of the targets. This method was introduced and demonstrated first by Garrington et al. (1999). In this case different target sources are observed between the regular observing intervals spent on the reference source. Each target source is observed several times during the observations which improves the (u, v)-coverage and hence the image quality. Potential calibrator sources are relatively common and well distributed, therefore practically the whole sky can be surveyed using similar method.

2.2. Initial target selection

In DEVOS, there are two types of considerations for source selection. First of all, we select sources that are bright and compact enough to be eventually detectable with VLBI with extremely high angular resolution, within a reasonable integration time – since the observing resources are limited. Therefore we applied the following selection criteria:

1. the target source is within 2° angular separation from the reference source selected

2. the integrated flux density at 1.4 GHz is $S_{1.4} > 30$ mJy using the VLA FIRST[4] catalogue data (White et al., 1997)

3. the angular size of the source is $\theta < 5''$ (i.e. unresolved in the FIRST)

Note that no criterion was set based on the spectra: radio sources with either steep or flat spectra could be selected as well.

Another important consideration follows from the fact that the sources must eventually have optical identification and redshift measured. This can be achieved by selecting sources from the fields which are surveyed in the SDSS. According to recent studies, optical counterparts of at least 30% of all FIRST radio sources are expected to be identified by SDSS (Ivezic et al., 2002). In our case this rate is likely to be higher because we select relatively bright and compact FIRST sources. The source sample for the DEVOS pilot observations described here was defined in 2000, before any of the SDSS data releases. Two bright quasars were selected as calibrators, keeping the potential survey areas in SDSS in mind. One of them (J1257+3229) is close to the North Galactic Pole (NGP), another quasar is an equatorial source (J1549+0237). The pilot DEVOS samples consist of 47 and 42 FIRST sources around the reference quasars J1257+3229 and J1549+0237, respectively.

2.3. MERLIN observations to refine the sample

After selecting the reference sources and the possible targets for VLBI observations, it was necessary to observe these sources at an intermediate angular resolution, to identify those that may be suitable for VLBI observations. Both the NGP and the equatorial fields have been observed with MERLIN[5] at 5 GHz in March 2001, with angular resolution of $\sim$ 50 mas (Frey & Mosoni, 2002). In the NGP field, 37 out of the total 47 sources have been detected and subsequently proposed for VLBI observations. In the case of these objects, the MERLIN image peak brightness exceeded $\sim$ 2 mJy/beam. The equatorial sample will be proposed for VLBI observations later.

2.4. VLBI observations

The 5-GHz global VLBI observations of the previously "filtered" NGP sample involving 19 antennas of the European VLBI Network[6] (EVN) and the US

[4]Faint Images of the Radio Sky at Twenty-centimeters, `http://sundog.stsci.edu/`
[5]Multi-Element Radio Linked Interferometer Network, `http://www.jb.man.ac.uk/merlin/`
[6]`http://www.evlbi.org/`

NRAO Very Long Baseline Array[7] (VLBA) took place on May 30, 2002. The total observing time was 21 hours (with 7 hours overlap between the two arrays). The participating radio telescopes were Effelsberg (Germany), Westerbork (the Netherlands), Jodrell Bank MK2 (UK), Medicina, Noto (Italy), Torun (Poland), Onsala (Sweden), Sheshan, Nanshan (China) from the EVN, and all the ten stations of the VLBA.

Each program source was observed in 10 scans of 2.5 min duration. Taking the slewing times of the antennas into account, it corresponds to 20 min accumulated observing time per source. In each observing loop (6 min), two program sources were observed, chosen with the minimum angular distance between them. The reference source was observed for 1 min. The estimated thermal noise (1σ) is ~ 100 μJy/beam.

3. Data reduction and imaging

The NRAO AIPS package (Diamond, 1995) was used for the initial data calibration. The reference source J1257+3229 was imaged first. The structural effects from the reference source were corrected for the target sources.

Initially, all target sources were imaged and investigated individually. Standard imaging procedures could be carried out. However, with the aim of observing and imaging thousands of radio sources in the full DEVOS survey, we certainly need an automated (fast and objective) procedure to prepare the source VLBI images. Relatively simple scripts have been developed for this task using the Caltech DIFMAP package (Shepherd et al., 1994). By the comparison with the images obtained "manually", it turned out that these scripts "detect" the same sources with qualitatively similar structures, peak brightnesses and noise level. The "manual" and "automatic" images of one of the sources can be compared in Fig. 2 as an example. The brightest (~ 100 mJy) and the faintest (mJy-level) sources demand somewhat different data reduction steps.

4. Results and discussion

17 sources were detected and imaged with VLBI from the originally selected 47 FIRST sources. Two interesting examples – the brightest and faintest sources in the FIRST sample – are shown in Fig. 3 and 4. The typical image noise level (3σ) achieved is $0.7-1$ mJy/beam, three times higher than the theoretical estimate. The distribution of peak brightnesses of the detected sources:

[7] http://www.aoc.nrao.edu/vlba/html/

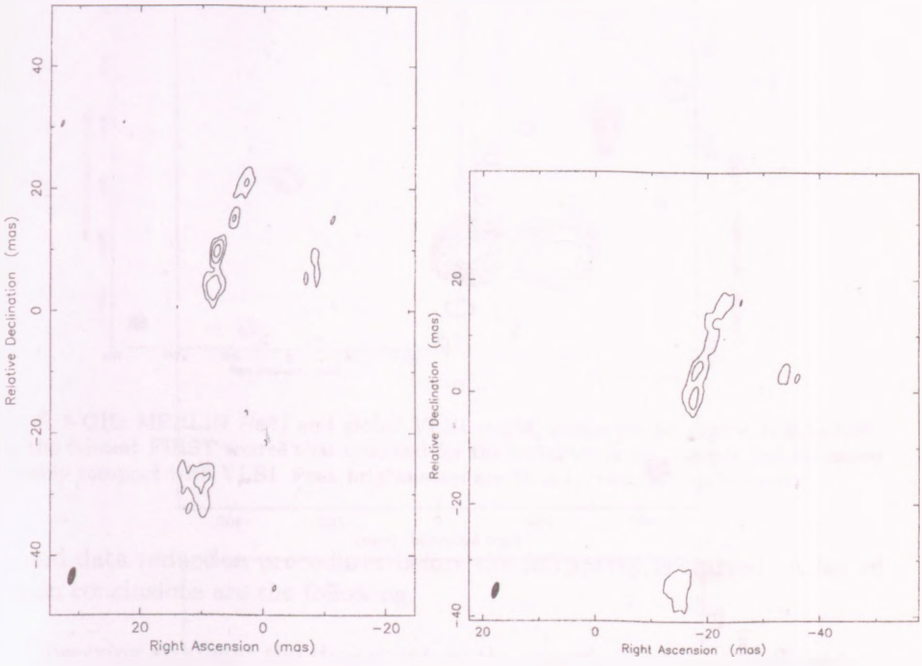

Figure 2: Comparison of the VLBI images of the source J1254+3118. The *left* image was obtained by "manual" data reduction, the *right* one was produced by an automated process. The contours are drawn at 0.4, 0.8 and 1.6 mJy/beam on the *left* image (peak brightness 2.2 mJy/beam) and at 0.6 and 1.2 mJy/beam on the *right* image (peak brightness 1.9 mJy/beam).

- 2 bright (80 and 120 mJy/beam)
- 6 with $\sim 10 - 20$ mJy/beam
- 5 with $\sim 1 - 4$ mJy/beam
- 4 marginally detected $(3 - 4\sigma)$ with ≤ 1 mJy/beam peak brightness

The initial analysis reveals a large variety of radio morphologies, ranging from those typical for compact AGNs through "classical" core–jet structures to well-resolved sources. Cross-identification of the pilot DEVOS sample with optical data from the SDSS will be done in the near future.

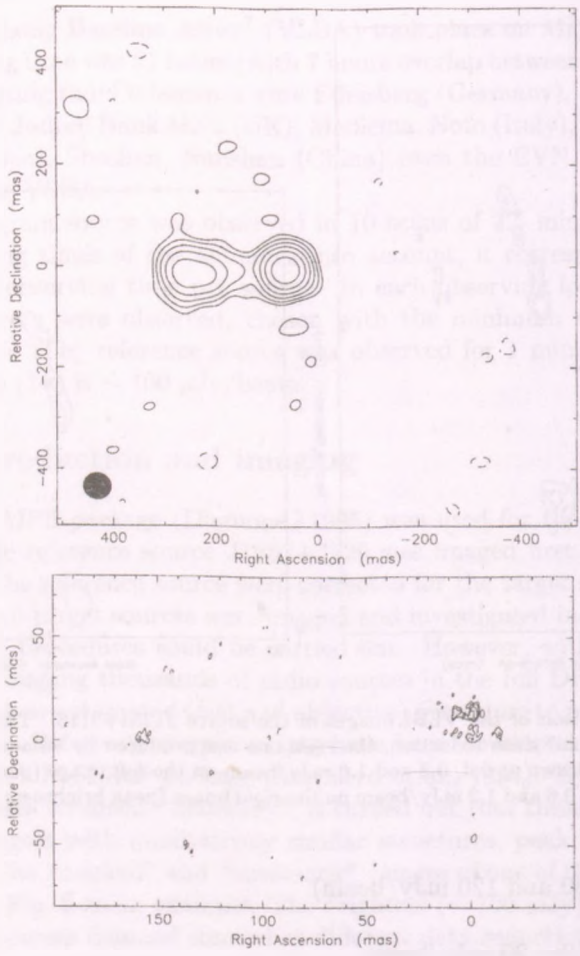

Figure 3: 5-GHz MERLIN (*top*, peak brightness 77.4 mJy/beam) and global VLBI (*bottom*, peak brightness 4.4 mJy/beam) image of the source J1259+3356. It was the strongest FIRST source in the pilot DEVOS sample (S = 478 mJy flux density). The source is resolved on MERLIN angular scales. Both components can be seen on the VLBI image.

5. Conclusions

The primary goal of the DEVOS pilot project was to verify and adjust sample selection criteria, observing strategies (network, observing time, image noise,

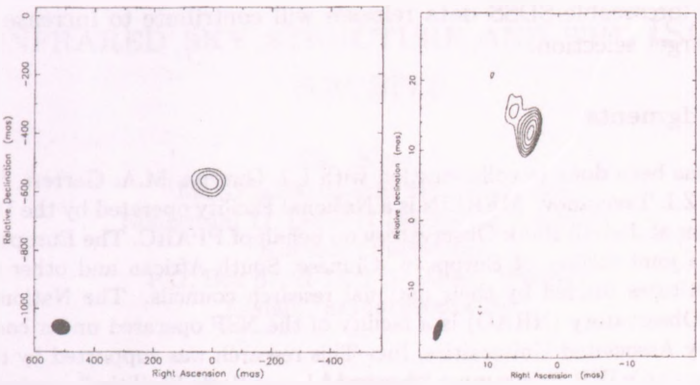

Figure 4: 5-GHz MERLIN *(left)* and global VLBI *(right)* image of the source J1303+3207. It was the faintest FIRST source that qualified for the inclusion in our sample and appeared remarkably compact with VLBI. Peak brightnesses are 13 mJy/beam in both images.

etc.) and data reduction procedures before the full survey is started. A list of our main conclusions are the following.

- Observing strategy: the time spent on the sources (20 min) is sufficient to image the compact objects with an image noise of 1 mJy/beam or better. The phase-referencing technique was successful. The 6-min long observing cycles are adequate at 5 GHz. Two target sources can be observed between the regular observations of the calibrator source.

- Source selection criteria (VLA FIRST): we might consider to set a lower flux density limit to pick up faint but very compact objects like J1303+3207 (see Fig. 3). However, this would demand more detailed filtering not to increase the ratio of non-detections and use the MERLIN and VLBI network resources effectively.

- Filtering criteria (MERLIN): based on the eventual VLBI detection ratio, we may want to set a higher peak brightness limit for filtering in the future (e.g. 4 mJy/beam).

- Data reduction: it can be done with an automated procedure or "pipeline".

After the completion of this pilot study, we will initiate further VLBI and target-finding filter observations in the framework of the DEVOS survey. The

recent and foreseeable SDSS data releases will contribute to increase the efficiency of target selection.

Acknowledgments

This work has been done in collaboration with L.I. Gurvits, M.A. Garrett, S.T. Garrington and Z.I. Tsvetanov. MERLIN is a National Facility operated by the University of Manchester at Jodrell Bank Observatory on behalf of PPARC. The European VLBI Network is a joint facility of European, Chinese, South African and other radio astronomy institutes funded by their national research councils. The National Radio Astronomy Observatory (NRAO) is a facility of the NSF operated under cooperative agreement by Associated Universities, Inc. This research was supported by the European Commission's IHP Programme "Access to Large-Scale Facilities", under contract no. HPRI-CT-1999-00045 and the OTKA project no. T046097. SF acknowledges the Bolyai Research Scholarship received from the Hungarian Academy of Sciences.

References

Abazajian, K., Adelman-McCarthy, J.K., Agüeros, M.A., et al. 2003, *AJ* , 126, 2081

Beasley, A.J., Conway, J.E. 1995, in: Zensus et al. (1995), 327

Dabrowski, Y., Lasenby, A., Saunders, R. 1995, *MNRAS* , 277, 753

Diamond, P.J. 1995, in: Zensus et al. (1995), 227

Frey, S., Mosoni, L. 2002, Publ. Astron. Dept. Eötvös Univ., Vol. 12, 53

Garrington, S.T., Garrett, M.A., Polatidis, A. 1999, New Astron. Rev., 43, 629

Gurvits, L.I. 1994, *ApJ* , 425, 442

Gurvits, L.I., Kellermann, K.I., Frey, S. 1999, A&A, 342, 378

Gurvits, L.I. 2003, in: *AGN Surveys*, eds. Green, R.F., et al. ASP Conf. Ser. 284, San Francisco, 137

Ivezic, Z., Menou, K., Knapp, G.R., et al. 2002, *AJ* , 124, 2364

Johnston, K.J., Fey, A.L., Zacharias, N., et al. 1995, *AJ* , 110, 880

Myers, S.T., Jackson, N.J., Browne, I.W.A., et al. 2003, *MNRAS* , 341, 1

Shepherd, M.C., Pearson, T.J., Taylor, G.B. 1994, BAAS, 26, 987

Wall, J.V., Jackson, C.A. 1997, *MNRAS* , 290, L1

White, R.L., Becker, R.H., Helfand, D.J., Gregg, M.D. 1997, *ApJ* , 475, 47

Zensus, J.A., Diamond, P.J., Napier, P.J. (eds.) 1995, *Very Long Baseline Interferometry and the VLBA*, ASP Conf. Ser. 82, San Francisco

The infrared sky structure and the ISO Data Archive

Cs. Kiss[1,2]
[1]Max-Planck-Institut für Astronomie
Königstuhl 17, D-69117 Heidelberg, Germany
[2]Konkoly Observatory
P.O. Box 67, H-1525 Budapest, Hungary
E-mail: [1]pkisscs@mpia.de

Abstract

ISO observations opened a new window to the studies of the infrared sky background, providing measurements surpassing IRAS in wavelength coverage, spatial resolution, number of photometric bands and spectroscopic capabilities. These measurements - which will remain unique for decades - provide an excellent data base to investigate the physical parameters, dust composition and spatial structure of the astrophysical components of the infrared sky background. In this proceeding I review the ISO Data Archive-based results and future capabilities of Zodiacal Light, Galactic cirrus and extragalactic background studies.
Keywords: *diffuse radiation, infrared*

1. Introduction

The ISOPHOT photopolarimeter (Lemke et al., 1996) was one of the four instruments on-board the ISO satellite (Kessler et al., 1996). Six years after the active measurement phase, the data collected by ISOPHOT and other ISO instruments is stored in the publicly available ISO Data Archive[1], which contains all raw and fully processed scientific, calibration, engeneering, up- and downlink data. It includes ~30000 individual scientific observations and ~110000 observations in the ISO parallel and serendipitious modes, which can be accessed through a user-friendly java interface.

In this proceeding I present our results and discuss the future possibilites on one selected topic, which was extensively studied by ISO, the infrared sky background.

[1] *http://www.iso.vilspa.esa.es/ida*

2. The infrared sky background

The definition of the background always depends on the basic characteristics of the actual measurement (wavelength, detector, configuration, etc.). However, there are two fundamental classes: a background can be intrinsically diffuse (e.g. the gas clouds of the Milky Way) or made of accumulated light of unresolved sources (e.g. the Mikly Way on the night sky for human's eye). In the former case the background remains diffuse regardless the resolution, while in the latter case higher resolution reduces the number of sources in the background (and therefore causes a lower background level as well). The background - either intrinsically diffuse or made of unresolved sources - limits the detectability of faint sources in two ways. First, the sources must be brighter than the limit set by the background ($S > S_{lim} = k \times \sigma_f$, photometric criterium) and, second, two sources (fulfilling the photometric criteria) must be sufficiently far from each other to be distinguishable ($\theta > \theta_{min}$, source density criterium, see Fig. 2.).

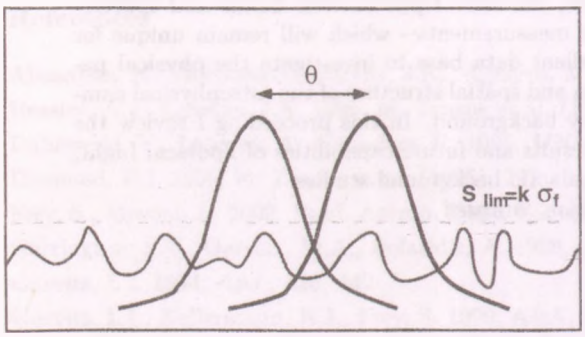

Figure 1: Schematic representation of the photometric and source density criteria for the detection of faint sources

The infrared background can be characterized by the following properties:

- absolute surface brightness
- spatial structure (or fluctuations). This is closely realted to the 'physics' of the background. There are to two main points-of-view one can consider the fluctuations:
  - Fourier power spectrum
  - confusion noise, i.e. uncertainty in the determination of the point source flux due to the unknown background. The analysis of the confusion noise at a specific spatial separation is equivalent to the selection of one spatial frequency from the power spectrum.

- colour (spectral energy distribution)

The sky background has several components, and it is a very challenging task to separate them. In Fig. 2 we present the contribution of different sky background components (without the extragalactic background, see later) from UV (0.1 μm) to $\sim 10^4$ μm (Leinert et al., 1998).

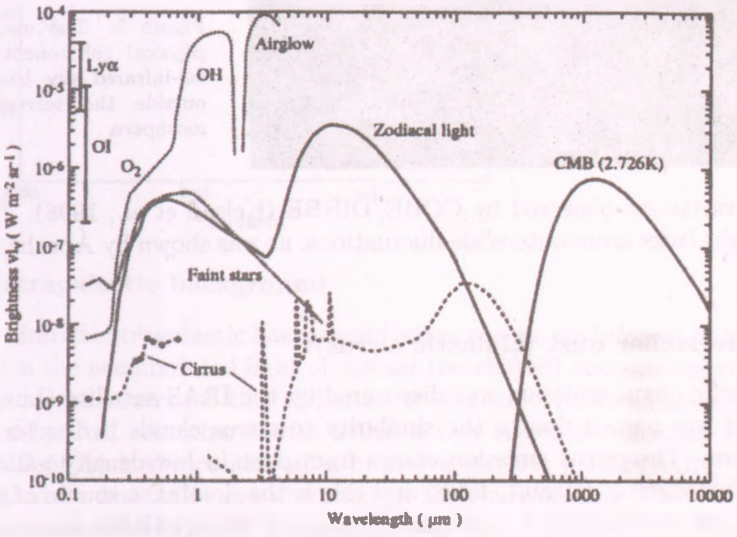

Figure 2: Sky brightness outside the lower terrestrial atmosphere at high ecliptic and galactic latitudes (from Leinert et al., 1998)

In the far-infrared the main astrophysical components of the sky background outside Earth's atmosphere are arisen from the following sources discussed below (see also Fig. 3).

2.1. Interplanetary dust (IPD, Zodiacal Light)

This component is a thermal ($T \approx 270$ K) emission from small, solid particles in the inner Solar System. The particles are originated from comets, collisions of asteroids and from interstellar dust. The IPD forms a flattened, lenticulary shaped dust cloud around the ecliptic plane, extending to the orbit of Jupiter. There are some local density enhancements due to comets and asteroids (Reach et al., 1997, Reach et al., 2000). The best IPD cloud model was derived from

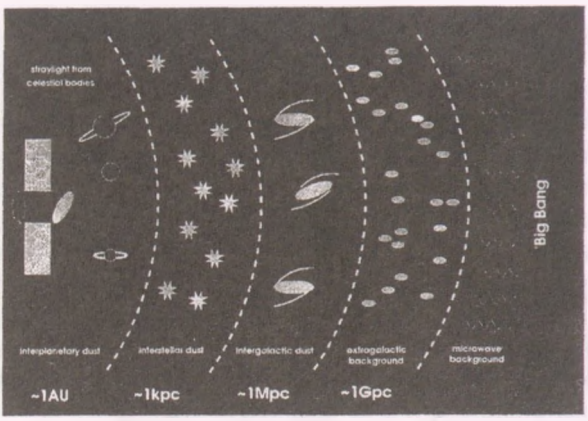

Figure 3: The main astrophysical components of the far-infrared sky background outside the terrestrial atmoshpere

seasonal variations observed by COBE/DIRBE (Kelsall et al., 1998). The Zodiacal Light lacks arcminute-scale fluctuations, as was shown by Ábrahám et al. (1997).

2.2. Interstellar dust (Galactic cirrus):

The Galactic cirrus emission was discovered by the IRAS satellite (Low et al., 1984) and was named due to the similarity to cirrus clouds in the terrestrial atmosphere. The cirrus emission comes from dust in low-density galactic HI coulds (Boulanger & Pérault, 1998) and this is the dominant source of the sky structure for wavelengths $\lambda > 60\,\mu$m. Based on COBE/DIRBE measurements a typical colour temperature of $T_d \approx 18\,$K has been derived, assuming $\beta = 2$ emissivity law ($F_\nu \propto \nu^2 B_\nu(T)$, Lagache et al., 1998).

The spatial structure of the cirrus emission is similar to that of a fractal. It can be described by a power-law power spectrum $P = P_0(f/f_0)^\alpha$. The first investigation of this structure was carried out by Gautier et al. (1992) based on IRAS 100 μm scans and resulted in $\alpha \approx -3$ (see Fig. 4).

2.3. Intergalactic dust

The first evidence for the existence of thermal dust distributed in the hot X-ray emitting plasma in the intergalactic medium of the Coma cluster was presented by Stickel et al. (1998). However, this emission was quite faint, despite the ongoing merging activity. Extrapolations based on this result suggest a negligible contribution to the infrared sky background from the intergalactic dust in the Local Group.

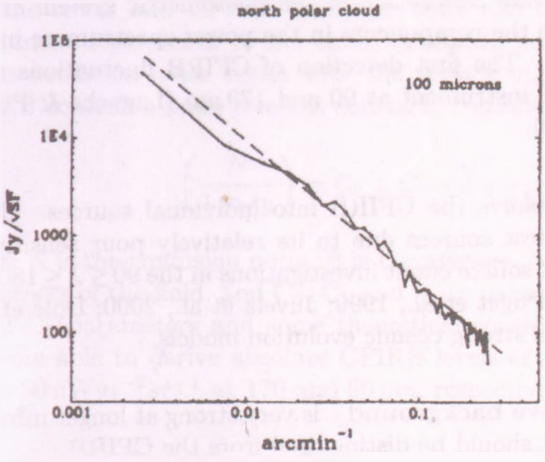

Figure 4: Annular power spectrum of the North Polar Cloud based on IRAS 100 μm scans (Gautier et al., 1992).

2.4. Extragalactic background

The far-infrared extragalactic background or the cosmic far-infrared background (CFIRB) is the accumulated light of distant (redshifted) and unresolved galaxies. The major source of this radiation is starlight, which is absorbed and re-emitted by dust. Another source which is very significant at some IR wavelengths (10–50%) is the contribution of active galactic nuclei (AGN). The X-rays emitted by the AGN are absorbed in the sourrounding dust thorus and re-emitted in the far-infrared. Mainly sources with $1 \leq z \leq 2$ contribute to the CFIRB, around the maximum of the global star-formation rate in the history of the Universe (see Hauser & Dwek, 2001, for a review). The CFIRB can be detected in three main ways.

Direct measurements of the CFIRB can be performed via very precise absolute surface brightness photometry. This is a very challenging task, mainly due to technical issues. This was done so far only using COBE/DIRBE measurements, after a careful removal of the Zodiacal Light and Galactic cirrus contributions (Hauser et al., 1998).

CFRIB fluctuations show a Poissonian spatial distribution for spatial frequencies $f \geq 1'$. Below this scale the clustering of galaxies becomes dominant resulting in a deviation from the Poissonian (Bond et al., 1986). Detection of CFIRB fluctuations are easier than the direct measurements since these do

not require the precise absolute calibration of the photometric system and it is relatively easy to separate the components in the power spectrum or in the confusion noise distribution. The first detection of CFIRB fluctuations were performed by the ISOPHOT instrument at 90 and 170 μm (Lagache & Puget, 2000; Matsuhara et al., 2000).

Source counts aims to resolve the CFIRB into individual sources. IRAS could only detect the brightest sources due to its relatively pour sensitivity. ISO performed extensive FIR source count investigations in the $90 \leq \lambda \leq 180\,\mu$m range (Kawara et al., 1998; Puget et al., 1999; Juvela et al., 2000; Dole et al., 2001). These results favoured strong cosmic evolution models.

Since the cosmic microwave background is very strong at longer infrared wavelengths ($\lambda \geq 300\,\mu$m), it should be distinguised from the CFIRB.

3. Results

3.1. Separation of CFIRB and cirrus confusion noise

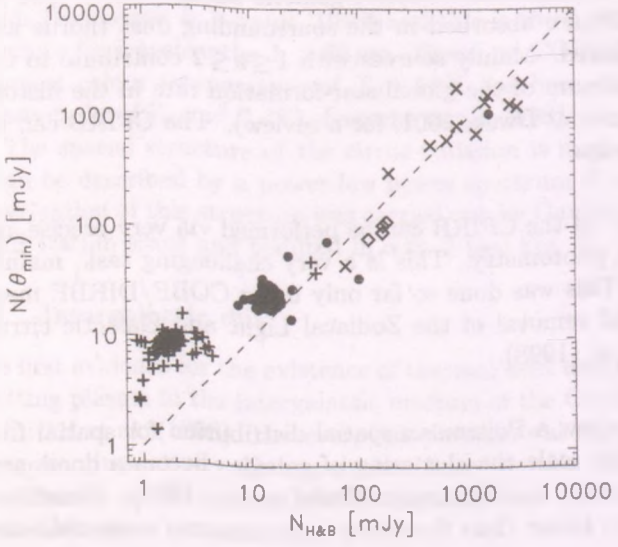

Figure 5: confusion noise at the resolution limit determined by our study as compared to the prediction of Helou & Beichman (1990). The deviation at low confusion noise values is caused by CFIRB fluctuations.

The cirrus and CFIRB confusion noise can be separated via the surface brightness dependence of the cirrus confusion noise (Kiss et al., 2001). The cirrus confusion noise scales with the surface brightness of the field, but the CFIRB confusion noise remains constant, regardless of the field brightness:

$$\left(\frac{N}{1\,mJy}\right) = C_0 + C_1 \times \left\langle \frac{B}{1\,MJysr^{-1}} \right\rangle^{\eta} \quad (1)$$

where N is the confusion noise, B is the average, zodiacal light removed surface brightness of the field, and C_0, C_1 and η are parameters to be fitted. Using the fitted C_0 parameters and some theoretical considerations (Bond et al., 1986) we were able to derive absolute CFIRB levels of $\nu I_\nu = 14 \pm 3\,\mathrm{nW}m^{-2}sr^{-1}$ and $\nu I_\nu = 30\,\mathrm{nW}m^{-2}sr^{-1}$ at 170 and 90 μm, respectively (the 90 μm value is upper limit).

3.2. Small scale structure of the Galactic cirrus emission

The spectral index α of the galactic cirrus emission is generally believed to be constant and wavelength-independent (Gautier et al., 1992; Helou & Beichman, 1990). We examined the Fourier power spectrum characteristics of cirrus structures in 13 sky fields with faint to bright cirrus emission observed with ISOPHOT in the 90–200 μm wavelength range in order to study the possible variations of α. We found that α varies from field to field with $-5.3 \leq \alpha \leq -2.1$. It also depends on the absoulte surface brightness and on the hydrogen column density. The results are presented in Fig. 6. The spectral indices of the same sky region were found to be different for different wavelengths as well (see Fig. 7). Longer wavelength measurements show steeper power spectra. This can be explained by the presence of dust at various temperatures, in particular of a cold, extended component. For the faintest areas of the far-infrared sky we derived a wavelength-independent spectral index of $\alpha = -2.3 \pm 0.6$ for the cirrus power spectrum. This is a precondition for the proper disantanglement of the cirrus foreground of the CFIRB.

3.3. Cirrus confusion noise redictions for current and future far-infrared space missions

Based on an extensive database of ISOPHOT measurement in the $90\,\mu\mathrm{m} \leq \lambda \leq 200\,\mu$m wavelength range we investigated the dependence of the confusion noise on the measurement configuration. We estimated the expected confusion noise

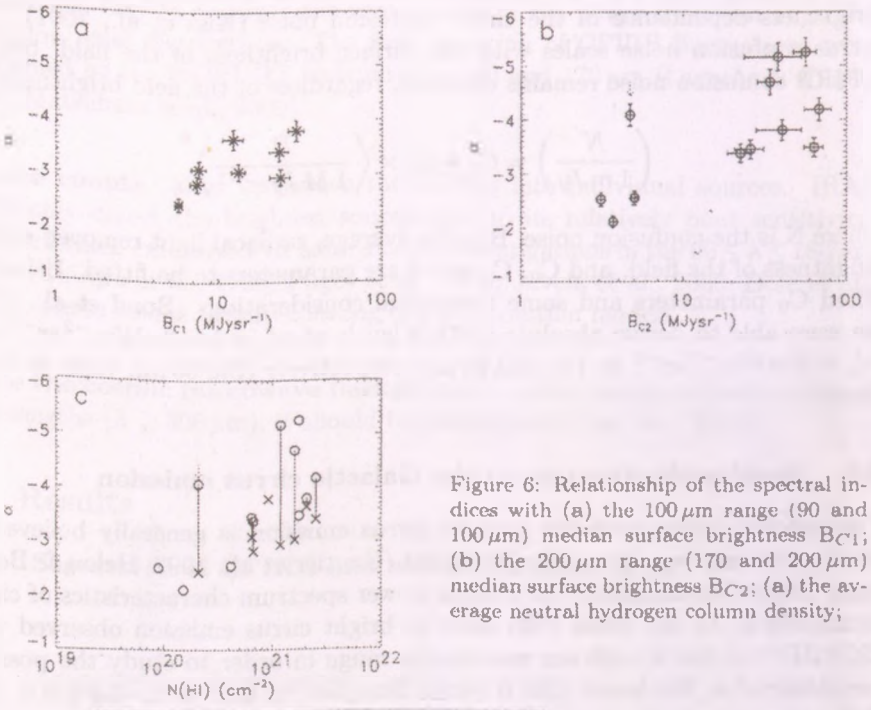

Figure 6: Relationship of the spectral indices with (a) the 100 μm range (90 and 100 μm) median surface brightness B_{C1}; (b) the 200 μm range (170 and 200 μm) median surface brightness B_{C2}; (a) the average neutral hydrogen column density;

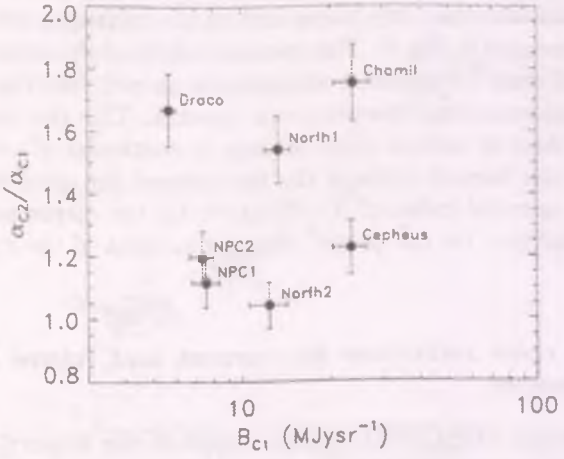

Figure 7: Ratio of the spectral indices of the sky regions measured in both a long (170–200 μm) and a short (90–100 μm) wavelength filter versus the averge surface brightness in the C100 filter B_{C1}.

at a specific surface brightness level for each possible measurement configura-

tion. We also found, that the favorable measurement mode is the mini-map (see Laureijs et al., 2003) for photometry of individual point sources. Fluxes of (point) sources in larger maps can be derived by the lowest confusion noise contamination using annular apertures ('aperture photometry'). This mode significantly reduces the confusion noise relative to flux extraction methods similar to triangular or rectangular chopping (Laureijs et al., 2003). A synthesis of these ISOPHOT results and the analysis of simulated fractal maps seen by various instruments provides cirrus confusion noise predictions for the current and future infrared space instruments (Spitzer/MIPS, ASTRO-F/FIS, Herschel/PACS). An example comparing the capabilities of four FIR instruments is presented in Fig. 8 (Kiss et al., 2004, in prep.).

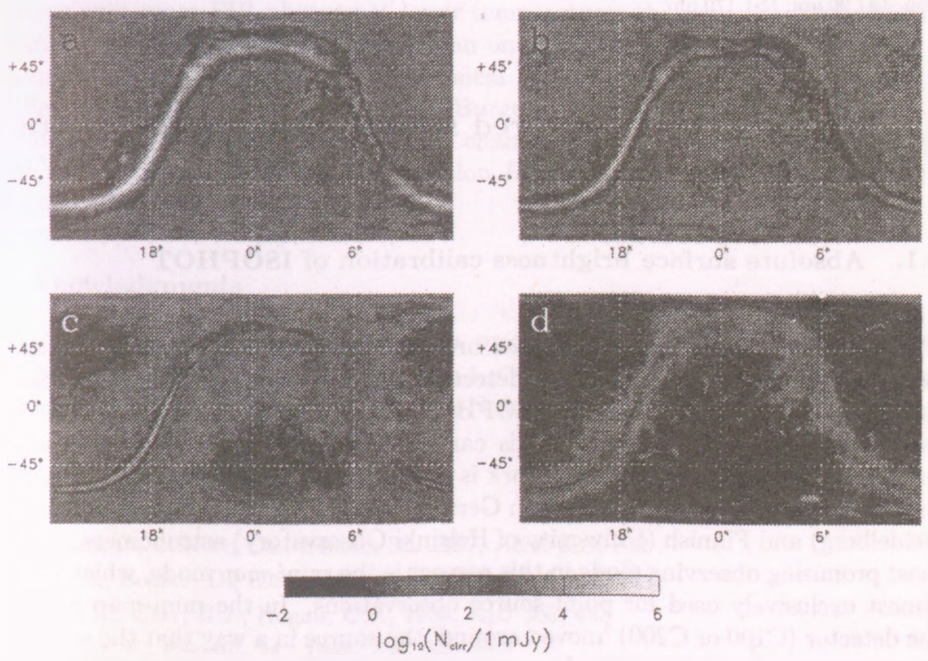

Figure 8: Comparison of predicted cirrus confusion noise levels of four different FIR instruments at their resolution limits. Confusion noise levels are displayed over the range as given by the colour-bar. All maps were constructed with *no* DIRBE–ISOPHOT conversion applied and with *constant* α and are presented in *equatorial* coordinate system (J2000). (a) ISO/ISOPHOT C200 camera, 170 μm; (b) ASTRO-F/FIS 170 μm; (c) Spitzer/MIPS 160 μm; (d) Herschel/PACS 175 μm

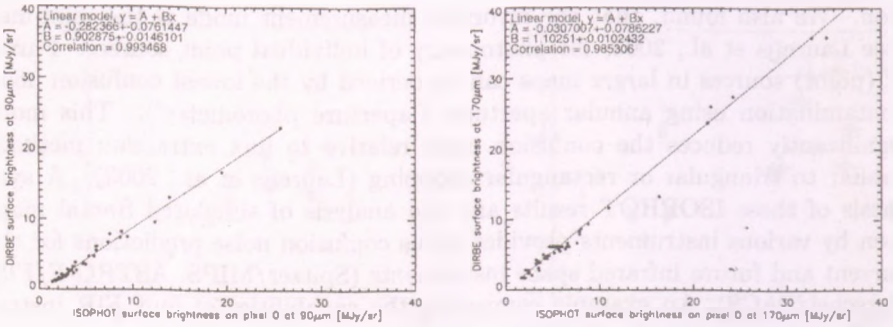

Figure 9: Example of the ISOPHOT–COBE/DIRBE surface brightness photometric comparison. (a) 90 μm; (b) 170 μm;

4. Prospects for far-infrared studies using the ISO Data Archive

4.1. Absolute surface brightness calibration of ISOPHOT

The final aim of the absolute surface brightness calibration – apart from its technical relevance – is the *direct* detection of the cosmic far-infrared background (see Sect. 2.4). Although ISOPHOT has not been especially designed to perform absoulte photometry, this can be done in some selected and well calibrated observing modes. This work is still in progress at Konkoly Observatory (Budapest) in collaboration with German (ISOPHOT Data Centre, MPIA, Heidelberg) and Finnish (University of Helsinki Observatory) astronomers. The most promising observing mode in this respect is the *mini-map* mode, which was almost exclusively used for point-source observations. In the mini-map mode the detector (C100 or C200) 'moved around' the source in a way that the source was centered at each detector pixel once during the measurement (see Laureijs et al., 2003, for a detailed explanation). This mode is proved to be very powerful for precise point-source photometry (Moór et al., 2004a, in prep.). However, in parallel with the flux of the point source, one can obtain a precised background value as well. In Fig. 9 preliminary results of the ISOPHOT – COBE/DIRBE photometric system comparison are presented (Moór et al., 2004b, in prep.).

4.2. Fluctuation studies:

Studies of the sky background fluctuations are limited at high spatial frequencies by the maximal achievable resolution of the instruments (46″and 92″for the C100 and C200 cameras, respectively) and by the final extension of the ISOPHOT maps at low spatial frequencies. This latter can be extended to lower spatial frequencies by the utilization of the ISOPHOT Serendipity slews which extend to a typical length of several degrees (Bogun et al., 1996).

4.3. Cirrus studies

A pilot study by del Burgo et al. (2003) has shown that the far-infrared emissivity of big dust grains (τ_{200}/A_V) changes with the ambient temperature: dust grains emit more FIR photons at lower temperatures than expected from their visual extinction properties. There is an on-going study looking for these variations in interstellar clouds with physical parameters (density, temperature) different from the ones analysed in del Burgo et al. (2003). Our study also aims to investigate the relation between the changes in dust emissivity τ_{200}/A_V and the ratio of total over selective extinction R_V (see e.g. Krügel et al., 2003, for a detailed description of R_V).

Acknowledgments

This work was partly supported by the ESA PRODEX programme (No. 14594/00/NL/SFe) and by the Hungarian Research Fund (OTKA, No. T037508). The author acknowledges the financal support of the N+N workshop.

References

Ábrahám, P., Leinert, Ch., Lemke, D., 1997, A&A 328, 702
Bogun, S., Lemke, D., Klaas, U., et al., 1996, A&A 315, 71
Bond, J.R., Carr, B.J., Hogan, C.J., 1986, ApJ 306, 428
Boulanger, F., Pérault, M., 1988, ApJ 330, 964
del Burgo, C., Laureijs, R.J., Ábrahám, P., Kiss, Cs., 2003, MNRAS 346, 403
Dole, H., Gispert, R., Lagache, G., et al., 2001, A&A 372, 364
Gautier III, T.N., Boulanger, F., Pérault, M., Puget J.L., 1992, AJ 103, 1313
Hauser,M.G., Arendt, R.G., Kelsall, T., et al., 1998, ApJ 508, 25
Hauser, M.G., Dwek, E., 2001, ARA&A 39, 249

Helou, G., Beichman, C.A., 1999, The confusion limits to the sensitivity of submilimeter telescopes, in: From Ground-Based to Space-Borne Sub-mm Astronomy, Proc. of the 29th Liège International Astrophysical Coll., ESA Publ., p. 117.

Juvela, M., Mattila, K., Lemke, D., 2000, A&A 360, 813

Kawara, K., Sato, Y., Matsuhara, H., 1998, A&A 336, L9

Kelsall, T., Weiland, J.L., Franz, B.A., 1998, ApJ 508, 44

Kessler, M.F., Steinz. J.A., Anderegg, M.F., et al., 1996, A&A 315, L27

Kiss, Cs., Ábrahám, P., Klaas, U., Juvela, M., Lemke, D., 2001, A&A 379, 1611

Kiss, Cs., Ábrahám, P., Klaas, U., et al., 2003, A&A 399, 177

Lagache, G., Abergel, A., Boulanger, F., Désert, F.X., Puget, J.-L., 1999, A&A 344, 322

Lagache, G., Puget, J. L., 2000, A&A 355, 17

Laureijs, R.J., Klaas, U., Richards, P.J., Schulz, B., Ábrahám, P., 2003, The ISO Handbook Vol. IV.: PHT – The Imaging Photo-Polarimeter, Version 2.0.1, ESA SP-1262, European Space Agency

Leinert, Ch., Bowyer, S., Haikala, L.K., et al., 1998, A&AS 127, 1

Lemke, D., Klaas, U., Abolins, J., et al., 1996, A&A 315, L64

Low, F., Beintema, D.A., Gautier, F.N., et al., 1984, ApJ 278, L19

Matsuhara, H., Kawara, K., Sato, Y., 2000, A&A 361, 407

Puget, J. L., Lagache, G., Clements, D. L., 1999, A&A 345, 29

Reach, T.W., Franz, B.A., Weiland, J.R., 1997, Icarus 127, 461

Reach, T.W., Sykes, M.V., Lien, D., 2000, Icarus 148, 80

Stickel, M., Lemke, D., Mattila, K., Haikala, L.K., Haas, M., 1998, A&A 329, 55

Infrared loops and the large scale structure of the diffuse interstellar matter in the Milky Way

V. Könyves[1], Cs. Kiss[2,3] and A. Moór[3]
[1]Eötvös University, Department of Astronomy
P.O.Box 32., H-1518 Budapest, Hungary
[2]Max-Planck-Institut für Astronomie
Königstuhl 17, D-69117 Heidelberg, Germany
[3]Konkoly Observatory of the Hungarian Academy of Sciences
P.O.Box 67., H-1525 Budapest, Hungary
E-mail: [1]v.konyves@astro.elte.hu, [2]pkisscs@mpia.de, [3]moor@konkoly.hu

Abstract

We performed an all-sky survey searching for loop- and arc-like intensity enhancements in the diffuse far-infrared emission using IRAS data. We identified 477 of these objects, analysed their indivual FIR properties and their distribution. Our objects trace out the spiral arm structure of the Galaxy in the neighbourhood of the Sun and their distribution clearly suggests that there is an efficient process that can generate loop-like features at high Galactic latitudes. We derived distances for 80 loops with the help of associated objects. Deviations in the celestial distribution of far-infrared loops clearly indicate, that violent events frequently overwrite the structure of the interstellar matter in the inner Galaxy. We obtained observational estimates of $f_{in} = 6.8\%$ and $f_{out} = 4.8\%$ for the hot gas volume filling factor of the immediate inward and outward Galactic neighbourhood of the Solar System.
KEYWORDS: *ISM, far-infrared: diffuse emission, Galactic structure*

1. Introduction

The large scale strucure of the cold interstellar matter can be siginficantly affected by violent events. This structure – which is diverse with the complex distribution of shells, cavities, filanemts, arcs and loops – is often referred to as the "Cosmic Bubble Bath" (Brand & Zealey, 1975). Observations of the soft X-ray background (Burstein et al., 1977) and of O VI absorpsion lines (Jenkins & Meloy, 1974) suggested the presence of hot gas in the Galactic disc. Studying supernova (SN) explosions in a uniform medium, Cox and Smith (1974) pointed out that if the Galactic SN rate is "sufficient" it can produce the "swiss

cheese morphology" of the cold diffuse ISM with hot coronal gas inside the bubbles. This was built into the model by McKee and Ostriker (1977) where SN explosions in the cloudy ISM produce a three-phase medium.

Shell- or arc-like intensity enhancements may be formed by various processes and are reported in many tracers of the ISM (see Kiss et al., 2004 (KMT04), and references therein, for an introduction).

Alhough these features can be detected by the X-ray and UV emission of the coronal gas inside the cavities seen by e.g. the ROSAT and EUVE satellites, most of them were detected at radio wavelengths, especiall in the HI 21 cm line (Heiles, 1979, 1980, 1984; Hu, 1981; Koo & Heiles, 1991).

Thilker et al. (1998) and Moshenko et al. (1999) made efforts to detect HI shells automatically, based on a model of the shell. Recently Ehlerová et al. (2004) developed a model-independent algorithm which automatically searches HI shells in data cubes and identified ~1000 structures in the Leiden-Dwingeloo Survey data. An artificial neural networks algorithm by Daigle et al. (2003) was succesfully applied to the Canadian Galactic Plane Survey data. However, these works intended to detect HI holes based on the velocity information of the 21 cm line data rather than on morphology.

The structure of the cold ISM is also well-represented in the far-infrared, mainly observed as the Galactic cirrus emission (Low et al., 1984). Although some studies reported loop features indentified in the far-infrared, these were either restricted to the Galactic midplane (Schwartz, 1987), performed for a special object type, e.g. Wolf-Rayer stars (Marston, 1996) or concentrated on individual loops (Kun, 1998; Meyerdierks et al, 1991). Most of the prominent HI loops are conspicious in the far-infrared, too.

In a recent work Kiss et al. (2004, hereafter KMT04) presented the results of a quest for far-infrared loop features in the 2nd Galactic Quadrant. They catalouged 145 loops and investigated their morphological and phyisical characteristics. This was the first study which was not restricted to the Galactic midplane, but reached even the vicinity of the Galactic poles and was able to give a comprehensive view on the distribution of large scale intensity enhancements.

In this present work – as a continuation of the KMT04 study – we extend the quest to the 1st , 3rd and 4th Galactic Quadrants. These two works together provide the *Catalouge of Far-InfraRed Loops in the Galaxy.*

2. Input data and data analysis

The data reduction steps and derived parameters are the same as described in KMT04. For the detailed description of the data reduction and derived parameters we refer to that paper. We summarize the comprised data reduciton steps below:

- indentification of the loop structure on ISSA 100 μm maps
- ellipse fitting (central Galactic coordinates, L_c and B_c; semi-major and semi-minor axis, a and b; position angle, PA)
- derivation of radial surface brightness profiles on ISSA 100 and 60 μm, Schlegel et al. (1998, hereafter SFD) 100 μm and SFD reddening maps
- derivation of significances from ISSA 100 and 60 μm, SFD 100 μm and SFD reddening intensity profiles (Ψ_{I100}, Ψ_{I60}, Ψ_{S100} and Ψ_{SEBV}, respectively)
- Determination of the position of the inner (a_{in}) and outer (a_{out}) edge of the loop wall, derivation of its relative width $W = 1 - a_{in}/a_{out}$
- calculation of the colour index of the loop wall $\Delta I_{60}/\Delta I_{100}$

3. Discussion of the main results

3.1. The catalogue

In our all-sky survey 477 FIR loops have been identified. The electronic version of the catalogue with many entries and images of the loops can be found at: "http://astro.elte.hu/CFIRLG".

3.2. Celestial distribution: dominated by confusion?

The celestial distribution of the identified loops is rather complex showing structures with scales even larger than the average loop diameters (Fig. 1). E.g. two remarkable features, forming 'chains' of loops can be found around l$\sim$125°, b$\sim$-15° and around l$\sim$175°, b$\sim$+10°. There are also some sky regions, which are overpopulated by loops, e.g. at l$\sim$135°, b$\sim$+55° and at l$\sim$180°, b$\sim$+0°. These regions coincide well with molecular complexes (Ursa Maior- and Taurus molecular clouds, respectively).

The Galactic longitude distribution of the GIRLs partly reflect the spiral structure of the Galaxy (Figs 2 and 3). Looking into the directions of the

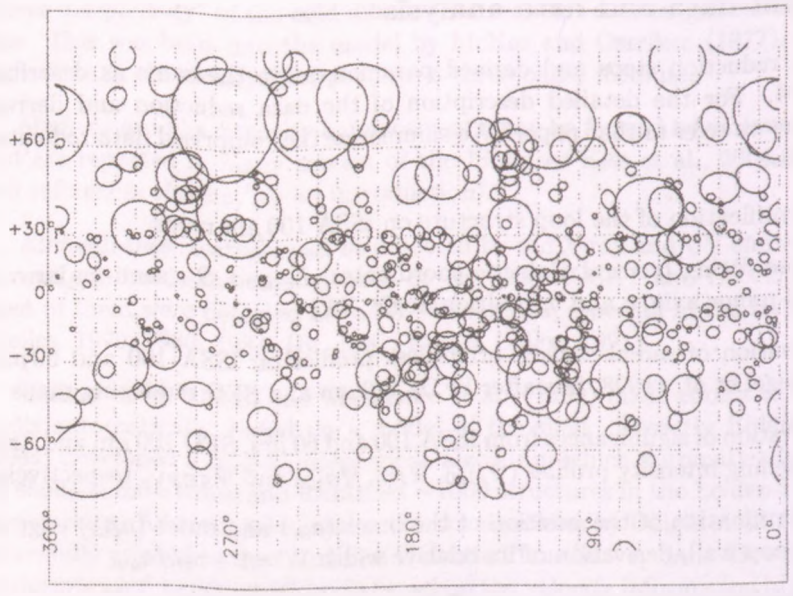

Figure 1: Distribution of GIRLs in the sky (Mercator projection), represented by the fitted ellipses. Note that Mercator projection causes a size distorsion around polar regions.

Local-, the inner Carina-Sagittarius- and the outer Perseus Arms, the increase in the loop-counts can be explained by a higher formation probability.

The distribution of GIRLs on the sky is expected to reflect the exponential disc distribution of the ISM. The Galactic Disc contains most of the ISM in the Galaxy, accordingly regions closed to the Galactic plane should be the most populated parts of sky in loops and this number should decrease nearly exponentially by the increasing Galactic latitude.

The distribution of loop centers in Galactic latitude (small panels in Fig. 3) shows a relatively large count at high $|b|$ values, especially for outward locations, which cannot be explained by projection and distance effects only, in agreement what was found for the 2^{nd} Galactic Quadrant (KMT04). To explain the formation of loops at high Galactic latitudes one needs an efficient process which is not sticked to the Galactic plane. This excludes SN-explosions and the stellar wind of massive stars as dominant effects. Although clouds infalling from the Galactic halo could create loops above the midplane, their infall rate is insuffi-

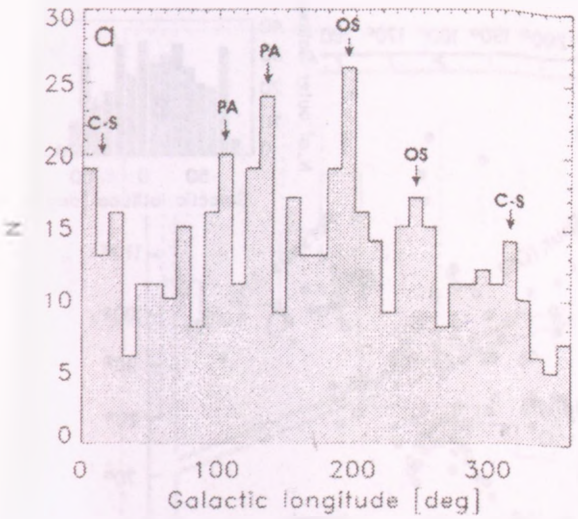

Figure 2: Distribution of loops in Galactic longitude. The peaks in loop-count are explained by the higher formation possibilities of loops in the Galactic arms (OS: Orion Spur, PA: Perseus Arm, C-S: Carina-Sagittarius Arm).

ciently small (Ehlerová & Palouš, 1996). Supersonic turbulence and nonlinear instabilities may be responsible for high latitude loops, however, further studies are necessary to identify the dominant formation scenarios.

To independently study the immediate inner and outer surrounding of the Solar System, we arbitrarily split the Galactic latitude space into two parts, approximately along the direction of the Local Arm, i.e. the $l = 70°–250°$ axis (see Fig. 3). The loop counts in the inner and outer part show significant differences: there are almost twice as many loops in the outer region than in the inner one. One possibility to explain this fact is the existence of strong confusion in the direction of central (inward) Galactic regions, as discussed below.

The possibility of a confusion-effected distribution for the 2^{nd} Galactic Quadrant was already mentioned in KMT04. In this study – due to the extension to the whole sky – this confusion effect is more expressed and manifests itself in two main ways:

i) As presented in Fig. 3, we see more loops in the 'outer' Galaxy than in the 'inner' regions for $|b| < 30°$, although probabilities of high-pressure events (e.g. SN-rates) must be significantly higher in the inner Galaxy.

ii) The size distribution is different in the inner and outer Galaxy for $|b| < 30°$, whereas this discrepancy in not observable elsewhere (see Fig. 4).

These two facts together indicate that (a) either we miss large loops at low $|b|$ in

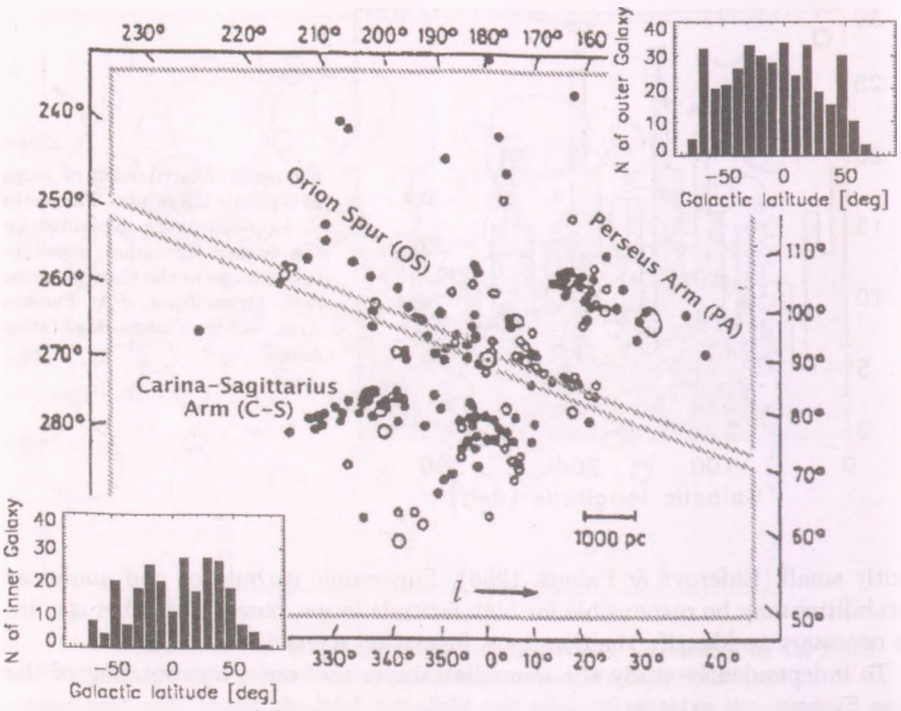

Figure 3: Galactic latitude distribution of loops in the inner- and outer Galaxy. The majority of the loops are probably located in the Local Arm (see text). The basic figure presenting the positions of OB associations around the Solar System is adapted from Humphreys (1970).

the inner Galaxy due to the strong IR background (small loops are still visible since they are smaller than the characteristic scale of background fluctuations) or (b) large loops are really distroyed by frequent violent events and are only visible at their early evolutionary phases (i.e. at small size). Unfortunately our present FIR data cannot distinguish between these possibilities.

3.3. Large scale distribution of ISM in the Milky Way

We made distance estimates for individual loops using associated objects projected to the loop walls and to the interior of loops as explained in KMT04. With the application of the same criteria when selecting distance indicators we were able to derive distances for 80 loops. Fig. 5a shows the distribution of

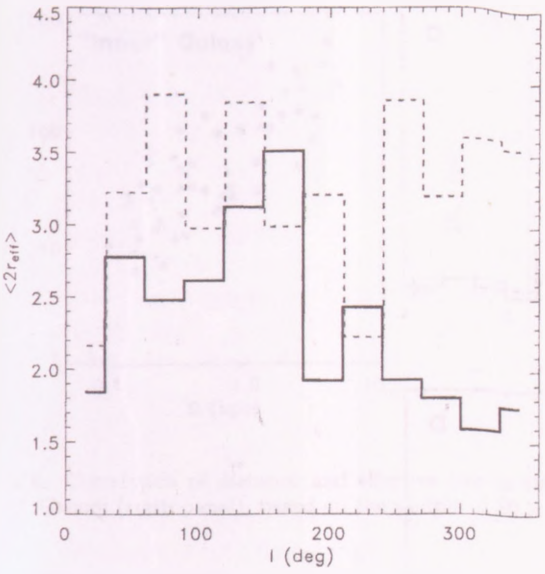

Figure 4: Average effective loop diameter vs. Galactic longitude; solid black line: $|b| \leq 30°$; dashed gray line: $30° \leq |b| \leq 60°$.

estimated distances and the distances projected to the Galactic plane. Loops closer than ~1 kpc belong to the Local Arm and to the nearby interarm region.

Oey & Clarke (1997) and Kim et al. (2003) investigated the large-scale structure of the interstellar medium in nearby galaxies. Using the transformation by Oey & Clarke (1997) we derived a power law index of $\beta_0 = 1.37 \pm 0.33$ from our sample for the interval $1.4 \leq \log(2R_{eff}) \leq 2.6$, which is similar to the value of the 2^{nd} Galactic Quadrant only ($\beta = 1.24 \pm 0.30$, KMT04) and lower than that of the other investigated galaxies (see Fig. 5c). The double-peaked distribution observed in KMT04 is less expressed in this larger sample and is closer to the expected power-law.

Figs 6 and 7 present the differences in distance distributions in the inward and outword Galactic regions. In the inward region most of the loops are squezzed into a small area around the Solar System, within a distance of ~0.2 kpc. In this region the distance and size are fully uncorrelated, i.e. this region is not affected by this selection effect (see KMT04 for a more detailed explanation). In the outward region (with a significant contribution of the 2^{nd} Galactic Quadrant) a close ($d \leq 800$ pc) and a far ($d > 800$ pc) group can be identified, as in KMT04.

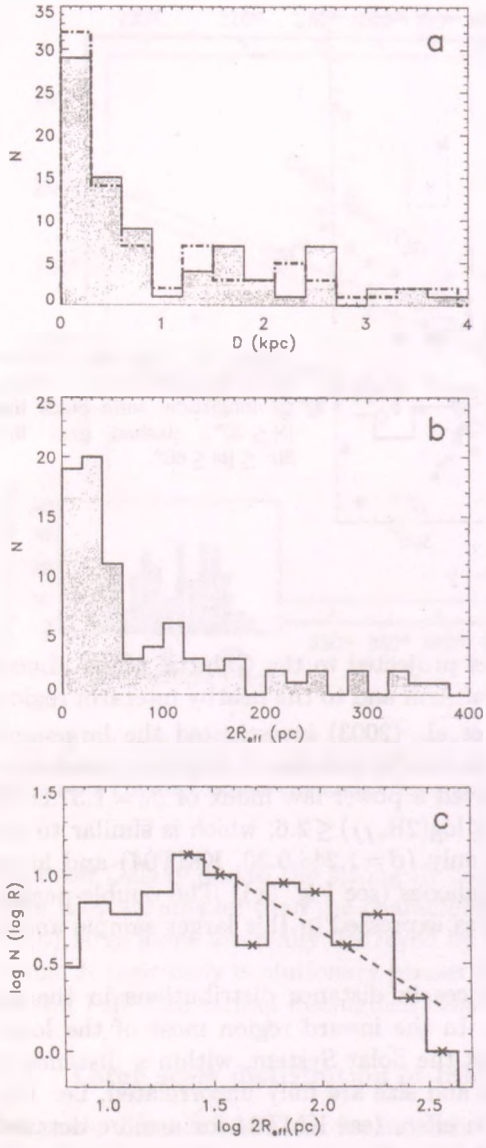

Figure 5: (a) Distribution of estimated distances for 80 loops. Dash-dotted line represents the distribution of distances *projected* to the Galactic plane.(b) Distribution of effective diameters in the same sample.(c) The same as b but displayed and binned on a logarithmic scale. The relation of $\log 2R_{eff}$-$\log N$ was fitted for $1.4 \leq 2R_{eff} \leq 2.6$. The points used for the fit are marked by asterisks (see Sect. 3.3 for details).

Applying the appropriate average size and distance values and biased/unbiased ratios (i.e. ratio of close/uncorrelated and far/correlated loop counts in the

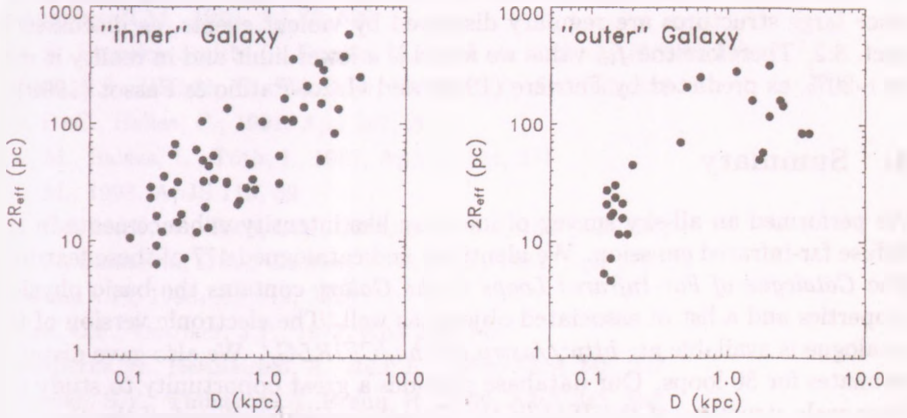

Figure 6: Correlation of distance and effective size in the "inner" (left panel) and in the "outer" Galaxy (rigth panel), based on the sample of 80 loops with known distances.

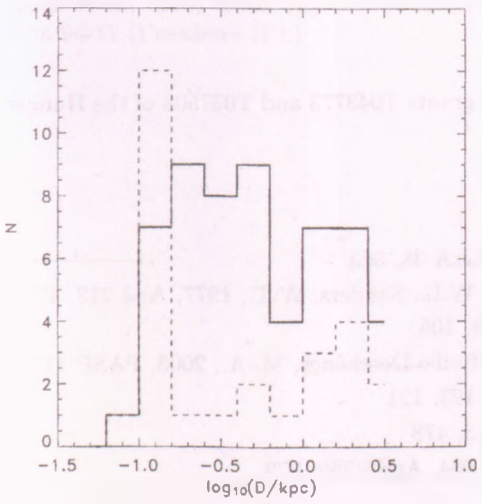

Figure 7: Distribution of distances in the 80 loop sample (logarithmic sampling). Gray dashed line: inner Galactic environment; Black solid line: outer Galactic environment.

distance-size relation) for the whole loop sample in the inward and outword regions separetely we obtained $f_{in} = 6.8\%$ and $f_{out} = 4.8\%$. This latter value is very similar to the one presented in KMT04 for the 2^{nd} Galactic Quadrant. However, the size distribution in the inward region almost completely misses

the large loops which might be due to confusion or they do not exist at all, since large structures are regulary distroyed by violent events, as discussed in Sect. 3.2. Therefore the f_{in} value we found is a lower limit and in reality it may be ~20%, as predicted by Ferrière (1998) and Gazol-Patiño & Passot (1999).

4. Summary

We performed an all-sky survey of loop/arc-like intensity enhancements in the diffuse far-infrared emission. We identified and catalogued 477 of-these features. The *Catalogue of Far-Infrared Loops in the Galaxy* contains the basic physical properties and a list of associated objects as well. The electronic version of the catalogue is available at: *http://astro.elte.hu/CFIRLG/*. We also gave distance estimates for 80 loops. Our database provides a great opportunity to study the large scale structure of the ISM in the Galactic neighbourhood of the Sun. As a first step we determined observational estimates for the hot gas volume filling factor of the the inner (f_{in}) and outer (f_{out}) Galactic environment of the Solar System: $f_{in} = 6.8\%$ and $f_{out} = 4.8\%$. The f_{in} value is probably a lower limit due to the strong IR background in the direction of central Galactic locations.

Acknowledgments

This work was partly supported by the grants T043773 and T037508 of the Hungarian Scientific Research Fund (OTKA).

References

Brand, P.W.J.L., Zealey, W.J., 1975, A&A 38, 363
Burstein, P., Borken, R.J., Kraushaar, W.L., Sanders, W.T., 1977, ApJ 213, 405
Cox, D.P., Smith, B.W., 1974, ApJ 189, 105
Daigle, A., Joncas, G., Parizeau, M., Miville-Deschênes, M.-A., 2003, PASP 115, 662
Jenkins, E.B., Meloy, D.A., 1974, ApJ 193, 121
Ehlerová, S., Palouš, S., 1996, A&A 313, 478
Ehlerová, S., Palouš, S., Wünsch, R., 2004, ApSS 289, 279
Ferriére, K., 1998, ApJ 503, 700
Gazol-Patiño, A., Passot, Th., 1999, ApJ 518, 748
Heiles, C., 1979, ApJ 229, 533
Heiles, C., 1980, ApJ 235, 833
Heiles, C., 1984, ApJS 55, 585

Hu, E.M., 1981, ApJ 248, 119
Humphreys, R.M., 1970, AJ 75, 602
Kim, S., Staveley-Smith, L., Dopita, M.A., et al., 2003, ApJS 148, 473
Klessen, R.S., Hietsch, F., Mac Low, M.M., 2000, ApJ 535, 887
Koo, B.-C., Heiles, C., 1991, ApJ 382, 204
Kun, M., Balázs, L., Tóth, I., 1987, Ap&SS 134, 211
Kun, M., 1998, ApJS 115, 59
Kiss, Cs., Moór, A., Tóth, L.V., 2004, A&A 418, 131 (KMT04)
Low, F., Beintema, D.A., Gautier, F.N., et al., 1984, ApJ 278, 19
Marston, A.P., 1996, AJ 112, 2828
McKee, C.F., Ostriker, J.P., 1977, ApJ 218, 148
Meyerdierks, H., Heithausen, A., Reif, K., 1991, A&A 245, 247
Moshenko, S.Y., Thilker, D.A, Braun, R., 1999, A&A 343, 352
Oey, M.S., Clarke, C. j., 1997, MNRAS 289, 570
Schlegel, D.J., Finkbeiner, D.P., Davis, M., 1998, ApJ 500, 525
Schwartz, P.R., 1987, ApJ 320, 258
Thilker, D.A, Braun, R., Wakterbos, R.A.M., 1998, A&A 332, 429
Wheelock et al., 1994, IRAS Sky Survey Atlas Explanatory Supplement, JPL Publication 94-11 (Pasadena:JPL)

Determination of dust temperature and emissivity from IRAS and ISO observations

Z T. Kiss[1], Cs. Kiss[2], P. Ábrahám[3] and L. V. Tóth[1]
[1]Eötvös University, Department of Astronomy
H-1518 Budapest, P.O.Box 32, Hungary
[2]Max-Planck-Institut für Astronomie
Heidelberg, Königstuhl 17, D-69117, Germany
[3]Konkoly Observatory of the Hungarian Academy of Science
H-1525 Budapest, P.O. Box 67, Hungary
E-mail: [1]Z.Kiss@astro.elte.hu, [2]pkisscs@mpia-hd.mpg.de

Abstract

We examined the dust emissivity at far-infrared wavelengths, and proposed a method to reproduce ISOPHOT temperatures combining ISOPHOT and IRAS far-infrared data. Analysing a sample of 13 clouds we compared the temperature values obtained using our method to those determined from ISOPHOT data and verified the accuracy of our method. This method allows one to extend the sample of clouds appropriate for investigation of dust emissivity properties. We applied our method for 8 clouds, for which ISOPHOT data are available olny at one wavelength, to determine dust colour temperature, and the dust emissivity was also computed for 3 of these clouds.
Keywords: *Infrared: ISM - ISM: dust, extinction*

1. Introduction

Even in moderately dense clouds optical properties of interstellar dust seem to be different from those of the diffuse ISM. Cambresy et al. (2001) have shown that estimating optical extinction from IRAS optical depth data results in 2 to 3 times higher values than does the optical star count method for the case of the Polaris molecular cloud. This finding can be explained by increased infrared emissivity of dust inside the cloud. As Dwek (1997) pointed out the increased emissivity may be caused by coagulation of dust grains. Based on ISOPHOT observations, del Burgo et al. (2003) studied 8 translucent clouds and found the emissivity of dust to be higher than in the diffuse interstellar matter. They obtained increasing emissivity with decreasing temperature. This trend suggests that the change of optical properties of dust may start even in moderately dense regions. del Burgo et al. (2003) proposed a two-component model to reproduce the derived emissivity values.

In order to determine the dust emissivity at a certain wavelength λ, defined by the ratio τ_λ/A_V, one needs to measure the optical depth, and consequently to measure dust temperature. del Burgo et al. (2003) computed dust colour temperatures by combining two ISOPHOT observations at different far-infrared wavelengths obtained with the same subinstrument (C200 camera). At these wavelengths the emission of big grains is not contaminated by smaller, transiently heated particles. The sample of clouds observed by ISOPHOT in at least two filters longwards of 100 μm is, however, rather limited. In order to involve more fields into the analysis we propose to combine ISOPHOT data with IRAS 100 μm observations for the computation of the temperature. In this paper we explore and test this possibility, and determine temperatures for 8 fields observed by ISOPHOT only at one wavelength. For three of the fields we also analyse the emissivity of dust and compare our results with those of del Burgo et al. (2003).

2. Temperature estimation from IRAS and ISOPHOT

2.1. The method

The most accurate way of determining dust temperatures is combining two (or more) measurements at different far-infrared wavelengths obtained with the same instrument. When this was not possible (e.g. fields mapped with ISOPHOT only at one wavelength) a straightforward idea is the utilization of the IRAS 100 μm database. However, due to the different calibration of the two instruments it is important to verify the accuracy of the temperature values determined in this way.

We searched the ISO data archive and selected 13 fields covered by ISOPHOT at 200 μm and at least one other far-infrared ($\lambda \geq 120\,\mu$m) wavelength. The observations were performed with the ISOPHOT C200 camera (2 $\times$ 2 pixel, 90″ pixel size) in the PHT22 raster mode. Data processing was performed in the standard way using the PHOT Interactive Analysis V.10.0. Processing of ISOPHOT maps was described in detail by del Burgo et al. (2003).

We extracted IRAS 100 μm data for all fields by sampling the ISSA maps precisely at the positions of the ISOPHOT grid points. Temperatures were determined from the slope of the correlation between (a) the two ISOPHOT filters, (b) the ISOPHOT 200 μm and IRAS 100 μm data. Figure 1 shows a comparison between the results of the two different methods. No systematic difference was found. The standard deviation of the temperature values around the $T_{IRAS} = T_{PHT}$ line is $\sigma_T = 0.78\,$K.

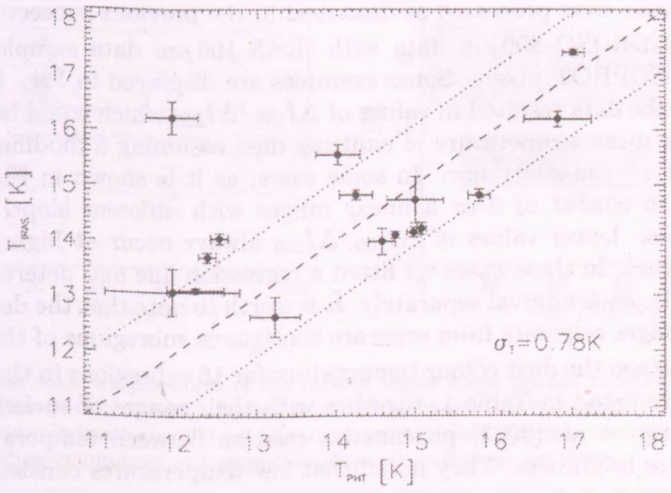

Figure 1: Correlation plot of temperatures determined from observations of ISOPHOT at two wavelengths (T_{PHT}) and ISOPHOT 200 μm data combined with IRAS 100 μm measurements (T_{IRAS}).

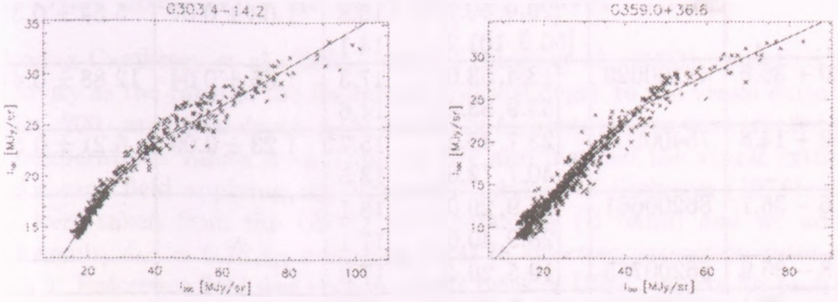

Figure 2: Examples for surface brightness I_{100} versus I_{200} of the regions with the fitted lines for subregions with different temperatures.

2.2. Application for 8 interstellar clouds

As an application of our method we examined 8 additional regions which were mapped with ISOPHOT at 200 μm only. The locations of the fields are listed in

Table 1. The extent of the fields is typically 30×30 arcmin with 90″ pixel size. The observations were processed as described in the previous subsection.

We correlated ISO 200 μm data with IRAS 100 μm data sampled at the positions of ISOPHOT pixels. Some examples are displayed in Fig. 2. Linear regression to the data resulted in values of $\Delta I_{100}/\Delta I_{200}$ which could be directly related to the mean temperature of emitting dust assuming a modified Planck function with ν^2 emissivity law. In some cases, as it is shown in Fig. 2, the trend seems to consist of 2 or 3 linear ranges with different slopes devided by breakpoints. Lower values of $\Delta I_{100}/\Delta I_{200}$ always occur at higher surface brightness values. In these cases we fitted a regression line and determined the temperature for each interval separately. It is worth to note that the data points of different ranges originate from separate contiguous subregions of the fields.

We determined the dust colour temperature for 15 subregions in the 8 fields. The results are listed in Table 1, together with their ranges of surface brightness. del Burgo et al. (2003) presented a relation between temperature and infrared surface brightness. They noted that low temperatures consistently occur at positions with higher infrared intensity than the characteristic value of

Name	ISO_{id}	$I(200)$	T_D[K]	$\frac{\tau_{200}}{A_V}[mag^{-1}]$	$\frac{I(200)}{A_V}$
$G303.4-14.2$	33300559	[15.0, 29.8]	17.4	1.54 ± 0.06	14.35 ± 0.50
		[29.9, 50.7]	15.8	1.03 ± 0.07	5.54 ± 0.35
		[50.9, 101.2	14.1		
$G359.0+36.6$	43100629	[13.4, 53.0	17.3	1.65 ± 0.04	12.88 ± 0.47
		[53.9, 83.4	14.6		
$G114.0+14.8$	75400905	[23.7, 40.2	15.7	1.23 ± 0.09	6.21 ± 0.50
		[40.7, 72.8	13.5		
$G211.6-36.7$	86200661	[21.9, 29.0	18.1		
		[28.9, 40.6	17.1		
$G210.8-35.9$	86200765	[19.5, 29.4	17.3		
$G211.3-36.1$	86201063	[19.8, 25.3	18.6		
		[24.2, 31.6	16.6		
$G211.8-36.2$	86201166	[16.4, 28.8	16.9		
		[28.9, 43.9	15.7		
$G210.3-36.7$	86201064	[16.4, 26.7	18.3		

Table 1: Colour temperatures obtained from ISOPHOT 200 μm and IRAS 100 μm data with the I(200) range of fit for each subregions of the 8 fields, and values of τ_{200}/A_V for 4 of these subregions.

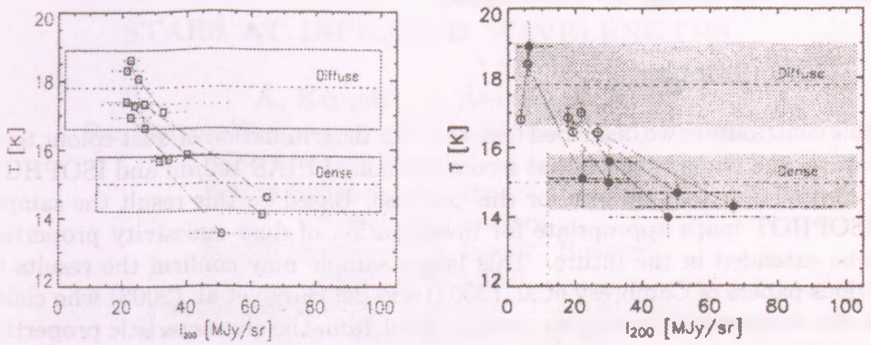

Figure 3: Dust color temperatures of the regions versus mean 200 μm emission compared to Fig. 7 of del Burgo et al. (2003). Dotted lines link temperatures belonging to the same field. Dot-dashed lines indicate the 200 μm emission ranges. Mean temperatures (dashed lines) of dense and diffuse component and respective dispersions (boxes) are shown.

diffuse matter. Fig. 3 demonstrates that our temperature values follow a similar relationship.

3. Dust emissivity at far-infrared wavelengths

Following Cambresy et al. (2001) and del Burgo et al. (2003) we defined dust emissivity as the ratio of the far-infrared optical depth to the visual extinction, τ_λ/A_V. 200 μm optical depth maps for all the separate ranges were created using the temperatures values from Table 1. We also mapped the visual extinction A_V for each field applying the star count method of Dickman (1978). Input data were taken from the USNO A2.0 catalogue (B band) and we adopted the formula $A_V = 0.76 A_B$ assuming total to selective extinction ratio to be $R = 3.1$. Reference field was chosen on the basis of USNO charts for each field. The extinction were sampled at the positions of ISOPHOT grid points with a circular beam of 8′ diameter.

For detailed analysis we selected those three fields where the τ_λ/A_V was determined with high enough signal-to-noise ratio. For these fields values of τ_{200}/A_V are given in Table 1. The four computed values are rather similar, and they are all systematically lower than the results of del Burgo et al. (2003). To trace the origin of this discrepancy requires further investigation. The narrow temperature range covered by our four subregions does not allow to check the

dependency of τ_{200}/A_V on temperature.

4. Conclusions

In this contribution we described a method for determination of dust colour temperature, and demonstrated that a combination of IRAS 100μm and ISOPHOT far-infrared data can be used for this purpose. Based on this result the sample of ISOPHOT maps appropriate for investigation of dust emissivity properties can be extended in the future. This larger sample may confirm the results of previous papers of Cambresy et al. (2001) and del Burgo et al. (2003) who claim that the dust optical properties seem to differ from the characteristic properties of the cirrus even in regions of moderately high density (A_V <2 mag). A possible reason of the increased infrared emissivity might be the coagulation of dust (Dwek, 1997). Examining a large sample covering various sky areas one could also investigate the generality of this phenomenon.

Acknowledgments

This work was partly supported by the grants OTKA T037508 and T043773 of the Hungarian Scientific Research Fund. P. Á. acknowledges the support of the Bolyai Fellowship.

References

Cambrésy, L., Boulanger, F., Lagache, G., Stepnik, B. 2001, *A&A* , 375, 999
Dickman, R. L. 1978, *AJ* , 83, 363
Dwek, E. 1997, *ApJ* , 484, 779
del Burgo, C., Laureijs, R. J., Ábrahám, P., Kiss, Cs. 2003, *MNRAS* , 346, 403

The long-term evolution of 7 FU Orionis-type stars at infrared wavelengths

Á. Kóspál[1], P. Ábrahám[2], Sz. Csizmadia[2], M. Kun[2], A. Moór[2] and T. Prusti[3]

[1]Eötvös University, Dept. of Astronomy,
H-1518 Budapest, P.O. Box 32., Hungary
[2]Konkoly Observatory,
H-1525 Budapest, P.O. Box 67., Hungary
[3]ESTEC/SCI-SAF,
Postbus 299, 2200 AG Noordwijk, The Netherlands
E-mail: [1]kospal@szofi.elte.hu

Abstract

We investigate the brightness evolution of 7 FU Orionis systems in the 1−100 μm wavelength range using data from the *Infrared Space Observatory*, 2MASS and MSX. The spectral energy distributions (SEDs) based on these data points are representative of the period 1996−2000. These SEDs were compared with earlier ones derived from the IRAS photometry and from ground-based observations carried out around 1983, in order to look for long-term evolution. Our data show that three objects have become fainter, while the others remained constant. We investigate the case of V1057 Cyg and discuss whether its observed fading could be understood in the framework of the existing models.

Keywords: *pre-main sequence stars, circumstellar matter, infrared astronomy, V1057 Cyg*

1. Introduction

FU Orionis objects are low mass pre-main sequence stars undergoing outburst in optical light of 4 mag or more, followed by a fading phase on the timescale of several decades. According to the most widely accepted picture the FU Orionis outburst is a consequence of a rapid temporal increase of the disk accretion rate (for a review see Hartmann & Kenyon 1996). Predictions of these types of models have to be confronted with multiwavelength monitoring observations of the outburst period and of the post-outburst phase. The fading phase is well documented in the optical/near-infrared, but very few data have been available so far at mid- and far-infrared wavelengths where thermal emission of the disk and of the circumstellar envelope can be observed. Recently the *Infrared Space Observatory* (ISO) provided new photometric data on FU Orionis-type stars

Instrument	Wavelengths [μm]	Aperture	Active period
ground-based	J, H, K, L, M, N, Q	$\leq 6''$	1970s –
IRAS	12, 25, 60, 100	$1 - 3'$	1983
MSX	4.25, 4.29, 8.28, 12.13, 14.65, 21.34	$18''$	1996 – 1997
2MASS	J, H, Ks		1997 – 2001
ISOPHOT	4.8, 12, 25, 60, 100, 120, 200	$43'' - 180''$	1995 – 1998

Table 1: Sources of infrared photometric data used in our study.

in the 4.8–200 μm range. In our study we search, for the first time, for systematic brightness variations during the post-outburst phases of 7 FU Orionis objects (V1515 Cyg, V1735 Cyg, V346 Nor, Z CMa, Parsamian 21, V1331 Cyg, V1057 Cyg) at mid- and far-IR wavelenghts.

2. Observations and data reduction

Tab. 1 lists the infrared photometric data used in this study. The time distribution of the active periods of the instruments/projects offers a possibility to check for long-term variations of the infrared fluxes between ~1983 (IRAS, ground-based data) and 1996–2000 (ISO, MSX, 2MASS).

We compiled a list of all confirmed/candidate FU Orionis objects from the literature and selected those 7 objects for further study where sufficient data were available at both epochs (1983 and 1996–2000) to create complete mid/far-infrared SEDs.

The data reduction was performed using the ISOPHOT Interactive Analysis Software Package V10.0 (PIA, Gabriel et al. 1997). After the corrections for non-linearities of the integration ramps, the signals were transformed to a standard reset interval. Then an orbital dependent dark current was subtracted and cosmic ray hits were removed. In case the signal did not fully stabilise during the measurement time due to detector transients, only the last part of the data stream was used. The calibration of the measurements was performed by using the default responsivity. As an error estimate we adopted an absolute calibration uncertainty of 25%, which represents well the sum of the random and systematic uncertainties. Colour corrections were applied for each measurement by convolving the observed SED with the ISOPHOT filter profiles in an iterative way.

3. Results

Our main results, the comparison of SEDs at the two different epochs, are presented in Fig. 1 and Fig. 2. The figures show that the sources exhibit wavelength dependent temporal behaviour, as summarised below:

- **Near-IR** ($\lambda \leq 5\,\mu$m): the sources show various trends: Par 21, V1331 Cyg and Z CMa are unchanged, V1057 Cyg, V1515 Cyg and V1735 Cyg have faded, V346 Nor have become slightly brighter.
- **Mid-IR** ($5 \leq \lambda \leq 20\,\mu$m): only V1057 Cyg shows systematic flux change: it faded by a factor of 2 during the period.
- **Far-IR** ($\lambda \geq 60\,\mu$m): five stars (V1057 Cyg, V1735 Cyg and Z CMa, Par 21 and V1515 Cyg) remained constant while V346 Nor seem to have become fainter. For V1331 Cyg there are no FIR data other than IRAS.

4. Discussion: the case of V1057 Cyg

In this section we investigate in details the case of V1057 Cyg, which shows the fastest evolution, and which has the best documented multiwavelength flux evolution following its outburst in 1970. There are two models of the circumstellar environment of this star to fit the complete $1-100\,\mu$m infrared SED. Kenyon & Hartmann (1991, hereafter KH) assume that a flat disk is embedded in a spherically symmetric envelope with a wind-driven polar hole. The envelope reprocesses the radiation from regions close to the central star. Turner et al. (1997, hereafter TBB) fitted the SED of V1057 Cyg by computing outbursting flared disk models in which the mass flux varies with radius. This model includes reprocessing of disk emission by other parts of the disk and an envelope of uniform thickness with a central hole exposing the inner disk. In the following we compare our observations with the theoretical results.

At $\lambda \leq 2.2\,\mu$m

Both models claim that emission of the central source (the star plus the innermost part of the accretion disk) dominates the observed flux. After the outburst the accretion rate close to the star decreases leading to a flux decrease in this wavelength range.

Our data confirm that the flux decay was observable between 1983 and 2000, too. The data reveal that in the R, J, H, and K bands the flux

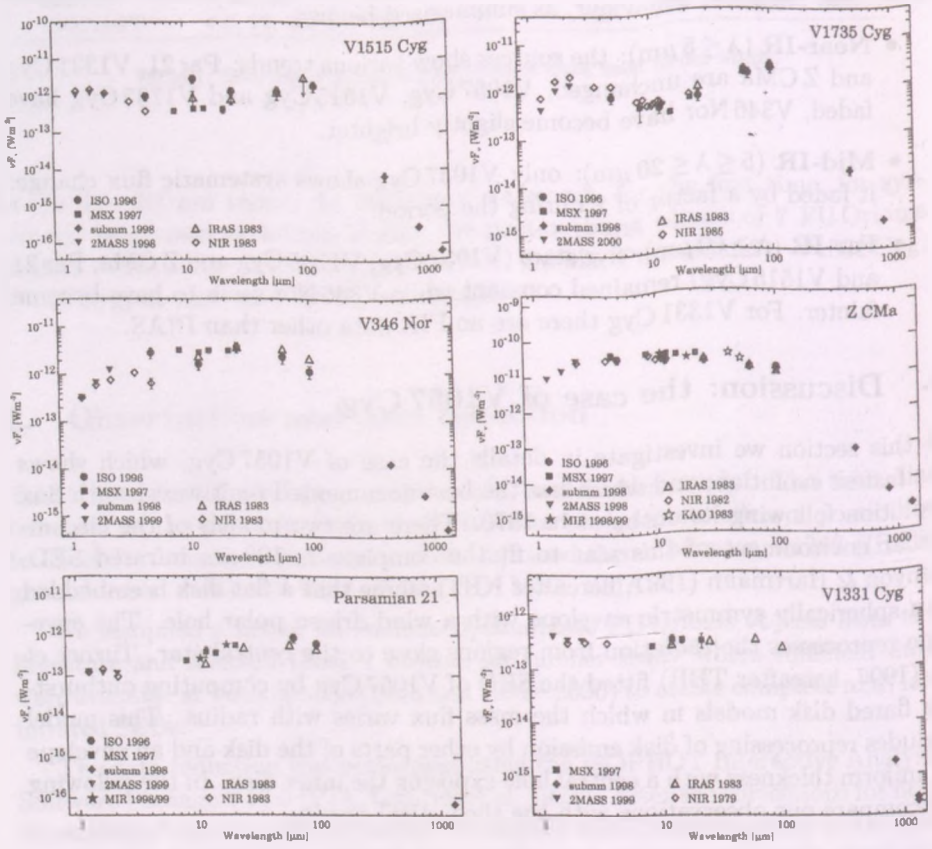

Figure 1: SEDs of FU Orionis objects. Open symbols represent data from ~1983 while filled symbols correspond to the epoch 1996–2000. The data are presented with no reddening correction. Error bars smaller than the symbol size were not plotted.

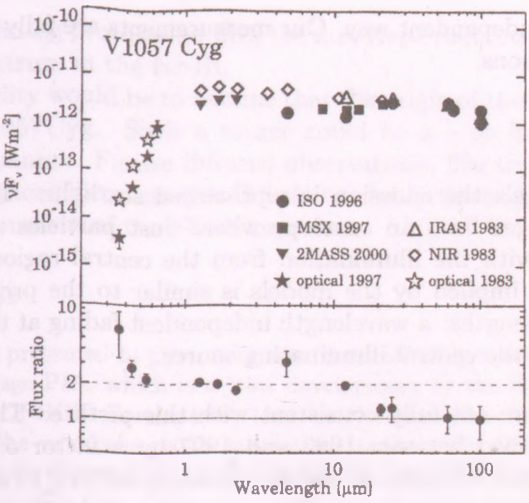

Figure 2: SED of V1057 Cyg at the two different epochs (1983 with open symbols and 1996–2000 with filled symbols). No reddening correction was performed. The lower panel shows the ratios of fluxes obtained in 1983 and in 1996–2000.

dropped by a constant factor of 2, while at B and V a larger decay was observed (lower panel of Fig. 2). This wavelength dependence may reflect the drop of effective temperature leading to a shift in the peak of the emission of the central source towards longer wavelengths.

At $3 \leq \lambda \leq 10\,\mu$m

According to the models the origin of the emission at these wavelengths is the release of accretion energy in the disk (KH, timescale: $\tau_{dyn} \sim$ several years at $\sim$1 AU, Pringle 1981), and also starlight reprocessed in the same part of the disk and in an envelope (TBB, timescale: $\tau_{th} \sim$ several years at $\sim$1 AU, Chiang & Goldreich 1997). Thus the prediction is that decreasing accretion rate at the centre would cause the drop of emission of all three components.

Since our temporal baseline (1983–1998) significantly exceeds the mentioned timescales and since the post-outburst evolution started already in the seventies, the model prediction is that the $3-10\,\mu$m emission is decreasing synchronised with the rate of the optical/near-IR decay in a

wavelength independent way. Our measurements are fully consistent with these predictions.

At $\lambda \geq 10\,\mu$m

In both models the emission is reprocessed starlight. The infrared radiation emerges from an envelope where dust particles are in radiative equilibrium with the illumination from the central region. The temporal evolution implied by the models is similar to the predicted trend at shorter wavelengths: a wavelength independent fading at the same rate as the fading of the central illuminating source.

Our results are not fully consistent with this picture. Though at $10\,\mu$m the flux dropped between 1983 and 1997 by a factor of 2 similarly to the optical/near-IR rate, at far-IR wavelengths ($\lambda \geq 60\,\mu$m) the flux of V1057 Cyg remained constant. The comparison of the IRAS 60 and $100\,\mu$m photometric points with our new ISOPHOT data at the same wavelengths clearly demonstrates that the far-infrared fluxes of V1057 Cyg showed no variation between 1983 and 1997.

Our results indicate that – unlike in the mentioned models – at $\lambda \geq 10\,\mu$m two important emission components have to be taken into consideration. Between 10 and $25\,\mu$m we probably observe the envelope, while at longer wavelengths the emission cannot be explained this way. In the following we discuss the possible nature of this second emission component.

The material responsible for the $\lambda > 25\,\mu$m emission apparently has the following properties: (1) relatively cold; (2) *not* reprocessing, *not* optically thin medium; and (3) has a flat spectrum below 100 μm, suggesting a $T \sim r^{-0.5}$ radial temperature profile. The first possibility for the source of emission would be an optically thick reprocessing medium (e.g. a flared disk), and its time-lag behind the fading of the central source significantly exceeds our temporal baseline. For certain reasons, however, both KH and TBB rejected this possibility.

The second possibility is that the energy source of the far-IR emission is accretion in the outer disk, where the accretion rate is constant and unrelated to the outburst of the central region. The spectral shape of such a standard accretion disk would follow the canonical $\nu F_\nu \sim \lambda^{-4/3}$ law, in contradiction with the observed flat spectrum (Fig. 2).

A solution for this problem could be if the temperature profile of the accretion disk differs from the standard one. Lodato & Bertin (2001) suggested

that in a self-gravitating protostellar disk the non-Keplerian rotation curve may result in a flat spectrum in the far-IR.

The last possibility would be to assume that the origin of the far-IR emission is unrelated to V1057 Cyg. Such a source could be a – so far undetected – embedded IR companion. Future infrared observations, like the ones expected from the Spitzer Space Observatory, will help to clarify the exact nature of the circumstellar environment of the FU Orionis stars.

Acknowledgments

The ISOPHOT data presented in this paper were reduced using the ISOPHOT Interactive Analysis package PIA, which is a joint development by the ESA Astrophysics Division and the ISOPHOT Consortium, lead by the Max-Planck-Institut für Astronomie (MPIA). The work was partly supported by the grant OTKA T 037508 of the Hungarian Scientific Research Fund. P.Á. acknowledges the support of the Bolyai Fellowship.

References

Chiang, E.I., Goldreich, P. 1997, in *Spectral energy distributions of T Tauri stars with passive circumstellar disks*, *ApJ* , 490, 368

Gabriel, C. et al. 1999, in *The ISOPHOT Interactive Analysis PIA, a calibration and scientific analysis tool*, Proceedings of the ADASS VI conference, ASP Conf. Ser. 125, p. 108

Kenyon, S.J., Hartmann, L. 1991, in *The dusty envelopes of FU Orionis variables*, *ApJ* , 383, 664 (KH)

Lodato, G., Bertin, G. 2001, in *The spectral energy distribution of self-gravitation protostellar disks*, *A&A* , 375, 455

Pringle, J.E. 1981, in *Accretion disks in astrophysics*, *Ann. Rev. Astron. Astrophys.* , 19, 137

Turner, N.J.J., Bodenheimer, P., Bell, K.R. 1997, in *Models of the spectral energy distributions of FU Orionis stars*, *ApJ* , 480, 754 (TBB)

Part Three

Solar Physics

Solar Oscillations and the Magnetic Atmosphere

Y. Taroyan and J.G. Doyle
Armagh Observatory
College Hill, Armagh BT61 9DG, Northern Ireland
E-mail: yat@arm.ac.uk, jgd@arm.ac.uk

Abstract

With the launch of the SOHO and TRACE satellites two new branches of solar physics have emerged: time-distance helioseismology and coronal seismology. In time-distance helioseismology the travel time of the acoustic waves between different points on the solar surface is measured to infer the local structure and properties of the subsurface layers of the Sun. The travel time changes due to the atmospheric magnetic field are evaluated theoretically. Coronal seismology utilises MHD waves in solar coronal structures as a tool to diagnose the physical parameters of the coronal plasma. Theoretical modelling of recently observed slow standing mode oscillations in stratified loops is presented. It is shown that large-amplitude resonant standing waves can be driven by small-amplitude oscillations at the chromospheric footpoints of the loops. The periods and the behaviour of these waves are different from those predicted by the classical theory of isothermal loop oscillations. The possible relationsheep between the oscillations studied in time-distance helioseismology and in coronal seismology is addressed.

KEYWORDS: *Sun, MHD waves, time-distance helioseismology, coronal seismology*

1. Introduction

Helisoeismology studies the internal structure and properties of the Sun by examining the oscillations observed on the solar surface. The measurements are carried out by observing oscillations in the intensity or in the Doppler shift of a spectral line as the plasma producing the spectral line oscillates back and forth along the line of sight. These oscillations are known to have a discrete spectrum. Each mode is described by the eigenfunctions of the wave equation of a spherical system and is characterised by a spherical harmonic degree l, radial order n, and azimuthal order m. In the $\omega - l$ diagram the eigenmodes reside on parabolic ridges. The p-mode ridges have been detected from degree $l = 0$ up to degree $l = 4000$. Radial orders up to $n = 35$ have been identified. The

lowest-order ridge with $n = 0$ is the fundamental, or the f-mode, characterised with the dispersion relation of surface waves in deep water $\omega^2 = gk_h$. The mode structure of the oscillations has been thoroughly investigated to study the internal structure of the Sun.

By measuring the frequencies of the modes, which extend over all longitudes, traditional helioseismology infers the rotation speed, sound speed and other parameters inside the Sun as functions of radius and latitude. However, it provides no information about the longitudinal variation of the properties of the Sun. Local helioseismology and, in particular, time-distance helioseismology has been useful for determining local properties of the subsurface layers of the Sun. The basic idea proposed by Duvall *et al.* (1993) is to measure the acoustic travel time between different points on the solar surface using the signal cross-correlation technique and then to use these measurements for inferring variations of internal properties such as the sound speed, magnetic field and flow velocities along the wave paths connecting the surface points. In the present paper the relationship between the time-distance helioseismology and the magnetic atmosphere of the Sun is clarified.

It is shown that the magnetic atmosphere of the Sun is expected to have a considerable effect on the solar oscillations. A natural question arises: could the opposite effect be true, i.e., could the solar oscillations propagate into the atmosphere and play any significant role in the corona?

Some 30-40 years ago the heating of the Sun's upper atmosphere by acoustic waves was a popular heating mechanism. Until recently these waves were thought to be important only below the lower chromosphere where they quickly dissipate due to shock formation. Since the launch of the SOHO and TRACE satellites there has been an explosion of observational evidences for MHD waves in solar coronal structures. In particular, different types of waves in coronal loops have been detected. In this report we are mainly interested in slow standing waves. The origin of these waves remains unclear in the majority of the examined cases. Here we consider the excitation of slow modes by a small-amplitude oscillation at the chromospheric footpoint of the loop. The period and the behaviour of the waves changes when gravity is introduced. The results are compared with recent SOHO/SUMER observations.

2. Signal Cross-Correlation

Time-distance helioseismology is based on the concept of signal cross-correlation. The cross-correlation function of the oscillation signal $f(t, \mathbf{r})$ between two dif-

ferent points on the solar surface is

$$\Psi(t, \Delta r) = \int f(\tau, \mathbf{r_1}) f^*(\tau + t, \mathbf{r_2}) d\tau, \tag{1}$$

where the integral is taken over the time of observation, Δr is the distance between the points $\mathbf{r_1}$ and $\mathbf{r_2}$ on the solar surface, t is the delay time, $*$ denotes the complex conjugate. The signal f is the velocity or intensity variation. One can represent it as a Fourier integral and use the cross-correlation theorem (the cross-correlation in the time domain is equivalent to the cross-correlation of the inverse Fourier transform of the signal in the frequency domain) to obtain

$$\Psi(t, \Delta r) = \frac{1}{2\pi} \int_{-\infty}^{\infty} A^2(\omega) e^{i\mu(\omega)t} d\omega, \tag{2}$$

where $A(\omega)$ is the distribution of the amplitude over the frequency,

$$\mu(\omega) = \omega - \frac{1}{t} \int \mathbf{k} \cdot d\mathbf{r} \tag{3}$$

and the line integral is taken along some path from $\mathbf{r_1}$ to $\mathbf{r_2}$. D'Silva (1996) has shown that if the amplitude $A(\omega)$ is a Gaussian, or, in general, is appreciable only in a small frequency interval around some ω_0, then the cross-correlation function attains its maximum along the path

$$\frac{d\mathbf{r}}{dt} = \left. \frac{\partial \omega}{\partial \mathbf{k}} \right|_0 \tag{4}$$

determined by the group velocity of the wave packet at the central frequency ω_0. In other words, the peak of the cross-correlation function obtained by cross-correlating the signals at two different points provides the time taken by the wave packet to travel between them. This time is called the group travel time and is given by

$$t_g = \int \left. \frac{\partial \mathbf{k}}{\partial \omega} \right|_0 \cdot d\mathbf{r}. \tag{5}$$

The phase peak of the cross-correlation function provides the phase travel time of the wave packet. In a nondispersive system these two times are equal to each other.

3. Travel Time in the Presence of an Atmospheric Magnetic Field

Below the photosphere the magnetic field resides in the form of flux tubes which form sunspots or remain isolated from one another as intense tubes. Above the photosphere the flux tubes spread out into the atmosphere creating a horizontal magnetic canopy. The field strength within such canopies and their heights depend upon whether they are in active regions associated with sunspots or in quiet regions far from them.

In order to evaluate the travel time in the presence of an atmospheric magnetic field consider a plane parallel equilibrium model in which the solar interior ($z > 0$) is permeated by a uniform equilibrium flow along the x-axis $\mathbf{V} = (V, 0, 0)$ and has a temperature profile linearly increasing with the depth. Gravity is constant in the interior and acts downwards. Above the interior lies the isothermal atmosphere permeated by a uniform magnetic field along the x-axis $\mathbf{B_0} = (B_0, 0, 0)$. The temperature in the isothermal atmosphere ($z < 0$) is assumed to be equal to that at the photospheric top of the field-free convection zone.

The travel time is determined by the dispersion relation of the medium in which the waves propagate and therefore in order to be able to calculate the travel time in the presence of an atmospheric magnetic field we have to derive a dispersion relation for such a medium. The wave motion is described by the linearised ideal MHD equations: These equations are reduced to second order ODEs and solved separately in the convection zone and in the atmosphere. The solutions must be selected in such a way that the wave energy density remains finite in the entire space. The normal component of the Lagrangian displacement and the total pressure perturbation are required to be continuous across the interface separating the two regions. The condition for the existence of a nontrivial solution yields the dispersion relation

$$2ak_x\omega_D^2\frac{U(1-a,m+3,2k_xz_0)}{U(-a,m+2,2k_xz_0)}+\frac{\gamma g\omega_D^2}{c_0^2}-k_x\omega_D^2$$
$$-gk_x^2=\frac{(\omega^2-k_x^2c_0^2)(g^2k_x^2-\omega_D^4)\rho_{0p}/\rho_{0c}}{gk_x^2c_0^2+(c_0^2+c_{Ac}^2)(\omega^2-k_x^2c_{Tc}^2)k_x\varphi}. \qquad (6)$$

where $\omega_D = \omega - k_xV$ is the Doppler shifted frequency, γ is the adiabatic index, c_0 is the sound speed at the temperature minimum, c_{Ac} and c_{Tc} are the Alvén

and cusp speeds in the atmosphere, the parameter a is given by

$$a = \frac{m+1}{\gamma}\frac{\omega_D^2}{2gk_x} + \left(m - \frac{m+1}{\gamma}\right)\frac{gk_x}{2\omega_D^2} - \frac{m}{2} - 1, \tag{7}$$

and

$$m = \frac{\gamma g z_0}{c_0^2} - 1 \tag{8}$$

is the polytropic index, $z_0/(1+m)$ is the pressure scale height at $z = 0$

$$\varphi = 1 - \frac{pqA_1A_3}{k_x r A_2}\frac{F\left(p+1, q+1, r+1, -\frac{A_1}{A_2}\right)}{F\left(p, q, r, -\frac{A_1}{A_2}\right)}, \tag{9}$$

$A_i = A_i(k_x, \omega)$ are some quantities depending on k_x and ω, U is the confluent hypergeometric function of the second kind, and F is the generalised hypergeometric function. In a static equilibrium the obtained dispersion relation is reduced to the one derived by Evans and Roberts (1990).

3.1. The Group Travel Time

We can now go back to the travel time formula and using the obtained dispersion relation calculate the travel time in the presence of a large-scale sub-surface flow and an atmospheric magnetic field. By representing $d\mathbf{r}$ in the form

$$d\mathbf{r} = \left(1, \frac{dz}{dx}\right)dx \tag{10}$$

we may rewrite the travel time formula in the form

$$t_g = \int_{x_1}^{x_2}\left(\frac{\partial k_x}{\partial \omega} + \frac{\partial k_z}{\partial \omega}\frac{dz}{dx}\right)dx, \tag{11}$$

where x_1 and x_2 are the x-coordinates of the two reflection points on the surface. The travel path is determined by $\dfrac{dz}{dx} = \dfrac{\partial k_x}{\partial k_z}$. Substituting this expression in the previous formula we arrive at

$$t_g = \frac{2\Delta x}{\partial \omega / \partial k_x}, \tag{12}$$

where $\Delta x = x_2 - x_1$ is the distance between the reflection points on the surface. The dependence of t_g on Δx is not linear since k_x is a function of Δx. By integrating the x-component of the equation $\frac{d\mathbf{r}}{dt} = \frac{\partial \omega}{\partial \mathbf{k}}$ and using the last expression for the group travel time t_g, we obtain

$$\Delta x = C t_g^2, \tag{13}$$

where C is a constant. This expression has been established observationally by Duvall *et al.* (1993) and it states that the cross-correlation between two intensity fluctuations on the solar surface separated by a distance Δx is maximal along a parabolic curve of the form $\Delta x \propto t_g^2$ known as the time-distance curve. In a nondispersive medium the expression for the travel time is reduced to

$$t = \frac{2\Delta x}{\omega / k_x}, \tag{14}$$

which has been derived by Price (2000).

3.2. An Asymptotic Result

In order to calculate the travel time in a dispersive medium one has to calculate the derivatives of the hypergeometric functions with respect to their parameters which is a difficult task. However, it turns out that under certain conditions it is possible to reduce the dispersion relation to a simple asymptotic expression for ω which is much easier to deal with than the original dispersion relation. The measure of the pressure scale height $z_0 \sim 0.25$ Mm, so that the argument of the hypergeometric functions $k_x z_0$ is small for not very large horizontal wavenumbers. This allows us to expand both sides of the dispersion relation into Taylor series. After some algebra we obtain the following asymptotic result

$$\begin{aligned} \frac{\omega}{\sqrt{g k_x}} &= \Omega_n + \frac{V}{c_0}\sqrt{\frac{\gamma}{2(m+1)}}(2k_x z_0)^{\frac{1}{2}} \\ &+ \frac{\gamma \Gamma(1+m+n)}{\Gamma(1+m)\Gamma(2+m)}\left[2\frac{m+1}{\gamma} - m\right]^{-1} \\ &\times \frac{(1-\Omega_n^4 - 2\Omega_n^2)(2k_x z_0)^{m+1}}{\Gamma(n)(\gamma + 2\beta)\Omega_n(1-\Omega_n^4)}, \end{aligned} \tag{15}$$

where Γ is the Γ-function, $\beta = c_{sc}^2/c_{Ac}^2$ is the square of the ratio of the sound and Alfvén speeds in the atmosphere and Ω_n is determined from the equation

$$\frac{m+1}{\gamma}\Omega_n^2 + \left(m - \frac{m+1}{\gamma}\right)\frac{1}{\Omega_n^2} - m = 2n. \tag{16}$$

The derivation details of the asymptotic result are given in Erdélyi and Taroyan (2001). The x-component of the group velocity can be determined as follows:

$$\frac{\partial\omega}{\partial k_x} = \frac{1}{2}\sqrt{\frac{g}{k_x}}\Omega_n + V + \frac{\gamma(m+3/2)\Gamma(1+m+n)}{\Gamma(n)\Gamma(1+m)\Gamma(2+m)} \times \frac{\sqrt{g}(1-\Omega_n^4-2\Omega_n^2)(2z_0)^{m+1}k_x^{m+\frac{1}{2}}}{(2\frac{m+1}{\gamma}-m)\Omega_n(1-\Omega_n^4)(\gamma+2\beta)}. \tag{17}$$

3.3. Numerical Results

In the numerical results presented below we have set $\gamma = 5/3$, $m = 3/2$, $g = 274$ km/s, $z_0 = 0.25$ Mm. The temperature at the interface separating the convection zone and the atmosphere is $T = 4170$ K. The angular frequency ω is replaced by the cyclic frequency $\nu = \omega/(2\pi)$. The solar radius is $R_\odot = 696$ Mm.

Fig. 1a shows the relative difference $D = 100(t_g^0 - t_g)/t_g^0$ of the travel times in nonmagnetic and magnetic states as a function of the magnetic field strength when the flow speed $V = 0$. Here $t_g^0 = t_g(V = 0, B = 0)$ is the travel time in the absence of the flow and magnetic field. The horizontal distance Δx is kept fixed and the values for the harmonic degree and frequency are $l = 200$, $\nu \approx 2.1 - 4.2$ mHz. It can be seen that the magnetic field reduces the travel time by speeding up the wave propagation. This effect is due to the greater rigidity at the interface introduced by the uniform magnetic field. For the given parameter values the magnetic field of $B_0 = 20$ G may shorten the travel time between two surface points by up to 10 %. For comparison we plot D as a function of the flow speed V measured in km/s (Fig. 1a). Here the atmosphere is assumed to be nonmagnetic while other parameters are the same. The flow is assumed to be in the same direction as the wave propagation and therefore reduces the travel time. The horizontal axis covers a wide range of flow speeds from those typical for meridional flows (10-20 m/s) to the equatorial rotation speed of 2 km/s.

We might have slightly overestimated the role of the magnetic field by taking zero height of the magnetic canopy and assuming a strongly uniform magnetic field. However, our results indicate that the magnetic field in the atmosphere should have a considerable effect on the travel time of the acoustic waves inside the Sun. The comparison with the flow effects shows that the magnetic effects should be measurable. On the other hand, by measuring and comparing the travel times in quiet and active regions one could in principle gain some infor-

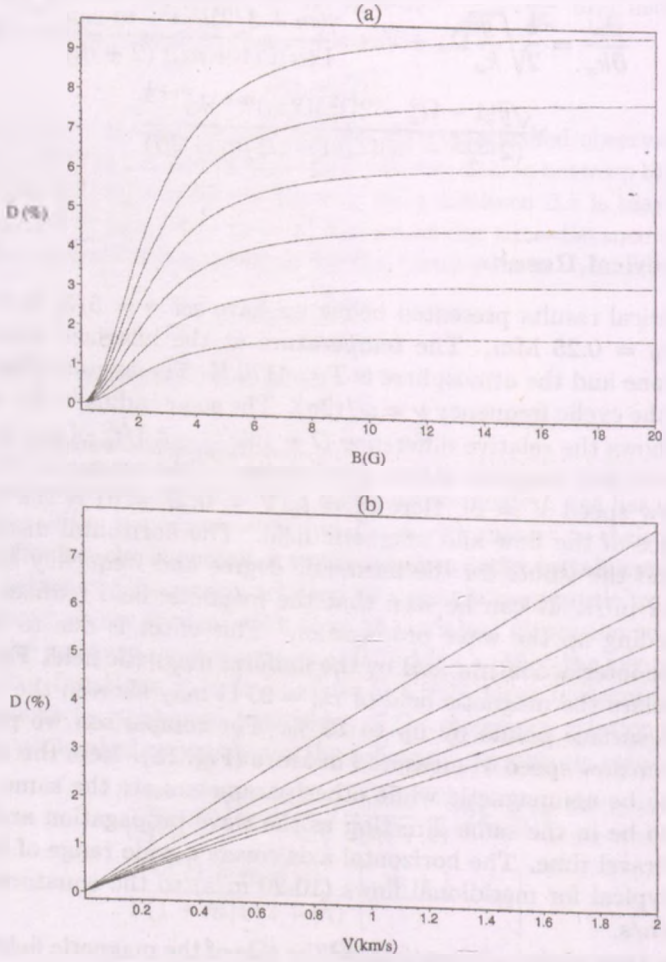

Figure 1: The relative time difference $D = 100(t_g^0 - t_g)/t_g^0$, where $t_g^0 = t_g(V = 0, B = 0)$; (a): as a function of the magnetic field strength when $V = 0$; (b): as a function of the flow speed when $B_0 = 0$. The horizontal distance Δx is kept fixed and the chosen parameter values are $l = 200$, $\nu \approx 2.1 - 4.2$ mHz.

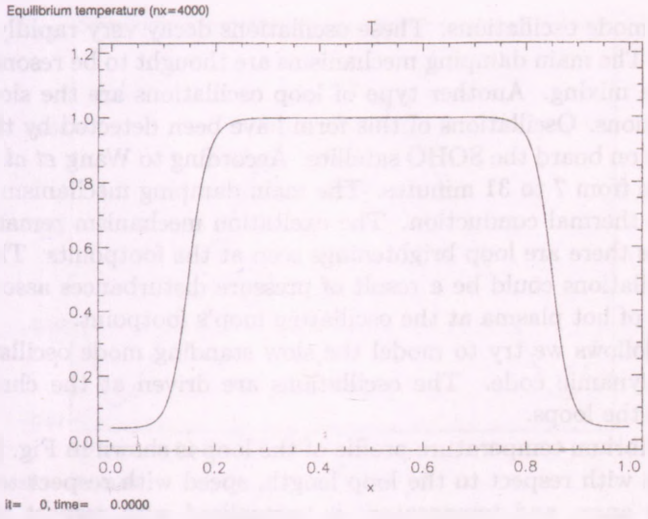

Figure 2: The equilibrium temperature along the loop between the two chromospheric footpoints.

mation about the structure and the properties of the magnetic field of the Sun's atmosphere.

4. Slow standing mode oscillations in coronal loops

The results of the preceding sections show that the magnetic atmosphere of the Sun is expected to have measurable effects on the solar oscillations. Could the opposite effect be true, i.e., could the solar oscillations affect the atmosphere. A particularly popular example is the p-mode dissipation by photospheric flux tubes (Erdélyi and Gossens 1994, 1996; Pintér et al. 1996). The idea of coronal heating by acoustic waves is not new. It has been a popular heating mechanism some 30-40 years ago. Later this idea was abandoned due to very strong arguments. The p-modes are expected to form shocks in the most lower parts of the solar atmosphere and could be important in the heating of the lower chromosphere (De Pontieu et al. 2003). Recent SOHO/TRACE observations have shown clear evidence of different types of oscillations in closed and open coronal structures (see the reviews in Aschwanden, 2003, Roberts and Nakariakov, 2003). Recently detected kink mode oscillations of coronal loops are most likely excited by flares. A flare triggers a CME, which, in turn, hits the loop and

excites kink mode oscillations. These oscillations decay very rapidly (Mendoza et al. 2004). The main damping mechanisms are thought to be resonant absorption or phase mixing. Another type of loop oscillations are the slow standing mode oscillations. Oscillations of this form have been detected by the SUMER spectrometer on board the SOHO satellite. According to Wang *et al.* (2003) the periods range from 7 to 31 minutes. The main damping mechanism is believed to be due to thermal conduction. The excitation mechanism remains unclear. In some cases there are loop brightenings seen at the footpoints. This suggests that the oscillations could be a result of pressure disturbances associated with the injection of hot plasma at the oscillating loop's footpoint.

In what follows we try to model the slow standing mode oscillations using a 1D hydrodynamic code. The oscillations are driven at the chromospheric footpoints of the loops.

The equilibrium temperature profile of the loop is shown in Fig. 3. Distance is normalised with respect to the loop length, speed with respect to the sound speed at the apex, and temperature is normalised with respect to the apex temperature. The approximate position and width of the transition region are $z_t = 0.15$ and $w_t = 0.1$, respectively.

Small amplitude oscillations are continuously driven at the left footpoint of the loop. The classical theory for isothermal loops predicts a slow standing wave period $P = 2L/c_s$, where L is the loop length and c_s is the sound speed. Here P is the period of the fundamental mode. Numerical simulations show that in a stratified loop the above expression for P is no loger valid. We modify this expression by subtracting $2z_t$ from L. The new period is

$$P = \frac{2(L - 2z_t)}{c_s(\frac{L}{2})} \tag{18}$$

and this turns out to be the fundamental slow mode period.

Small amplitude oscillations with the fundamental mode period and a velocity amplitude $v = 0.001$ are continually driven at the left boundary and are reflected from the right boundary. A snapshot of the resulting slow standing mode oscillation is shown in Fig. 4. The amplitude reaches its maximum at the apex. There are two nodes located at the centre of the transition region on the corresponding side of the loop. The period of the driver matches the natural oscillation period of the loop. Therefore, the amplitude of this resonant oscillation gradually increases until the ideal MHD approximation is no longer valid. Fig. 4 shows that the amplitude at the apex can be as large as 50 times the amplitude at the left footpoint.

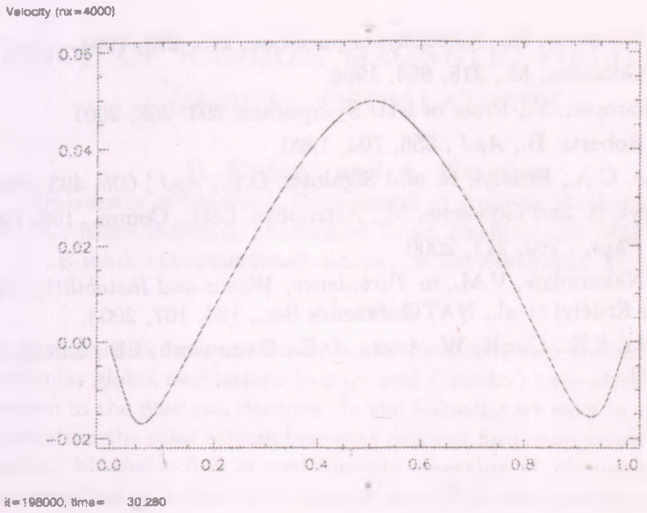

Figure 3: The oscillation velocity along the loop. Oscillations with an amplitude $v = 0.001$ are continually driven at the left footpoint.

The presented simulations show that the slow standing mode oscillations observed by SUMER could be driven by small amplitude oscillations at the chromospheric footpoints of the loops. In a future study the effects of thermal conduction radiative losses and compressive viscosity must be taken into account.

Acknowledgments

Y. Taroyan acknowledges the organisers of the meeting and the British Council for the invitation and financial support.

References

Aschwanden, M., in *Turbulence, Waves and Instabilities in the Solar Plasma*, (eds.) Erdélyi et al., NATO, Science Ser., 124, 215, 2003.

D'Silva, S., *ApJ* , 469, 964, 1996.

De Pontieu, B., Erdélyi, R. and De Wijn, A.G., *ApJ* , 595, L63, 2003.

Duvall, T.L., Jr., Jefferies, S.M., Harvey, J.W., Pomerantz, M.A., *Nature*, 362, 430, 1993.

Erdélyi, R. and Goossens, M., Astrophys. Space Sci., 213, 273, 1994.

Erdélyi, R. and Goossens, M., 313, 664, 1996.

Erdély, R. and Taroyan, Y., Proc. of IAU Symposium 203, 208, 2001.

Evans D.J. and Roberts, B., *ApJ* , 356, 704, 1990.

Mendoza-Briceno, C.A., Erdélyi, R. and Sigalotti, D.L., *ApJ* , 605, 493, 2004.

Pintér, B., Erdélyi, R. and Goossens, M., Astrophys. Lett., Comm., 169, 1996.

Price, G., *Solar Phys.* , 192, 211, 2000.

Roberts, B. and Nakariakov, V.M., in *Turbulence, Waves and Instabilities in the Solar Plasma*, (eds.) Erdélyi et al., NATO, Science Ser., 124, 167, 2003.

Wang, T., Solanki, S.K., Curdt, W., Innes, D.E., Dammasch, I.E., Kliem, B., *A&A* , 406, 1105, 2003.

INFLUENCE OF RANDOM MAGNETIC FIELD ON SOLAR GLOBAL OSCILLATIONS

R. Erdélyi and A. Kerekes
University of Sheffield, Department of Applied Mathematics
SPARC, Hicks Building, Hounsfield Road, Sheffield, S3 7RH, England
E-mail: robertus@shef.ac.uk, a.kerekes@shef.ac.uk

Abstract

The discrepancies between theoretically predicted and observed frequencies of solar global oscillations (e.g. p- and f-modes) have attracted major attention in the past two decades. In the following we wish to explore further whether the solar atmosphere may account for the apparent frequency paradox. Magnetic flux is continuously emerging at photospheric levels and expanding into the solar atmosphere. We investigate the possible effects of an atmospheric *random* magnetic field on the solar fundamental mode (f-mode).

1. Introduction

Magnetic field observations revealed that the photospheric field of the Sun is a complex, mixed polarity network with flux continuously emerging and disappearing nearly homogeneously over the surface. The replacement time of the quiet Sun flux is about 40 hours, the mean absolute flux density is estimated to be about 2 Gauss (Title & Schrijver, 1998).
The large-scale magnetic patterns vary with the solar cycle but the magnetic carpet is always present on the solar surface. The structure of this small-scale field is locally changing on very short spatial and time scales but its overall features seem to be persistent. The temporal evolution of the whole complex structure is so far unknown, the detailed mapping of the small-scale field requires more sensitive magnetograms.

2. Random magnetic field and the f-mode

The f-mode is essentially a surface gravity wave with the simple dispersion relation $\omega^2 = gk_h$, where g is the surface gravity and k_h is the horizontal wavenumber. It is apparent from this relation that the mode frequencies are independent of the internal stratification, so this mode is expected to be sensitive to near-surface effects only, like interaction with granular flows (Murawski &

Roberts, 1993) and the emerging photospheric magnetic fields. The observations reveal a significant departure from the quoted parabolic dispersion relation; for low-degree modes the measured frequencies are slightly higher (Antia, 1998), for high-degree modes *several percent lower* (Libbrecht et al., 1990) than the theoretical values.
In the present paper we investigate the effect of the magnetic carpet on the solar global f-mode. The magnetic carpet is modelled as random magnetic field. We analyse whether such magnetic fluxes can reduce the eigenfrequencies of surface gravity waves as observed by Libbrecht et al., (1990).

3. The model

We model the magnetic carpet as a random magnetic field overlaying a field-free interior.

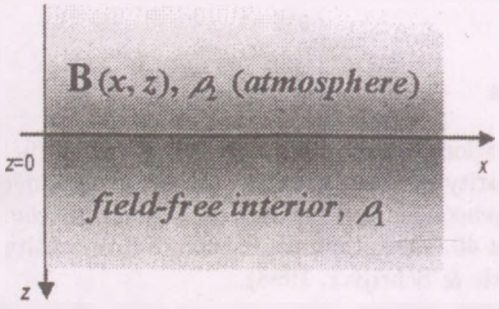

Figure 1: *Model*

We consider a simple 2D Cartesian model where the $z = 0$ surface corresponds to the solar surface (i.e. the bottom of the photosphere) with the z-axis directed downwards so that the gravitational acceleration is in the positive z-direction. In the initial state the magnetic field is random and time-independent. This is a good approximation as long as the characteristic time scales of the considered observations (i.e. f-modes) are much smaller than the replacement time for the magnetic carpet. Both the interior and the atmosphere are taken to be incompressible. No flows are present in the system.
The equation describing the undisturbed state is

$$0 = -\nabla p_{0,1} + \rho_1 g \hat{z}, \tag{1}$$

in $z > 0$, and

$$0 = -\nabla p_{0,2} + \frac{1}{\mu}(\nabla \times \mathbf{B_0}) \times \mathbf{B_0} + \rho_2 g \hat{\mathbf{z}} \tag{2}$$

in $z < 0$, where p is the fluid pressure, ρ_1, ρ_2 are the constant densities and g is the local gravitational acceleration.

Because of the $\nabla \cdot \mathbf{B} = 0$ condition the magnetic field $\mathbf{B}$ can always be represented by a vector potential $\mathbf{A}$, so that $\mathbf{B} = \nabla \times \mathbf{A}$. In the current two-dimensional case the vector potential can be chosen to have only one component

$$\mathbf{B}_0(x, z) = \nabla \times \mathbf{A}_0(x, z), \qquad \mathbf{A}_0 = [0, A_0, 0]. \tag{3}$$

The profile of the initial magnetic field is given by

$$A_0 = -z \exp(\alpha^2 z) b(x). \tag{4}$$

This way the field is confined to a layer whose thickness is determined by the α^2 decay factor. In the case of the solar atmosphere it is about a few 100 km which corresponds to $\alpha^2 = O(10^{-2})$. The ensemble average of the magnetic field is taken to be zero,

$$\langle b(x) \rangle = 0. \tag{5}$$

The observed oscillations have extremely small amplitudes compared to the background so they can be described as linear perturbations about the equilibrium state.

3.1. Wave perturbation

Let us now study the perturbed state of the equilibrium described in §3. The governing equations of the dynamical behavior of the perturbations are

$$\nabla \cdot \mathbf{v}_2 = 0, \tag{6}$$

$$\rho_1 \frac{\mathrm{d}\mathbf{v}_1}{\mathrm{d}t} = -\nabla p_1 + \rho_1 g \hat{\mathbf{z}} \tag{7}$$

in $z > 0$, and

$$\nabla \cdot \mathbf{v}_2 = 0, \tag{8}$$

$$\rho_2 \frac{\mathrm{d}\mathbf{v}_2}{\mathrm{d}t} = -\nabla p_2 + \frac{1}{\mu}(\nabla \times \mathbf{B}) \times \mathbf{B} + \rho_2 g \hat{\mathbf{z}}, \tag{9}$$

$$\partial_t \mathbf{B} = \nabla \times (\mathbf{v}_2 \times \mathbf{B}) \tag{10}$$

in $z < 0$. In the presence of a wave motion the generated velocity fields have to satisfy the incompressible condition ($\nabla \cdot \mathbf{v}_i = 0$, i=1,2) so they can be represented by the following vector potentials

$$\mathbf{v}_1(x,z,t) = \nabla \times \Psi_1(x,z,t), \tag{11}$$

$$\Psi_1 = [0, \Psi_1, 0], \tag{12}$$

$$\mathbf{v}_2(x,z,t) = \nabla \times \Psi_2(x,z,t), \tag{13}$$

$$\Psi_2 = [0, \Psi_2, 0]. \tag{14}$$

The boundary conditions at the perturbed surface, $z = \eta(x,z,t)$ are the kinematic boundary condition

$$\eta_t = (v_i \cdot \nabla)(z - \eta), \tag{15}$$

and the continuity of total pressure

$$p_1 = p_2 + \frac{B^2}{2\mu}. \tag{16}$$

Since we are interested in surface waves only, where the energy of waves is localized to the surface discontinuity, far away from the interface we require that

$$\Psi_1 \longrightarrow 0 \qquad \text{as} \qquad z \longrightarrow +\infty, \tag{17}$$

$$\Psi_2 \longrightarrow 0 \qquad \text{as} \qquad z \longrightarrow -\infty. \tag{18}$$

3.2. Dispersion relation

In order to obtain the dispersion relation we integrate the z-components of the equations of motion and substitute $z = \eta$. Using the continuity of pressure we can eliminate the pressure terms and the resulting equation connects the scalar functions Ψ_1, Ψ_2 and A on the perturbed surface $\eta(x,z,t)$.
It is straightforward to decompose the magnetic and velocity fields into equilibrium and perturbation values, and linearize them around the stationary state. No flow is present in the equilibrium state so $\Psi_1 = \delta\Psi_1$, $\Psi_2 = \delta\Psi_2$, $\eta = \delta\eta$, and for the magnetic field $A = A_0 + \delta A$. Discarding the δ-notations, at $z = 0$ we obtain the kinematic boundary condition

$$\eta_t = \Psi_{1x} = \Psi_{2x}\,, \tag{19}$$

the induction equation

$$A_t = \Psi_{2x} b\,, \tag{20}$$

and from the equation of motion

$$\kappa \int \Psi_{2xt}\,\mathrm{d}z - \int \Psi_{1xt}\,\mathrm{d}z + a\eta =$$
$$\frac{1}{\mu\rho_1}\int e^{\alpha^2 z}(1+\alpha^2 z)[bA_{xx} - b_x A_x]\,\mathrm{d}z + \frac{1}{\mu\rho_1}\int e^{\alpha^2 z} z[b_{xx}A_z - b_x A_{xz}]\,\mathrm{d}z. \tag{21}$$

Because of the presence of the random fields we divide the wave fields into random and coherent components

$$A = \langle A\rangle + A' \qquad \langle A'\rangle = 0, \tag{22}$$
$$\Psi_1 = \langle\Psi_1\rangle + \Psi_1' \qquad \langle\Psi_1'\rangle = 0, \tag{23}$$
$$\Psi_2 = \langle\Psi_2\rangle + \Psi_2' \qquad \langle\Psi_2'\rangle = 0, \tag{24}$$
$$\eta = \langle\eta\rangle + \eta' \qquad \langle\eta'\rangle = 0. \tag{25}$$

Substituting (22)–(25) into (19)–(21) and taking ensemble averages we obtain the equations for the coherent field at $z = 0$. Subtracting these equations from (19)–(21) and using the the *binary collision approximation* (Howe, 1971) we arrive at the equations governing the evolution of the random field at $z = 0$. From the induction equation we obtain

$$\langle A\rangle_t = \langle b\Psi_{2x}'\rangle, \tag{26}$$
$$A_t' = b\langle\Psi_{2x}\rangle, \tag{27}$$

from the kinematic boundary condition

$$\langle\eta\rangle_t = \langle\Psi_{1x}\rangle = \langle\Psi_{2x}\rangle, \tag{28}$$
$$\eta_t' = \Psi_{1x}' = \Psi_{2x}', \tag{29}$$

and from the equation of motion

$$\kappa\int\langle\Psi_2\rangle_{xt}\,\mathrm{d}z - \int\langle\Psi_1\rangle_{xt}\,\mathrm{d}z + a\langle\eta\rangle =$$
$$\frac{1}{\mu\rho_1}\int e^{\alpha^2 z}(1+\alpha^2 z)[\langle bA\rangle_{xx} - \langle b_x A\rangle_x]\,\mathrm{d}z$$
$$+\frac{1}{\mu\rho_1}\int e^{\alpha^2 z} z[\langle b_{xx}A_z\rangle - \langle b_x A_{xz}\rangle]\,\mathrm{d}z, \tag{30}$$

and

$$\kappa \int \Psi'_{2xt}\,\mathrm{d}z - \int \Psi'_{1xt}\,\mathrm{d}z + a\eta' = \frac{1}{\mu\rho_1}\int e^{\alpha^2 z}(1+\alpha^2 z)[b\langle A\rangle_{xx} - b_x\langle A\rangle_x]\,\mathrm{d}z + \frac{1}{\mu\rho_1}\int e^{\alpha^2 z} z[b_{xx}\langle A\rangle_z - b_x\langle A\rangle_{xz}]\,\mathrm{d}z. \quad (31)$$

It is straightforward to write the perturbation quantities in Fourier form, so that

$$A'(x,z,t) = F^{-1}\{e^{|k|z}\hat{A}(k,\omega)\} = \frac{1}{4\pi^2}\int \hat{A}(k,\omega)e^{|k|z}e^{i(kx-\omega t)}\,\mathrm{d}\omega\,\mathrm{d}k, \quad (32)$$

$$\langle A\rangle(x,z,t) = F^{-1}\{e^{|k|z}\bar{A}(k,\omega)\} = \frac{1}{4\pi^2}\int \bar{A}(k,\omega)e^{|k|z}e^{i(kx-\omega t)}\,\mathrm{d}\omega\,\mathrm{d}k, \quad (33)$$

$$\Psi'_2(x,z,t) = F^{-1}\{e^{|k|z}\hat{\Psi}_2(k,\omega)\}, \quad (34)$$

$$\langle\Psi_2\rangle(x,z,t) = F^{-1}\{e^{|k|z}\bar{\Psi}_2(k,\omega)\}, \quad (35)$$

$$\Psi'_1(x,z,t) = F^{-1}\{e^{-|k|z}\hat{\Psi}_1(k,\omega)\}, \quad (36)$$

$$\langle\Psi_1\rangle(x,z,t) = F^{-1}\{e^{-|k|z}\bar{\Psi}_1(k,\omega)\}, \quad (37)$$

$$\eta'(x,t) = F^{-1}\{\hat{\eta}(k,\omega)\}, \quad (38)$$

$$\langle\eta\rangle(x,t) = F^{-1}\{\bar{\eta}(k,\omega)\}. \quad (39)$$

We introduce the following new notations

$$\kappa = \rho_2/\rho_1 \quad \text{density ratio,} \quad (40)$$

$$a = (1-\kappa)g \quad \text{effective gravity,} \quad (41)$$

$$R_0(x-x_1) = \langle b(x)b(x_1)\rangle \quad \text{correlation function,} \quad (42)$$

$$\bar{R}_0(k-k_1) \quad \text{Fourier transformed correlation function.} \quad (43)$$

Applying Fourier transforms to (26)–(31) leads to the dispersion relation

$$(\kappa+1)\omega^2\,\mathrm{sgn}(k) - ak = \frac{k^2}{\mu\rho_1}\int_{-\infty}^{\infty}\frac{|k_1|(2k_1-k)\hat{R}_0(k-k_1)}{(\alpha^2+|k_1|)^2}\,\mathrm{d}k_1. \tag{44}$$

Let us intorduce a Gaussian correlation function and its Fourier transform

$$R_0(x) = {\sigma_0}^2 e^{-x^2/4l^2}, \qquad \hat{R}_0(k) = {\sigma_0}^2\frac{l}{\sqrt{\pi}}e^{-k^2l^2}, \tag{45}$$

where σ_0 and l are the characteristic strength and length-scale of the magnetic field. Substitution of the correlation function into (44) yields

$$(\kappa+1)\omega^2\,\mathrm{sgn}(k) - ak = \frac{\sigma_0^2k^2l}{\mu\rho_1\sqrt{\pi}}\int_{-\infty}^{\infty}\frac{|k_1|(2k_1-k)}{(\alpha^2+|k_1|)^2}e^{-(k-k_1)^2l^2}\,\mathrm{d}k_1. \tag{46}$$

In the field-free case the right-hand side of the dispersion equation is zero returning the standard dispersion relation for surface waves. The random magnetic field results in a correction to this dispersion relation where the correction is given by the RHS of (46).

3.3. Numerical results

In order to analyze the dispersion relation (46) we introduce the following non-dimensional variables:

$$K = kl\,, \qquad K_1 = k_1l$$

denote the dimensionless wavenumbers,

$$p = \alpha^2 l$$

is the dimensionless decay factor,

$$\Omega^2 = (\kappa+1)\frac{l}{a}\omega^2$$

is the dimensionless frequency, and finally

$$\frac{\sigma_0^2}{\mu\rho_1} = \kappa\frac{\sigma_0^2}{\mu\rho_2} = \kappa v_{Ar}^2,$$

where v_{Ar}^2 is the (random) Alfvén speed. Equation (46) can be written in its dimensionless form

$$\Omega^2 \operatorname{sgn}(K) - K = \kappa v_{Ar}^2 \frac{K^2}{al\sqrt{\pi}} \int_{-\infty}^{\infty} \frac{|K_1|(2K_1 - K)}{(p + |K_1|)^2} e^{-(K-K_1)^2} \, \mathrm{d}K_1 . \tag{47}$$

The dispersion relation is solved numerically in the form

$$\Omega^2 \operatorname{sgn}(K) - K = 2\frac{\kappa v_{Ar}^2}{al\sqrt{\pi}} K^2 (2I_1 - KI_2) , \tag{48}$$

where

$$I_1 = \frac{\sqrt{\pi}}{2} \operatorname{erf}(K) - 2p \int_0^{\infty} \frac{e^{-s^2}}{s+p} \sinh(2Ks) \, \mathrm{d}s + p^2 \int_0^{\infty} \frac{e^{-s^2}}{(s+p)^2} \sinh(2Ks) \, \mathrm{d}s , \tag{49}$$

and

$$I_2 = \int_0^{\infty} \frac{e^{-s^2}}{s+p} \cosh(2Ks) \, \mathrm{d}s - p \int_0^{\infty} \frac{e^{-s^2}}{(s+p)^2} \cosh(2Ks) \, \mathrm{d}s . \tag{50}$$

Numerical results are displayed in Figs. 2-5. Figure 2a-d shows the *f*-mode frequencies and the obtained correction to the field-free case. It is apparent from these figures that the presence of the random magnetic field results in a frequency *increase* for all wavenumbers. Our findings are in qualitative agreement with the observations for small wavenumbers, but for large values of k the f-mode frequencies should be appreciably reduced. Similar results were obtained by Campbell & Roberts (1989) and Evans & Roberts (1990) who investigated the effects of the magnetic canopy (modeled as uniform or exponentially decaying magnetic field) on global oscillations. This clearly shows that magnetic fields alone cannot be responsible for the observed discrepancies.

The influence of the magnetic field is assessed by applying different Alfvén speed values (Fig. 3). Due to observational difficulties, the small-scale field cannot be measured accurately because the diameter of the individual flux tubes is probably much smaller than the resolution of MDI (on-board SOHO). The instrument measures the magnetic flux of an area about 1 *arcs* x 1 *arcs*. The average flux density, which is the measured flux divided by the corresponding area, can only give a lower limit for the value of the magnetic field concentrated in thin flux tubes (so-called filling factor problem). On the other hand, it has been established that the sound speed and the Alfvén speed are at the same order in the solar atmosphere, so we considered v_{Ar} values between 0.1 and 10

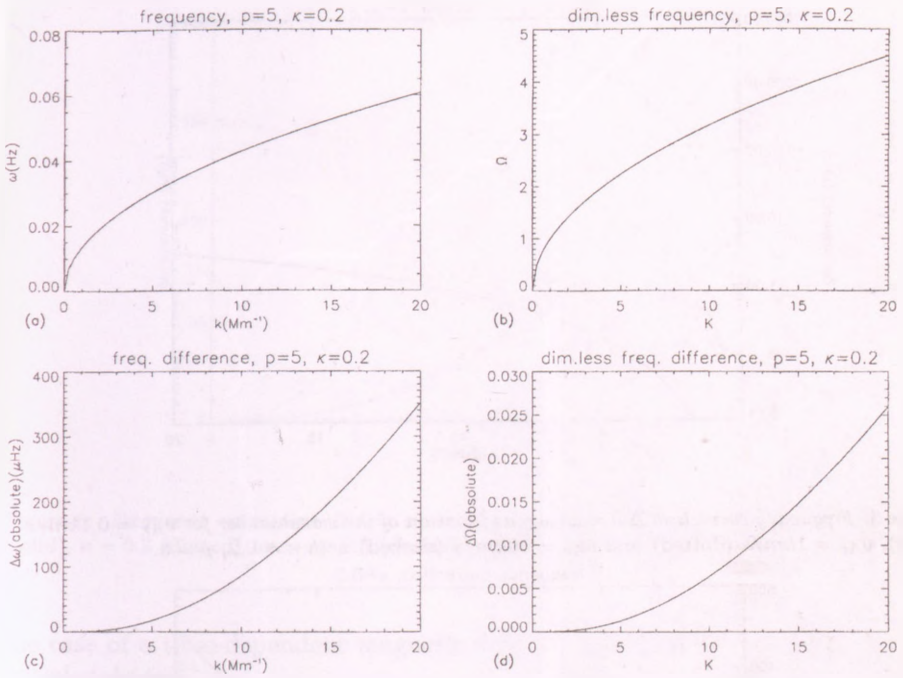

Figure 2: *Frequency and frequency difference as functions of the wavenumber for* $\kappa = 0.2$, $l = 1000km$, $p = 5$ *and* $v_{Ar} = 1km/s$

km/s.

The random magnetic field in the model is confined to a layer whose thickness is prescribed by the parameter p. The small-scale field is expected to dominate a few 100 km in the photosphere, which places p between 2 and 10. Fig. 4 shows the variations of the correction within this range.

We also found (Fig. 5), that the frequency correction largely depends on the ratio of the density of the atmosphere and the density of the interior, apparently, for high contrast the correction is decreased.

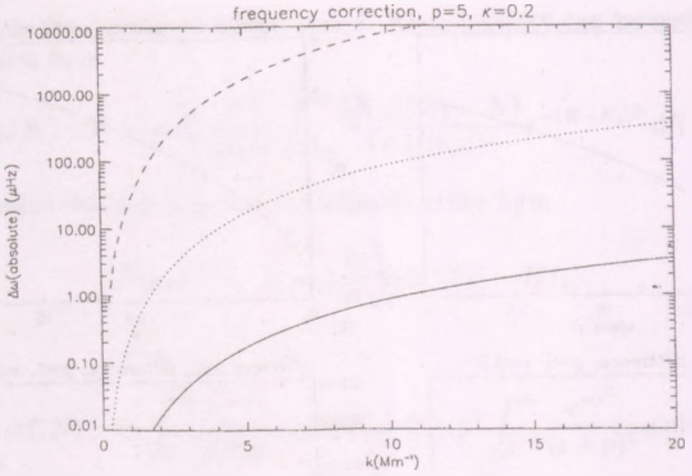

Figure 3: *Frequency correction* $\Delta\omega = \omega - \omega_0$ *as function of the wavenumber for* $v_{Ar} = 0.1km/s$ (solid), $v_{Ar} = 1km/s$ (dotted) *and* $v_{Ar} = 10km/s$ (dashed) *with* $\kappa = 0.2$, $p = 5$.

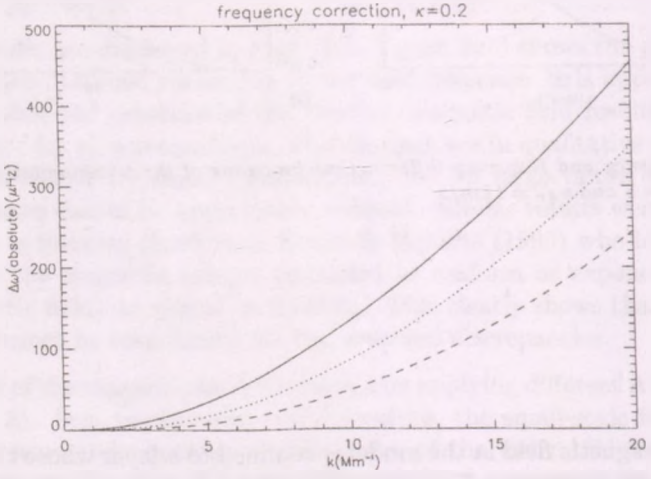

Figure 4: *Frequency correction* $\Delta\omega = \omega - \omega_0$ *as functions of the wavenumber for* $p = 2$ (solid), $p = 5$ (dotted) *and* $p = 10$ (dashed), *with* $v_{Ar} = 1km/s$, $\kappa = 0.2$

4. Future Plans

We are planning to extend our study to more realistic models. Observations clearly reveal that the real structure shows temporal variations as well, and in

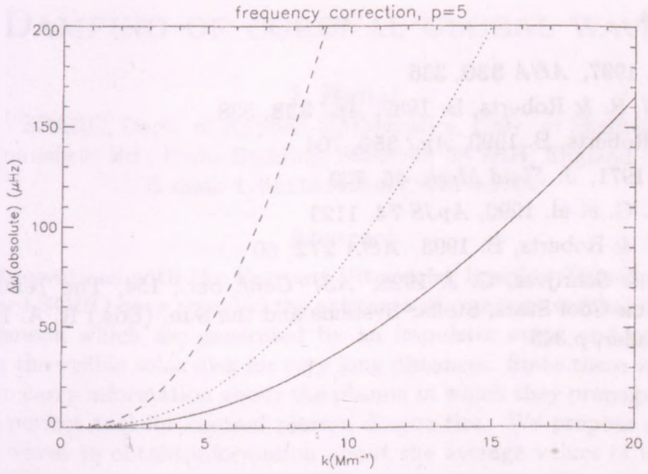

Figure 5: *Frequency correction* $\Delta\omega = \omega - \omega_0$ *as function of the wavenumber for* $\kappa = 0.1$ (solid), $\kappa = 0.2$ (dotted) *and* $\kappa = 0.5$ (dashed), *with* $v_{Ar} = 1km/s$, $p = 5$.

the case of a time-dependent magnetic field we have to deal with flows in the initial state too.

The assumption of the 2D Cartesian geometry removes some of the substantial physical effects, so we also have to improve our geometry.

Our knowledge of the small-scale solar magnetic field is rather limited and so far there is no information about any possible connection between the solar cycle and the evolution of the magnetic carpet. One way to model the observed changes with the solar cycle is to include coherent background field in the current model.

Acknowledgments

The authors acknowledge the financial support from NFS Hungary (OTKA, TO43741) and the support received from The British Council for the N+N Workshop for Young Researchers on 'Computer processing and use of satellite data in astronomy and astrophysics' held in Budapest, Hungary, where this work was presented. RE acknowledges M. Kéray for patient encouragement. This work was supported in part by the OTKA under grant no. T043741 and by the Royal Society grant no. 15599.

References

Antia, H. M. 1997, *A&A* **330**, 336

Campbell, W. R. & Roberts, B. 1989, *ApJ* **338**, 538

Evans, D & Roberts, B. 1990, *ApJ* **356**, 704

Howe, M. S. 1971, *J. Fluid Mech.* **45**, 769

Libbrecht, K. G. et al. 1990, *ApJS* **74**, 1129

Murawski, K. & Roberts, B. 1993, *A&A* **272**, 60

Title, A. M. & Schrijver, C. J. 1998, *ASP Conf. Ser.*, 154, The Tenth Cambridge Workshop on Cool Stars, Stellar Systems and the Sun, (Eds.) R. A. Donahue & J. A. Bookbinder, p.345

Damping of coronal global waves

I. Ballai
SPARC, Dept. of Applied Mathematics, Univ. of Sheffield
Hounsfield Rd., Hicks Building, Sheffield, S3 7RH, England (UK)
E-mail: i.ballai@sheffield.ac.uk

Abstract

Observations with the Extreme Ultraviolet Imaging Telescope (EIT) onboard SOHO have revealed the existence of transient coronal wave-like phenomena which are generated by an impulsive event and propagate across the visible solar disk for very long distances. Since these waves are able to carry information about the plasma in which they propagate, they are a perfect tool for coronal plasma diagnostics. We propose to study these waves to obtain information about the average values of the magnetic field and mean transport coefficients in the quiet Sun. The present contribution answers some of the 'hottest' questions related to the nature and proprieties of EIT waves.

KEYWORDS: *Coronal waves, EIT waves, TRACE waves, global MHD waves*

1. Introduction

Recent high-resolution observations with space telescopes have made it very clear that eruptive events such as flares or coronal mass ejections (CME) can generate waves and/or shocks which can propagate over very long distances, sometimes comparable to the solar radius. These waves and shocks have a very important role in, e.g. acceleration of particles, could give valuable information about the flare/CME process and can also carry information about the plasma in which they propagate. Due to their propagation characteristics, these waves are different from those observed recently in coronal structures (e.g. loops), called local coronal waves.

A pressure pulse generated by a flare or CME gives rise to a circularly propagating wave-front seen by SOHO/MDI (hence their name, *MDI waves*) in the solar photosphere (Kosovichev and Zharkova, 1998). These waves propagate as concentric rings, similar to the ripples spreading from a stone dropped into water, with a speed of 30-100 km/s over distances of the order of 120 Mm.

Going up in the solar atmosphere, a flare or a CME can generate *Moreton waves* (Moreton and Ramsey, 1960) seen in the wings of H_α propagating with a speed of 400-2000 km/s in an arc rarely exceeding 120° in the hot chromosphere

at 10^4 K. Based on their propagation character, Moreton waves are considered fast super-Alfvénic shock waves.

In the corona a sudden energy release can generate so called *X-ray waves* seen by YOKHOH in soft X-ray images propagating in the corona with a speed of 600 ± 100km/s. They are interpreted as fast shock waves. Narukage et al. (2002) studied the Moreton and X-ray waves which appeared on 3 November 1997 and concluded that the two types of wave are very well correlated in time and space, so (i) they must have the same origin and (ii) X-ray waves are the coronal counterpart of the chromospheric Moreton waves.

Probably the most studied global waves in the solar atmosphere are the recently observed *EIT waves* seen by the Extreme-ultraviolet Imaging Telescope (EIT) onboard SOHO (Thompson et al, 1999). These images have provided unambiguous evidence of large-scale coronal impulses initiated during the early stage of a flare and/or CME. Although it seems that all EIT waves are correlated to CMEs, there are some reports stating that in fact EIT waves would be correlated to flares, at least those waves which are very bright. EIT waves are typically of lower amplitudes than Moreton waves and they can extend 360° surrounding the initiating region. EIT waves are seen in a bandpass dominated by the emission lines of Fe XII at 192.3, 193.5 and 195 A (the 195 A bandpass). The emission lines exhibit a peak emission near 1.6 MK at coronal densities. EIT waves propagate in the quiet Sun with a speed of 200-400 km/s at almost constant altitude. Observations show that an EIT wave has two stages: first there is an early (driven) stage where the wave is associated with radio II type burst. In the initial stage the propagating wave could excite plasma radiation by accelerating electrons and creating an energized population that serves as the source of the radio emission. The second stage consists of a freely propagating wavefront. The correlation between EIT and Moreton waves is still subject to a strong debate in the literature. For a very long time, EIT waves have been considered as the coronal counterpart of chromospheric Moreton waves. This was based on studies by Thompson et al. (2000) and Warmuth et al. (2002) where the authors found a very good correlation between Moreton and EIT waves. In contrast, Shibata et al. (2001) and Eto et al. (2002) found that EIT and Moreton waves are not correlated, so the terminology of coronal Moreton waves attributed to EIT waves may seen to be inappropriate. This conclusion is further supported by the differences of their characteristic properties (different generating event, different propagation speed, different propagation character, etc.).

2. Open questions about EIT waves

One of the most basic questions about EIT waves is related to their nature. There is still a strong debate in the literature whether EIT waves are shock or fast magnetoacoustic waves. EIT waves were first seen by SOHO/EIT which has a very limited spatial and temporal resolution (in spite of this, more than 200 events have been studied). Therefore, data provided by this instrument are not enough for a very precise analysis (e.g. wavelet analysis) of the exact nature of EIT waves.

New methods of studying the nature of EIT waves became possible after the launch of the TRACE satellite which has a much better resolution than SOHO. The first EIT event studied, based on TRACE data was carried out first by Wills-Dawey and Thompson (1999). The event appeared at 15:23 UT and it is associated with a GOES C2.9 class flare with a maximum intensity at 15:33 UT that occurred between the eastern part of the main negative polarity and the western part of the main positive polarity regions of AR 8237 at S25 deg, W04 deg where a filament was observed in H_α. The data containing these observations were taken as part of JOP082 designed for flux cancellation in magnetic bipoles. The dataset covers a period of 20 minutes and 13 seconds (15:24:58 - 15:45:11 UT) after which the TRACE satellite entered the South Atlantic Anomaly. In our study we investigate the 195 $\AA$ data. The field of view (FOW) of TRACE images is $511'' \times 511''$ and it points to the centre of the solar disk. Since the studied event seems to be generated by a flare situated below the FOW, TRACE images show the propagating front coming from below and propagating northward. The 195 $\AA$ data are binned 2×2 pixels (1 pixel=0.998″). The TRACE data has been analyzed using a wavelet-transform which, unlike the Fourier transform, returns wavelet scale information as a function of time. The wavelet analysis allows us to investigate the time dependence of periods within the observed data, i.e. the time localisation of periods. A detailed description of the wavelet transform and wavelet analysis is given by Torrence and Campo (1998). Using the 195 $\AA$ TRACE data, we were able to perform a wavelet analysis for the pixel location (100,100) with a superpixel size of 1" × 1", shown in Fig. 1.

This result clearly shows that EIT waves are indeed waves and their period is about 250-255 s.

Another concern related to EIT waves is the propagation height at which these waves propagate. This quantity is essential in other applications as the propagation height will determine many other parameters (density, temperature, etc.). In order to find the role of the propagation height, we take a quiet Sun

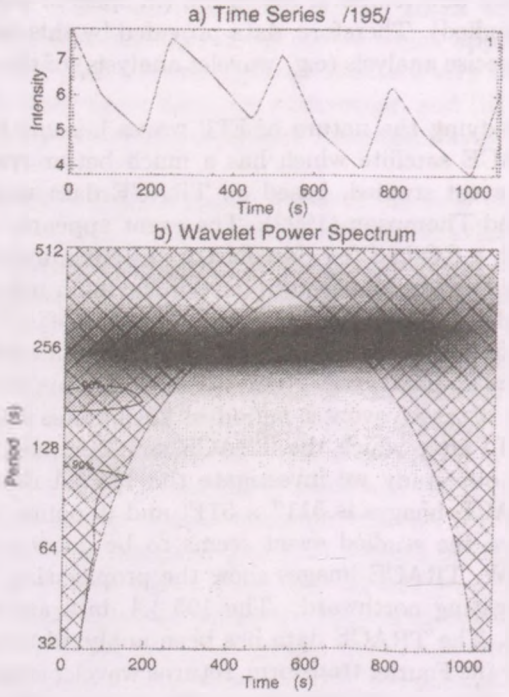

Figure 1: Wavelet power spectrum as a function of time and period. The curved line shows the "cone of influence" (COI), a measure of where edge effects of the finite wavelet transform may already dominate. The area between the COI and the horizontal axis does not contain edge effects caused by the discontinuity at the ends of time series (see Torrence and Compo, 1998).

model and investigate the change in physical quantities for different propagation heights. We consider the atmospheric model proposed by Sturrock et al. (1996). The temperature profile above a region of the quiet Sun where the magnetic field is radial is given by

$$T(x) = \left[T_0^{7/2} + \frac{7R_\odot F_0}{2a}\left(1 - \frac{1}{x}\right)\right]^{2/7}, \tag{1}$$

where F_0 is the inward heat flux (1.8×10^5 erg cm^{-2} s^{-1} in Wheatland et al. (1997)), x is the normalized coordinate defined by $x = r/R_\odot$, T_0 is the temperature at the base of the model (1.3×10^6 K) and a is the coefficient of thermal conductivity (Spitzer, 1962). The quantity a is weakly dependent on pressure and atmospheric composition; for the solar corona a value of 10^{-6} (in cgs units) is appropriate (Nowak & Ulmschneider, 1977). Assuming a model atmosphere in hydrostatic equilibrium we obtain that the number density based on the temperature profile supposed in Eq. (1) is

$$n(x) = \frac{n_0 T_0}{T(x)} \exp[-\delta(T(x)^{5/2} - T_0^{5/2})], \tag{2}$$

where

$$\delta = \frac{2\mu G M_\odot m_p a}{5 k_B R_\odot^2 F_0}, \tag{3}$$

and $\mu = 0.6$ is the mean molecular weight, G is the gravitational constant, $M_\odot$ is the solar mass, m_p is the proton mass and k_B is the Boltzmann constant. Having the variation of the density, we can calculate the magnetic field using $B = v_A(4\pi m_p n)^{1/2}$ where v_A is the Alfvén speed. The variation of temperature, sound speed, density are shown in Figs. 2-4. We can see, the increase in the height does not induce a significant change in the physical parameters, therefore we conclude that the propagation height of EIT waves is not an important quantity.

3. Applications

Since EIT waves are able to propagate over very large distances, they can carry information about the plasma they pass through. The method of extracting this information from waves is called *coronal seismology* (Roberts et al., 1984, Aschwanden et al. 1999, 2002, Nakariakov et al., 1999, Ruderman and Roberts, 2002, Roberts and Nakariakov, 2003, Mendoza-Briceno and Erdélyi,

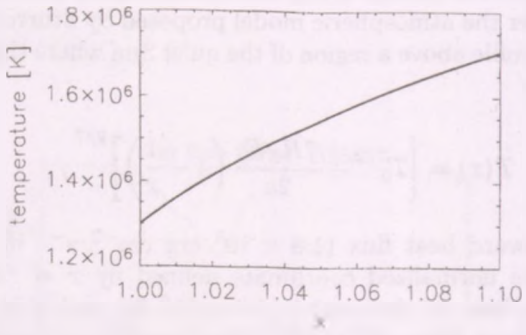

Figure 2: The variation of temperature with height in the quiet Sun based on the quiet Sun model developed by Wheatland et al. (1997). The height is expressed in units of solar radius. The temperature at the base of the corona has been chosen 1.3×10^6 K.

2003, Erdélyi & Mendoza-Briceno, 2004 Mendoza-Briceno et al., 2004) and it became possible after the launch of high resolution space telescopes. Coronal seismology is a perfect tool to obtain physical parameters such as magnetic field and transport coefficients (volume and shear viscosity, resistivity, thermal conduction). Measuring the properties of MHD waves and oscillations (periods, wavelengths, amplitudes, attenuation lengths) combined with theoretical models (dispersion relations, evolutionary equations) we can determine coronal parameters.

Since EIT waves are observed to propagate in every direction almost isotropically, we suppose that they are fast magnetoacoustic waves (FMW) propagating perpendicular to the vertical equilibrium magnetic field. The temperature forming the 195 $\AA$ wavelength is $T = 1.6 \times 10^6$K corresponding to a sound speed c_S $(= (\gamma k_B T/\mu m_p)^{1/2})$ of 190 km/s ($\gamma = 5/3$, ratio of the specific heats; k_B, Boltzmann's constant; $\mu = 0.6$, the mean molecular weight; m_p, proton mass). Since the FMW propagate perpendicular to the field, their phase speed is given by $(c_S^2 + v_A^2)^{1/2}$, where v_A is the Alfvén speed. The value of the magnetic field can be determined using $B = v_A(4\pi m_p n)^{1/2}$, where n denotes the full particle number density, which is related to the electron number density by $n_e = 0.52n$ for $\mu = 0.6$. Using the same quiet Sun model as before, we obtain that the variation of the magnetic field with height for an EIT wave propagating with a speed of 250 km/s is given in Fig. 5. Further analysis shows that the quantity

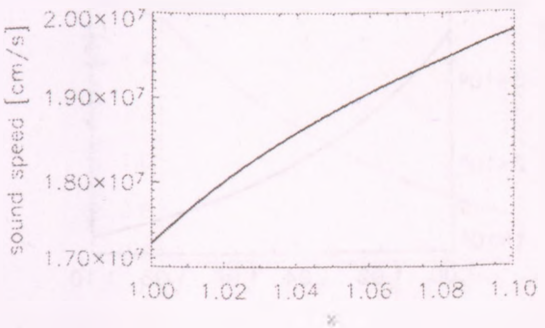

Figure 3: The variation of the sound speed with height in the quiet Sun.

which has the most significant effect on the final results is the propagation speed of EIT waves.

For an average value of EIT waves' speed of 300 km/s we find that the magnetic field is 3 G. From $Br^2 = const.$, i.e. the magnetic flux is constant, we obtain that the magnetic field is 3.5 G in the photosphere and at 1 AU it is 7.5×10^{-5} G, respectively. These values are in very good agreement with the averages of the magnetic field found in the literature .

Another way to exploit EIT waves is to extract information about its wave behaviour and to use this to obtain averaged values for transport coefficients. This task is rather difficult due to the poor temporal resolution of the EIT instrument. In the best case, the waves disappear after 3 or 4 images were taken. Unfortunately, this is far from being enough to carry out wavelet analysis. Luckily, this problem can be overcome by TRACE and the forthcoming space satellites with a much better time resolution.

One piece of information which can be extracted from the propagation of EIT waves is related to their velocity attenuation. For illustration, we consider the event which occurred on 7 April 1997, at the start of solar cycle 23. The EIT wave has been generated by a flare situated in NOAA AR 8027, located slightly southeast of disk center. The peak emission occurred at 14:07 UT and images were taken at 14:12 UT; 14:21 UT; 14:35 UT; 14:53 UT after which the leading front of the wave disappeared. Fig. 6 shows a series of running difference images taken at the above times. These images were obtained by rotating each image back to the time of the pre-event (at 13:28 UT). White regions show an increase

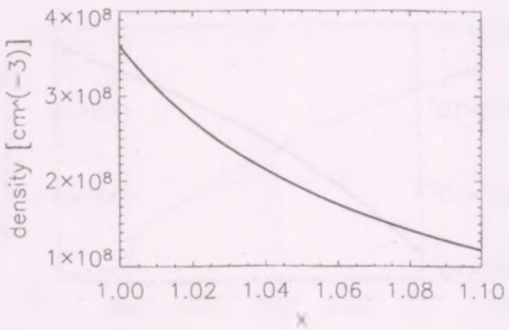

Figure 4: The variation of particle number density with height in the quiet Sun.

in emission while dark regions indicate a decrease.

Measuring the distance travelled by the leading front at the given times, we were able to obtain that the average distance over which the speed of the wavefront decreased e-fold is roughly 365.3 Mm, i.e. $0.52R_{\odot}$. This result, determined from the observations is going to be used in conjunction with theoretical modelling in order to determine plasma parameters.

We suppose the coronal plasma is isotropic, inhomogeneous and viscous. For the sake of simplicity we choose a constant background magnetic field parallel to the z-axis in Cartesian geometry. Although a spherical geometry and a viscosity described by the Braginskii tensor (usually in coronal plasmas the viscosity due to ions is a tensorial quantity, see, e.g. Erdélyi and Goossens 1994ab) would be a better approximation, we suppose a simplified case in order to get a first insight into the problem of EIT wave attenuation. The equilibrium density and pressure decrease with height according to an $exp[-z/H]$ law where H is the pressure scale height ($H = 79$ Mm) in an isothermal plasma. For horizontal propagation perpendicular to the ambient magnetic field, the evolution of the velocity is given by an equation of the form $\sum_{n=1}^{4} a_n(\eta) d^n v/d\eta^n = 0$, where $\eta = \omega^2 H^2/v_A^2 exp[z/H]$. Considering weak dissipation (valid approximation for coronal plasmas), the solution of this ODE with non-constant coefficients can be found in form of Fröbenius series, provided $\eta < 1$, resulting in

$$v \sim e^{-z/l_a} cos(Az), \tag{4}$$

where l_a is the attenuation length and A is a constant expressed in terms of equi-

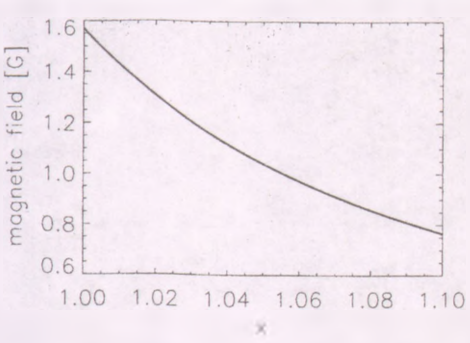

Figure 5: The variation of particle number density with height in the quiet Sun.

librium quantities. The dependence of the damping length on the coefficient of kinematic viscosity is given in Fig. 7 for three propagation heights of EIT waves ($1.05R_{\odot}$, $1.07R_{\odot}$ and $1.09R_{\odot}$) Using the attenuation length obtained earlier, we found that the coefficient of kinematic viscosity derived, based on the attenuation of an EIT wave is two orders of magnitude larger than the value obtained using Spitzer's formula (Spitzer, 1962) for every considered propagation height.

4. Conclusions

The quality of the TRACE data (cadence, resolution) allows us to perform a much better analysis of propagating events in the solar coronal than using SOHO/EIT data. Using a very simple theoretical model, we were able to derive average values for magnetic field in quite Sun and transport coefficients (kinematic viscosity, in particular). Using TRACE data, we have shown that the 'EIT wave' phenomenon is indeed a wave with a precise period and frequency.

So far there is only one TRACE dataset investigated but a systematic analysis of the data provided by this spacecraft could give as a better image on the propagation characteristics of EIT waves. The main conclusion which can be obtained based on the above analysis is that *EIT waves are a suitable tool for global coronal seismology.* For the studied event which occurred on 7 April 1997, we obtained an average value of the magnetic field strength and the coefficient of kinematic viscosity. While the value of the magnetic field strength is in very good agreement with other previous studies, the value of the coefficient

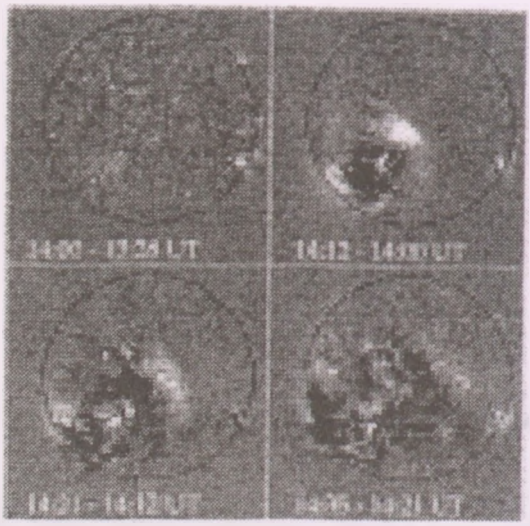

Figure 6: SOHO/EIT 195 Å running difference images showing the propagation of an EIT wave on 7 April 1997. The images were taken at 14:00 UT, 14:12 UT, 14:21 UT, 14:35 UT, with image at 13:28 subtracted from them (courtesy of B. Thompson)

of viscosity is larger than the value obtained from Spitzer's classical formula.

Obviously, the calculations in the present paper are just a first approximation where we used the simplest model possible. For a better understanding of the phenomena of global coronal waves one should include other important effects, such as the effect of curvature, the anisotropic character of the coronal plasma, the effect of geometrical dispersion, waveguide, etc. The results obtained by the EIT camera should be combined with observations by TRACE which has a better temporal resolution. Of course, the field of view of TRACE is limited compared to EIT. These new effects will be addressed in our forthcoming investigations.

Acknowledgments

The author acknowledges the financial support by the NFS Hungary (OTKA, TO43741) and British Council.

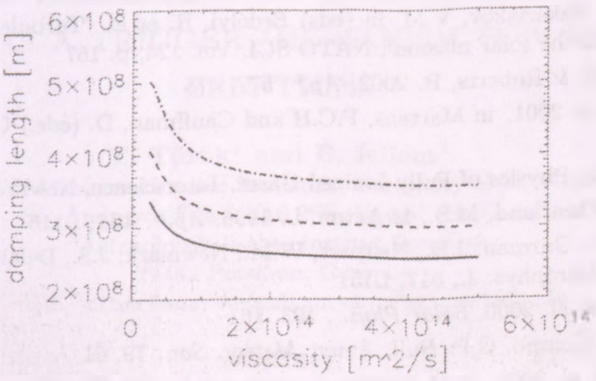

Figure 7: Variation of the damping length with respect to the coefficient of kinematic viscosity for different values propagation height of EIT waves, $1.05R_{\odot}$ (solid line), $1.07R_{\odot}$ (dashed line) and $1.09R_{\odot}$ (dash-dotted line), respectively. The coefficient of viscosity varies between 5×10^{13} and and 5×10^{14} m^2 s^{-1}.

References

Aschwanden et al. 2002, Sol. Phys., 206, 99

Aschwanden et al. 1999, *ApJ* , 520, 880

Erdélyi, R. & Mendoza-Briceno, C. ESA-SP, 547, p.1-6

Erdélyi, R. & Goossens, M. 1994a, in Rusin, V. et al. 'Solar Coronal Structures', IAU Coll. 144, p. 503

Erdélyi, R. & Goossens, M. 1994b, in Rusin, V. et al. 'Solar Coronal Structures', IAU Coll. 144, p. 506

Eto, S. et al. 2002, *PASJ*, 54, 481

Kosovichev, A.G. & Zharkova, V. 1998, *Nature*, 393, 317

Mendoza-Briceno, C., Erdélyi, R. Sigalotti, D.L. 2004, *ApJ* , **605**, DOI: 10.1086/382182

Mendoza-Briceno, C. & Erdélyi, R. 2003, *PADEU*, 13, 171

Moreton, G.F. & Ramsey, H.E. 1960, *PASP*, 72, 357

Nakariakov, V. M. et al. 1999, Science, 285, 862

Narukage, N. et al. 2002, *ApJ* , 572, L109

Nowak, T. & Ulmschneider, P. 1977, *A&A* , **60**, 413

Roberts, B., Edwin, P.M. & Benz, A. O. 1984, *ApJ* , 279, 857

Roberts, B. & Nakariakov, V.M. in (eds) Erdélyi, R. et al. 'Turbulence, Waves and Instabilties in the solar plasma', NATO SCI, Vol. 124, p. 167

Ruderman, M.S. & Roberts, B. 2002, *ApJ* , 577, 475

Shibata, K. et al. 2001, in Martens, P.C.H and Cauffman, D. (eds.) Cospar Series, p. 103

Spitzer, L. 1962, Physics of Fully Ionized Gases, Interscience, New York

Sturrock, P., Wheatland, M.S., & Acton, L. 1996, *ApJ* , **461**, L115

Thompson, B.J., Gurman, J.B., Neupert, W.M., Newmark, J.S., Delaboudinière, J.-P. et al. 1999, Astrophys. J., 517, L151

Thompson, B. et al. 2000, *Solar Phys.* , 193, 161

Torrence, C. & Compo, G.P. *Bull. Amer. Meteor. Soc.*, 79, 61

Warmuth, A. et al. 2002, in H. Sawaya-Lacoste (ed.), ESA SP-477, p. 198

Wheatland, M.S., Sturrock, P.A. & Acton, L.W. 1997, *ApJ* , **482**, 510

Wills-Davey, M.J. and Thompson, B.J. 1999, *Solar Phys.* , 190, 467

THE KINK INSTABILITY OF A CORONAL MAGNETIC LOOP AS A TRIGGER MECHANISM FOR SOLAR ERUPTIONS

T. Török[1] and B. Kliem[2]
[1]School of Mathematics and Statistics, University of St. Andrews
St. Andrews, Fife KY16 9SS, United Kingdom
[2]Astrophysikalisches Institut Potsdam
14482 Potsdam, Germany
E-mail: [1]tibor@mcs.st-and.ac.uk, [2]bkliem@aip.de

Abstract

The kink instability of twisted magnetic flux tubes in the solar corona is regarded as a possible initiation process of solar eruptions. We study the stability properties and the dynamic evolution of such coronal magnetic loops using 3D numerical simulations within the framework of ideal MHD. The analytical force-free coronal loop model by Titov and Démoulin (1999) is used as initial condition in the simulations. The loop model is found to be kink-unstable if a critical twist is exceeded. The growing kink perturbation leads to the formation of current sheets, which steepen exponentially and define the locations of plasma heating. Due to the twist in the magnetic field, the heated structures are S shaped – in very good agreement with soft X-ray observations of solar eruptions. The model, however, does not yet show a successful eruption, rather the kink instability starts to saturate. We present an improvement of the model which is promising with regard to eruption: a modification of the equilibrium so that the magnetic field surrounding the loop decreases more rapidly with height above the photosphere. Furthermore, we briefly discuss how the simulation results can be related to observations of solar eruptive phenomena.

KEYWORDS: *Sun, corona, solar eruptions, magnetic flux tubes, kink instability, MHD*

1. Introduction

Large-scale solar eruptive events like flares, eruptive prominences, and coronal mass ejections (CMEs) are the most violent energy release processes in the solar system. They occur in the coronal layers of the solar atmosphere and

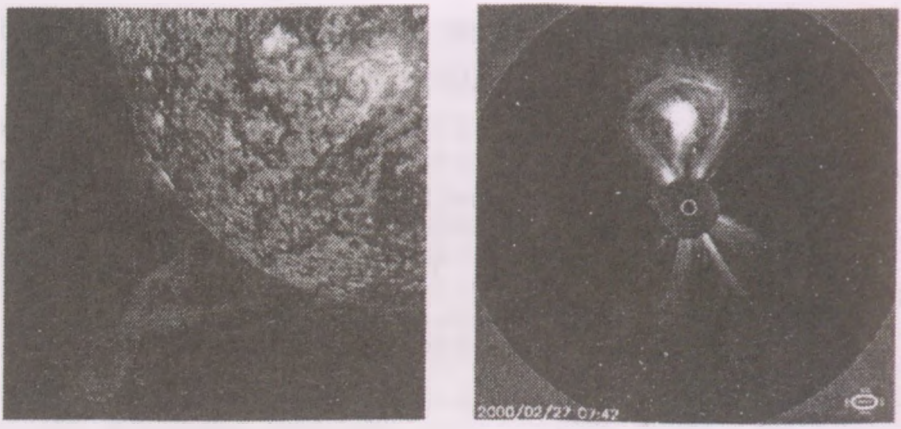

Figure 1: Solar eruptions observed by different instruments onboard the SOHO satellite (http://sohowww.nascom.nasa.gov) during the latest solar maximum. *Left:* eruptive prominence as seen by the EIT telescope in extreme ultraviolet light. *Right:* CME as seen by the LASCO/C3 coronagraph. The solar surface (indicated by the white circle) is covered by an occulting disc in order to allow the observation of regions of enhanced coronal density in the scattered white light.

preferentially in the active regions formed by sunspot groups, i.e., mainly at the maxima of the solar activity cycle. The time-scale of such events is within the range of minutes up to several hours and the length-scale is comparable to the sunspot distance but can finally exceed the size of the Sun (Fig. 1). During the largest solar eruptions, more than ten billion tons of solar plasma are ejected into interplanetary space and particles are accelerated to energies of up to $\sim$ 20 GeV. The interaction of the ejected magnetic plasma clouds with the terrestrial magnetosphere can lead to the malfunctioning of satellites, affect telecommunication and navigation systems, and be hazardous to astronauts. Understanding the physics behind solar eruptions, a prerequisite to forecasting, is therefore of broad significance.

Since the magnetic energy dominates all other forms of energy in the solar corona, there is general consensus that solar eruptions must be primarily driven magnetically (Forbes, 2000). However, their initiation process as well as their dynamic evolution are still not very well understood. The most commonly accepted models of solar eruptions ("storage and release models") assume that the energy released during the eruption is stored in the coronal magnetic field prior to the eruption, and that a loss of stability, or even a "loss of equilibrium"

(a catastrophe) of the magnetic configuration leads to the eruption and energy release.

Solar eruptions often show the phenomenology of a loop-shaped magnetic flux system with fixed footpoints at the coronal base and signatures of magnetic twist (Fig. 1). A single twisted magnetic flux tube appears to contain essential elements of the magnetic topology of the unstable erupting flux. The canonical instability of a twisted magnetic flux tube is the kink instability. A small perturbation of a kink-unstable flux tube leads to an exponentially growing helical distortion ("kink") of the tube. The simplifying assumption of straight, cylindrically symmetric flux tube equilibria has been used in all previous investigations of the kink instability for conditions appropriate to the solar corona (Hood, 1992). The instability occurs if the twist, $\Phi(r) = lB_\phi(r)/(rB_z(r))$, exceeds a critical value, Φ_c. Here l is the length of the cylinder, r is the radial coordinate, and B_z and B_ϕ are the axial and azimuthal field components, respectively. In most configurations of practical interest the twist depends on r and a critical average twist is then determined. The threshold for the onset of the instability depends on the details of the considered equilibrium, in particular on the radial profile, $\Phi(r)$, and on the aspect ratio, l/a, where a is the characteristic radius of the configuration (the minor radius in a toroidal system). For a straight, cylindrically symmetric flux tube with fixed (line-tied) ends and uniform twist (the Gold and Hoyle (1960) force-free equilibrium), which is the simplest possible model of a coronal loop, the threshold was numerically determined (Hood and Priest, 1981) to be $\Phi_c \approx 2.49\pi$.

In this paper we present some of the results of our ongoing studies (Török and Kliem, 2003; Török et al., 2004; Kliem et al., 2004) of the stability properties and the dynamic evolution of arched twisted magnetic flux tubes. Our simulations are directed at the question whether coronal loops become unstable if they are sufficiently twisted and whether such an instability can lead to the eruption of the configuration. We briefly discuss how the simulation results can be related to observations of solar eruptive phenomena.

The analytical coronal magnetic loop model recently developed by Titov and Démoulin (1999) (hereafter T&D) is used as initial condition in the simulations presented here. This approximate force-free equilibrium contains a twisted and curved current-carrying flux tube. The amount of magnetic twist in this model is a parameter that can be freely chosen. For small amounts of twist, we show that the analytical configuration is very close to a stable numerically obtained equilibrium. If a critical twist is exceeded, the loop is always found to be kink-unstable. The kink mode with rising loop apex forms structures that are in excellent agreement with X-ray observations of solar eruptions. On the other

hand, we have not yet been able to find an eruption of the loop in our simulations, rather the very strong surrounding magnetic field in the T&D model eventually slows down the rise of the loop. First tests indicate, however, that a modified T&D model with a more realistic surrounding field can permit a full eruption of the twisted loop.

2. Numerical Model

We integrate the compressible ideal one-fluid MHD equations given by the continuity equation, the momentum equation, the induction equation (which follows from Ohm's law), Ampere's law for the current density, and an energy equation as follows

$$\partial_t \rho = -\nabla \cdot (\rho \mathbf{u}), \tag{1}$$

$$\rho \partial_t \mathbf{u} = -\rho (\mathbf{u} \cdot \nabla) \mathbf{u} + \mathbf{j} \times \mathbf{B} - \nabla p + \nabla \cdot \mathsf{T}, \tag{2}$$

$$\partial_t \mathbf{B} = \nabla \times (\mathbf{u} \times \mathbf{B}), \tag{3}$$

$$\mathbf{j} = \mu_0^{-1} \nabla \times \mathbf{B}, \tag{4}$$

$$\partial_t U = -\nabla \cdot \mathbf{S}. \tag{5}$$

Here $\mathbf{B}$, $\mathbf{u}$, ρ, p, $\mathbf{j}$, U, $\mathbf{S}$, and T are the magnetic field, velocity, mass density, pressure, current density, total energy density, energy flux vector, and viscous stress tensor, respectively. The latter are defined by

$$U = \rho w + \frac{\rho}{2} u^2 + \frac{B^2}{2\mu_0},$$

$$\mathbf{S} = (U + p + \frac{B^2}{2\mu_0}) \mathbf{u} - (\mathbf{u} \cdot \mathbf{B}) \frac{\mathbf{B}}{\mu_0},$$

$$\mathsf{T}_{ij} = \rho \nu \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} - \frac{2}{3} \delta_{ij} \nabla \cdot \mathbf{u} \right),$$

where w is the internal energy per unit mass and ν is the kinematic viscosity. The system is closed by an adiabatic equation of state, $p = (\gamma - 1) \rho w$, which is appropriate for the nearly collisionless corona. The ratio of specific heats is $\gamma = 5/3$. Viscosity is included only to improve the numerical stability and to facilitate relaxation towards equilibrium for the stable configurations.

The system of equations has been simplified for most of the calculations presented here by adopting the zero-beta assumption. The plasma-beta, $\beta = p/(B^2/2\mu_0)$, is of order $10^{-3} \ldots 10^{-2}$ in the inner solar corona where the eruptions are initiated. This eliminates the pressure in Eq. (2) and decouples the

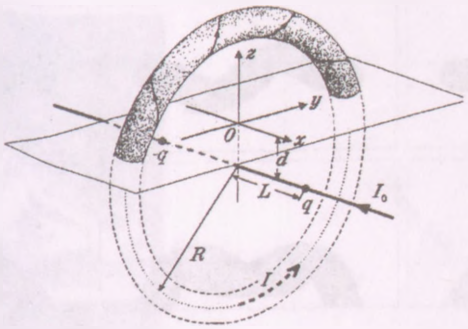

Figure 2: Schematic of the coronal magnetic loop model by Titov and Démoulin (1999). The plane $\{z = 0\}$ corresponds to the solar photosphere. The subphotospheric part has no physical significance.

energy equation from the system. In order to investigate the scaling of kink mode growth rates with the plasma-beta, some runs were perfomed with the full system of equations.

We use a standard normalization of the variables. Lengths, velocities and times are normalized, respectively, by a characteristic length of the initial configuration, Λ (taken here to be L, see Fig. 2), the initial maximum Alfvén velocity, $v_{a0} = B_0^2/(\mu_0/\rho_0)^{1/2}$, and the corresponding Alfvén time $\tau_a = \Lambda/v_{a0}$. Here B_0 is a characteristic field strength, it is chosen to be the total field on the loop axis.

For a detailed description of the numerics used in our simulations see Török and Kliem (2003) and Török et al. (2004).

3. Kink Instability of Magnetic Loops and Sigmoid Formation

In this section we study the magnetic loop model by Titov and Démoulin (1999) (Fig.2). This approximate, cylindrically symmetric, force-free equilibrium consists of a toroidal ring current of major radius R and minor radius a, whose outward-directed Lorentz self-force is balanced with the help of a field by two fictitious magnetic charges of opposite sign which are placed at the symmetry axis of the torus at distances $\pm L$ to the torus plane. That axis lies below the photospheric plane $\{z = 0\}$ at a depth d. The resulting field outside the torus

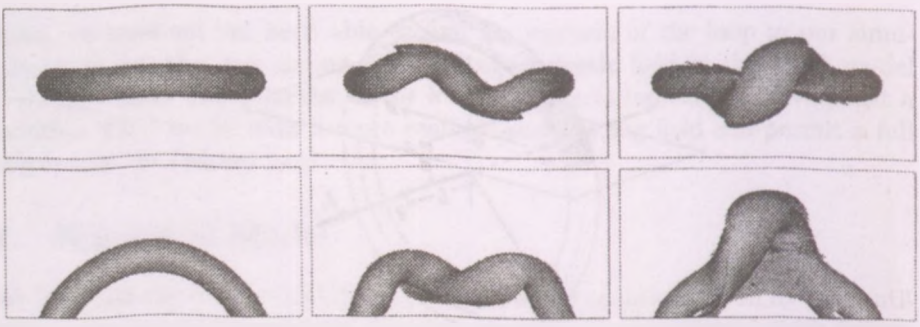

Figure 3: Top and side view on iso-surfaces of the current density at the beginning of the simulation (*left*) and after the development of the kink instability for the downward kinking case (*middle*) and for the upward kinking case (*right*), for a left-handed twist of 4.9π. The iso-surfaces outline the magnetic loop and the newly formed current sheets.

is current-free. A toroidal field component created by a fictitious line current running along the symmetry axis is added. Its strength controls the twist of the field in the torus, $\Phi_{\rm loop}$. The section in the volume $\{z > 0\}$ is a model of a coronal magnetic loop. The parameters of the model are chosen according to observed values of a typical active region on the Sun. The twist is varied in the range $(2\text{–}8)\pi$.

3.1. Parametric Study of the Ideal Kink Instability

For small values of the twist, the T&D loop model is found to be stable; the numerically obtained equilibrium is then always very close to the approximate analytical equilibrium. The kink instability clearly occurs if the critical twist is exceeded, which has a value of $\Phi_{\rm c} \approx 3.5\pi$ for the particular set of parameters considered. The two principal cases of the instability, the downward and the upward kink of the loop apex, are shown in Fig. 3 for $\Phi_{\rm loop} = 4.9\pi$. For the given set of parameters, the downward kink develops spontaneously, while the upward kink is obtained by applying a small initial velocity perturbation at the loop apex. (This is reversed for some of the other parameter combinations studied.) See Török et al. (2004) for a full parametric study of the kink instability growth rates, which also comprises the stabilizing effect of increasing $\beta \neq 0$.

The kink perturbation consists in a helical distortion of the loop. Accordingly, in both cases a helical current sheet wrapped around the loop is formed as the loop pushes into the surrounding field. This is analogous to the kink

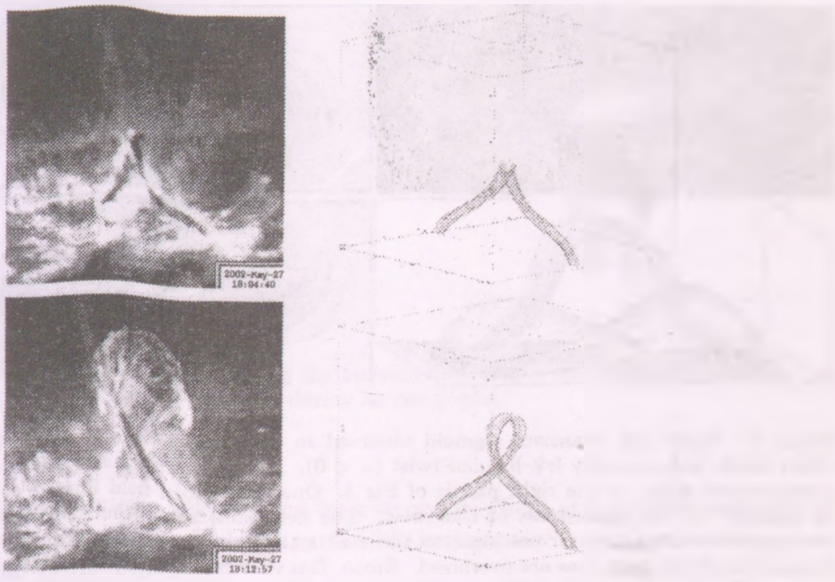

Figure 4: *Left:* two stages of a "failed" prominence eruption as observed by the TRACE satellite (http://vestige.lmsal.com/TRACE) in extreme ultraviolet light (Ji et al., 2003). The lower image shows the prominence after its rise came to a stop. *Right:* field lines of the current density showing the unstable loop in the exponential rise phase (*upper plot*) and in the saturation phase (*lower plot*) of the kink instability.

instability in straight (cylindrical) loop models. The upward kink perturbation additionally leads to the formation of a vertical current sheet below the loop apex which has no counterpart in the cylindrical case. Such a vertical current sheet is the main hypothesis in the so-called standard model of solar eruptions (e.g., Shibata, 1999). In the T&D model, it is a natural consequence of the pinching of the hyperbolic field structure (a generalized X-type field configuration) below the loop, which is trigged by the instability (Török et al., 2004).

In the linear phase of the instability, the apex velocity, the current densities in the formed sheets, and the total kinetic energy in the box grow exponentially. In the nonlinear regime, however, the perturbations are slowed down by the strong magnetic field that surrounds the loop in the T&D model, and the instability starts to saturate without leading to a full eruption. Even this apparently unsatisfactory result is consistent with part of the observations: many

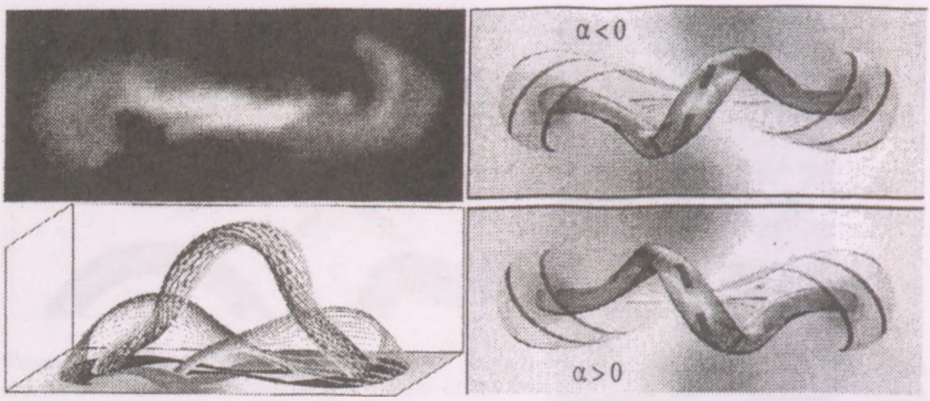

Figure 5: *Upper left*: transient sigmoid observed in an eruptive flare (Manoharan et al., 1996) which had primarily left-handed twist ($\alpha < 0$). *Lower left*: magnetic field lines of the configuration shown in the right panels of Fig. 3. One of the loop field lines is emphasized to indicate the left handedness of the twist. The field lines below the loop pass through two symmetrical vertical stripes close to the z axis that bracket the vertical curent sheet. Projections of the field lines are overlayed. *Right*: Top view on selected field lines and on the surfaces that they form. The field lines for $\alpha < 0$ are identical to those plotted in the lower left panel. The bottom panel shows similarly selected field lines of the system with reversed α at the same time.

flare events and some filament activations start with exponential characteristics but do not lead to escape of plasma from the Sun. Such an example is compared with our simulations in Fig. 4.

3.2. Formation of Transient Sigmoids

Sigmoidal (S or inverse-S shaped) structures, which are suggestive of current-carrying, i.e., helical field lines, are often seen in soft X-ray images of solar active regions. A particularly bright and sharp form occurs transiently at the onset of eruptions. It is of practical relevance as an indicator for geoeffective eruptions. Sigmoids show a strong tendency to obey a "helicity rule": S shapes occur preferably in the southern solar hemisphere, where active regions possess preferably right-handed twist ($\alpha = \mathbf{j} \cdot \mathbf{B}/B^2 > 0$), and the inverse is true for the northern hemisphere, where the twist is preferably left-handed ($\alpha < 0$). There is a long-lasting debate whether sigmoids consist in kink-unstable loops (Rust and Kumar, 1996) or whether they show so-called quasi-separatrix layers, i.e.,

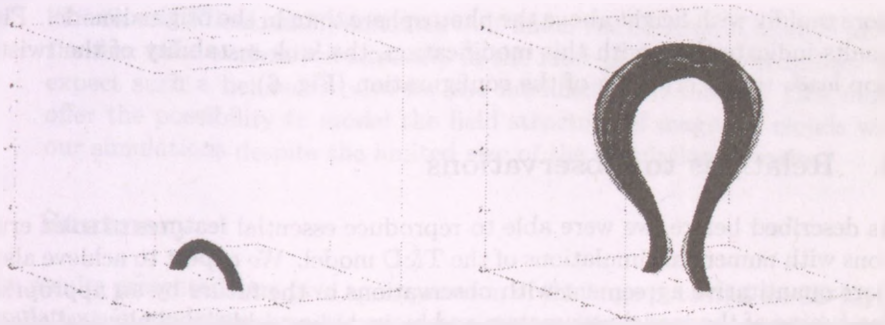

Figure 6: Magnetic field lines showing the unstable loop in the modified T&D model at the beginning of the simulation (*left*) and during its rise (*right*).

layers in the magnetic field defined by a strong gradient of footpoint mapping of penetrating field lines, which are preferable locations of current sheet formation (Titov and Démoulin, 1999).

Our simulations show that upward kinking loops develop a sigmoidal photospheric projection that is opposite to the helicity rule (see Fig. 3 [*right*] which shows the S shape for a left-handed twist). This is a general property of the kink instability, in which the handedness of the perturbed loop axis equals the handedness of the field line twist. Figure 5 shows that the field lines passing through the vertical current sheet below the rising loop agree with the sense of observed sigmoids and reproduce many aspects of their detailed shape as well. This current sheet is formed at the hyperbolic flux tube made of two crossing quasi-separatrix layers (Titov and Démoulin, 1999). See Török et al. (2004) and Kliem et al. (2004) for further details.

3.3. Modification of the T&D equilibrium

The eruption of the kink-unstable loop in the T&D model is prevented by the very strong surrounding magnetic field which is dominated by the contribution of the subphotospheric line current (denoted by I_0 in Fig. 2). Roussev et al. (2003) showed that an eruptive behaviour of the configuration can be obtained if the line current is removed from the system. However, removing the line current leads to an infinite twist at the loop surface.

We have begun to study a modification of the T&D equilibrium in which the twist is finite everywhere in the system and the surrounding field decreases much

more rapidly with height above the photosphere than in the original model. First results indicate that, with this modification, the kink instability of the twisted loop leads to the eruption of the configuration (Fig. 6).

4. Relations to Observations

As described before, we were able to reproduce essential features of solar eruptions with numerical simulations of the T&D model. We expect to achieve also a more quantitative agreement with observations in the future by an appropriate fine-tuning of the model parameters and by including additional physical effects, in particular magnetic reconnection. In the following we briefly discuss how the simulations could be related to observations and how they might help explaining certain aspects of solar eruptions which are difficult to observe directly.

- The determination of twist in emerging, stable or eruptive coronal structures from observations is difficult (see, e.g., Vršnak et al. (1991) for measurements of twist in prominences). The excellent agreement of our simulations with the "failed filament eruption" shown in Fig. 4 indicates that future simulations could become a useful tool to estimate the twist in such structures. This could be used, in turn, to check nonlinear force-free extrapolations of vector magnetograms into the corona.

- An interesting and important question strongly debated in the literature is the rate of acceleration of CMEs. One reason for the existence of different acceleration models (e.g., St. Cyr et al., 1999; Alexander et al., 2002; Gallagher et al., 2003) is the lack of observations of CMEs close to the Sun. First tests of the modified T&D model described in Sect. 3.3. show that the rise characteristic of the eruptive loop is in very good agreement with observed height-time profiles of certain CMEs. Since the acceleration of the eruptive magnetic flux can be measured easily in the simulations, the model might help to bridge the gap left by the observations.

- Magnetic clouds are regarded as the manifestation of CMEs in interplanetary space. They are believed to be twisted magnetic flux tubes, but a specific field model is needed to extract the magnetic field configuration from the measured data (e.g., Dasso et al., 2003). In-situ measurements of magnetic clouds are available only at distances from the Sun which are far beyond the extent of our simulation boxes. However, our study of an eruptive magnetic flux tube which was twisted by photospheric motions

(Török and Kliem, 2003) indicates that under the conditions of ideal MHD the flux tube evolves self-similarly in the later phase of the eruption. We expect such a behaviour also for the modified T&D model. This might offer the possibility to model the field structure of magnetic clouds with our simulations despite the limited size of the simulation domain.

5. Summary

The results presented here give an impression of the huge potential for 3D MHD simulations in modeling dynamic processes in magnetized astrophysical plasmas. The kink instability of twisted flux ropes was identified as a prime candidate for the initiation of solar eruptions. Essential features of solar eruptions were reproduced so that the results could be used for diagnostic purposes in future applications. Additional physical effects, in particular magnetic reconnection, should be included in future simulations.

Acknowledgments

The simulations presented here were performed on the Cray SV1 supercomputer of the John von Neumann-Institut für Computing, Jülich. The investigation was supported by BMBF/DLR grants No. 50 OC 9901 2 and 01 OC 9706 4. The authors also acknowledge the financial support provided through the European Community's Human Potential Programme under contract HPRN-CT-2000-00153.

References

Alexander, D., Metcalf, T. R., and Nitta, N. 2002, Geophys. Res. Lett., 29(10), 1403

Dasso, S. et al. 2003, J. Geophys. Res., 108, 1362

Forbes, T. G. 2000, J. Geophys. Res., 105, 23 153

Gallagher, P. T., Lawrence, G. R., and Dennis, B. R. 2003, Astrophys. J., 588, L53

Gold, T. and Hoyle, F. 1960, Monthly Not. RAS, 120, 89

Hood, A. W. and Priest, E. R. 1981, Geophys. Astrophys. Fluid Dynamics, 17, 297

Hood, A. W. 1992, Plasma Phys.&Contr. Fusion, 34, 411

Ji, H. et al. 2003, Astrophys. J., 595, L135

Kliem, B., Titov, V. S., and Török, T. 2004, Astron.&Astrophys., 413, L23

Manoharan, P. K. et al. 1996, Astrophys. J., 468, L73

Rust, D. M. and Kumar, A. 1996, Astrophys. J., 464, L199

Roussev, I. I. et al. 2003, Astrophys. J., 588, L45

Shibata, K. 1999, Astrophys. Space Sci., 264, 129
St. Cyr, O. C. et al. 1999, J. Geophys. Res., 104, 12 493
Titov, V. S. and Démoulin, P. 1999, Astron.&Astrophys., 351, 707
Török, T. and Kliem, B. 2003, Astron.&Astrophys., 406, 1043
Török, T., Kliem, B., and Titov, V. S. 2004, Astron.&Astrophys., 413, L27
Vršnak, B., Ruždjak, V., and Rompolt, B. 1991, Solar Phys., 136, 151

Structure of the solar tachocline: confrontation of theoretical models with SOHO/MDI helioseismic constraints

E. Forgács-Dajka[1,2]
[1]Eötvös University, Department of Astronomy
H-1518 Budapest, Pf. 32, Hungary
[2]Konkoly Observatory of HAS
H-1525 Budapest, Pf. 67, Hungary
E-mail: E.Forgacs-Dajka@astro.elte.hu

Abstract

Helioseismic measurements indicate that the solar tachocline is very thin, its full thickness not exceeding 4 % of the solar radius. The mechanism that inhibits differential rotation to propagate from the convective zone to deeper into the radiative zone is not known, though several propositions have been made. We present detailed numerical calculations of the fast solar tachocline based on the assumption that the dynamo field dominates over the dynamics of the tachocline.
KEYWORDS: *Sun: interior, MHD, Sun: rotation*

1. Introduction

Helioseismic inversions of the solar internal rotation during the past decade invariably showed that the surface-like latitudinal differential rotation, pervading the convective zone, changes to a near-rigid rotation in the radiative zone. The change takes place in a very thin layer known as the tachocline. The existence and properties of this layer have been known from helioseismic studies, recently reviewed by Corbard et al. (2001). The tachocline is known to be extremely thin; the central radius of the tachocline appears to lie within the range $0.69 - 0.71\,R_\odot$, and its thickness is estimated to be $< 0.05\,R_\odot$ (Kosovichev, 1996; Corbard et al., 1999; Schou et al., 1998; Charbonneau et al., 1999).

Precise values for the parameters of the solar tachocline as such mean position, thickness, latitudinal and temporal variations, depend on the inversion techniques used for helioseismic data. Elliott & Gough (1999) suggest that there exists a difference in the sound speed profile between the observations and the Standard Solar Models. By including an additional mixing layer below the convection zone, they calibrate the thickness of the tachocline and they find that the mean tachocline thickness is $0.019\,R_\odot$ with a formal standard error of about

5%. The observed light-element abundances at the surface also suggest that the mixing process in the solar tachocline is confined to a shallow layer (Brun et al., 1999, 2002).

Making use of the most recent helioseismic results one can investigate the temporal variation and the latitudinal structure of the tachocline. Basu & Antia (2001) found that the tachocline is prolate, with the difference between the tachocline position at 0° and that at 60° latitude being about $0.02\,R_\odot$. This is in agreement with results obtained by Charbonneau et al. (1999). Basu & Antia (2001) also report an increase in thickness of the tachocline with latitude, but this increase is less significant, though still at the 3σ level. Regarding the dynamical aspect, one would expect temporal variations associated with the solar cycle to appear in the tachocline (Howe et al., 2000; Basu & Antia, 2001; Corbard et al., 2001). Howe et al. (2000) report that a large-scale oscillation may be taking place in this layer with a period about 1.3 yr. It is not clear whether this period is associated with solar-cycle-related variations, and Basu & Antia (2001) did not find any periodic or systematic changes in rotation rate in the tachocline region. This fact may be the consequence of the insufficiency of the helioseismic data.

The extreme thinness of the tachocline implies a strongly anisotropic transport of angular momentum (Spiegel & Zahn, 1992). Several different mechanisms have been proposed for this transport, but it is now widely believed that the magnetic field is instrumental in its origin. It is known that an oscillating magnetic field cannot propagate far into the radiative zone. The extent of the penetration of this magnetic field is the skin depth $H_{\rm skin} = \sqrt{2\eta/\omega_{\rm cyc}}$, where η is the magnetic diffusivity and $\omega_{\rm cyc}$ is the frequency of the cycle. On the one hand, for $\eta \lesssim 10^8\,{\rm cm}^2/{\rm s}$ the dynamo field cannot penetrate the tachocline, and we can expect the tachocline to be pervaded by the internal primordial field. On the other hand, for $\eta \gtrsim 10^9\,{\rm cm}^2/{\rm s}$ the tachocline dynamics should be governed by the dynamo field. As the associated diffusive and Lorentz timescales are also very different, these two cases basically correspond to the case of "slow" and "fast" tachocline, discussed in the literature (see esp. Table I in Gilman, 2000).

In recent years the case of a slow tachocline has been investigated extensively by a number of authors. Rüdiger & Kitchatinov (1997) and MacGregor & Charbonneau (1999) studied the interaction between a large-scale field with fixed poloidal component and differential rotation without taking into consideration the meridional circulation. They found that the internal poloidal field of $10^{-3}\,{\rm G}$ is sufficient to confine the tachocline to its observed thickness if the internal field is fully contained within the radiative zone. Gough & McIntyre (1998) presented a model for the solar tachocline which allows the nonlinear

interaction between the meridional flows and a large-scale magnetic field in self-consistent way. Garaud (2001) performed calculations taking into account the meridional flow and the self-consistent evolution of the poloidal field.

In Forgács-Dajka & Petrovay (2002, hereafter Paper I) and Forgács-Dajka (2004, hereafter Paper II), we studied the alternative case of a fast, turbulent tachocline with a turbulent diffusivity of about $\eta = 10^{10}\,\mathrm{cm}^2/\mathrm{s}$ and we presented detailed numerical calculations allowing for the self-consistent evolution of the poloidal field. It was found that a sufficiently strong oscillatory poloidal field with dipolar latitude dependence at the tachocline-convective zone boundary is able to confine the tachocline to its observed thickness. This is in good agreement with our analitical estimate in Forgács-Dajka & Petrovay (2001).

2. The model

2.1. Equations

In order to describe the differential rotation and the evolution of the large-scale magnetic field in the solar interior it is useful to write the equation of motion and the hydromagnetic induction equation in a frame rotating with the fixed internal rotation rate. These equations can be found in Paper I, Sect. 2.1.

2.2. Poloidal Field

It is well known that sunspot activity occurs in the form of waves with periods ~ 11 yr, which propagate from middle latitudes towards the solar equator. In the extended butterfly diagram (Makarov & Sivaraman, 1989), one can see the poleward migration of the magnetic fields indicated by the migration pattern of a number of tracers such as quiescent prominences or the coronal green line.

Stenflo (1988, 1994) analyzed 33 yr of synoptic observations of the Sun's magnetic field carried out daily at the Mount Wilson and Kitt Peak observatories over all solar latitudes and longitudes. The observed longitudinal magnetic field has been converted to a radial magnetic field assuming that the field direction is on average radial in the layers where the field is measured. This provides an opportunity to analyse the global modes of the Sun. Stenflo (1988, 1994) expanded the radial magnetic field in spherical harmonics and he found that the zonal magnetic field pattern can be represented as a superposition of N discrete modes with purely sinusoidal time variations with frequency $\omega_{\mathrm{cyc}} = 2\pi/22\ \mathrm{yr}^{-1}$. Consequently, the evolutionary pattern of the axisymmetric radial magnetic field

can be written in the following form:

$$B_r(t,x) = \sum_{l=1}^{N} |a_l| \cos\left[\omega_{\rm cyc}\,(t - t_l)\right] P_l(x), \tag{1}$$

with

$$\Phi_l = -\omega_{\rm cyc} t_l, \tag{2}$$

where B_r is the radial, axisymmetric magnetic field, $|a_l|$ is the amplitude, Φ_l is the phase lag, t_l is the time lag and $x = \cos\theta$.

Petrovay & Szakály (1999) found that the latitudinal distribution of the field at the surface reflects the conditions at the bottom of the convective zone, i.e. in this regard the convective zone behaves as a "steamy window". Thus, if we assume that the above formula describes the time-evolution of the radial magnetic field at the base of the convective zone, then the development of the vector potential is given by:

$$A(r,\theta,t) = \frac{\int B_r(r,\theta,t)\sin\theta r {\rm d}\theta + F(r,t)}{\sin\theta}, \tag{3}$$

where F is an arbitrary function. Stenflo (1988, 1994) also studied the dominance of the odd and even modes and he showed that the odd modes prevail in the evolution of the radial, axisymmetric magnetic field. Thus, the development of the vector potential is represented as a superposition of the sinusoidal, 22 yr variations for the 7 odd modes with $l = 1, 3, \ldots, 13$.

2.3. Boundary Conditions for the Physical Parameters

Our model does not include the convection zone, so the computational domain for the present calculations consists of just an annular meridional cut in the northern hemisphere ($0 \le \theta \le \pi/2$) in the upper part of the radiative interior, between radii $r_{\rm in}$ and $r_{\rm bcz}$, where $r_{\rm bcz}$ is the radius of the bottom of the convection zone. For the solution of our equations we use the boundary conditions,, which can be found in Paper II, Sect. 2.4.

3. Results

As stated above, we examine the influence of a migrating field on the radial spreading of the differential rotation into the radiative interior. Prescribing a

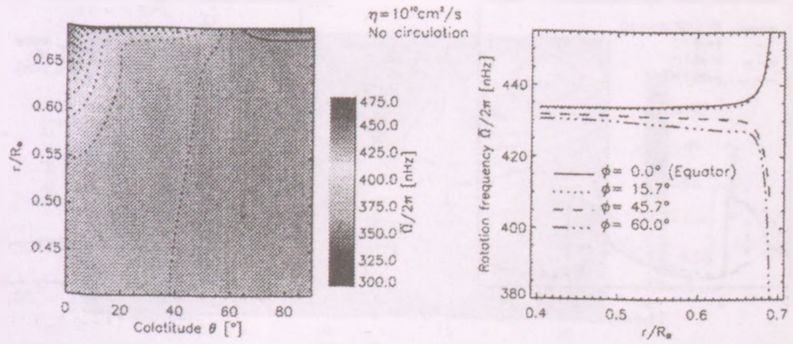

Figure 1: Spreading of the differential rotation into the radiative interior for $\eta = \nu = 10^{10}\mathrm{cm}^2/\mathrm{s}$. *Left-hand panel:* contours of the time-average of the rotation rate under one dynamo period, with contour spacing 8.75 nHz; the solid lines represent the regions rotating faster and the dotted lines show the regions slower than $\Omega_0/2\pi = 437\mathrm{nHz}$. *Right-hand panel:* the rotation rate at different latitudes as a function of radius. The peak amplitude of the poloidal magnetic field and the toroidal field are $B_\mathrm{p} \sim 2400$ G and $B \sim 12000$ G, respectively.

migrating magnetic field at the top we define the form of the vector potential from the observed radial magnetic field.

The values of the viscosity and magnetic diffusivity chosen for these simulations are identical, $\eta = \nu = 10^{10}\mathrm{cm}^2/\mathrm{s}$, i.e. the magnetic Prandtl number is one $\mathrm{Pr_m} = \nu/\eta = 1$ and the meridional circulation is neglected. Several calculations were run with different field strengths and we chose the most suitable case for us, where the sufficiently strong magnetic field is able to interfere the spreading of the differential rotation into the radiative interior and to reproduce the observed thickness of the tachocline. This case is shown in Fig. 1 after relaxation.

In the left-hand panel we plotted the contours of the time-average of the rotation rate, which is defined as

$$\overline{\Omega}(r,\theta) = \overline{\omega}(r,\theta) + \Omega_0 \tag{4}$$

$$\overline{\omega}(r,\theta) = \frac{1}{P_\mathrm{cyc}} \int_t^{t+P_\mathrm{cyc}} \omega(r,\theta,t)\,\mathrm{d}t, \tag{5}$$

where $P_\mathrm{cyc} = 22\mathrm{yr}$ is the period of the cycle.

In the right-hand panel we plotted the rotation rate at different latitudes as a function of radius. We note that in Paper I we plotted the differential rotation amplitude $\Delta\omega$ in right-hand panels, where ω was weighted with latitude and $\Delta\omega$ was normalized (see Eq. (23) in Paper I). However, in this paper we retained

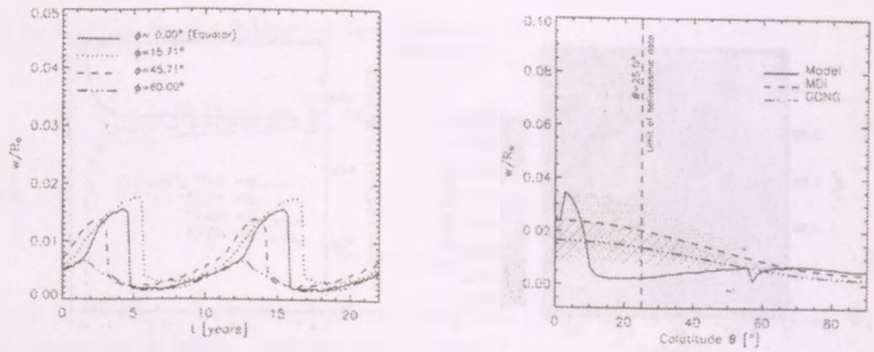

Figure 2: *Left-hand panel:* The thickness of the tachocline at different latitudes as a function of time for the case in Fig. 1. *Right-hand panel:* Latitudinal variation of the tachocline thickness for the case in Fig. 1. The solid line represents the thickness from the model and the dashed and dashed-dot lines show w from MDI and GONG data respectively (Basu & Antia, 2001). The grey and the hatched areas show the error estimates on these.

the values of the frequency, so in this case the results can be compared easier with helioseismic observations and even the more insignificant variations of the rotation rate are more conspicuous. Based on the helioseismic observations we expect that the differential rotation changes to the rigid rotation in a thin transition region. It is well visible that the prescribed differential rotation at the base of the convection zone changes to the near-uniform rotation in a thin layer.

We also study in detail the dynamics of the tachocline and compare to the observations. Accordingly, the thickness of the tachocline and the change in the rotation rate across the tachocline are defined as in Basu (1997). Thus, the jump in the rotation rate across the tachocline $\delta\Omega$ is the difference between the rotation rate at the base of the convection zone and the rotation rate in the interior. The thickness of the tachocline w is defined as the rotation rate increases from the factor $1/(1+\epsilon)$ of its maximum value to the factor $1-1/(1+\epsilon)$ of its maximum value in the range $r = r_t - w$ to $r = r_t + w$.

The fast tachocline shows a quite marked latitude-dependence in our results, as illustrated in Fig. 2 on the right-hand side. In both figures we also plot the results of the helioseismic measurements from MDI and GONG (Basu & Antia, 2001). It should be noted that, according to the helioseismic results, the latitudinal variation in position of tachocline is significant, while the variation in width is not clear, but the results tend to suggest that the thickness increases

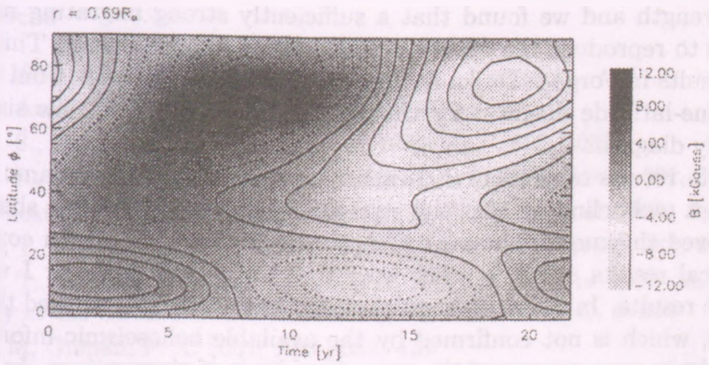

Figure 3: Time-latitude diagram for the toroidal field (*shaded region*) and the radial field (*contours*) at the base of the convection zone. Equidistant contour levels of the radial field are separated by intervals of 30 G.

with latitude. Since the top of the computational domain is fixed at $0.69R_\odot$ in our model, therefore the position of the tachocline cannot be directly compared with the observations. Notice that the thickness of the tachocline is close to zero at around the colatitude of 60°, because at this colatitude the rotation rate is equal to the rigid rotation velocity. This region is the so-called corotation zone (Fig. ??). In addition to latitudinal variations, we also investigate the temporal variations in the properties of the tachocline, and we find that the thickness of the tachocline depends on cycle phase (Fig. 2 on the left-hand side). This variation is not confirmed by the observations, but the helioseismic data set is relatively short. The development of the toroidal magnetic field and the radial magnetic field shown in Fig. 3 correspond to the observed evolutionary pattern of the magnetic field, i.e. the butterfly diagram.

4. Discussion and Conclusion

The models discussed in this paper show that the dynamo field dominates over the dynamics of the fast solar tachocline provided that the turbulent diffusivity in the tachocline region exceeds $\eta \gtrsim 10^9\,\mathrm{cm}^2/\mathrm{s}$.

For the more general magnetic field structures, we use the results of Stenflo (1988, 1994), where the evolutionary pattern of the axisymmetric magnetic field can be written as a superposition of the sinusoidal, 22 yr variations for the 7 odd modes spherical harmonic components. Based on our previous works we studied

the field strength and we found that a sufficiently strong migrating magnetic field is able to reproduce the observed thickness of the tachocline. This agrees with the results in Forgács-Dajka & Petrovay (2001, 2002), apart from the fact that the time-latitude diagram for the toroidal field (Fig. 3) is now similar to the butterfly diagram.

The main results of present and earlier models is that the dynamo field in the turbulent tachocline by Maxwell stresses is capable to limit the shear layer to its observed thickness. The next step of this analysis consists in comparing the numerical results in the present work and the model in Paper I with the helioseismic results. In Paper I the thickness of the tachocline showed the "pit" at the pole, which is not confirmed by the available helioseismic information, which may be in consequence of the latitudinal limit of observations. In Paper I the poloidal field was a simple oscillating dipole, and the reality of this pit needs to be verified with other poloidal field geometries and other parameters. In this paper we found that in case of the more general magnetic field structure the polar pit is reduced to its half. If we use the more general field geometry reminiscent of the butterfly diagram, the penetration depth of the differential rotation into the radiative interior in the polar region is decreased owing to the fact that at higher latitudes the stonger magnetic fields stay up.

Other important points what we should consider in the future are the boundary conditions on the magnetic field and the complete dynamo-generated poloidal field, because the conditions in the present treatment poorly represent the dynamical interaction between the flow and the magnetic field in the tachocline and the convection zone (Petrovay, 2000). Many theoretical models suggest that the toroidal magnetic field is up to 10^5 G in the tachocline, required to produce sunspots in low latitudes (Petrovay & Moreno-Insertis, 1997). In our model the peak toroidal field is above about 10^4 G, but this may change if different boundary conditions are used or the dynamo-generated field is included. Dikpati & Gilman (2001); Gilman & Dikpati (2002) suggest that the dynamo-generated toroidal field should play an important role in the global instability of tachocline differential rotation. In future work we plan to develop our model along the lines described above.

Acknowledgments

I would like to thank J. O. Stenflo for providing the parameters of the evolutionary pattern of the axisymmetric radial magnetic field. We would like to thank K. Petrovay for fruitful discussions on the manuscript. This work was funded by the OTKA under grant no. T043741.

References

Basu, S. 1997, *MNRAS* , 288, 572

Basu, S., Antia, H. M. 2001, *MNRAS* , 324, 498

Brun, A. S., Turck-Chièze, S., Zahn, J.-P. 1999, *ApJ* , 525, 1032

Brun, A. S., Antia, H. M., Chitre, S. M., Zahn, J.-P. 2002, *A&A* , 391, 725

Charbonneau, P., Christensen-Dalsgaard, J., Henning, R., et al. 1999, *ApJ* , 527, 445

Corbard, T., Blanc-Fraud, L., Berthomieu, G., et al. 1999, *A&A* , 344, 696

Corbard, T., Jiménez-Reyes, S. J., Tomczyk, S., et al. 2001, in Helio- and Asteroseismology at the Dawn of the Millennium (ESA Publ. SP-464), p. 265

Dikpati, M, Gilman, P. A. 2001, *ApJ* , 559, 428

Elliott J. R., Gough D. O. 1999, *ApJ* , 516, 475

Forgács-Dajka, E., Petrovay, K. 2001, *Solar Phys.* , 203, 195

Forgács-Dajka, E., Petrovay, K. 2002, *A&A* , 389, 629 (Paper I)

Forgács-Dajka, E. 2004, *A&A* , 413, 1143 (Paper II)

Garaud, P. 2001, *MNRAS* , 329, 1

Gilman, P. A. 2000, *Solar Phys.* , 192, 27

Gilman, P. A., Dikpati, M. 2002, *ApJ* , 576, 1031

Gough, D. O., McIntyre, M. E. 1998, Nature, 394, 755

Howe, R., Christensen-Dalsgaard, J., Hill, F., et al. 2000, Science, 287, 2456

Kosovichev, A.G. 1996, *ApJ* , 469, L61

MacGregor, K. B., Charbonneau, P. 1999, *ApJ* , 519, 911

Makarov, V. I., Sivaraman, K. R. 1989, *Solar Phys.* , 123, 367

Petrovay, K. and Moreno-Insertis, F. 1997, *ApJ* , 485, 398

Petrovay, K. and Szakály, G. 1999, *Solar Phys.* , 185, 1

Petrovay, K. 2000, in The Solar Cycle and Terrestrial Climate, ESA Publ. SP-463, p. 3; also astro-ph/0010096

Rüdiger, G., Kitchatinov, L. L. 1997, *Astr. Nachr.* , 318, 273

Schou, J., Antia, H. M., Basu, S., et al. 1998, *ApJ* , 505, 390

Spiegel, E. A., Zahn J.-P. 1992, *A&A* 265, 106

Stenflo, J. O. 1988, *Ap&SS* 144, 321

Stenflo, J. O. 1994, in R. J. Rutten and C. J. Schrijver (eds.), Solar Surface Magnetism, p. 365

Fine structure of the butterfly diagram

B. Major
Eötvös University, Department of Astronomy
H-1518 Budapest, P.O.Box 32., Hungary
E-mail: B.Major@astro.elte.hu

Abstract

Fine-structure of the butterfly diagram has been studied using the Greenwich Photoheliographic Results (GPR) between the years of 1874-1976. Two main activity zones have been found situated between $5° - 17°$ and $17° - 35°$, respectively. KEYWORDS: *Sun, sunspots*

1. Introduction

One of the most important observational results in solar research - especially in solar dynamo theory - is the time-latitude distribution of sunspot groups on the surface of our Sun which was first drawn by Maunder and is known as butterfly diagram (Fig. 1). It is apparent sunspot groups appear mainly in a particular belt between the heliographic latitude of 5° and 35°, out of these belts there are much less groups and above 40° there is almost no appearance. However, the appearance changes during a solar cycle that lasts about 11 years on average. Some essential properties of this change are shown in Fig. 2 and Fig. 3. First of all, the activity migrates equatorward during a whole cycle (Spörer's law). Another aspect is that the extension of the activity on the solar surface is growing to one third of the cycle and after that this extension is shrinking until the cycle finishes. This asymmetry can also be seen in sunspot numbers (e.g. Wolf Sunspot Number or Group Sunspot Number).

Though we usually mention 11 years in connection with solar cycle, we should note the following. If one considers a cycle that begins when groups and spots start to appear at high latitudes and ends when no more features are visible at low latitudes, one will obtain about 15-17 years for the duration of the cycle. Certainly this is because two successive cycles overlap each other (Fig. 4, for details see Forgacs-Dajka et al. (2004)). Hence, the average will be less than this value, approximately 11 years (see e.g. Usoskin & Mursula (2003)). From theoretical point of view it can be interpreted as the toroidal magnetic field originated from the previous cycle is still there when the new cycle begins.

When plotting the standard butterfly diagram we use only the latitude and time data and do not consider the area of sunspot groups. In the following

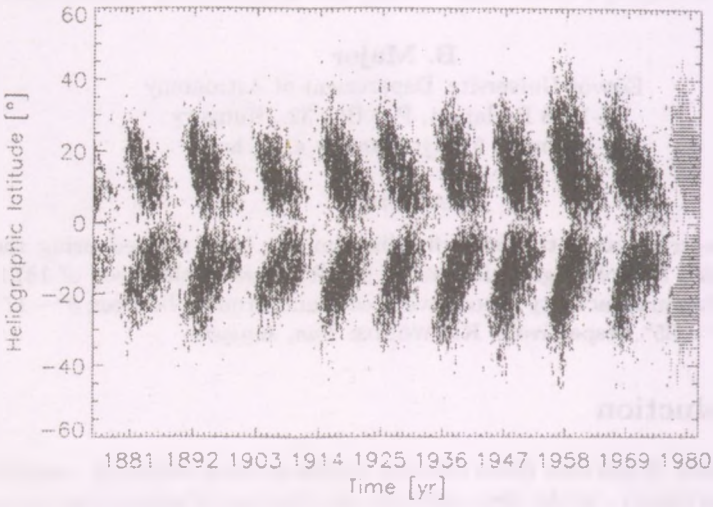

Figure 1: The butterfly diagram for sunspot groups from the GPR data.

we use a simple method to show what can be obtained if the examination is extended to these data, too. A similar investigation was performed by Antalová & Gnevysev (1985), who found that solar activity in certain latitudinal intervals is higher than in other heliographic latitudes. The width of these intervals is some tens of degrees in latitude. Our goal is to confirm these results, and make more quantitative statements concerning their significance and the position of the peaks.

2. Data and method

From 1874 to 1981 the visible solar hemisphere was captured every day mainly in the Greenwich Royal Observatory (there were two similar observational setups in Kodaikanal and Cape Town, in order to get pictures even if it was cloudy in Greenwich). They produced a huge database called GPR (Greenwich Photoheliographic Results) by processing these images. There are exactly 180022 measurements for 30401 groups transiting on the surface. However, in practice we have analyzed data till the year 1976, because after that time the value of

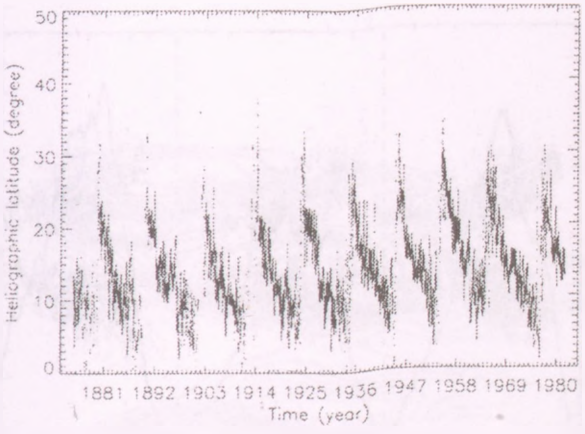

Figure 2: Equatorward migration of activity drawn by area-weighted latitudes of sunspot groups smoothing with a one-month sliding average.

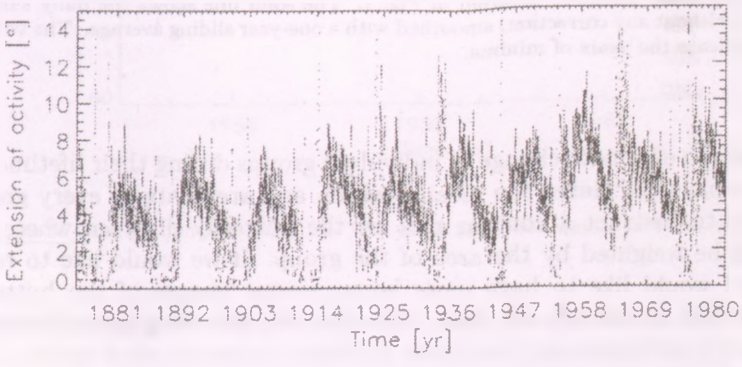

Figure 3: Variation of the mean extension of activity drawn by area-weighted latitudes of sunspot groups smoothing with a one-month sliding average.

heliographic latitude are given to a precision one order of magnitude less than before. It means 161714 measurements of 29850 individual groups. GPR consists data of sunspot groups, there are their heliographic positions, areas and time of observations, as well as some other data which are not important in this

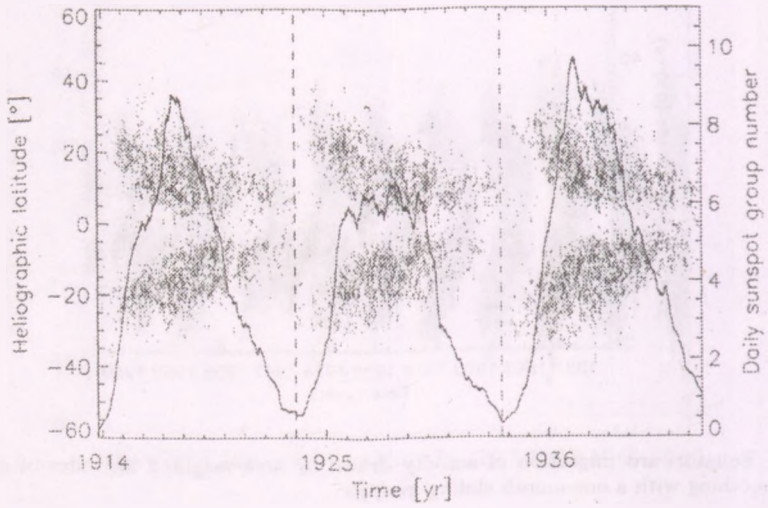

Figure 4: A part of the butterfly diagram in Fig. 1. The solid line shows the daily sunspot group number (without any correction) smoothed with a one-year sliding average. The vertical dashed lines indicate the years of minima.

study.

First of all, to eliminate change of individual groups during their lifetime (or visibility) we have calculated the mean position and area data of every group. Our aim was to costruct a contour plot on the butterfly diagram where the points would be weighted by the area of the group. If we would like to reach this goal and would like to have more homogeneous images of the butterfly diagram, we have to smooth our data. Therefore the following spread function was used:

$$f(t,B) = \sum_i \left[A(t_i, B_i) \cdot \exp\left[-\left(\frac{(t-t_i)^2}{2\sigma_t^2} + \frac{(B-B_i)^2}{2\sigma_B^2} \right) \right] \right],$$

where f means the smoothed value, A is the area of the ith group, t and B refer to time and latitude, respectively. σ_t and σ_B are free parameters. By changing their value we can vary the width of the spread function in time and in latitude. In the present study the value of σ_t was 1° and of σ_B was half a year. In principle, the summation should go over all groups but in practice only

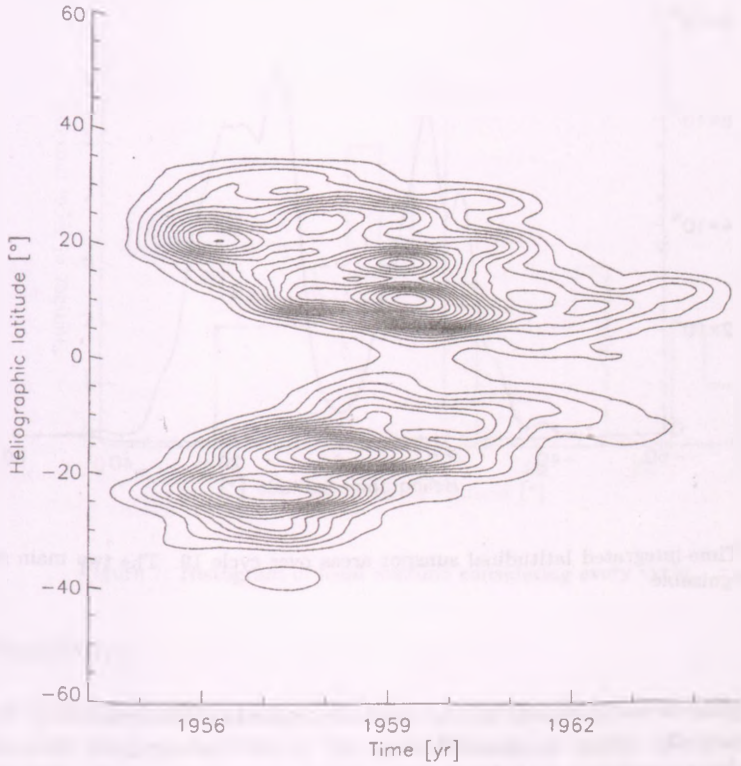

Figure 5: Contour plot for cycle 19. Two main zones are visible on both hemispheres

extend it to groups for which $|t - t_i| < 2\sigma_t$ and $|B - B_i| < 2\sigma_B$. Finally we have $f(t, B)$ which is the smoothed and area-weighted representation of the butterfly diagram. Obviously, structures of a scale smaller than σ_B and σ_t cannot be detected by this method.

After this process we follow a simple contour process to show what sort of structures we may recognise (the idea of contour method is from Becker (1959)). Actually, we try to draw the main activity cores or zones considering that the cadence of activity depends on the area of an active region.

Results of the above procedure for one part of the butterfly diagram, cycle 19 are shown in Fig. 5 These figures indicate that there are two main latitudes

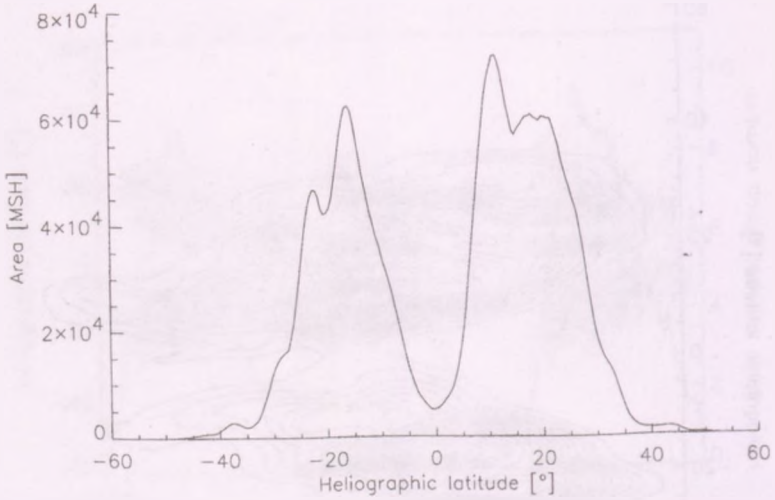

Figure 6: Time-integrated latitudinal sunspot areas over cycle 19. The two main zones are clearly recognizable

in the region of sunspot activity in each hemisphere. The separation between the two activity zones is situated near $15° - 20°$ heliographic latitude. To validate these suspicion we take the time-integrated (and smoothed) latitudinal sunspot areas over every cycle. In Fig. 6 the values obtained for cycle 19 are plotted. Similar features can be found in other cycles. Next, the locations of local maximum and minimum values are studied.

First we consider the histogram of absolute values of local maxima in the area-latitude plots (like Fig. 6), drawn for each cycle. In Fig. 7 two main groups of local maxima are shown. For a crude estimate of the significance if the bimodal character of the distribution we assume that the histogram results from a probability density function with an approximately constant value in the interval $10° < |B| < 24°$. Assuming a Poisson distribution of the actual values around of this probability function, we find that the probability of having zero maxima in the bin around 16° is $P = 0.038$. Thus, the bimodality of the histogram is significant at the 96% level. The mean latitude of the first group is 12.3° and of the second one is 21.2°.

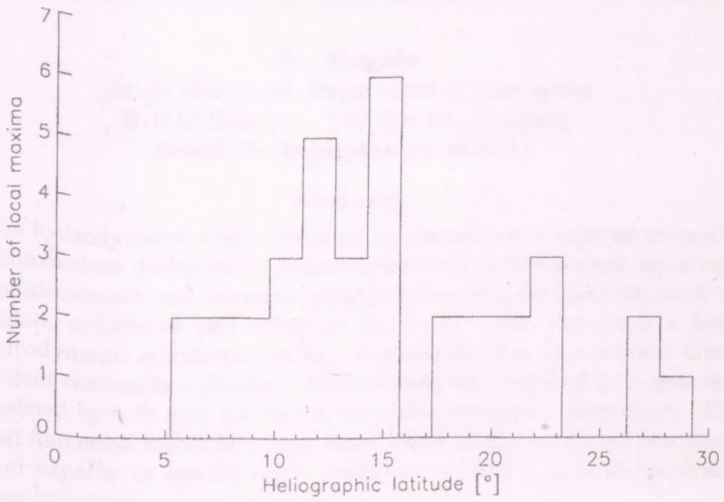

Figure 7: Histogram of local maxima considering every cycle

3. Summary

We have found indication that, typically, a solar cycle has two main zones of activity. One of them is situated about 12.3° and the other one about 21.2°. These values are averaged, so in certain cycles these zones are in higher or lower latitudes, but they are clearly recognizable. This result confirms and make more quantitative previous similar results by Antalová & Gnevysev (1985).

Acknowledgments

The author thank Kristóf Petrovay for useful discussions and proposals.

References

Antalová, A., Gnevysev, M. N. 1985, Contr. of the Astr. Obs. Skalnate Pleso, 11, 63
Becker, U. 1959, Zeitschrift für Astrophysik, 48, 88
Forgacs-Dajka, E., Major, B., Borkovits, T. 2004, *A&A* , in press
Usoskin, I. G., Mursula, K. 2003, Solar Phys., 218, 319

General Purpose Hydrodynamic Code

Zs. Regály
Eötvös University, Department of Astronomy
H-1518 Budapest, P.O.Box 32., Hungary
E-mail: Zs.Regaly@astro.elte.hu

Abstract

The hydrodynamic code developed by the author is capable to model fluid convections described by conservative form hydrodynamic equations in multidimension and versatile orthogonal coordinate systems, such as Cartesian, cylindrical and spherical geometry. The conservative form of hydrodynamic equations can be rewritten in form of nonlinear time-dependent continuity equations. Multidimensional physical problems can be modeled by split step method in versatile orthogonal geometries. The applied numerical algorithms were three types of flux-corrected transport method capable to handle steep gradients of fluid properties generally appears in astrophysical phenomena.

1. Introduction

A general purpose hydrodynamic code developed by author to solve hydrodynamic (HD) and magneto hydrodynamic equations numerically. The code easily configured to solve versatile HD and MHD problems. The base idea was the generalization of differential equations written in conservative form (Borris et.al, 1996).

The general hydrodynamic code (GHYCO) was designed to solve hydrodynamic equations in multidimension and in versatile orthogonal geometries. The GHYCO was developed in object oriented C++ programming language, using free ware numerical package Blitz++ Numerical Library. The C++ programming language offers many features useful for generalization of complex scientific computing problems. Unfortunately, these advanced features came with a hefty performance pricetag. The goal of the Blitz++ project was to develop techniques which will enable C++ to rival the speed of Fortran for numerical computing, while preserving an object-oriented interface. The Blitz++ Numerical Library is being constructed as a testbed for these techniques (Blitz++ NL, 2001).

In this paper I present a description of a well known numerical method the flux-corrected transport method, and some test run results. In Sec. 2, one can read how the hydrodynamic differential equations written in conservative form

can be transform into the generalized continuity equation. In Sec. 3 multi-dimensional generalization, the so called split step method, will be described. In Sec. 4 the short description of the flux-correction process presented. Some numerical tests and results on the success of shock capturing placed in Sec. 5. In Sec. 6 some concluding remarks on developed code will be described.

2. Generalized Continuity Equation

The well known hydrodynamic equations of inviscid fluid flow in conservative forms, deduced from conservations of mass, impulse and energy, are the followings:

$$\partial_t \rho = -\partial_i \rho v_i, \tag{1}$$

$$\partial_t \rho v_i = -\partial_j \rho v_j v_i - \partial_i P + \rho \partial_i \Phi, \tag{2}$$

$$\partial_t \rho e = -\partial_i \rho e v_i - P \partial_i v_i, \tag{3}$$

where ρ is the density, v_i is the i-th velocity vector component, and e is the internal energy density of the fluid. In the Euler equation (2) the P is the pressure of fluid, which is equivalent to the equation of state of fluid, could be a function of density, temperature etc. (in this case $P = P(\rho, T)$), while Φ denotes the gravitational potential. In a more complex problem additional source terms could be appeared in the Euler equation (2), like stress tensor derivatives ($\partial_i \tau_{i,j}$) or heat production ($\rho \dot{q}$) and heat transfer ($\partial_i k \partial_i T$). The same as well in the energy conservation equation (3), see (Wendt, 2000).

Analyzing the system of hydrodynamic equations (Eqs. 1-3), one can release that all equations can be written in the same form (Borris et.al, 1996) called generalized 1D continuity equation, which is the following

$$\partial_t A = -\partial_i A v_i - \partial_i S_1 + S_2 \partial_i S_3 + S_4, \tag{4}$$

where A can be interpreted as the density, impulse or internal energy density. The terms denoted by S_i are called as source terms: ∂S_1 is the divergence of a vector, $S_2 \partial S_3$ is the gradient of a scalar, and the last term, S_4 could be any function of known fluid variables. So, solving Eq (1-3) in a versatile (but orthogonal, seel latter) geometry is equivalent to solving coupled generalized continuity equations in form of Eq. (4). With this generalization more complex problems can be handled. In the description of an non-inviscid, or elastic flow the Eq. (2) must be replaced by the Navier–Stokes equation, see (Wendt, 2000). The Navier–Stokes equation involves velocity dependent source terms, which also can be rewritten into the Eq. (4) form.

While hydrodynamic equations Eqs. (1-3) can be expressed by same form Eq. (4), we need a numerical algorithm capable to solve this generalized 1D continuity equation.

3. Split Step Method

In many physical problems, it is practical to write fluid dynamics equations in other than Cartesian geometry. For example, a spherically symmetric body, such as a star, can be described in spherical geometry easier, than in Cartesian geometry. If the coordinate system is orthogonal, the 2D or 3D systems of equations can be expressed by decoupled equations, Borris et.al (1996). The hydrodynamic equations (1-3) in 2D Cartesian geometry are

$$\partial_t \rho = -\partial_x \rho v_x - \partial_y \rho v_y, \tag{5}$$

$$\partial_t \rho v_x = -\partial_x \rho v_x v_x - \partial_y \rho v_x v_y - \partial_x P + \rho \partial_x \Phi, \tag{6}$$

$$\partial_t \rho v_y = -\partial_y \rho v_y v_y - \partial_x \rho v_y v_x - \partial_y P + \rho \partial_y \Phi, \tag{7}$$

$$\partial_t \rho e = -\partial_x \rho e v_x - \partial_y \rho e v_y - P \partial_x v_x - P \partial_y v_y. \tag{8}$$

Here we applied that the mixed derivative (like $\partial_x A_y$, where A denotes ρv or $\rho v v$ or $\rho e v$) disappears in an orthogonal coordinate system. All of this equations can be split to two orthogonal direction flow equations. The change in density, impulses and energy can be split into the change due to x and y direction flows. So, we can rewrite the Eqs. (5-8). The change in fluid variables into direction x flow are

$$\partial_t \rho = -\partial_x \rho v_x, \tag{9}$$

$$\partial_t \rho v_x = -\partial_x \rho v_x v_x - \partial_x P + \rho \partial_x \Phi, \tag{10}$$

$$\partial_t \rho v_y = -\partial_x \rho v_y v_x, \tag{11}$$

$$\partial_t \rho e = -\partial_x \rho e v_x - P \partial_x v_x, \tag{12}$$

while into direction y

$$\partial_t \rho = -\partial_y \rho v_y, \tag{13}$$

$$\partial_t \rho v_x = -\partial_y \rho v_x v_y, \tag{14}$$

$$\partial_t \rho v_y = -\partial_y \rho v_y v_y - \partial_y P + \rho \partial_y \Phi, \tag{15}$$

$$\partial_t \rho e = -\partial_y \rho e v_y - P \partial_y v_y. \tag{16}$$

The Eqs. (9-16) can be rewritten in the form of generalized 1D continuity equation, Eq. (4). The full change in density, impulses and energy can be calculated by the sum of changes into direction x and y convection, that called as split step method. This method can be applied in any orthogonal geometry, where the hydrodynamic equations can be decoupled into equations of orthogonal direction convection (where mixed derivatives disappeared).

At this point we know that how to describe the convection of fluid by the system of differential equations in multidimension and in arbitrary orthogonal geometry. To model fluid convection we need an appropriate method solving the 1D generalized continuity equation (4) and adopt the split-step method.

4. Numerical Algorithm

There is many astrophysical phenomena, like the evolution of a star in binary system, or the supernova explosions, or a shock wave propagation in the interstellar media etc., where the physics can be described by fluid mechanics. In this phenomena the convection is not a simple inviscid flow, rather it has detonation fronts, flame fronts and shock waves, where the profile of fluid properties have strong discontinuity or steep gradients. Therefore, we need a numerical algorithm that can handle the strong discontinuity well.

Positivity is a property satisfied by the continuity equation Eq. (4). When density is everywhere positive and the source terms are zero, it is a mathematical consequence of the continuity equation that the density can never become negative anywhere. The retain this mathematical and physical property in numerical convection through the grid involve certain amount of numerical diffusion. The numerical diffusion arises as a consequence of the physical requirements that the profiles being convected remain stable while remaining positive (Borris et.al, 1996).

In an example shown in Fig. 1 the initial discontinuity erodes rapidly. This process looks like physical diffusion, but it arises due to numerical errors. The numerical diffusion occurs because material that has just entered the cell, and should still be near the left boundary, is smeared over the whole cell when the transported fluid elements are interpolated linearly back onto the grid. According to Borris et.al (1996), let us write first order three point explicit finite-difference formula for advancing the new value of the fluid element, for

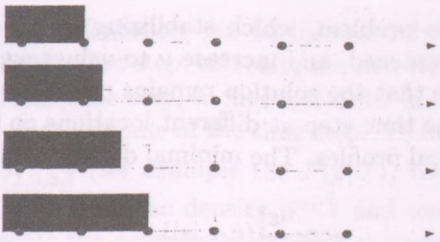

Figure 1: The picture shows the propagation of density discontinuity through four step of integration. When the fluid element enters a new cell it smears out, and the density discontinuity rapidly diffuse in the forthcoming steps (Borris et.al, 1996). The numerical diffusion appears due to the finite grid resolution.

example the density

$$\begin{aligned}\rho_i^{n+1} &= \rho_i^n + \frac{1}{2}[\varepsilon_{i+\frac{1}{2}}(\rho_{i+1}^n + \rho_i^n) - \varepsilon_{i-\frac{1}{2}}(\rho_i^n + \rho_{i-1}^n)] + \\ &+[\nu_{i+\frac{1}{2}}(\rho_{i+1}^n + \rho_i^n) - \nu_{i-\frac{1}{2}}(\rho_i^n + \rho_{i-1}^n)], \end{aligned} \tag{17}$$

$$\varepsilon_{i+\frac{1}{2}} = v_{i+\frac{1}{2}}\frac{\Delta t}{\Delta x}, \tag{18}$$

where ρ_i^{n+1} is the new value, at the end of time step $n+1$. $\nu_{i+\frac{1}{2}}$ are the non-dimensional numerical diffusion coefficients. Note that, subscript $i+\frac{1}{2}$ denotes the properties calculated at the i'th cell interface, while at the cell centered values denoted by i. In the Eq. 18 $v_{i+\frac{1}{2}}$ is the velocity of fluid elements at the cell interface, Δt is the time step, while Δx is the cell size. While the solution must be suffice the positivity requirements, $\varepsilon_{i+\frac{1}{2}}$ has to be chosen to a certain value. According to Eq. 18, Δt has to be chosen enough small, the fluid element can not to travel through the cell during one time step. This criteria called as Courant condition. Furthermore, positivity condition requires that $\nu_{i+\frac{1}{2}}$ must be positive and large enough. So, the conditions to assure positivity are

$$|\varepsilon_{i+\frac{1}{2}}| < \frac{1}{2} \tag{19}$$

$$\frac{1}{2} \geq \nu_{i+\frac{1}{2}} \geq |\varepsilon_{i+\frac{1}{2}}|. \tag{20}$$

When this conditions is guaranteed, the numerical diffusion decreases, but unphysical overshoots in fluid properties can appear. Nonlinear monotone method

need to circumvent this problem, which stabilizing $\nu = \frac{1}{2}\varepsilon^2$ diffusion where monotonicity is not threatened, and increase ν to values approaching $\nu = \frac{1}{2}|\varepsilon|$ when required to assure that the solution remains monotone. Different criteria are imposed in the same time step at different locations on the computational grid according to the local profiles. The minimal diffusion that assure positivity and stability is

$$\nu \simeq \frac{|\varepsilon|}{2}(c + |\varepsilon|), \qquad (21)$$

where c is a clipping factor, $0 < c < 1 - |\varepsilon|$, that controls how much extra diffusion must be added to ensure positivity over that required for stability, $\varepsilon^2/2$. In the vicinity of steep discontinuities, $c \simeq 1 - |\varepsilon|$, and in smooth regions away from local maxima and minima, $c \simeq 0$.

The new value of density, ρ_i^{n+1} must be strongly diffused to ensure positivity. The major point of nonlinear flux-correction is to remove the strong diffusion involves an additional antidiffusion stage. The new density values corrected by antidiffusion will be

$${}^*\rho_i^{n+1} = \rho_i^{n+1} - \mu_{i+\frac{1}{2}}\left(\rho_{i+1}^{n+1} - \rho_i^{n+1}\right) + \mu_{i-\frac{1}{2}}\left(\rho_i^{n+1} - \rho_{i-1}^{n+1}\right), \qquad (22)$$

where $\mu_{i+\frac{1}{2}}$ is the positive antidiffusion coefficients

$$\mu_{i+\frac{1}{2}} < \nu_{i+\frac{1}{2}} - \frac{1}{2}|\varepsilon_{i+\frac{1}{2}}|. \qquad (23)$$

to assure positivity. Antidiffusion reduces the strong diffusion, but also reintroduces the possibility of negative values or non physical overshoots in the corrected profile. The antidiffusive fluxes $\mu_{i+\frac{1}{2}}\left(\rho_{i+1}^{n+1} - \rho_i^{n+1}\right)$ should be limited so minima or maxima in the profile are made no deeper or larger by the antidiffusive stage. The correction of antidiffusive flux is

$$\mu_{i+\frac{1}{2}}\left(\rho_{i+1}^{n+1} - \rho_i^{n+1}\right)_{corr} = S * max\{0, min\left\{\begin{array}{c} S,\left(\rho_{i+2}^{n+1} - \rho_{i+1}^{n+1}\right) \\ \left|\mu_{i+\frac{1}{2}}\left(\rho_{i+1}^{n+1} - \rho_i^{n+1}\right)\right| \\ S,\left(\rho_i^{n+1} - \rho_{i-1}^{n+1}\right) \end{array}\right\}\} \qquad (24)$$

$$|S| = 1,\ S = sign\left(\rho_{i+1}^{n+1} - \rho_i^{n+1}\right). \qquad (25)$$

The antidiffusion stage should not generate new maxima and minima in the solution, nor accentuate already existing extrema. This nonlinear filtering is the base idea of flux correction (Borris et.al, 1996).

The accuracy of the calculation can be increased by two-step integration (Borris et.al, 1996). In the first step, calculate the new fluid properties (such as density, $\rho^{n+\frac{1}{2}}$) at the half time step, to find first-order accurate approximations to the fluid variables at the middle of the time step. Than, calculate the source terms of Eq. 4, like $S_{1,2,3,4}^{n+\frac{1}{2}}$ (for example the $P(\rho, T)$) using the time-centered half values of fluid properties (like density $\rho^{n+\frac{1}{2}}$ and temperature $T^{n+\frac{1}{2}}$). In the second step, integrate the generalized continuity equations, Eq. 4, using the time-centered source terms to find results which are second-order accurate in time at the end of the full time step.

Numerous flux correction schema have been developed yet. In GHYCO three type of flux correction process adopted, which have different diffusion and antidiffusion coefficients, and has different correction processes of antidiffusive flux. In the methods called ETBFCT and LCPFCT have somewhat simpler flux correction method. While the method called YDFCT is rather complicated, because the determination of the diffusion and antidiffusion coefficients differs at the two step integration process, and an additional correction coefficients used to determine the adequate correction at the antidiffusion stage. The detailed description of this methods can be found in Tóth & Odstrcil (1996).

5. Test Calculations and Results

Simplest test of shock capturing and non physical diffusion is propagate shock wave through the computational domain. As can be read in Sec. 4, the numerical diffusion erodes the discontinuity rapidly, which is an inherent problem of numerical simulations. The flux-corrected methods applied in GHYCO have different success in shock capture, but each of them have proper behavior than a simple upwind scheme.

In 1D test calculation a shock wave (strong discontinuity in density) with constant propagation velocity ($v = 1$) starts from the center of the domain. The boundary condition was periodic, therefore the property leave the domain at the right side boundary will enter at the left side boundary. After appropriate steps the shock wave has to arrive to its original position. The time step was constant, and its value set to large enough to show significant difference in methods, $\Delta t = 0.8$. The cell mesh extents 0 to 1, the cell number was 100. With this conditions the shock wave traveled fourth times across the simulation domain. Two kind of shock wave have chosen: a simple square wave and a semicircular shock wave. The flux-correction methods have different success on shock capturing, but each of them are better than a simple upwind scheme,

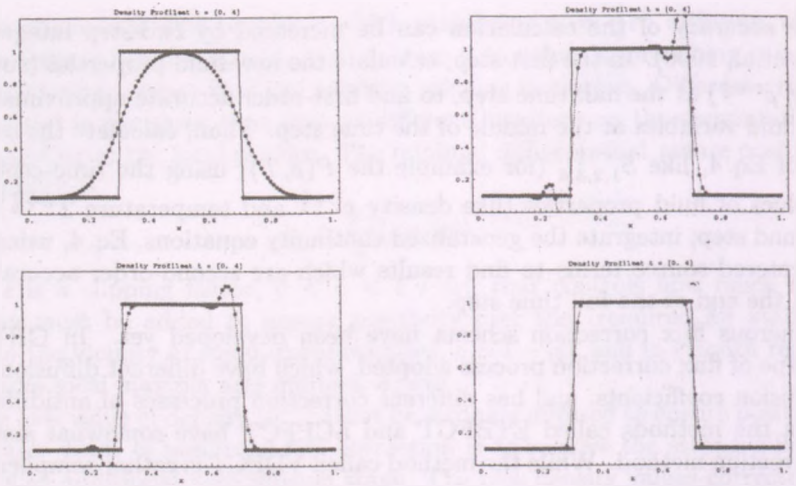

Figure 2: 1D square wave test calculations. Continuous line shows the square shock wave at the initial position, dotted lines shows it after 400 step ($t = 4$). As can be seen in the pictures at the top left, the upwind scheme unsuccess, the shock wave diffused strongly. In pictures at the top right LCPFCT method and bottom left ETBFCT method shows weak shock wave diffusion, but it has large phase errors. As can be seen in the picture bottom right, with YDFCT method the square wave has very small diffusion and phase error disappeared.

see the the density profiles in Fig. 2. It seems that the square shock wave can be handled by flux-correction method well (weak wave diffusion), but the phase error is large in ETBFCT and LCPFCT methods. In the simulation run with YDFCT method the phase error disappeared completely, thanks to the sophisticated flux-correction process. Note that, the phase error disappeared using enough small time step, $\Delta t = 0.2$ in ETBFCT and LCPFCT methods too.

The semicircle wave tests shows success in the shock capture also, see Fig. 3. While the non physical diffusion is large (the shock wave diffused strongly) in upwind scheme, it is rather small in the ETBFCT and the LCPFCT methods. Unfortunately strong phase errors generated in shock wave after 100 step, using even such a small time step as $\Delta t = 0.2$. But the YDFCT method succeed, diffusion and phase errors disappeared even using as large time step, as $\Delta t = 0.8$. In 2D test calculation a semicircular shock wave with constant propagation velocity ($v_x = 1, v_y = 1$) starts from the center of the domain. The boundary

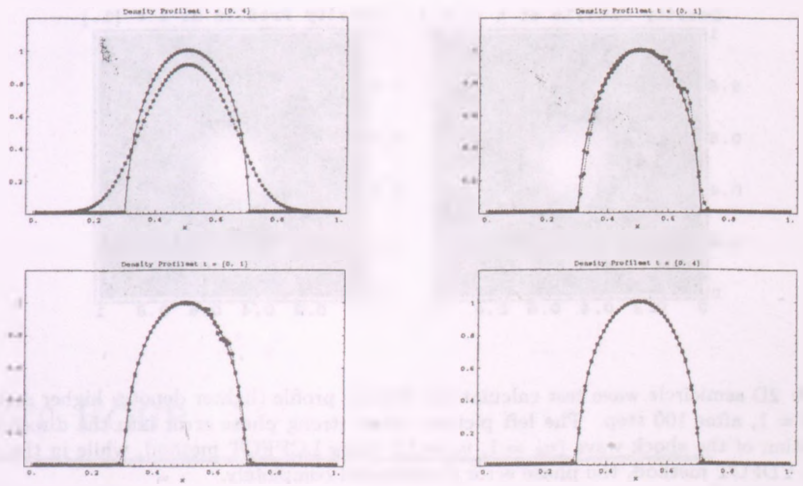

Figure 3: 1D semicircle wave test calculation. Continuous line shows the semicircle shock wave at the initial position, dotted lines shows it after 100 step ($t = 1.$). Same behavior can bee seen as in the Fig. 2. The upwind completely unsuccess, the LCPFCT and ETBFCT has phase errors, even such a small time step as $\Delta t = 0.2$, while the YDFCT method succeeded on shock capturing, even large time step as $\Delta t = 0.8$, after 400 step.

condition was periodic as well as in 1D tests. The time step was constant, $\Delta t = 0.2$. The cell mesh extents 0 to 1 into each directions, and domain has 100x100 cells. The ETBFCT and LCPFCT methods showed same misbehavior likewise in 1D tests, while YDFCT worked well. In Fig. 4, one can see the comparison of density profiles of LCPFCT and YDFCT methods, after 100 time step. The density profiles have the same phase error using of LCPFCT method likewise in 1D. The phase error disappeared completely simulation run with YDFCT method.

6. Conclusion

The 1D and 2D test calculations showed that, all flux-corrected methods are successful in shock capturing. The dispersion of shock waves can be reduced strongly by flux correction methods, but large phase errors occurred in simpler methods as LCPFCT and ETBFCT. In case of a square shock wave propagation, the phase error can be reduced by choosing enough smaller time step. Unfortu-

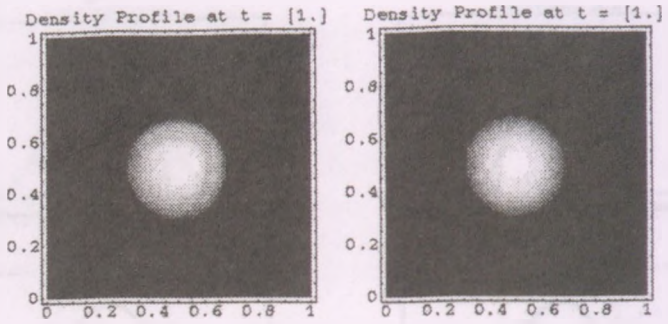

Figure 4: 2D semicircle wave test calculations density profile (lighter denotes higher density) at the $t = 1$, after 100 step. The left picture shows strong phase error into the direction of propagation of the shock wave ($v_x = 1$, $v_y = 1$), using LCPFCT method, while in the right picture, YDFCT method, the phase error disappeared completely.

nately, the same is not true for semicircle shock wave propagation, smaller the time step will not reduce the phase error significantly. By the advance of sophisticated antidiffusion stage of YDFCT method, the phase errors disappeared completely. So, there is no guarantee to ensure non physical phenomena, using a simpler flux-correction method, but YDFCT could be served veritable results.

Although the presented test calculations are made in 1D and 2D, the GHYCO code is capable to solve 3D hydrodynamic systems. In the future, the author try to run 3D tests and use GHYCO in real hydrodynamic problems, such as SN II explosions, evolution of compact binary stars, and simulation of magnetohydrodynamic phenomena in the Sun.

References

Blitz++ Numerical Library, 2001 http://www.oonumerics.org/blitz/

Borris, J. P., Landsberg, A. M., Oran, E. S., Gardner, J. H., *LCPFCT-Flux-Corrected Transport Algorithm for Solving Generalized Continuity Equations*, Naval Research Laboratory, 1993

Tóth, G., Odstricil, D., 1996, J. Comp. Phys., 128, 82

Wendt, J. F. von (Ed.),*Computational Fluid Dynamics:An Introduction*, New York: Springer-Verlag, 1996

Part Four

Optical studies and distant space

SHORT TERM TIDAL AND THIRD BODY PERTURBATIONS IN CLOSE HIERARCHICAL TRIPLE STELLAR SYSTEMS

T. Borkovits
Baja Astronomical Observatory of Bács-Kiskun County
H-6500 Baja, Szegedi út, Kt. 766, Hungary
E-mail: borko@alcyone.bajaobs.hu

Abstract

We introduce our first results with a new numerical integrator which was developed for studying the orbital and spin evolution of hierarchical triple stellar systems. (The code includes equilibrium tide approximations with arbitrary direction of rotational axes.) The variation of the orbital elements (e.g. the inclination of the close -eclipsing- binary) and its observational consequences according to the distorted models with different mass-distributions of the stars, as well as with and without dissipation is studied in the case of the well-known eclipsing triple system Algol. We found that in lack of the stellar dissipation, the presence of the third star may cause sudden fluctuations in the orbital elements and the stellar rotation of the binary members even in the previously synchronized case, too. The dissipation can eliminate these fluctuations, nevertheless some variations which would produce observable effects in the same order which have been measured in several eclipsing binaries are also present. We also studied the perturbations of a hypothetical third companion on the eccentric eclipsing binary AS Cam. We found that the complex influence of the variation of the orbital elements on the occurence of the minima events can make it impossible to obtain the real apsidal motion rate from an O–C curve which covers only a small fraction of the period of the rotation of the apsidal line.

KEYWORDS: *methods: numerical – celestial mechanics – stars: binaries: close – stars: individual: Algol – stars: individual: AS Cam*

1. Introduction

Several close binary systems have third (or further) more distant companions. As it is well-known the dynamics of the members of such systems fairly depart from the pure Keplerian motion. Considering only the gravitational perturbations, these arise from two different aspects. On one side the gravitational force of the more (but not so) distant tertiary perturb the orbital motion of the close

binary, while on other side the shape of the stars will not be spherical due to purely the strong gravitational interaction between the two stars. Furthermore, the members of the very close binaries usually rotate so fast that their oblateness gives another non-spherical contribution to their gravitational field.

Both these effects may be strong enough to cause observable orbital changes on a time scale of years, or decades. Some well-known examples are the variation of the eclipse depths in some eclipsing binaries due to the orbital precession caused by an inclined third star, as well as the apsidal motion forced by the distortion of the binary members which is observed in already more than a hundred systems.

In most of the previous works these two different categories of the orbital perturbations were studied separately. On one hand, the perturbations of a third body were considered in the frame of the general three body problem, i.e. all three stars were treated as mass-points. Departing from the lunar theories Harrington (1968, 1969) gave the first formulae for the perturbations for both the orbital elements of the binary orbit and the orbit of the tertiary. Söderhjelm (1975, 1982) reformulated these equations into a more explicit form, and used them to compare their predictions with the observational results for two well-known close triples, Algol (which is subject of the present study, too) and λ Tau for the first time. Recently Ford et al. (2000) published the third order theory of the perturbations in the hierarchical triple systems.

On the other hand, the tidal interaction was mainly studied in the case of close binaries without further companion(s). The main aim of such studies is to understand how the circularization and (both rotational and orbital) synchronization of the close binaries works, as well as to interpret those orbital perturbations which cannot be discussed in the frame of the mass-point models. (We mainly refer to the apsidal motion in close, eccentric binaries.) While most recent works (e.g. Claret, 1999;Witte & Savonije, 1999; Smeyers & Willems, 2001;Claret & Willems, 2002) in this field are based on the most recent stellar models, they do not consider the possible perturbative effects of a close third companion which can make e.g. some precession.

While this second treatment seems to be correct at least in those cases where the system really does not contain relatively close third companion, the mass-point approximation of the hierarchical triple systems containing distorted components should be revised. As the observational interval of the closest triple systems becomes longer and longer, and the precision of the measurements in the future will grow even several orders (we refer to the planned space missions, e.g. GAIA) it will be necessary to examine the different perturbations simultaneously.

Up to know only a very limited number of papers were published where the dynamics of the hierarchical triple systems was discussed in the tidally distorted body model. After the first works of Kopal (1978, Chapter V) and Söderhjelm (1984) recently Kiseleva et al. (1998); Eggleton & Kiseleva-Eggleton (2001) studied the effects of tidal friction in triple stars with numeric integration based on the model of Eggleton et al. (1998), while Beust et al. (1997) used a simplified method to explain the observed properties of the circularized but not synchronized eclipsing binary TY CrA, which is also a member of a close hierarchical triple system. Finally, Borkovits et al. (2002) integrated numerically the orbital motion of the eclipsing triple system IM Aur, and compared the perturbations in the orbital elements in the frame of the mass-point model and the disturbed star model. This integration was already based on that model and integrator code which is described in the present article.

In this work we use the improved version of the method which was used already in Borkovits et al. (2002). The model itself is described somewhere else (Borkovits et al., 2004a). Here we concentrate on two interesting systems, namely the Algol and AS Camelopardalis.

2. Algol

The well-known triple system Algol was also considered in the similar work of Kiseleva et al. (1998). This gives the opportunity of checking of equations on one side, and shows the differences between the methods on the other side. So for the first runs we used the same parameters for the β Persei system which were described in Kiseleva et al. (1998). On the other hand, we have to note that setting the mutual inclination of the orbits to 100°, as it was done in the previously mentioned work, results in a too fast variation in the observable inclination of the close system which is not supported by the observations (see e.g. Söderhjelm, 1980). For these first runs we used three different sets of the k_2, k_3 apsidal motion constants (see Table 2.). In the first run, (denoted by "Pol") a simple $n = 1.5$ politrop model was applied for the primary, and $n = 3$ one for the secondary. The k_2 values which can be calculated from the Q parameters of Kiseleva et al. 1998 ("K98" run) differ from the politropic ones with a factor of two, which is a 'result' of some mistakes in the equation of motion in Kiseleva et al. 1998, as it was pointed out in Eggleton & Kiseleva-Eggleton 2001. Nevertheless, these values seem to be unrealistically large comparing them with the tabular data of Claret & Gimènez (1992), whose tables were used to run "CG92". Fig. 1 is analogous with Figs. 3b–3e of Kiseleva et al. (1998). Compar-

Code	$k_2^{(1)}$	$k_3^{(1)}$	$k_2^{(2)}$	$k_3^{(2)}$
Pol	0.0144	0.00368	0.1433	0.0529
K98	0.0288	0.00736	0.2866	0.1058
CG92	0.0038	0.0011	0.0240	0.0087

Table 1: The apsidal motion constants for the Algol AB stars in three different runs.

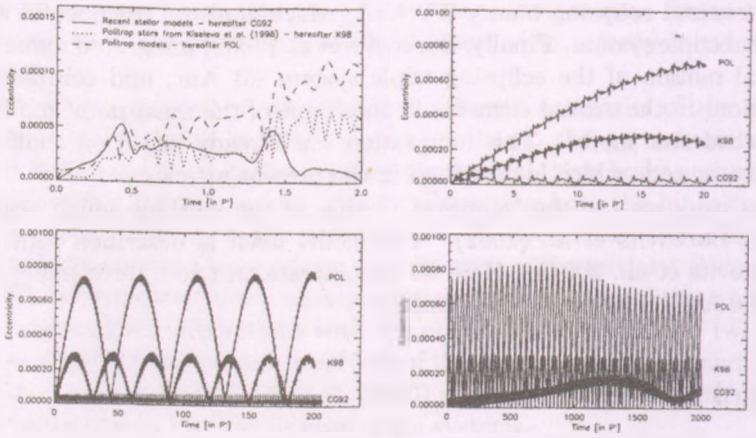

Figure 1: The variation of the 'super-osculating' eccentricity of the close binary of Algol during 2, 20, 200, and 2000 revolutions of the third companion, with three different sets of apsidal motion constants. (See text for details.) The eccentricity was sampled at the integration step which is the closest to the pericenter of the inner binary. In the first 10000 days the sampling was done at every pericentrum passage, then at every tenth one.

ing the "K98" curve in our Fig. 1a to Fig. 3b of the aforementioned paper the agreement is almost perfect (apart from the different sampling). However, the next panels which cover an approx. 37 year-long time interval show completely different behaviour in the eccentricity variation. This is because the stellar equators are 'frozen' into the orbital plane of the close pair in Kiseleva et al. (1998), i.e. no stellar precession is allowed. In our physically more realistic case the orbital precession of the plane of the close binary driven by the effect of the aligned tertiary forces a slight precession in the rotation of the members of the close binary which reacts to the orbital eccentricity (as well as the other orbital elements). This effect in the present configuration can grow up the eccentricity variation even by one order of magnitude in the politropic cases, although it

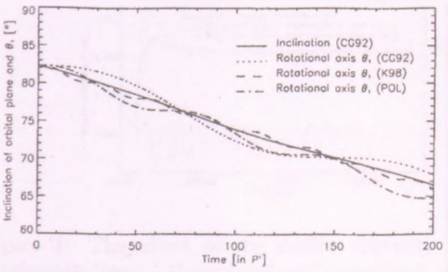

Figure 2: The precession of the component A of Algol. We plotted the variation of the i inclination of the binary together with the variation of θ_1, the inclination of the rotational axis of Algol A, at different mass distributions.

remains almost inefficient in the recent centrally more condensed stellar models ("CG92" run). In Fig. 2. we plotted the variation of the inclination-equivalent Eulerian angle θ_1 of the component A of Algol during 200 revolutions of the outer body in the different models. As a reference, the variation of the visible orbital inclination i of the close pair is also shown in the "CG92" run. (The variation of θ_2 as well as the inclination in the other runs are very close to the reference inclination line, so they are not drawn.) We carried out the integration on a longer time-scale, too. As it can be seen in Fig. 2., that in the case of the physically more realistic "CG92" run, departing from the synchronized rotation and revolution the rotational angular velocity of the components of Algol begin to show sudden fluctuations within 50 000 years, which react on some of the orbital elements, too.

In the following we 'switched on' the dissipation. We set to $\lambda_1 = 10^{-4}$, $\lambda_2 = 10^{-3}$, which correspond to a tidal lag time of $\Delta t_1 \approx 2\overset{\mathrm{d}}{.}5 \times 10^{-6}$, and $\Delta t_2 \approx 5\overset{\mathrm{d}}{.}8 \times 10^{-5}$, respectively. The other stellar parameters were identical with the values of "CG92" run, while the orbital and rotational parameters were chosen to be equal with the corresponding parameters of "CG92" run at $t = 0$, and $t \approx$ 104 000 years, i.e. in the first case ("CG92d1") the close system departed from a synchronized state, while in the second one ("CG92d2") the above mentioned fluctuations were already present. In the first case the large fluctuations of the non-dissipative run did not appear in the orbital elements, only small quasi-cyclic variations were found in the order of 10^{-5} in the eccentricity, and $10^{-5}\mathrm{R}_\odot$ in the semi-major axis. (Note that the perturbation of the semi-major axis of the binary in every pericenter and/or nodal-line passage of the tertiary may be larger even by two orders.) The second case can model the relaxation of such a triple system after some moderate perturbations (e.g. mass-transfer, or engulfing a giant planet, etc). As it is well seen in Fig. 4 in this latter case the small lag nearly circularized the orbit, and re-synchronized the rotation of the more dissipative secondary on a time-scale of 10 000 years, while for

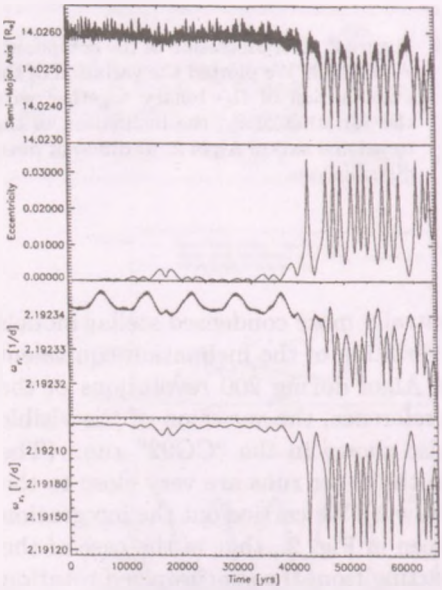

Figure 3: The variation of the semi-major axis, eccentricity, and the rotational angular velocities of the components in the close pair of Algol in the "CG92" run.

the less-dissipative primary the circularization time is naturally longer, in the present run we extrapolated it as some 10^6 years. Nevertheless, even after the whole time-interval of our some 10^5 year-long integration small periodic variations both in the orbital as well as in the rotational angular momenta were present, with a period which was exactly the half-period of the precession of the orbital plane. This change in the orbital angular momentum gives a period change rate $\dot{P}/P$ in the order of $10^{-10}\mathrm{d}^{-1}$. This magnitude of secular period change is found in several Algol-type eclipsing binaries, and usually explained as a consequence of some mass-transfer or mass-loss events, although in some systems this explanation does not seem to be plausible. (A recent case is e.g. IM Aur, see Borkovits et al., 2002.)

3. AS Cam

While in the previously studied case the inner binary had an almost circular orbit with approximately synchronized stellar rotation, in the present subsection we concentrate on one member of the small group of eccentric binaries with remarkable relativistic apsidal motion contribution, which show a significantly lower apsidal motion rate than which is calculated from theory. Since that

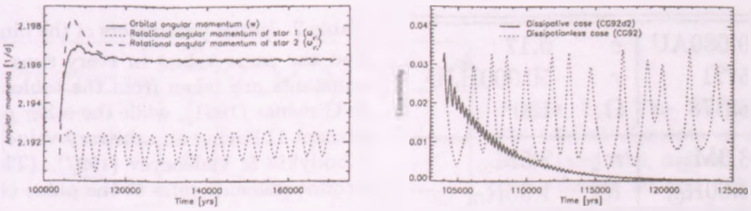

Figure 4: The effect of the stellar viscosity in the "CG92d2" run. Left panel: the fast synchronization of the orbital and rotational angular velocities. Right panel: circularization of the inner orbit. Here we also plotted the variation of the eccentricity in the non-dissipative case as well.

disrepancy first time was found at DI Herculis (Semeniuk, 1968), several authors have examined this phenomenon. A summary of the efforts can be found in Claret (1998). One of the possible explanation is the perturbation of a third component.

AS Cam is an Algol-type eclipsing binary formed by two main sequence stars of spectral types late B. The discrepancy in the apsidal motion rate for the first time was found by Maloney et al. (1989). According to the most recent analysis of Wolf et al. (1996) the velocity of the apsidal advance is $\dot{\omega}_{\rm obs} = (18.3 \pm 2.6)^\circ/100{\rm y}$, while the theoretical value is $\dot{\omega}_{\rm theo} = 44^\circ/100{\rm y}$. Khodykin & Vedeneyev (1997) calculated the possible mass and orbital elements of such a perturbing body which could have explained the measured anomaly, and could have remained undetectable up to now. They found, that supposing a circular outer orbit, the third companion having a mass between $0.78 - 1.45{\rm M}_\odot$ might revolve on an inclined ($i_{\rm m} > 32^\circ$) orbit of period $P' = 0.6 - 2.9{\rm y}$. Two years later Kozyreva et al. (1999) really reported that after the removal of the apsidal motion terms the O–C curve shows small amplitude periodic variations. They fitted light-time solution, and they found a third body of mass ($m_3 = 1.1 - 1.7{\rm M}_\odot$) orbiting in an eccentric ($e' = 0.5$) orbit with period $P' = 805{\rm d}$. We also carried out the analysis of the O–C diagram, but only a very poor, and consequently very questionable fit was found. Nevertheless, we used the calculated orbital elements as input parameters in our numerical studies.

We carried out several sets of integrations, from which here we present five different ones. The (fixed) initial parameters of the binary are given in Table 2, while the initial orbital elements (and mass) of the tertiary are listed in Table 3. All columns of this latter table refer four individual runs. One of them is the

a	0.080AU	e	0.17
ω	57?1	τ	50 000HJD
i	88?78	Ω	130°
m_1	$3.3M_\odot$	m_2	$2.5M_\odot$
R_1	$2.60R_\odot$	R_2	$1.96R_\odot$
$k_2^{(1)}$	0.0049	$k_2^{(2)}$	0.0038
$k_3^{(1)}$	0.0011	$k_3^{(2)}$	0.0008

Table 2: Initial parameters of the binary which had the same values in every runs. The $k_j^{(i)}$ constants are taken from the tables of Claret & Gimènez (1991), while the other parameters (except Ω having an arbitrary value) are from Khodykin & Vedeneyev (1997). (The angular orbital elements refer to the plane of the sky.)

	AS1	AS2	AS3	AS4	AS5
P'	882^d				
e'	0.41				
ω'	279°	279°	261°	279°	261°
τ'	$2\,085^d$				
i'	88?0	88?0	88?0	30?0	30?0
Ω'	130°	85°	220°	130°	0°
m_3	1.1	1.1	1.1	2.5	2.5

Table 3: The initial orbital elements (and mass) of the tertiary components at different integration runs. (The angular elements refer to the plane of the sky with an arbitrary initial direction.)

dissipationless case, starting the binary at its periastron, and three of them with dissipation ($\lambda_1 = \lambda_2 = 10^{-4}$), with initial position of the binary in its periastron, apastron, and at true anomaly $f \approx 104°$, where the instantaneous orbital angular velocity was equal with the averaged one. (As we used that option of our code which makes equal the initial instantaneous orbital angular momentum with the same quantities of the stellar rotation, or more precisely sets $\vec{w}_{\rm initial} = \vec{\omega}_{1,\rm initial} = \vec{\omega}_{2,\rm initial}$ on an iterative way, it means that the initial rotation of the stars were synchronized on this three different ways.) The mass of the third star was also changed at the different sets of runs (in a way, that the amplitude of the light-time curve was constant). Furthermore, the relative spatial orientation of the two orbits was varied also. The variation of the orbital elements during the integration AS5, where the perturbations are the larger can be seen in Fig. 5. This figures represent the dissipationless cases. The left panel covers a time interval of approx. 72 years, which is naturally less than a moment from dynamical point of view, but this is the time interval (in the best lucky cases) which is usually can be reached for the present studies. Additional integrations were performed for the close binary without tertiary component both in the quasi-synchronized rotator, and the precessing secondary component case.

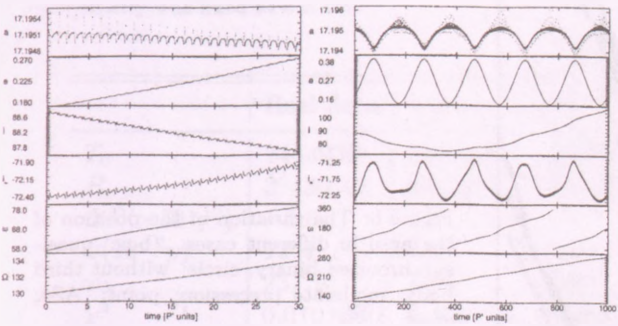

Figure 5: The variation of the orbital elements of the binary (and the mutual inclination) during 30 and 1000 revolutions of the outer body. (See Col. AS5 of Table 3 for the initial values.)

In the quasi-synchronized binary case $U = 381\,800\text{d} \approx 1\,046\text{y}$ was found for the period of the apsidal motion without the relativistic contribution. This yields $\dot{\omega}_{\text{cl}} = 34°4/100\text{y}$ for the classical parts of the apsidal advance rate, which is in good agreement with the theoretical value of Maloney et al. (1989) $\dot{\omega}_{\text{cl}} = 35°.8 \pm 5°.8/100\text{y}$.

We plotted in Fig. 6 the variation of the apsidal line for four different configurations of the system. Comparing the three perturbed case with the quasi-synchronized binary configuration (marked with boxes) it can be seen well that there are no significant differences in the apsidal motion periods. In the case of a binary containing an inclined component (denoted with circles) a somewhat but not significantly larger U value was found. Nevertheless, in the third body perturbed cases the speed of the apsidal advance varied according to the variations of the inner eccentricity (the corresponding variations of the orbital elements in the case of the AS5 run are plotted in Fig. 5). This is because for the observable apsidal advance rate $\dot{\omega} \sim e^{-1}$. (Strictly speaking this would be correct only if the perturbing force had no component perpendicular to the orbital plane. Nevertheless, this latter component multiplied by $\cot i$. As the inclination is close to 90° in the case of the eclipsing systems interesting from the present point of view this contribution can be safely eliminated.)

Although, as it was shown above, we cannot explain the observed discrepancies with the perturbations only in the period of the apsidal advance rate U, however, we shall illustrate that the net effect of the different kind of perturbations may result in the observed phenomenon, especially in the case when only a small portion of the total apsidal motion period is observed. As it is well known the departure of the secondary minima from the phase 0.5 can be expressed by

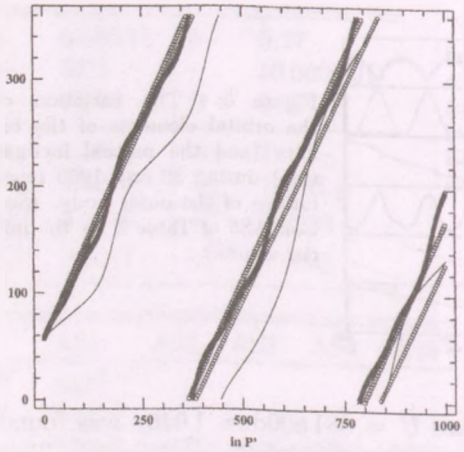

Figure 6: The variation of the position of the apse in different cases. (box: quasi-synchronized binary; circle: without third body, moderate precession; point: AS4; cross: AS5)

the following quantity:

$$D = \frac{P}{\pi}\left\{\tan^{-1}\left[\frac{e\cos\omega}{(1-e^2)^{1/2}}\right] + (1-e^2)^{1/2}\frac{e\cos\omega}{1-e^2\sin^2\omega}\right\}. \qquad (1)$$

It is easily can be seen that

$$\frac{\mathrm{d}D}{\mathrm{d}t} = \frac{P}{\pi}\frac{2(1-e^2)^{1/2}}{1-e^2\sin^2\omega}\left[\frac{\mathrm{d}(e\cos\omega)}{\mathrm{d}t} + \frac{e\sin\omega}{1-e^2\sin^2\omega}\frac{\mathrm{d}(e\sin\omega)}{\mathrm{d}t}\right] + \frac{D}{P}\frac{\mathrm{d}P}{\mathrm{d}t}, \qquad (2)$$

consequently the most important effect is the above mentioned simultaneous variation of the eccentricity and the apsidal advance rate which can be descripted by one common term, i.e. the variation of $e\cos\omega$. Furthermore, although according to (2) the variation of the orbital period is directly not crucial, due to its cummulative effect on the O–C diagram, it can still be very important. It can be seen well in Fig. 5 that the semi-major axis of the binary no longer will be constant. This variation is very small, e.g. in the case of the AS5 run, where its rate is the largest in our integrations, it remains in the order of 100km within a century (see the left panel of Fig. 5). Nevertheless, due to its cumulative effect (via the orbital period) on the occurence of the eclipsing minima, this variation strongly affects the shape of the O–C curve which is the only source of the apsidal motion rate information.

In order to study these aforementioned effects in details we integrated the configuration AS3 on a longer time-scale with all the three different initial rotational velocities. Then we calculated the eclipsing minima times on different

Table 4: Apsidal motion constants from the best fits of the sample data series. The inner eccentricity was fixed at $e = 0.17$.

	Real data	unpert.	#1	#2	#3
T_0	2444939	2444942	2860025	3500023	5440340
P_s	$3^d.430968$	$3^d.428330$	$3^d.428306$	$3^d.428292$	$3^d.428256$
$\Delta P(\times 10^{-10})$	-2.6	0.0	8.2	8.5	4.4
$\dot{\omega}(10^{-2}y)$	$12^\circ.4$	$34^\circ.5$	$52^\circ.0$	$53^\circ.2$	$45^\circ.0$
ω_0	$237^\circ.0$	$232^\circ.3$	$210^\circ.9$	$28^\circ.2$	$233^\circ.7$
χ^2	0.010108	0.000002	0.003244	0.003554	0.001257

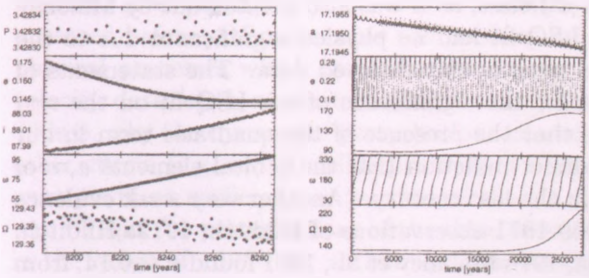

Figure 7: The variation of the orbital elements of the binary in the dissipative case. See text for details.

century-long intervals on that part of the integrations where the apsidal line of the binary lies in similar position than this century, furthermore, eccentricity also takes its present value of $e \approx 0.14 - 0.17$. We analysed the O–C curves calculated from this century-long data sets by the least-quares method. We fitted simultaneously a second order polinom, and Fourier-terms of the light-time orbit and the theoretical apsidal advance up to the third order in e. Our results are listed in Table 4, together with the analysis of the real O–C data, and the data produced from the unperturbed two-distorted body integration.

As it is well can be seen, in all third body perturbed cases a significant quadratic term were found. In the case of the #1 data series an eyeball estimation gives a $\delta P \approx -2 \times 10^{-6}$d/100y period decrease (see left panel of Fig. 7, where we plotted directly the instantaneous Keplerian period instead of the semi-major axis). This gives that $\Delta P \approx -1.9 \times 10^{-10}\mathrm{d}/P$. This is significantly less (although is in the same order) than the rate which was deduced from the O–C diagram, moreover, it has different sign. The same is the situation in the

other two cases, too. This suggests that the parabolic term mainly affected by the non-constancy of the e, and $\dot{\omega}$. Furthermore, although in the cases #1–3 for $e = 0.17$ we obtained smaller period for the apsidal advance rate, i.e. larger $\dot{\omega}$ than the real one, fixing e at larger values better fits were obtained (e.g. for $e > 0.5$ the χ^2 was less even by an order), and in parallel with the increasing of e, $\dot{\omega}$ decreased. (Here we refer again the form of Eq. [2].) In conclusion we were not able to reproduce from that small portion of the numerically produced O–C data the real orbital parameters in the perturbed cases, although our series were well-sampled and noiseless.

Analysing the real O–C data in the same manner a quadratic term in the same order of magnitude, but with different sign was found also. In that case a smaller value of the χ^2 was found by decreasing the (fixed) eccentricity e which corresponded to a smaller period U (i.e. larger $\dot{\omega}$). The theoretical apsidal motion period was found at $e \approx 0.10$. The same value was suggested by Krzesinski et al. (1990), nevertheless, as it was also pointed out by Maloney et al. (1991) this result of the LSQ fit had no physical significance due to the presence of the large noise and gaps in the observed data. The statements of this latter paper are also valid for the significance of our LSQ fit on the real data, nevertheless, we suppose that the presence of the quadratic term in our solution might be a further indirect indiction that the orbital elements a, e of the real binary may be varied in the last century. (Another very weak evidence for this that while from the 1969-1971 observations of Hilditch, 1972a;Hilditch, 1972b and Padalia & Srivastava, 1975 Maloney et al., 1991 found $e \approx 0.14$, from the December 1981 photometry of Khaliullin & Kozyreva, 1983 they deduced already $e \approx 0.16$. Although this variation seems to be unrealistically large it would be very important to obtain new precise light curves from the system.)

4. Conclusion

We studied the motion of close binaries in triple stellar systems combining the effects of the gravitational perturbations of the tertiary as well as the influence of the stellar distortion, arbitrary stellar rotation, and tidal friction. Besides the questions of the dynamical evolution of such systems we concentrated mainly those short-term effects which might be observable already or at least in the near future. This is the main reason while eclipsing systems were chosen as examples. From that point of view the short fluctuations found in the system Algol in some of our integrations might be of importance. On one side, they show that chaotic variations may be also present even in dynamically relaxed

systems in the presence of a tertiary component, and so a new explanation of the several non-understood eclipsing O–C diagrams may arise. On the other side, as it was shown, the presence of the dissipation eliminated this phenomenon. Nevertheless, in the present case we used a large, perhaps unrealistically large dissipation rate. We plan to perform further tests to study the effect of smaller dissipation rate for such kind of perturations. Such studies would be especially important as the presence of this kind of perturbations could give a possibility to measure the dissipation rate of stars in close triple systems.

Finally, we show a further aspect, why the short term, great accuracy integration (which requires more and more complex dynamical models) of motion of the concrete systems will be important in the near future. The small astronomical satellites of the next decades will produce enormous amount of high precision data for millions of stars. From these satellites a precision of micromagnitude is expected. This (together with the relatively long observational interval) gives possibility to observe so small physical effects which could not be observed before. Of course this can be done only if all of the presently known physical processes which can influence on the observed light-curve (or radial velocity data, etc.) at the available accuracy are carefully taken into account. This is, why the dynamical perturbations caused by a close third body, or some other effects (e.g. small stellar precession) have to be already considered. For example in an earlier paper we showed that in the close hierarchical triple system IU Aur the perturbations of the third companion may affect the light-curve in the range of 10^{-4} mag within some months (Borkovits et al., 2004b). Perhaps more important is the indirect problem, when the perturbations might be deduced from the light curve, and consequently we can get more exact third body parameters, or, what is more important, further information on the tidal terms. Keeping in mind this possibility we plan to improve the above described dynamical model with more accurate stellar models.

Acknowledgments

I would like to thank the help and comments of Dr. Emese Forgács-Dajka.

References

Beust, H., Corporon, P., Siess, L., Forestini, M., & Lagrange, A-M., 1997, *A&A* , 320, 478

Borkovits, T., Csizmadia, Sz., Hegedüs, T. et al., 2002, *A&A* , 392, 895

Borkovits, T., Érdi, B., Forgács-Dajka, E., & Kovács, T., 2003, *A&A* , 398, 1091

Borkovits, T., Forgács-Dajka, E., & Regály, Zs., 2004, *A&A* , accepted
Borkovits, T., Könyves, V., & Hegedüs, T., 2004, Proc. 2nd Eddington workshop "Stellar Structure and Habitable Planet Finding", Palermo, 9–11 April 2003 (ESA SP-539, F. Favata, S. Aigrain eds.), in press
Claret, A., 1998, *A&A* , 330, 533
Claret, A., 1999, *A&A* , 350, 56
Claret, A., & Gimènez, A., 1991, *A&AS* , 91, 217
Claret, A., & Gimènez, A., 1992, *A&AS* , 96, 255
Claret, A., & Willems, B., 2002, *A&A* , 388, 518
Eggleton, P. P., & Kiseleva-Eggleton, L., 2001, *ApJ* , 562, 1012
Eggleton, P. P., Kiseleva, L. G., & Hut, P., 1998, *ApJ* , 499, 853
Ford, E.B., Kozinsky, B., & Rasio, F.A., 2000, *ApJ* , 535, 385
Harrington, R.S., 1968, *AJ* , 73, 190
Harrington, R.S., 1969, Cel. Mech., 1, 200
Hilditch, R.W., 1972a, MemRAS, 76, 141
Hilditch, R.W., 1972b, *PASP* , 84, 519
Khaliullin, Kh., F., & Kozyreva, V.S., 1983, Ap&SS, 94, 115
Khodykin, S.A., & Vedeneyev, V.G., 1997, *ApJ* , 475, 798
Kiseleva, L. G., Eggleton, P. P., & Mikkola, S., 1998, *MNRAS* , 300, 292
Kopal, Z., 1978, Dynamics of Close Binary Systems, D. Reidel, Dordrecht
Kozyreva, V.S., Zakharov, A.I., & Khaliullin, Kh.F., 1999, IBVS, No. 4690
Krzesinski, J., Kuczawska, E., & Pajdosz, G., 1990, IBVS, No. 3495
Maloney, F.P., Guinan, E.F., & Boyd, P.T., 1989, *AJ* , 98, 1800
Maloney, F.P., Guinan, E.F., & Mukherjee, J., 1991, *AJ* , 102, 256
Padalia, T.D., & Srivastava, R.K., 1975, Ap&SS, 38, 87
Semeniuk, T., 1968, Acta Astr., 18, 1
Smeyers, P., & Willems, B., 2001, *A&A* , 373, 173
Söderhjelm, S., 1975, *A&A* , 42, 229
Söderhjelm, S., 1980, *A&A* , 89, 100
Söderhjelm, S., 1982, *A&A* , 107, 54
Söderhjelm, S., 1984, *A&A* , 141, 232
Witte, M.G., & Savonije, G.J., 1999, *A&A* , 350, 129
Wolf, M., Šarounová, L. & Diethelm, R., 1996, *A&AS* , 116, 463

Differential rotation of LQ Hya and IL Hya from Doppler imaging

Zs. Kővári[1] and M. Weber[2]
[1]Konkoly Observatory of the Hungarian Academy of Sciences
H-1525 Budapest, P.O.Box 67., Hungary
[2]Astrophysical Institute Potsdam
An der Sternwarte 16, 14482 Potsdam, Germany
E-mail: [1]kovari@konkoly.hu, [2]mweber@aip.de

Abstract

We introduce two techniques of detecting surface differential rotation on stars, *cross-correlation* of consecutive Doppler images and the *sheared-image method*. We test the methods using artificial and real data. As a result, for the rapidly rotating K2 dwarf LQ Hya we find an almost rigid rotation, with weak (α=0.0056) equatorial acceleration. On the giant component of the RS CVn-type star IL Hya we detect solar-type differential rotation with α=0.03.
Keywords: *stars: activity of – stars: imaging – stars: individual: LQ Hydrae, IL Hydrae – stars: late-type – starspots*

1. Introduction

The primary source of all kind of solar magnetic phenomena is the dynamo mechanism. In our time, the Sun is the only star for which one of the conditions of a working dynamo, the surface differential rotation can be measured directly. A century ago, in the morning of extensive solar surface observations, tracing of sunspots at different latitudes was used to measure the surface shear (e.g. Maunder & Maunder, 1905). Although, today's advanced techniques can measure the solar surface (and radial) rotation law with much higher accuracy, on stars the detection of the known effects of the surface differential rotation is still a difficult observational task. Direct tracing of starspots is a promising tool to observe surface shear, however, requires high resolution spectroscopic data and reliable surface reconstruction technique.

Notwithstanding the hightech solar observations, the magnetic dynamo is still poorly understood. The surface differential rotation measurements for stars of different types, thus, would mean useful input for developing (M)HD models and solar/stellar dynamo theory.

2. Doppler imaging of stars

Since today's most advanced telescope cannot even resolve the nearest stellar surface, to get sufficiently detailed maps from stars, we are restricted to use indirect mapping techniques. The one called Doppler imaging is an inversion to recover the 2-D image of a rapidly rotating star from a series of high-resolution spectral line profiles. The term 'Doppler imaging' was introduced by Vogt & Penrod (1983) who recognized that for a spotted star, a photospheric absorption line, in which rotation is the dominant broadening mechanism, displays variable irregularities with rotation (see Fig. 1). The authors demonstrated on HR 1099, that a series of spectral lines covering the rotation cycle could have been used to recover the image of the stellar surface.

Our Doppler imaging program TEMPMAP is based on the code by Rice et al. (1989). The program performs a full LTE spectrum synthesis by solving the equation of transfer through a set of Kurucz (1993) model atmospheres, at all aspect angles, and for a given set of chemical abundances. Simultaneous inversions of the spectral lines are then carried out using a maximum-entropy regularization. A more detailed description of the code and additional references regarding the inversion technique can be found in Piskunov & Rice (1993), Rice & Strassmeier (2000) and most recently in Rice (2002).

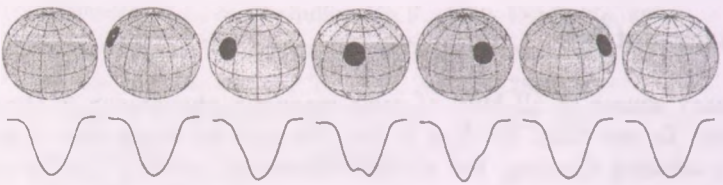

Figure 1: Basic principle of Doppler imaging: a photospheric spectral line of a spotted star performs changes with rotation.

3. Detecting differential rotation

3.1. Cross-correlation of Doppler images

Cross-correlation of Doppler images is among the most reliable techniques used to detect differential rotation on stars. Two independent images are taken, which should be close in time since changes in the spot configuration wash out

the differential rotation pattern. Each latitude strip of the first image is cross-correlated with the same strip of the second image. This way the latitudinal cross-correlation functions are assembled into a cross-correlation image, in which the best correlating pattern can be fitted by the differential rotation law in its usual form:

$$\Omega(\theta) = \Omega_{\rm eq} - \Delta\Omega \sin^2 \theta \ , \tag{1}$$

where θ is the stellar latitude, $\Omega_{\rm eq}$ the equatorial angular velocity, and $\Delta\Omega = \Omega_{\rm eq} - \Omega_{\rm pole}$ is the angular velocity difference between the equator an the pole. The differential rotation parameter α is then defined as

$$\alpha = \Delta\Omega / \Omega_{\rm eq} \ . \tag{2}$$

AB Dor (Donati & Collier Cameron, 1997) was the first star for which differential rotation was measured using this method. Recently, the technique has been succesfully applied to other stars, including IL Hya (Weber & Strassmeier, 1998), HD 218153 (Weber & Strassmeier, 2001), PZ Tel (Barnes et al., 2000), Speedy Mic (Barnes et al., 2001), and σ Gem (Kővári et al., 2002).

3.2. Sheared-image method

Another possibility to reconstruct the differential rotation using Doppler imaging is to include the latitude-dependent rotation pattern in the inversion code. The differential rotation changes the line-profile in two ways. The overall shape of each line-profile is changed to a flatter shape for solar-type differential rotation with the pole rotating slower than the equator, or to a more pointed one in the case of a pole rotating faster than the equator. This can also be observed on unspotted stars. The second effect is the different rotation period for spots found at different latitudes. However, due to the already large number of unknowns in Doppler imaging, the image shear is introduced as a fixed parameter, and calculations are done for a range of meaningful $\Omega_{\rm eq} - \alpha$ parameter pairs. Each parameter pair results into a Doppler image with a χ^2-value describing the goodness-of-fit to the spectral line profiles. Plotting these χ^2-values in a period$-\alpha$ diagram, the most probable surface rotation law can be determined. The big advantage of the sheared-image method compared to the cross-correlation technique is, that spectra to compute one single Doppler image are already enough, in contrast with the minimum of two independent data sets in the second case.

4. Simulating differential rotation

To check the ability of the two techniques in recovering surface differential rotation we carry out a test on artificial data. Using the test star from Rice & Strassmeier (2000) (imitating II Peg with a rotation period of 6.72 days) with the spot configuration shown in the top of Fig. 2 we rotated it differentially according to a differential rotation law of $\Omega(\theta) = 53.537 - 2.677 \sin^2 \theta$ °/day (α=0.05). Then, two datasets were constructed for two consecutive rotation cycles. For each day of the two rotations one spectral line is synthesized (13 Ca I-6439 spectra for both cycles).

Our Doppler imaging code TEMPMAP is then applied for the two artificial datasets separately, ignoring differential rotation. The resulting images are shown in the bottom of Fig. 2. Since realistic conditions corresponding to real observations are used, the imperfect phase-coverage combined with the limited spectral resolution and signal/noise ratio produces elongated spots in the reconstructed images. The phases are calculated using the equatorial rotation period, thus, the features close to the pole are distorted most. At this point, the cross-correlation technique is applied to the two images. The resulting correlation

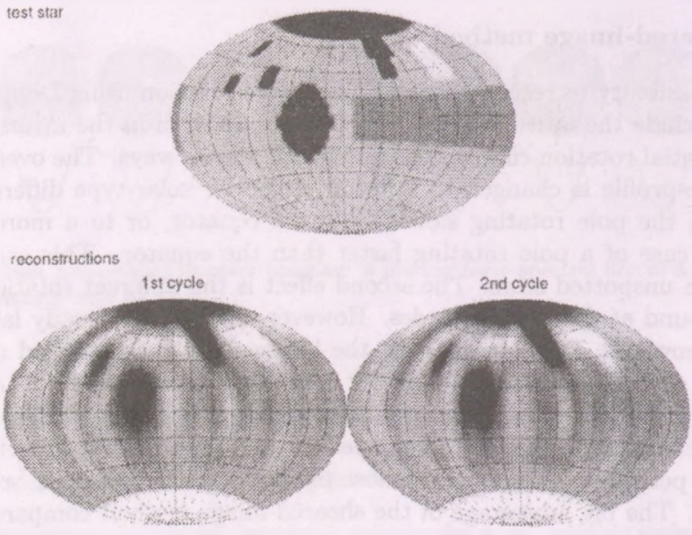

Figure 2: Our test star (**top**) and its reconstructed Doppler images for two consecutive rotations (**bottom**). See the text for further explanation.

pattern is seen in the left side of Fig. 3. The elongations of the recovered spots combined with the north-south mirroring effect of the Doppler imageing technique affect especially the equatorial regions, an thus the equatorial rotation rate is underestimated (see Table 1 for the summary of the differential rotation parameters).

Table 1: Differential rotation parameters derived from the tests

	Ω_{eq} [°/day]	$\Delta\Omega$ [°/day]	α
Original values	*53.537*	*2.677*	*0.05*
Cross-correlation	53.295±0.303	1.865±0.660	0.035±0.012
Sheared-image	53.41±0.31	2.67±1.06	0.05±0.02

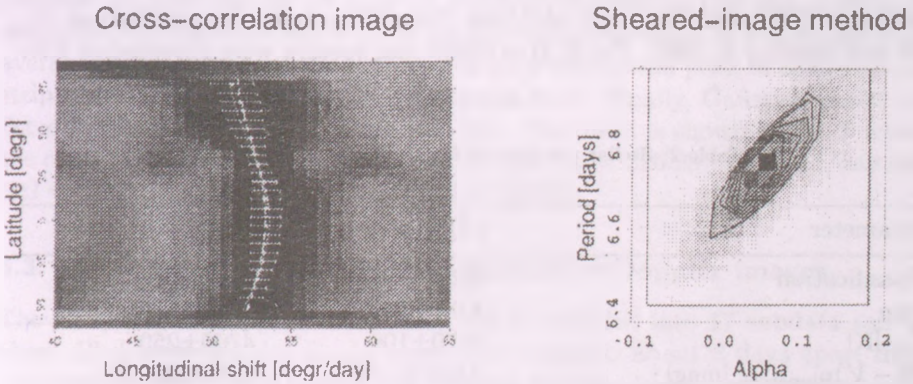

Figure 3: Re-modelling the differential rotation with cross-correlating the test images (**left**) and with using the sheared-image method (**right**). See the text for details and Table 1 for the retrieved parameters.

Then the sheared-image method is applied to the synthetic data. We use the 'combined image', i.e., all the 26 spectra are included in the inversion. The resulting χ^2 map in the $P_{eq} - \alpha$ frame is plotted in the right panel of Fig. 3, where darker pixels mean better fits and the solid lines are contours of constant χ^2. The method performed an almost perfect reconstruction of the original differential rotation law (see also Table 1). This good performance is

in part due to the added restriction of a $\sin^2\theta$ differential rotation law, while in the cross-correlation technique, this restriction is not applied before fitting the points of the individual latitudes to obtain the overall differential rotation law.

5. Application

The methods introduced in Sect. 3. and tested in Sect. 4. are applied now to the young, rapidly-rotating, single K2 dwarf LQ Hydrae (HD 82558) and the evolved component of the RS CVn-type spectroscopic binary IL Hydrae (HD 81410), both stars known from stong cromospheric activity. The corresponding astrophysical data are summarized in Table 2 (Weber & Strassmeier, 2001, Kővári et al., 2004).

5.1. Data

The bulk of our spectroscopic data were collected at the National Solar Observatory (NSO) with the 1.5-m McMath-Pierce telescope between October 31, 1996 and January 8, 1997. For IL Hya additional spectra were obtained at Kitt

Table 2: Stellar parameters for LQ Hya and IL Hya.

Parameter	LQ Hya	IL Hya
Classification	K2V	K0III-IV
$\log g$	4.0±0.5	2.5±0.5
$T_{\rm eff}$ [K]	5070±100	4700±250
$(B-V)_{\rm Hipparcos}$ [mag]	0.933±0.021	1.012±0.010
$(V-I)_{\rm Hipparcos}$ [mag]	1.04±0.02	0.99±0.01
$\rm Distance_{Hipparcos}$ [pc]	18.35±0.35	119.6±12.4
$v \sin i$ [km s^{-1}]	28.0±1.0	26.5±1.0
Inclination, i [°]	65 ±10	55±5
Rotation period, $P_{\rm rot}$ [days]	1.60066±0.00013	12.730±0.004
Radius [$R_\odot$]	0.97±0.07	8.1±0.9
Microturbulence, ξ [km s^{-1}]	0.5	2.0
Macroturbulence, $\zeta_{\rm R} = \zeta_{\rm T}$ [km s^{-1}]	1.5	4.0
Chemical abundances	solar	0.9 dex below solar
Mass [$M_\odot$]	≈0.8	≈2.2

Peak National Observatory (KPNO) with the Coudé feed telescope in December 1997/January 1998. The wavelength range of our spectra includes the mapping lines Fe I-6393, Fe I-6430 and Ca I-6439 used here as input for Doppler imaging.

5.2. LQ Hya: cross-correlation of time-series Doppler images

From the 35 Fe I-6430 spectra taken during the 1996 observing run (covering 35 consecutive stellar rotations), we built 28 data subsets with eight spectra each. The first subset consists of the first eight observations. The next subset is formed by omitting the first spectrum from the beginning and adding the ninth to the end. This pattern is continued until the last eight spectra are included, then Doppler image is reconstructed for each subset. Fig. 4 (available electronically in the on-line version of the paper) shows a movie from the time-series images.

In the next step we cross correlate consecutive and adjacent maps, i.e. we compute a cross-correlation-function map from image #1 and #9, then one from #2 and #10 and so forth, until the end of the time series. This gives 20 ccf maps. Because the time baseline for these ccf maps varied between 10.6 and 17 days, we normalize the longitude shifts to the average time interval of 12.8 days, and then average the 20 ccf maps. We then search for a correlation peak in each latitude strip of the average map and fit a Gaussian to it. Finally, Gaussian peaks are fitted with a $\sin^2\theta$ differential rotation law. The result is shown in Fig. 5 where the continuous line represents the best fit with $\Omega_{\rm eq} = 225.287 \pm 4.011$ °/day and $\Delta\Omega = 1.261 \pm 0.458$ °/day ($\alpha = 0.0056 \pm 0.0022$).

5.3. IL Hya: cross-correlation of time-series Doppler images

The 35 spectra from the NSO (1996) run are divided into 17 subdata sets of about 13 days length ($\approx P_{\rm rot}$), where each subset is about 3 days apart from the previous one. Then, using these subsets, Doppler images are reconstructed for both the Fe I-6430 and the Ca I-6439 lines. The resulting time series of 17 Doppler images are used to create the iron and the calcium movies shown in Fig. 6*a-b* (available on-line only!).

Those Doppler images that are approximately one rotation apart, i.e. #1 and #4, #2 and #5, #3 and #6, etc., are then cross-correlated and the resulting 14 correlation maps are corrected to match exactly $P_{\rm orbit}$. The maps are averaged and the correlation pattern is fitted as described in the preceding subsection. The result shown in Fig. 7 suggests solar-like equatorial acceleration. Fitting the best correlated regions we estimate the differential rotation law as $\Omega(\theta) = 28.44 - 0.76\sin^2\theta$ °/day for the average of both mapping lines.

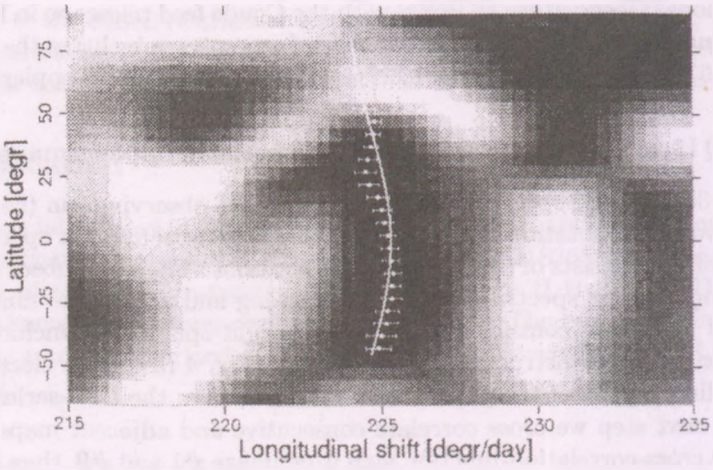

Figure 5: Cross-correlation function (ccf) image for LQ Hya from the NSO 1996 data. The image is the average of 20 individual ccf maps obtained for the time-series Fe I-6430 Doppler images . The best correlated (dark) regions are fitted by a solar-like differential rotation law (solid line). See Table 3 for the parameters.

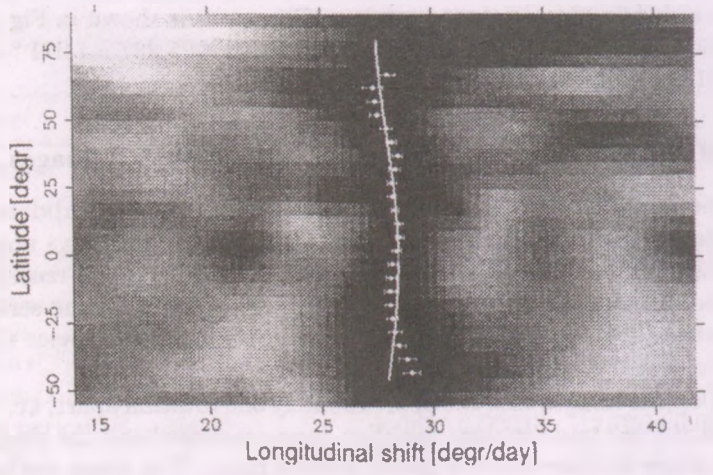

Figure 7: Cross-correlation function (ccf) image for IL Hya from the NSO 1996 data. The image is the average of 14 individual ccf maps obtained for the time-series Fe I-6430 and Ca I-6439 Doppler images. For the parameters of the fitted differential rotation law (solid line) see Table 3.

5.4. IL Hya: application of the sheared-image method

The 13 spectra from the 1997/98 KPNO run span roughly one rotation, thus they are suitable only for reconstructing one single Doppler image. The spectral resolution and the signal-to-noise ratio of this dataset, however, is the highest available to us. We therefore used this to test the line-profile inversion code which involves the latitude-dependent rotation (see Sect. 3.2.). The sheared-image method is applied to the Fe I-6393, the Fe I-6430, and the Ca I-6439 lines. The χ^2-maps of the three mapping lines are then averaged (see Fig 8). The most probable differential rotation value read off the average χ^2-map is $\Omega(\theta) = 28.02 - 0.84 \sin^2 \theta$ °/day.

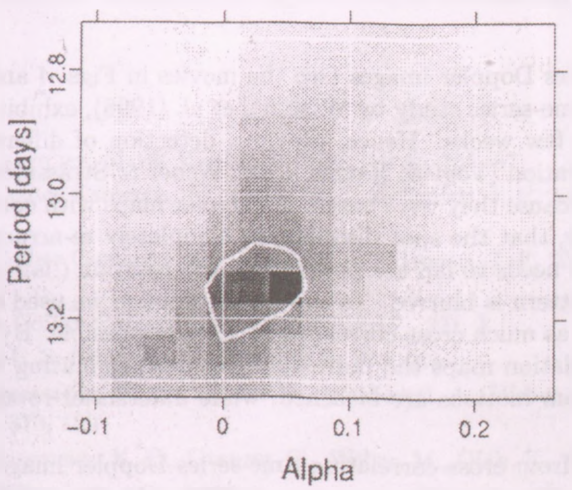

Figure 8: Results from the sheared-image method for the IL Hya 1997/98 (KPNO) data. The average of the χ^2-maps obtained for Fe I-6393, Fe I-6430, and Ca I-6439 lines is plotted. Darker pixels mean better fits, the contour represents constant χ^2. For the parameters of the best-fitting differential rotation see Table 3.

6. Summary and conclusions

We demonstrated the ability of two methods of detecting differential rotation from Doppler images. The summary of our results compared with the solar rotation law are found in Table 3.

Table 3: Comparison of solar-type differential rotation laws obtained using cross-correlation of Doppler images (ccf) and sheared-image method (sim). Lap time (LT) is the time the equator needs to lap the pole by one full rotation.

star	method	line	$\Omega(\theta)$ [°/day]	α	LT [d]
LQ Hya	ccf	Fe I-6430	$225.287 - 1.261 \sin^2\theta$	0.0056	285
IL Hya	ccf	average	$28.44 - 0.76 \sin^2\theta$	0.027	474
IL Hya	sim	average	$28.02 - 0.84 \sin^2\theta$	0.03	429
Sun†			*14.37-2.86 sin²θ*	*0.2*	*126*

† Solar values are derived from tracing long-lived spots by Balthasar et al. (1986).

Our time-series Doppler images (see the movies in Figs. 4 and 6), in agreement with the time-series study by Washüttl et al. (1998), exhibit spot lifetimes of as short as a few weeks. Hence, previous detection of differential rotation from cross-correlation (Vogt & Hatzes, 1996, Weber & Strassmeier, 1998) may be questioned because they were based on Doppler maps with considerable time delay. It is likely, that the spot distribution completely re-arranges within the time the equator needs to lap the pole by one full rotation ('lap time'), and the coherent spot pattern is blurred. To avoid this problem we used our time-series spectra to make as much cross-correlation figures as possible. By averaging the time-series correlation maps the individual features originating from spot evolution and random motions are bleached, while differential rotation pattern is amplified.

For LQ Hya, from cross-correlating time-series Doppler images we find that the star rotates almost rigidly, i.e. $\alpha = 0.0056$. On the other hand, $\Delta\Omega = 1.261$ °/day translates into a lap time of ≈285 days, i.e. the surface shear at a rotation rate roughly 15 times higher than the Sun's, turns out to be effective to produce the observed strong magnetic activity.

The (sub-)giant component of the RS CVn system IL Hya exhibits differential rotation behaviour. We get a small solar-like differential rotation of $\alpha = 0.03$. Due to the relatively slow rotation of this star, this translates to a lap time of 470 days, about four times as on the Sun. On the other hand, this star is significantly more evolved than the Sun. Some luminosity class III stars (thus a little more evolved than IL Hya, since its radius hints towards an evoloutionary stage between IV and III) even show differential rotation in the opposite direction (e.g. HD 31993, Strassmeier et al., 2003, or UZ Lib, Oláh et al, 2003). This

could be due to a larger meridional circulation which could induce differential rotation in the anti-solar sense (Kitchatinov 2003, private communication).

Taking our new detections of differential rotation for LQ Hya and IL Hya, the theoretical model developed by Kitchatinov & Rüdiger (1999) predicting a linear dependence of α with rotational period can be confirmed.

Acknowledgments

ZsK is very grateful to the Hungarian Science Research Program (OTKA) for support under grants OTKA T-038013 and OTKA T-043504. The authors are grateful to the German-Hungarian exchange program for grant D-21/2001. ZsK would like to express his thanks to Collegium Budapest IAS for providing an outstanding milieu throughout his fellowship.

References

Balthasar, H., Vazquez, M., & Woehl, H. 1986, *A&A* , 155, 87

Barnes, J. R., Collier Cameron, A., James, D. J., & Donati, J.-F. 2000, *MNRAS* , 314, 162

Barnes, J. R., Collier Cameron, A., James, D. J., & Donati, J.-F. 2001, *MNRAS* , 324, 231

Donati, J.-F. & Collier Cameron, A. 1997, *MNRAS* , 291, 1

Kitchatinov, L. L. & Rüdiger, G. 1999, *A&A* , 344, 911

Kővári, Zs., Strassmeier, K. G., Bartus, J., Washuettl, A., Weber, M., & Rice, J. B. 2001, *A&A* , 373, 199

Kővári, Zs., Strassmeier, K. G., Granzer, T., Weber, M., Oláh, K., & Rice, J. B. 2004, *A&A* , in press

Kurucz, R. 1993, ATLAS9 Stellar Atmosphere Programs and 2 km/s grid. Kurucz CD-ROM No. 13. Cambridge, Mass.: Smithsonian Astrophysical Observatory, 1993., 13,

Maunder, E. W. & Maunder, A. S. D. 1905, *MNRAS* , 65, 813

Oláh, K., Jurcsik, J., & Strassmeier, K. G. 2003, *A&A* , 410, 685

Piskunov, N. E. & Rice, J. B. 1993, *PASP* , 105, 1415

Rice, J. B. 2002, Astronomische Nachrichten, 323, 220

Rice, J. B. & Strassmeier, K. G. 2000, *A&AS* , 147, 151

Rice, J. B., Wehlau, W. H., & Khokhlova, V. L. 1989, *A&A* , 208, 179

Strassmeier, K. G., Kratzwald, L., & Weber, M. 2003, *A&A* , 408, 1103

Vogt, S. S. & Hatzes, A. P. 1996, IAU Symp. 176: Stellar Surface Structure, 176, 245

Vogt, S. S. & Penrod, G. D. 1983, *PASP* , 95, 565

Washuettl, A., Strassmeier, K. G., & Collier-Cameron, A. 1998, ASP Conf. Ser. 154: Cool Stars, Stellar Systems, and the Sun, 10, 2073

Weber, M. & Strassmeier, K. G. 1998, *A&A* , 330, 1029

Weber, M. & Strassmeier, K. G. 2001, *A&A* , 373, 974

Adventures in the World of Pulsating Variable Stars: Multisite Photometric Campaigns

R. Szabó
Konkoly Observatory of the Hungarian Academy of Sciences
H-1525 Budapest, P.O.Box 67., Hungary
E-mail: rszabo@konkoly.hu

Abstract

Advantages of photometric multisite campaigns are discussed, then published or prospective results of four - preceding and ongoing - observing runs are presented. Prospects and limits of photometric observations carried out from Hungary with small ($\lesssim 1m$) telescopes are also outlined.
Keywords: *variable stars, nonradial pulsation, asteroseismology*

1. Introduction

Asteroseismology nowadays tries to find its way to apply the already existing solar seismology notions and techniques to pulsating variable stars (Fig. 1). The difficulties are numerous. From observational point of view: the low amplitude of the modes ($0.01 - 0.001^m$) and the complex frequency patterns require long and precious observing runs with optimal instrument performance and careful planning. From theoretical side: the essentially low number of observable modes (up to some tenths at maximum) prevents meaningful seismology, mode identification is sometimes questionable (photometric and spectroscopic methods exist) and no effective nonradial mode selection mechanism is known, therefore no firm explanation emerged to account for the relatively small number of observable (excited) eigenmodes. The ultimate goal would be to test the stellar structure and improve our models and physical understanding of the widest possible spectrum of different stars. δ Scuti stars, pulsating B subdwarfs, stars exhibiting solar-like oscillations were the target objects of the presented asteroseismical studies, though the list is far from complete, eventually it can be extended for the whole HRD.

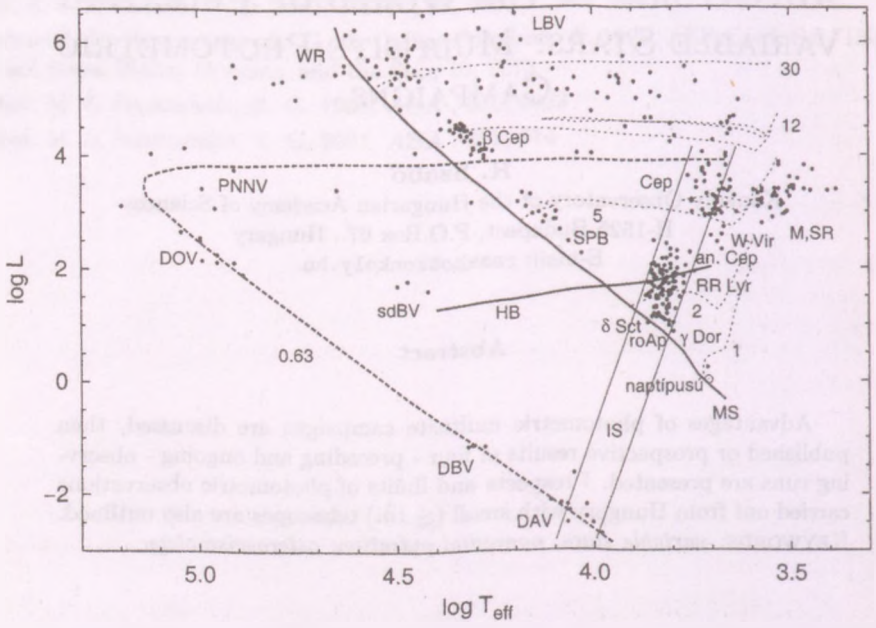

Figure 1: Pulsating variables on the HRD based on Fig. 1. of Gautschy & Saio, 1995. Main Sequence and Horizontal Branch are indicated, as well as sample evolutionary tracks, labeled with the corresponding solar masses. One representative white dwarf cooling sequence is shown ($M = 0.63 M_{\odot}$). The classical instability strip is delineated by parallel thin lines. As only a limited number of individual objects is drawn, the relative density of the dots is not significant.

2. Motivations and Requirements

What are the requirements for the observers to achieve the scientific objectives outlined in the previous section? One needs good data sampling, i.e. good photometric coverage by minimizing gaps to avoid spurious peaks (aliases) in the Fourier-spectrum. Excellent frequency resolution is a must to resolve and determine the frequencies of the multimode pulsators. This invokes long, uninterrupted observations. To attack the outlined problems there are two ways: space-based instruments and multisite campaigns. We only deal with latter here. Fig. 2. shows the striking improvement of the frequency resolution and peak structure for a short and a longer single-site, as well as a continuous multisite

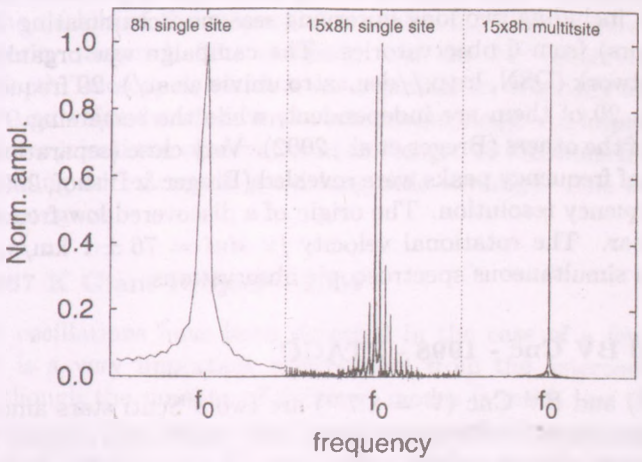

Figure 2: Differences between the structure of the frequency peak of a single noise-free sine wave sign from an 8-hour long, single site observation; from a 15-day single-site observing run (consecutive nights), and from a 15-day multisite (continuous) observing run.

observation of the same length. Further requirements can be: good signal-to-noise ratio, sometimes good time resolution (the bigger telescope the better), use of reliable instruments and infrastructure, proper coordination, much telescope time (proposal writing, lobbying) and finally cooperation of the weather and good luck. It is important to homogenize data from different sites and instruments, but without exterminating real variations!

3. Multisite Campaigns

Four multisite campaign are described here briefly; the author participated in these ones. The Hungarian contribution (to these and many other campaigns) was made from Piszkéstető Mountain Station (Konkoly Observatory), with the 60/90/180cm Schmidt and the 1m RCC telescopes, equipped with various CCD-cameras.

3.1. BI CMi - 1997-2000 - DSN

BI CMi is a bright ($V = 9.2^m$) δ Scuti star. The assumption, that this star may show both δ Scuti and γ Doradus type pulsation, has led to a resurgence of

interest in organizing a dedicated observing campaign. It was realized between 1997 and 2000, including two long observing seasons, accumulating 177 nights' data (1024 hours) from 6 observatories. The campaign was organized by the Delta Scuti Network (DSN, http://dsn.astro.univie.ac.at/). 29 frequency peaks were identified, 20 of them are independent, while the remaining 9 are linear combinations of the others (Breger et al., 2002). Very close (separation less then 0.7 μHz) pairs of frequency peaks were revealed (Breger & Bishof, 2002), thanks to the high frequency resolution. The origin of a discovered low-frequency peak remained unclear. The rotational velocity (v $sini$ = 76 $\pm$ 1 km/s) could be derived, from a simultaneous spectroscopic observations.

3.2. BN and BV Cnc - 1998 - STACC

BN ($V = 7.8^m$) and BV Cnc ($V = 8.7^m$) are two δ Scuti stars among the 14 known such variables in the Preasepe (M44, NGC 2632) open cluster. This arsenal of specific types of stars provides the enormous potential to get a step closer to the understanding of the mode selection, as these stars presumably share the same age, distance and chemical composition. These constraints reduce significantly the parameter space of the theoretical models. With 277 hours of observation from 10 different locations, 6 and 4 frequencies could be identified above the detection limit (0.5 mmag) in the case of BN and BV Cnc, respectively (Frandsen et al., 2001). It was shown, by comparing the dataset with earlier observations, that substantial amplitude variations occurred between 1997 and 1998 in the frequency spectrum of the very low-amplitude pulsator BV Cnc. The photometric campaign was initiated by the STACC (Small Telescope Arrays with CCD Cameras), and it was accompanied by a spectroscopic monitoring campaign.

3.3. KPD 1930+2752 - 2003 - WET Xcov23

KPD 1930+2752 ($V = 13.8^m$) is a short period ($\sim$ 2h17m) close binary system, as revealed by the presence of an ellipsoidal variation in the lightcurve (Brassard et al., 2001). The companion is almost certainly a white dwarf as there are no indications of eclipse or reflection effect. Radial velocity measurements indicate that the total mass of the system exceeds the Chandrasekhar limit, making KPD 1930+2752 the first known (non-X-ray source) candidate for a Type Ia supernova progenitor (Maxted et al., 2000). The pair should merge within an astrophysically interesting timescale ($\sim$ 200 Myr) due to gravitational wave radiation. The goal of the international collaboration was to fully resolve the

complex pulsation pattern of the primary component - a pulsating sdB star, as well as the details of the ellipsoidal modulation. (There were two additional, lower priority targets.) The campaign was the 23^{rd} observing run of WET (Whole Earth Telescope), which is an organization originally established to observe and use white dwarfs to make asteroseismology. An impressive number of observatories joined the observational challenge: 18 locations from 5 continents, thus providing an extremely good longitude coverage. The evaluation of the data is in progress.

3.4. M67 K Giant Project - 2004

Solar-like oscillations have been detected in the case of a few main-sequence stars. It is a very important achievement from the asteroseismical point of view. Although the number of detected modes is much less than in the solar case, the global properties of the frequency spectrum can already be compared to theoretical predictions. To test various stellar models, it would be crucial to find solar like oscillations in evolved stars, too. However, up to the present, only one such object (ξ Hya) has been discovered (Frandsen et al., 2002). The ultimate goal of the M67 K Giant campaign is to reveal solar like oscillations in a multitude of evolved object (possibly of similar age, composition, distance etc.). The ideal selection was M67 (NGC 2682), a famous open cluster with a well-populated giant branch. The ambitious aim of the ongoing project is to supersede the previous results, as 4m class telescopes have not succeeded in detecting unambiguously solar like oscillations in a limited number of M67 giants, providing only lower amplitude limits (Gilliland et al., 1993).

4. Conclusions

The decline of small telescopes (1-4m) is a world-wide tendency. In addition, observational asteroseismology is being transposed to satellite missions (WIRE, MOST, COROT). Despite these circumstances, observations with small telescopes do have their own, indispensable role. E.g. supporting and complementing space missions: ground-based initiative and follow-up observations, target selection, spectroscopy are pivotal and based entirely on small telescopes. What can Hungary add to this research? As previous sections have hopefully demonstrated, we provide knowledge, know-how and expertise, longitude and time coverage, availability and reliability. The main conclusion of this contribution is, that it is possible to carry out state-of-the art, competitive observing projects with current telescopes, instruments and infrastructures at the beginning of the

21^{st} century. Geographic and economical circumstances of Hungary also force this kind of cooperation and collaboration with international organizations and observing campaigns ...

Acknowledgments

The efforts and the hard work of many colleagues who organized and added to the success of the campaigns, are acknowledged. Because of the long list I recite here only the Hungarian astronomers who contributed by observing, organizing, camera software development or any other way: Sz. Barcza, J. Benkő, Z. Csubry, L. Kiss, J. Nuspl, M. Paparó and D. Zsuffa.

References

Brassard, P., Fontaine, G., Billeres, M. et al. 2001, *ApJ* , 563, 1013
Breger, M., Garrido, R., Handler G. et al. 2002, *MNRAS* , 329, 531
Breger, M., Bishof, K. M., 2002, *A&A* , 385, 537
Frandsen, S., Carrier, F., Aerts, C. et al. 2002, *A&A* , 394, 5
Frandsen, S., Pigulski, A., Nuspl, J. et al. 2001, *A&A* , 376, 175
Gautschy, A., Saio, H, 1995, *A&A Rev.* , 33, 75
Gilliland, R. L., Brown, T. M., Kjeldsen, H. et al. 1993, *AJ* , 106, 2441
Maxted, P. F. L., Marsh, T. R., North, R. C., 2000, *MNRAS* , 317, L41
Nather, R. E., Winget, D. E., Clemens, J. C. et al. 1990, *ApJ* , 361, 309

Fermion Stars in $3+1$ vs. $4+1$ Dimensional Space-Time

G.G. Barnaföldi[1,2], P. Lévai[1] and B. Lukács[1]
[1] Research Institute for Particle and Nuclear Physics
H-1525 Budapest, P.O.Box 49, Hungary
[2] Laboratory for Informational Technology, Eötvös University
H-1117 Budapest, Pázmány P. sétány 1/A.
E-mail: [1]bgergely@rmki.kfki.hu, plevai@rmki.kfki.hu, lukacs@rmki.kfki.hu

Abstract

We introduce the description of a fermion star in $3+1_c+1$ dimensional space-time, where 1_c is a 1-dimensional compactified spatial dimension with size $\sim 10^{-13}$ cm. This compactification size might be interpreted as the strangeness quantum number, but in a geometrical way. A neutron moving into the extra dimension would behave as a Λ_s or a barion with even higher strangeness content. We display the Einstein equations for a static, spherically symmetric fermion star and discuss a special solution, which leads to the standard Tolman – Oppenheimer – Volkov equation and an extra condition for the extra dimensional contribution.
KEYWORDS: *compact star, neutron star, compactified extra dimensions*

Introduction

There are many astrophysical objects in the surrounding Universe, which have extreme physical conditions inside, e.g. millisecond pulsars. Here strong, electroweak and gravitational interactions interplay with each other. It is a very interesting challenge to describe the properties of these extreme compact objects. In our approach we introduce microscopical extra dimensions (see Ref. Randall and Sudrum, 1999) and use the frame of a Kaluza – Klein (K-K) like model, where the excited states of particles appear as geometrical degrees of freedom, see Refs. Kan and Shiraishi, 2002 and Barnaföldi et al., 2003.

In this short contribution we present a few of our ideas about these space-like extra dimensions, their possible connection to particle physics and their appearance in the core of fermion stars. We display some of our numerical results and discuss the observability of extra dimensions in these objects.

1. Motivation for the 5th Dimension

Although the symmetries of the observable physical world suggest a space-time in $3+1$ macroscopical dimensions, we can not exclude the existence of compactified extra dimensions at microscopical scales. Here we follow the idea described in Ref. Barnaföldi et al., 2003 and introduce a $3+1_c+1$ dimensional space-time, including one compactified dimension.

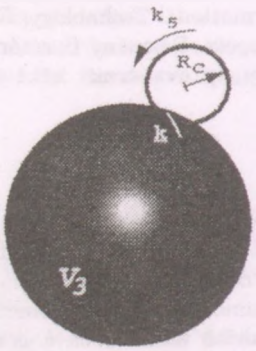

Figure 1: The $3+1_c+1$ dimensional space-time, where the space has a $V_3 \times S^1$ structure with a compactified radius R_c of the S^1 circle. The momentum for moving in the 5th dimension is k_5.

Particles can move freely along the extra x^5 direction, however we require a periodic boundary condition, which results a Bohr-type quantization condition for k_5 momentum component. This condition induces an uncertainty in the position with the size of $2\pi R_c$, where R_c is the compactification radius. Motion into the 5th dimension generates an extra mass term appears in pure 4D description:

$$k_5 = \frac{n\,\hbar}{R_c} \quad \longrightarrow \quad \widehat{m} = \sqrt{m^2 + \left(\frac{n\,\hbar}{R_c}\right)^2}\,. \tag{1}$$

Considering compactified radius $R_c \sim 10^{-12} - 10^{-13}$ cm this extra "mass" is $\widehat{m} \sim 100$ MeV (see Ref. Arkhipov, 2003 on hadron spectroscopy).

One can introduce a "pseudo charge" generated by k_5 with its $\pm$ sign:

$$\widehat{q} = n \cdot \frac{2\hbar\sqrt{G}}{c\,R_c}\,, \tag{2}$$

which acts in a vector-scalar interaction. We can directly see that $\widehat{q}$ is *not* the electric charge. Indeed $\widehat{q}^2 < 16\pi G m_0^2$, where G is the gravitational constant

and m_0 is the rest mass (Lukács and Pacher, 1985). This is either a familiar quantum number (e.g. strangeness, S), or some other not yet observed charge.

2. Field Equations in 5D

Our goal is to describe the inner structure of a *spherical* symmetric object in $3 + 1_c + 1$ D space-time. These object are *static* and are in equilibrium, at close to *zero temperature*. We are using a perfect fluid description with *isotropy* in the standard 3 spatial dimensions. *Killing-symmetries* are still exist in 5D space-time, but an *anisotropy* needs to be chosen for the fluid in the microscopical dimension. These symmetries yields to the metric, see Ref. Notations,

$$\mathrm{d}s^2 = e^{2\nu}\,\mathrm{d}t^2 - e^{2\lambda}\,\mathrm{d}r^2 - r^2\mathrm{d}\Omega^2 - e^{2\Phi}(\mathrm{d}x^5)^2, \tag{3}$$

where the quantities ν , λ and Φ depend only on r. $\mathrm{d}\Omega^2$ is the usual spherical elementary surface. The 5D Einstein equation is written for the energy-momentum tensor with ε energy density, P and P_5 pressures ($P \neq P_5$):

$$T^{ik} = \mathrm{diag}\,(\varepsilon\, e^{2\nu},\ P\, e^{2\lambda},\ P\, r^2,\ P\, r^2 \sin^2\theta,\ P_5\, e^{2\Phi}) \tag{4}$$

$$-\gamma\, T_{ik} = R_{ik} - R^l{}_l\, g_{ik}/2\ . \tag{5}$$

Substituting (3) and (4) into the 5D Einstein equation is read as:

$$-\gamma\,\varepsilon = e^{-2\lambda}\left[\Phi'' + \Phi'^2 - \lambda'\Phi' + \frac{2\Phi'}{r} - \frac{2\lambda'}{r} + \frac{1}{r^2}\right] - \frac{1}{r^2}, \tag{6}$$

$$-\gamma\, P = e^{-2\lambda}\left[-\nu'\Phi' - \frac{2\Phi'}{r} - \frac{2\nu'}{r} - \frac{1}{r^2}\right] + \frac{1}{r^2}, \tag{7}$$

$$-\gamma\, P = e^{-2\lambda}\left[-\nu'' - \nu'^2 + \nu'\lambda' - \Phi'' - \Phi'^2 - \nu'\Phi' + \lambda'\Phi' - \right.$$
$$\left. - \frac{2\Phi'}{r} - \frac{\nu'}{r} + \frac{2\lambda'}{r}\right], \tag{8}$$

$$-\gamma\, P_5 = e^{-2\lambda}\left[-\nu'' - \nu'^2 + \nu'\lambda' - \frac{2\nu'}{r} + \frac{2\lambda'}{r} - \frac{1}{r^2}\right] + \frac{1}{r^2}, \tag{9}$$

where $\gamma = 8\pi G/c^4$ and G is the gravitational constant.

Assuming cold nuclear matter inside the fermion star ($T = 0$ approximation), all local material characteristics of the fluid depend on one thermodynamic quantity, e.g. particle density ρ. For the local matter we can write

$$\varepsilon = \varepsilon(\rho); \quad P = P(\rho); \quad P_5 = P_5(\rho)\ . \tag{10}$$

Furthermore, one can consider the appropriate Bianchi identity and obtain

$$T^{ir}_{\ ;r} = 0 \qquad \longrightarrow \qquad P' = -\nu'(\varepsilon + P) + (P_5 - P)\Phi' \ . \tag{11}$$

This equation clearly demonstrates the influence of the extra dimensional behavior on the normal isotropic pressure, $P = P_1 = P_2 = P_3$.

In the 5D Einstein equations (6-9) two extra variables appeared compared to the 4D formalism, namely P_5 and Φ. However, $P_5(\rho)$ is a known function of the particle number density and specified by the actual interaction in the matter. Thus $\Phi(r)$ is the *only new degree of freedom* determined by eq. (9). Details can be found in Ref. Lukács et al., 2003.

3. A Special Solution of the Field Equations in 5D

We can recognize the existence of similar solutions both in 5D and in 4D. For specially chosen pressure P_5 there is a unique solution of the Einstein equations in eqs. (6-9), namely $\mathrm{d}\Phi/\mathrm{d}r = 0$. In this case eqs. (6-8) lead to the standard Tolman – Oppenheimer – Volkov equation (see Ref. Glendenning, 1997) and can be solved separately with $\Phi = \text{const.}$ condition. The eq. (9) gives P_5. Although these solutions do not differ formally from the 4D neutron star solution (except for P_5), but the extra dimension will have its influence on $\varepsilon(n)$ and $P(\rho)$, as shown also by Kan and Shiraishi, 2002. For $\Phi \neq \text{const.}$ one must turn to numerics and solve directly eqs. (6-9).

Neglecting the effects of electromagnetic charge, let us start with a neutral, single massive fermion ("neutron", N) as elementary building block of a compact star. Since the minimal nonzero fifth momentum component is given by Bohr-type quantization (1), then the extra direction of the phase space is not populated until the Fermi-momentum $k_F < \hbar/R_c$. However, at the threshold both $k_5 = \pm\hbar/R_c$ states appear. One can represent this as another ("excited") particle with mass $m_X = \sqrt{m_N^2 + (\hbar/R_c)^2}$ (with a nonelectric "charge" $\widehat{q} = \pm\frac{2\hbar\sqrt{G}}{cR_c}$ as well). The equations obtain a form as if this second particle appears in complete chemical equilibrium with the neutron: $\mu_X = \mu_N$.

We note, assuming higher excitations, the above mentioned phenomena is repeated in any case, when k exceeds a threshold $n\hbar/R_c$, with integer n.

We apply our model for hybrid stars with strangeness content and display on Fig. 2 the mass and radius dependence on the central energy density (ε_{cent}) and the $M(R)$ relation. We consider hyperon stars in 4D with N and Λ_s content as reference object (solid lines). In parallel we introduce 1_c dimension with $R_c = 0.33$ fm, when Λ_s can be interpreted as the first excited state of N.

The higher excitations are named as $E_1(\equiv \Lambda_s)$, E_2, E_3, ... (dashed lines). Varying the compactification radius to $R_c = 0.66$ fm, we obtain the excitations F_1, F_2, F_3, ... (dotted lines).

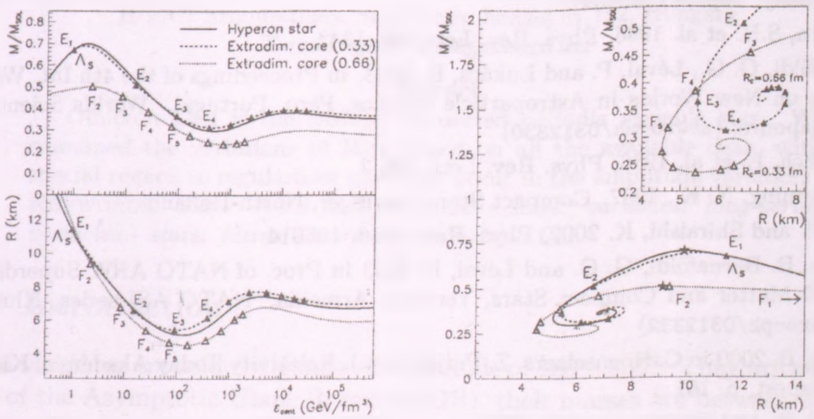

Figure 2: $M(\varepsilon_{cent}), R(\varepsilon_{cent})$ and $M(R)$ diagrams for 4D and 5D hybrid stars.

4. Discussion

Introducing extra dimensions into the description of hybrid stars led to similar results in $3+1$ and $3+1_c+1$ space-time, if the compactification radius is at fermi scale. In this case extra dimensional excitations can be interpreted similarly to the strangeness appearance in the hadron spectroscopy. We are aware of the complexity of this question, and do not want to claim the extra dimensional origin of the strangeness. However such an astrophysical object as a hybrid star serves as "natural laboratory" to investigate these problems (see Ref. Weber, 1999) or other exotic ones, e.g. the flavour dependence of gravity (Fishbach et al., 1986, Aronson et al., 1986).

Acknowledgments

The speaker (G.G.B.) would like to thank for the invitation to this workshop. This work was supported by OTKA T043455 grant.

References

Ambartsumyan, V. A. and Saakyan, G. S. 1960, Astron. Zh. **37**, 193

Arkhipov, A. A. 2003, in Proceedings of HADRON '03 Conference, Aschaffenburg, Germany, hep-ph/0309327

Aronson, S.H. et al. 1986, Phys. Rev. Lett. **56**, 1342

Barnaföldi, G. G., Lévai, P. and Lukács, B. 2003, in Proceedings of the 4th Int. Workshop on New Worlds in Astroparticle Physics, Faro, Portugal, Worlds Scientific, Singapoure; (astro-ph/0312330)

Fischbach, E. et al. 1986, Phys. Rev. Lett. **56**, 3

Glendenning, N. K., 1997, Compact Stars, Springer, North-Holland

Kan, N. and Shiraishi, K. 2002, Phys. Rev. **D66**, 105014

Lukács, B. Barnaföldi, G. G. and Lévai, P. 2003 in Proc. of NATO ARW Superdense QCD Matter and Compact Stars, Yerevan, Armenia, NATO ASI series, Kluwer; (astro-ph/0312332)

Lukács, B. 2000 in C. Hoenselaers, Z. Perjés (eds.), Relativity Today Akadémiai Kiadó, Budapest, p. 161

Lukács, B. and Pacher, T. 1985, Phys. Lett. **A113**, 200

Notations: we are using coordinates x^i with latin indices ($i = 0, 1, 2, 3$ and 5) for $3 + 1_c + 1$ D space-time with the Einstein convention.

Randall, L. and Sudrum, R. 1999, Phys. Rev. Lett. **83**, 3370, *ibid.* 4690

Weber, F. 1999, Astrophysical Laboratories for Nuclear and Particle Physics IOP Publishing, Bristol

Visual light variation of Mira Ceti

Gábor Marschalkó
Eötvös University, Gothard Astrophysical Observatory
H-9707 Szombathely, Szent Imre herceg út 112, Hungary
E-mail: gmarscha@gothard.hu

Abstract

Omicron Ceti is the earliest discovered periodic variable star. We examined the variations of Mira based on all the available data, with special regard to regularities that may occur in the amplitude variations.
KEYWORDS: *stars: AGB and post-AGB - stars: variables: long-period variables - stars: Miras - stars: individuals: o Cet*

1. Introduction

Mira variables are cold red giants with periods of $100 - 500^{\mathrm{d}}$. They are near the top of the Asymptotic Giant Branch (AGB), their masses are between 0.5 and 5 $M_{\odot}$. The amplitude of the visual light variation is huge, exceeding 2.5 mag, though sometimes it can be as large as 6 mag.

The atmosphere of Mira variables is convective; its outer layers consist of rising and descending clouds which are independent of the pulsation (Murdin, 2001). We know from 11-micron interferometric observations that the circumstellar dust is also concentrated in clouds (Lopez et al., 1997).

Though *o* Cet is the oldest known variable star, it is still not clear whether it pulsates in its fundamental mode or in overtone (Barthes & Tuchmann, 1994). The first possibility is supported by the shape of the observed radial velocity curves (Willson & Hill, 1979, Hill & Willson, 1979, Wood, 1990). On the other hand the hydrodynamic models seem to indicate overtone pulsation (Wood, 1974, Tuchmann et al., 1979, Perl & Tuchmann, 1990).

Mira's period is $331\overset{\mathrm{d}}{.}96$, the ascending branch of the light curve is steeper; it reaches the light maximum in approximately 110 days. The visual light amplitude is in the range 2.0-10.1 mag, the spectral type varies between M5e and M9e. Its distance is $\sim$ 420 light-years, estimated absolute magnitude is $1\overset{\mathrm{m}}{.}0$. (Murdin, 2001) Already in 1920 Joy (Joy, 1922; Joy, 1926) supposed that Mira has a thirteenth magnitude secondary component. This star, named VZ Ceti, was first optically identified by Aitken (1923) with the 36-inches refractor of Lick Observatory in 1923.

Here we investigate the light and amplitude variations of the star based on the longest possible time base.

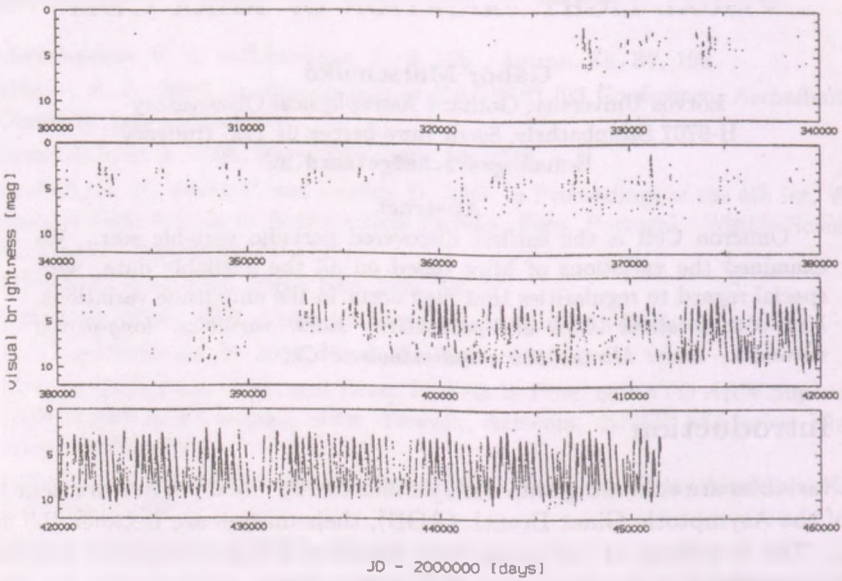

Figure 1: The complete light curve of Mira between 1596 and 2000.

1.1. Data and results

The data were gathered from the literature (Müller & Hartwig, 1918; Prager, 1934; Schneller, 1957); those after 1994 were taken from the AFOEV (Association Francaise des Observateurs d'Etoiles Variables) homepage[1]. (See Fig. 1) Because these were visual observations 5 and 10-day averages were used to minimize the errors. I have a continuous light curve since 1850, these were used for the study of amplitude changes. I used the Mufran package (Kolláth, 1990) which is based on the Deeming method (Deeming, 1974). Fig 2. shows the power spectrum and spectral window of Mira. The main frequency and some of its harmonics are identifiable. The main frequency is at 0.0030123 c/d ($331^d.98$). Whitening with this frequency did not clear the spectrum (right side of Fig.2.). This might indicate amplitude and/or frequency variability; see Fig. 3. To illustrate the change of period I divided the data into fifty-years long intervals and analysed them separately. The resulting periods since 1650: 330.6; 332.2;

[1] http://cdsweb.u-strasbg.fr/afoev

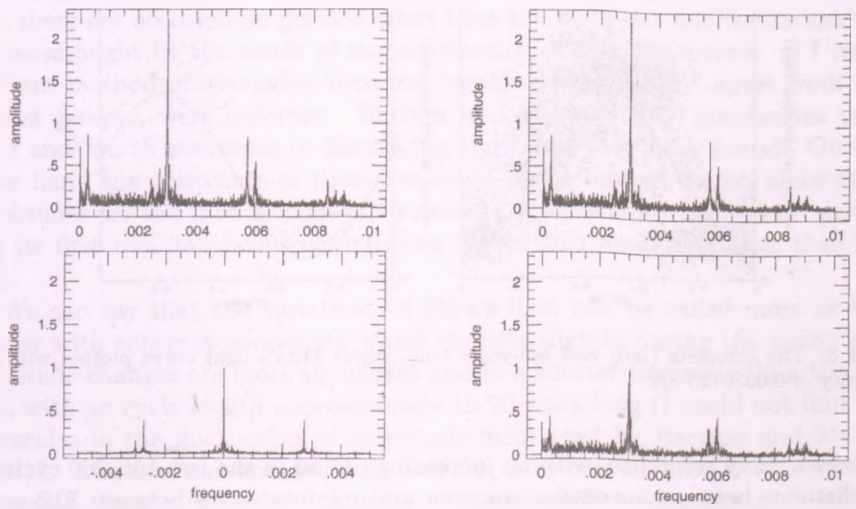

Figure 2: *Left:* The power spectrum and the spectral window of Mira. *Right:* The spectrum before the whitening with the main frequency and after it.

332.6; 330.8; 331.5; 331.6; 334.0 days (the error is $\sim 0.6^{\mathrm{d}}$).

The power spectrum shows several peaks. The main frequency and two of its harmonics can be regarded as certain. This is clear for the spectrum and it agrees with the usually accepted period of the star. Barthes and Mattei (1997) found a frequency at 0.000665 c/d which they considered to be the frequency of the fundamental mode. I could not find this frequency with any reasonable significance. The secondary peaks visible near the main frequencies are probably due to amplitude and/or period variability.

To study the possible variability of the amplitude I divided the light curve between 1850 and 2000 into overlapping 2, 4, 6, 8 and 10-years long intervals (e.g. 1850-1860, 1855-1865, 1860-1870 etc.). The possible amplitude variation was studied by fitting the light curve with the main period in each section. The resulting amplitudes did not, however, show any reasonable variability. A periodicity at 6250^{d} is suspected when using the 4, 6 and 8-years long sections (and there is a slight possibility for its existence in the 10 years case).

I redetermined the maxima and minima of Mira. The O-C diagrams (Fig.

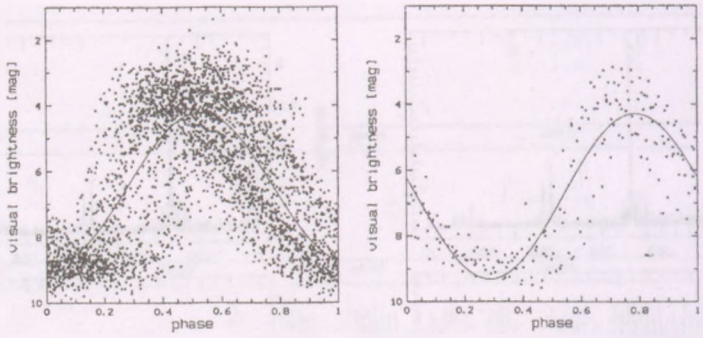

Figure 3: The complete (*left*) and ten-years long (*right*) Mira's light curve plotted with a frequency of 0.0030123 c/d.

4.) show a wavy structure, with an increasing period in the last 150-200 cycles. The distance between successive maxima and minima varies between 312 and 357 days, and 309 and 360 days, respectively. From minima to maxima it takes 97-148 days and from maxima to minima 191-233 days.

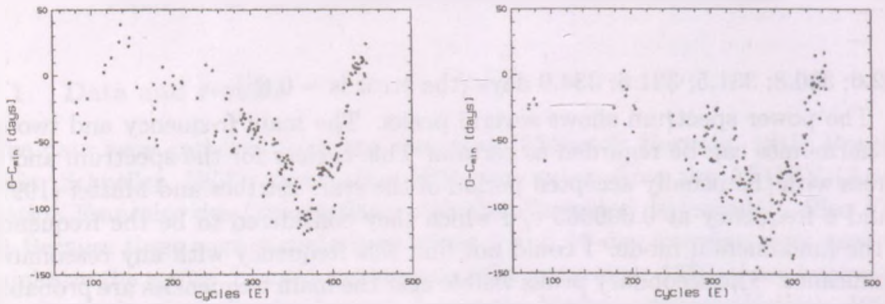

Figure 4: Mira's O-C diagrams computed with the maxima (*left*) and the minima (*right*). $\mathrm{JD_{max}} = 2304089.8 + 331.97564E$ and $\mathrm{JD_{min}} = 2303974.8 + 331.97564E$.

1.2. Discussion

The period analysis of the 400 years long light curve of Mira Ceti yielded only one period ($331\overset{\mathrm{d}}{.}98$) and its harmonics. The secondary structure around these peaks are quite possibly due to amplitude and period changes.

Comparing it with the results of Barthes and Mattei (1997) it is interesting that there are no common periods apart from the main one and its harmonics. Its cause might be the result of the distribution of the data points. If I use a different method of averaging then the resulting frequencies - apart from the highest peaks - were different. Barthes and Mattei (1997) themselves used third and fourth overtones in identifying the peaks they have found. On the other hand the existences of these overtones is not beyond doubt, since their amplitudes are less then 0.05 mag. Therefore I accepted only the $\sim 332^{\mathrm{d}}$ period and its first two harmonics disregarding peaks with amplitudes less then 0.1 mag.

We can say that the variations of Mira's light can be called more or less regular with concrete periodicity which changes slightly during the time. The amplitude changes are more significant and in a shorter timescale than the period, with an cycle length approximately 10-20 years long (I could not find the timescales in the fluctuation of amplitude mentioned by Barthes and Mattei (1997)). Extreme maxima occur in every 2, 3 or 5 cycles which are difficult to explain. In any case, the reason for irregular amplitude variations (and frequency) may be the strongly turbulent and convective atmosphere of Miras, the shock waves in the atmosphere; the clouds in the dust shells, the distribution and movement of which effect variations in the star's light; and also coupling processes on the surface of the star which are caused e.g. by the strong stellar wind. There are some theories that these variations may be explained dynamically by chaotic pulsation. (Barthes and Mattei, 1997; Barthes & Tuchmann, 1994; Murdin, 2001)

Acknowledgments

I am very grateful to E. Zsoldos (Konkoly Observatory, Budapest) who drew my attention to Mira Ceti and came to my assistance during my work, and to J. Kovács (Gothard Astrophysical Observatory, Szombathely) for help in preparing the figures.

References

Aitken, R.G. 1923, *PASP* , 35, 323
Barthes, D., Mattei, J.A. 1997, *AJ* , 113, 373
Barthes, D., Tuchmann, Y. 1994, *A&A* , 289, 429
Deeming, T.J. 1975, *Ap&SS* , 36, 137
Hill, S.J., Willson, L.A. 1979, *ApJ* , 229, 1029

Joy, A.H. 1922, Publications of the American Astronomical Society, twenty-ninth meeting, p.14

Joy, A.H. 1926, *ApJ* , 63, 281

Kolláth, Z. 1990, Konkoly Observatory Occasional Techn. Notes, No. 1

Lopez, B., Danchi. W.C., Bester, M., et al. 1997, *ApJ* , 488, 807

Murdin, P. (ed.), 2001, Encyclopaedia of Astronomy and Astrophysics, Nature Publishing Group, London, New York, Tokyo

Müller, G., Hartwig, E. 1918, Geschichte und Literatur des Lichtwechsels der bis Ende 1915 als sicher veränderlichen anerkannten Sterne nebst eine Katalog der Elemente ihres Lichtwechsels, Leipzig, Erster Band p.57

Perl, M., Tuchmann, Y. 1990, *ApJ* , 360, 554

Prager, R., 1934, Geschichte und Literatur der Lichtwechsels der veränderliche Sterne, Berlin, Zweite Ausgabe, Erster Band, p.286

Schneller, H., 1957, Geschichte und Literatur der Lichtwechsels der veränderliche Sterne, Berlin, Zweite Ausgabe, Vierter Band, p.190

Tuchmann, Y., Sack, N., Barkat, Z. 1979, *ApJ* , 234, 217

Willson, L.A., Hill, S.J. 1979, *ApJ* , 228, 854

Wood, P.R., 1974, *ApJ* , 190, 609

Wood, P.R. 1990, in Mennessier, M.O., Omont, A. (eds.), From Miras to Planetary Nebulae, Frontieres Ed., Gif-sur-Yvette

Old questions reassessed by modern cosmology and astrobiology

Zs. Hetesi and Zs. Regály
Eötvös University, Department of Astronomy
H-1518 Budapest, P.O.Box 32., Hungary
E-mail: [1]Zs.Hetesi@astro.elte.hu, [2]rezso@general.elte.hu

Abstract

We critically investigate some aspects of modern cosmology and astrobiology, which are a central place for contemporary research. Three important fields are investigated: the quantum cosmology, the concept of anthropic principle and the SETI[1] within the context of the combined methods of physics and philosophy.
Keywords: *cosmology: theory, quantum theory, SETI, philosophy of science*

1. Introduction

During the last centuries of human history, there were many questions in connection with the great problems of the existence and origin of human beings, and also of the Universe. Since the emergence and progress of modern natural science mainly physics, these puzzles appear to be solved. But are they really solved? We do not think so. The old questions of common sense and philosophy have not been solved in spite of the indisputable results of modern natural sciences. These questions remain to be unsolved both in physics and also in philosophy.

In this paper we investigate some important results of the modern cosmology and astrobiology and investigate their consequences with respect to the corresponding questions of philosophy and logic. There are three baselines in this paper. The first one touches the branch of questions that are connected with the quantum cosmology, and the multiple-world hypothesis[2]. First, we investigate those formulae of quantum theory, which are supposed to be the ground for the theory of "quantum-creation", and the many worlds-hypothesis. We are going to criticize these theories, and expound our own view.

[1]Search for Extraterrestrial Intelligence
[2]Often called ensemble of universes.

In the second place we are going to examine the question of the anthropic principle, which is likely to be one of the most puzzling theories in the contemporary astronomy and cosmology.

Finally we are throwing a glance at the possible future of astronomy when we investigate the Drake's equation and limit its real implication.

Although our method is based partly on philosophy and logic, it is mainly grounded in the results and methods of the natural sciences. So we need both physics and philosophy to go in our way.

2. Is there quantum-creation? The critique of the modern interpretation of quantum cosmology

It is indisputable, that there has been many successes in the history of the quantum theory, since it was constructed in the early decades of the 20th century. Without claiming full description, we can mention the results of spectroscopy, the explanation of the periodic table, lately the discoveries of new elementary particles, their interaction, and their systematization. In spite of these good results, there are very serious problems with the philosophical meaning of some concepts and phenomena of the theory. These questions generated many debates without much hope of solutions.

There were two ways of presentation in the early quantum theory, which were different in their mathematical methods, but their results were totally equivalent[3]. The first was Schrödingers wave mechanics, and the second was the matrix mechanics constructed by Heisenberg. If we compose the difference of products of two matrices in the second theory, $pq - qp$ i.e. the so called commutator, the result usually will not be zero in accordance with the theory of matrices. As we can find in every textbooks of quantum theory, one can calculate from this commutator the so-called uncertainty relation:

$$\Delta p \cdot \Delta q \geq \frac{\hbar}{2} \tag{1}$$

where p and q are matrices of canonically conjugated physical quantities, and Δp and Δq are the uncertainties referring to the measurements of p and q respectively. This equation shows nothing more that there is an a priori unavoidable uncertainty in the measurement of two physical quantities which are conjugated canonically. From the equation (1) has emerged the most disputable

[3]It was first shown by John von Neumann

result of quantum field theory: the particle creation from the empty space - the so-called quantum vacuum. What is the logic of this indisputedly surprising conclusion?

In 1927, Heisenberg wrote his famous article in the Zeitschrift für Physik, in which he published the above mentioned uncertainty relation, and gave the philosophical interpretation of it. According to this, one cannot achieve perfectly accurate measurements. Heisenberg concluded that the strong causality has finally ended in physics. Since then there is a widespread agreement among theoretical physicists, that the existence of a physical process depends on its precise measurability. Nevertheless this conclusion is false and contains a basic logical error. In spite of the perspicuity of this error, only a little number of scientists noticed it (Jaki, 1990).

So, according to the Copenhagen interpretation of quantum mechanics, the interaction, which is not measurable accurately, will not take place accurately. One can often meet this opinion in contemporary articles written about quantum cosmology, especially in connection with the formation of the Universe: e.g. MacRobert (1983). As we mentioned, this conclusion is false, because of the so-called equivocation fallacy. If one investigates the meaning of 'accurately' in the statement above, he will find that it is different in the first part from the second one. In the first part, the word 'accurately' stands in an operational meaning (accurate measurability), but in the second part it has ontological meaning (accurate existence). Namely, the fault is that the accurate measurability is not a necessary condition for existence!

Due to the above train of thought it is quite a big mistake to try to prove that any kind of matter can be produced by fluctuations of the quantum field. I do not think that it is impossible at all, but a correct deduction could be more reassuring. Moreover, a universe that was created by a quantum fluctuation, is (beyond doubt) virtual, so there cannot be any observer noticing it, because an observer must be real [4].

Finally, the dimensionless density of the Universe does not correspond to the expected value of the quantum cosmology. In a Universe, made by a fluctuation, the density parameter is larger than zero, because this type of universes is closed.

To be honest, there is little chance for a universe made by a quantum fluctuation, and much less for an ensemble of universes that grew up from quantum foam, after we have pointed out the inconsistency. The idea of an ensemble

[4]There is no experimental evidence of matter coming out from fluctuations, if only one can detect Hawking-radiation, which might appear around black holes (Hawking, 1974). Anyway, virtual particles, can not become real, because this particles do not satisfy the relation of dispersion i.e. $p^2 \neq m^2$ (in natural units).

directly implies the next problem: the anthropic principle.

3. Questions of fine-tuning. The anthropic principles

The universe, as one can see is comfortable for mankind, i.e. the constants and laws of nature are fine tuned to the effect that we can exist. There are three types of anthropic principles that explain this fine-tuning of the Universe.

The first called the weak anthropic principle which points out that every new theory must be compatible with the existence of life. Therefore this principle is a methodological one, and so has no strict philosophical content.

The strong anthropic principle - at the other hand - tries to explain that the Universe is fine-tuned in order that human existence could emerge. Some physicists argue that there must be an observer in order that the wave function of the Universe should collapse by an observation. Others think that the there is an ensemble of universes with various physical constants in each one, so there is no wonder, that our universe bears life and we do not need further explanation of life and fine-tuning. There are here some ideas, that could not be proved, as we can see later.

The third type of these principles is the so-called final anthropic principle, which contains the idea that life is a necessary thing in the Universe, i.e. if there is any chance for life formation, it will surely happen.

The last two principles have heavy philosophical contents, while the first remains in the filed of physics and its methodology. It is well known, that there were many debates around these principles. We mention only some problems with them but the list will not be complete.

Let us see first the problem connected with the theory of the ensemble of universes. If there are many universes (with small u) they will or will not touch each other. If yes, there is only one Universe (with capital U), and if not, then it is not worth mentioning[5].

Moreover we do not think that it is reasonable to apply a statistical description to universes that have only one real sample. We can add that how can two universes interact if there are different natural laws and/or constants in them?

The next problem is the collapse of the wave function. As we mentioned above, some physicists thought that an observer was necessary in order to col-

[5]We are physicists at first with the claim of experimental control and only secondly philosophers but not *bad* philosophers! Lots of counter-examples exist. It is worth to note that the distinction between the Universe, and universes, is a mere obscurity and abuse of notion.

lapse the wave function of the Universe. It is known that the Schrödinger-equation could be converted to this form:

$$-i\hbar\frac{\partial|\psi|^2}{\partial t} = -\frac{\hbar^2}{2m}\nabla\left(\psi\nabla\psi^* - \psi^*\nabla\psi\right) \tag{2}$$

that is the so-called equation of continuity. It is clearly seen from the lhs that $|\psi|^2$ acts the role of density. To be precise $|\psi|^2$ bears the meaning of the likelihood-density, or likelihood of finding. But it is clear, that there is no wave function in the reality because it is a mere mathematical construction, and a construction cannot collapse. Another serious trouble comes from the theory of quantum measurements. As it is known, we are part of the Universe, and cannot investigate it from an outer point of view, but it would be necessary to collapse its wave-function. That is, one must be outside of the investigated system to perform it[6].

We can impugn the theory of the ensemble from another point of view, too. Can we justify the existence of an ensemble with hypothetical universes? There is only one Universe from the ensemble what we can investigate. In other words: is it rational to apply statistical description to the one real Universe[7]? It could be noticed, that there were possibilities in the initial conditions of the Universe, and one can substitute these possibilities for an ensemble, but it would be an unjustified and illegal step[8]. The real goal of an ensemble hypothesis is to avoid the thought of a fine-tuned Universe, i.e. the idea of a Creator. Some researchers admit this (e. g. F. Hoyle), others do not[9]. One may object, that there is no need to explain the fine tuning in an ensemble. But even then it is necessary to explain why there is an ensemble. This is why we refere to the Occam's razor.

Some cosmologists think, that the anthropic principles and the problem of fine-tuning will be worked out if the GUT[10] is discovered: e.g.(Linde, Kallosh, 2002). As Hawking noticed in "The brief history of time", if there was a final theory, it would not be more than a number of equations and rules. This

[6]There is a question that is more important than the others above: if the existence of an observer is needed for the existence of the Universe what was the situation before the human life?

[7]As C. Isham noticed: "The interpretation of probability as applied to a single electron remains as mysterious as ever (Isham, 1988)".

[8]"There is clearly a danger that the original problem of the multiplicity of possible initial states will be replaced by an equivalent multiplicity of theories claiming to determine it uniquely." Isham (1988), p.405.

[9]But these scientists always forget to think over the origin of this hypothetical ensemble.

[10]Grand Unified Theory, or Theory Of Everything

is an important warning, because an equation is contingent, (not necesary) consequently it cannot explain itself[11], and the question "Why is it such, and why not different?" remains an open question and it will be asked until there are thinkers.

4. Questions of a possible future. The critique of the Drake's equation and SETI

One of the most puzzling questions is whether there are other civilizations in the Cosmos. This research raises many questions. The well-known Drake's equation raises some questions about its accuracy and validity. Other questions could be asked about the evolutionary theory connected to the question of extraterrestrial life. Last but not least there is the question of communication between ETI and ourselves, from which emerges the question of epistemology.

Drake's equation, in its best known form,contains some important factors among the necessary components of intelligent life:

$$N = R_* n_\mathrm{p} f_\mathrm{pl} f_\mathrm{l} f_\mathrm{civ} L \tag{3}$$

where N is the number of intelligent civilizations in the galaxy, R_* means the rate of star formation, n_p is the number of planets in the habitable zone around a star, f_pl is the fraction of habitable planets, f_l is the fraction of life-bearing planets, f_civ is the fraction of civilized ones, and finally L is the average lifetime of an intelligent civilization. As we shall see, someone, who thinks that equation (3) contains all important parameters which are necessary for intelligent life, is wrong. Besides, the members of the equation (3) do not stand in their final form. Let us investigate the component L. If that civilization of which average lifetime we investigate could colonize a habitable planet, then it could exist forever, i.e. $L \to \infty$. Then N will tend to infinity too. But if there is a finite Universe, that will be impossible[12]. This line of argument could be repeated with other factors of eq. (3).

As it was mentioned above, the equation (3) is not adequate in our view. For example there is not an f_jup multiplier in it, which denotes the fraction of the

[11]Every non-trivial consistent arithmetic group of statements suffers from the lack of completeness as Gödel has proved, so does the hypothetical GUT too. Consequently, there is no theory, which is able to explain why the Universe is such and why cannot be different.

[12]We neglected the supposition that the need for the energy in the far future could be more than now.

planetary systems, which contains a Jupiter-like giant planet in the right place. The presence of a huge gas-planet in the right place seems to be necessary from the view of dangerous minor planets and comets, which will be captured by it, and these hazardous objects cannot extinguish the civilization we are referring to.

Or we can mention also the lucky role of our Moon. It is well-known that our Moon has a very beneficial role in the stabilization of the Earths inclination, without which our civilization may not survive. Only a few scientist know that how small was the probability of the collision which has created our Moon! Let us see further examples about the Moon!

In 1178 a friar saw an impact to the Moon, which formed the crater on the far side of the Moon and was later called Giordano Bruno. If this object had not collided with the Moon, but had reached the Earth it would have wiped out the human civilization.

The Moon was important in the history of the natural science too. It is a lucky chance, that the apparent diameters of the Sun and Moon are equal. So the cycle of solar and lunar eclipses allow the early men to reach exact natural laws. The most important solar eclipse was in 1919, which allowed to check the predictions of the general relativity about the light diffraction near the Sun. One must be convinced, that there is a need of an f_{moon} multiplier in the equation (3)[13].

Those scientists, who are optimists that we could contact ETI, almost always forget that the development of science was very hard on Earth. They think that the emergence of a civilization is certain whenever the intelligent life appears. These thinkers ought to read some textbooks about the history of science. Some shocking examples: Kepler's three laws would have been constructed later if the sword which has cut Tycho de Brahes nose during a duel had not stopped then, and would have killed the famous astronomer. If Einstein had not started to speak when he was five, but would remain dumb, we would not have the general relativity till now. Only the blind can say that the way of science is certain.

If the evolution is the correct theory of the life on Earth, then SETI will be one of the weakest and most dangerous enterprise of mankind. Weak, because the line of evolution is so circuitous, that the expectations for the existence of other intelligent beings in our Galaxy are not well-founded. Dangerous, because the natural selection offers nothing good for us in an encounter.

Last but not least we should mention that who is committed to SETI, ought to think through the problem of the communication. It is a fact, that our species

[13]Further examples: Benn (2001)

cannot communicate with other species intelligently. We do not understand the process of the communication well or the process of thinking. How dare we assume that we shall communicate with an extraterrestrial species?

To make a long story short, there are several unsolved problems with Drake's equation and SETI and we have to work harder and more advisedly in order to solve them.

5. Conclusion

We investigated some problems in cosmology, that lead us out from the realm of physics, but their importance is indisputable. We investigated the idea of the creation out of a quantum process, the anthropic principle and finally some questions induced by SETI and have given our opinion about them.

As the proverb says: "many men many minds", and we can add that there are as many questions. Obviously, there are more questions than answers, but asking questions is our duty by profession with the hope of good results.

Acknowledgments

The authors thank Béla Balázs, Stanley L. Jaki, and Ervin Nemesszeghy for useful and friendly discussions and proposals. Technical help of András Pál is also kindly appreciated.

References

A. MacRobert: Beyond the Big Bang. Sky & Telescope, **63** (1983) p. 211-213

S. Hawking: Black Hole Explosion. Nature, **248** (1974) p. 30-31

S. L. Jaki: Determinism and Reality. Great Ideas Today, 1990 p. 277-302

C. Isham: Creation as a quantum process. Physics, Philosophy, & Theology. A Common Quest for Understanding. (1988) p. 375-408.

C. R. Benn: The moon and the origin of life, astro-ph/0112399

A. Linde - R. Kallosh: The beginning of the end of the anthropic principle. New Astronomy, **7** Issue 1,(2002) p. 45-53

Part Five

Celestial mechanics

PROTOPLANETS IN THE PROTOPLANETARY DISC

I. Nagy
[1]Eötvös University, Department of Astronomy
H-1518 Budapest, P.O.Box 32., Hungary
E-mail: imre@ludens.elte.hu

Abstract

This paper is an overview of the formation of planetary systems. We discuss the star formation with the X-mechanism. After that, we review the two main planet formation processes. At the end, we give a summary of planet-disc interaction with planet migration, which can cause the evolution of resonant orbits. These resonant orbits in the protoplanetary disc change when the disc disappears.
KEYWORDS: *planet formation, resonances, accretion disc, star formation*

1. Introduction

Nowadays more than a hundred exoplanets are known. Some of them have very strange orbits (high eccentricity, very small semimajor axis) and others are in resonance (e.g. 2:1). To explain these phenomena we have to understand the formation of these planets. However, the origin of planets and the origin of stars are related, we cannot understand planet formation without understanding star formation. The foremost question that we have to answer is, what kind of circumstances rule under star formation. The investigation of the planet formation mechanism is also an important question, as well as the investigation of the future of protoplanets.

2. Star formation in brief

Stars form in giant molecular clouds, like the Orion Nebula. First a globula evolves, which is 0.1 pc in diameter, the density of this core is 10^4 molecule/cm^3 and the temperature is 10-100 K. These cores very often contain an IR source, which implies the existence of a star or protostar in the centre.

The globulas are spherical, but their collapses are not spherical. The cosmic radiation ionize the interstellar matter in the globula and the ions move around the local component of the galactic magnetic field. The ions collide with neutral atoms, so the globula begins to rotate around an axis. The rotating matter can

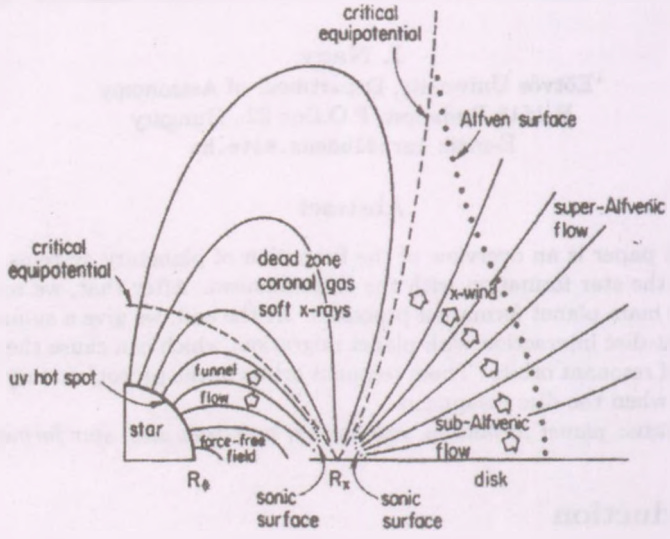

Figure 1: The X-mechanism Shu et al. (1994)

collapse along its axis, so an accretion disc grows up. The gas does not fall directly into the protostar, first it falls into the disc.

The disc does not extend to the protostar, due to the interaction of the protostellar magnetic field and the disc matter. The protostar ionizes the surface of the disc, thus the disc disturbs the magnetic field (Figure 1). At a radius (R_X) the magnetic field opens a gap between the star and the disc. Moreover, angular momentum flows from the protostar to the disc through the magnetic field, so a part of the disc matter is thrown out. This theoretical outflow is the so-called X-wind, which we see as a protostellar stellar wind (Shu et al., 1994). This mechanism unfortunately cannot exactly explain the observed jets in the protostellar systems, because the angle of the X-wind and the disc is relatively small, but the observed jets are perpendicular to the accretion discs.

max/min density	100:1	30:1	10:1
M $[M_J]$	15	1.9	0.63
M_{Jeans} $[M_J]$	0.39	0.57	0.44
R $[AU]$	1.0	0.61	0.35
T $[K]$	97	103	104
E_T/E_{grav}	0.03	0.44	1.0
a $[AU]$	11.8	11.8	11.8
e	0.08	0.1	0.1

Table 1: Structure of a clump. In all the cases the clump mass is larger than the Jeans mass. The free-fall time is 1 year in all cases.The integration time was 345 years (Boss, 2002).

3. Planet formation in brief

3.1. The core collision mechanism

There are two different types of planet formation. The first one is the so-called core collision mechanism. In the accretion (or protoplanetary) disc the temperature is relatively low and the density is high, so heavier elements can condense. Since the disc temperature depends on the distance from the star, the high melting-point components (e.g. silicates) condense in the inner system. These cores collide with each other and sometimes they cohere creating larger and larger cores occasionally.

This is a slow process with timescale of 10^6 years.We think that less massive planets are produced by this mechanism. Due to the interaction with the gas, the small particles spiral into the star very fast. For example the semimajor axis of a particle with 1 cm diameter will decrease 10 cm in each second.

3.2. Planet formation by disc instability

The gas flows in a protoplanetary disc are turbulent according to numerical simulations. Turbulence can be caused by viscosity of the gas and/or thermodynamical effects (Boss, 2002). Turbulence creates gravitationally unstable clumps on a timescale of 100-1000 years. These clumps are at least 5 AU away from the star and their mass is 0.5 or more Jupiter-mass. The structure of a typical clump is shown in Table 1.

One can see that this mechanism can produce giant planets. These planets perturb the disc strongly. This strong perturbation can even cause fragmenta-

tion of the disc. The fragmentation can cause the formation of some additional giant planets. We have to note, that the circumstances are more prosperous for clump creation at the beginning of the star formation, since the mass of the future star is still in the disc. But how does a system evolve after an early fragmentation?

4. Interaction between planets and the protoplanetary disc

The future of the protoplanets in the disc is strongly influenced by the disc. The planets interact with the disc, which cause angular momentum transport. Because of this transport, the semimajor axis of the planets decrease, the planets migrate toward the star (Lin & Papaloizou , 1986). The speed of the migration is proportional to the mass of the planet. The planet mass is also important because of another aspect. When the mass is greater than one Jupiter-mass, a gap opens in the disc around the planet (Ward , 1997). The migration across the gap is slow. Migration without gap is called type I migration, and with gap type II migration.

The disc-planet interaction helps to evolve resonant orbits in planetary systems. Some examples are shown in Table 2. There are 2:1 resonance in two of the systems. About the stability of them see Érdi & Pál (2003), and about the origin of this 2:1 resonance see Kley (2003).

A protoplanetary disc contains about 0.1 $M_{\odot}$ of matter at the main accretion phase. After the source of the falling matter runs out, the disc disappears slowly. The X-wind blows out one third of the disc mass, and just two third fall into the star. The fotoevaporation and stellar wind reduce it also, so at the end the disc disappears. This process takes about 10^6-10^7 years. Because the mass inside a given radius around the star depends on the distance from the star and density distribution of the disc, the periods of the planets do not change in the same way during the disappearing. Moreover the disc disappears not simultaneously everywhere. When one of the resonant planets gets out of the disc, the other one is still in the disc and interacts with it, while the first one doesn't. So the existing resonances change.

5. Summary

The structure of planetary systems is strongly influenced by their original protoplanetary disc. Streams in the disc can produce giant planets. Interaction between the planets and the disc result type I or II migration, that can cause

	Per $[d]$	M *sini* $[M_J]$	a $[AU]$	e	ω	M
GJ 876 (2:1)						0.32
c	30.569	0.766	0.13	0.24	159	
b	60.128	2.403	0.21	0.04	163	
HD 82943 (2:1)						1.05
b	221.6	0.88	0.73	0.54	138	
c	444.6	1.63	1.16	0.41	96	
55Cnc (3:1)						0.95
b	14.65	0.84	0.11	0.02	99	
c	44.26	0.21	0.24	0.34	61	
d	5360	4.05	5.9	0.16	201	

Table 2: Resonant orbits in exoplanetary systems.

the planet falling into the star. On the other hand the disc may help to create resonances (for example 2:1) as well. When the disc disappears the resonances will be modified.

I would like to investigate the effect of the disappearance of a disc for the orbits of planets. This may be especially important in case of resonant orbits. First I would like to use an approximate formula for the variation of the semimajor axis.

Acknowledgments

I thank Bálint Érdi for the useful consultations.

References

Érdi, B. & Pál, A. 2003, in F. Freitetter, R. Dvorak, B. Érdi (eds.), Proceedeings of the 3rd Austrian Hungarian workshop on Trojans and related topics, Eötvös Univ. Press, Budapest, p. 3

Shu et al. 1994, The Astrophysical Journal, 429, 781

Boss, A. P. 2002, The Astrophysical Journal, 576, 462

Lin, D. N. C. & Papaloizou, J. 1986, The Astrophysical Journal, 309, 846

Ward, W. R. 1997, Icarus, 126, 261

Kley, W. 2003, Celestial Mechanics and Dynamical Astronomy, 87, 85

On the comparison of the different methodes to calculate the derivatives of the Laplace-coefficients

Á. Süli
Eötvös Loránd University, Department of Astronomy
H-1518 Budapest, P.O.Box 32., Hungary
E-mail: A.Suli@astro.elte.hu

Abstract

The determination of the secular variations of the orbital elements of objects in N-body systems is based on the literal development of the perturbing function. The development makes use of the Laplace coefficients and their derivatives. In this paper a new method is shortly described for the analytical computation of the derivatives of the Laplace coefficients. It is an explicit formula in the sense that it only contains the Laplace coefficients and the parameter α on which the Laplace coefficients depend. The advantage of this method is that it directlly gives the desired numerical results. Easy coding is a further benefit of the method and it provides more accurate numerical results. The paper describes in detail the application of the method through an example and gives a detailed comparison with former methods.

KEYWORDS: *analytical methods, N-body systems, perturbation theory, Laplace coefficients*

1. Introduction

Many studies on the long term variations of the orbital elements of objects in N-body systems have been carried out. The most natural example of an N-body system is the Solar System and on larger scale the Milky way. The problem of the secular variations of the orbital elements of the major planets was first tackled in 1781 by Lagrange and later by Laplace.

The Laplace-Lagrange approach took into account only linear terms. Later on Hill (1897) was the first to introduce degree 3 and 5 of the eccentricities into the development of the perturbing function of the 2nd order with respect to the masses of Jupiter and Saturn. Brouwer & van Woerkom (1950) used the perturbing function of Hill and incorporated it into their secular theory of the eight planets. This theory has been extensively used in the studies of motion of Solar System bodies.

Finally, Laskar (1984, 1985) developed a solution which uses terms in the perturbing function up to the second order in the masses and degree 5 for the eccentricity-longitude of perihelion, and degree 3 for the inclination-longitude of node couples. This theory has been made in a very extensive way: none of the terms which can appear in the secular system has been neglected and each coefficient of the system has been computed.

Clearly, all these works are based on the literal development of the disturbing function. The ultimate purpose is to create an analytical theory for the secular variations of the major planets which is valid for a time span comparable with the age of the Solar System. To fulfill this purpose on one hand each significant term of the disturbing function has to be taken into account, on the other hand the coefficients of the disturbing function have to be determined as precise as possible. The latter one is more important, since it improves the accuracy of the most significant terms too and also because small errors in the terms become quite large on million-years timescale and distort the secular solution.

It is well known that the literal expansion of the disturbing function makes use of the Laplace coefficients and their derivatives. These coefficients and their derivatives are functions of the semimajor axes of the planets. The Laplace coefficients $b_s^{(j)}(\alpha)$ have been defined in two different ways, but these definitions are equivalent. One way to define them is to use series expansion:

$$\frac{1}{2}b_s^{(j)}(\alpha) = \frac{s(s+1)\dots(s+j-1)}{1\cdot 2\cdot\dots\cdot j}\alpha^j\left[1+\frac{s(s+j)}{1(j+1)}\alpha^2+\right.$$
$$\left.+\frac{s(s+1)(s+j)(s+j+1)}{1\cdot 2(j+1)(j+2)}\alpha^4+\dots\right], \qquad (1)$$

another one is to use integral representation:

$$\frac{1}{2}b_s^{(j)}(\alpha) = \frac{1}{2\pi}\int_0^{2\pi}\frac{\cos(j\psi)}{(1-2\alpha\cos(\psi)+\alpha^2)^s}d\psi, \qquad (2)$$

where s is a positive half integer, j is an integer, and $\alpha = a/a'$, a and a' being the semi-major axes of the orbit of two planets (for details see Hagihara (1973) and Stumpf (1974)). It has been shown that the series defining the Laplace coefficients is always convergent for $\alpha < 1$.

It is easy to derive an explicit formula for the derivatives of the Laplace coefficients from the series definition where one knows how to calculate any term from the previous one. If A_j and B_i denote the quantities

$$A_j = \begin{cases} 1 & \text{if } j = 0 \\ \frac{s(s+1)\ldots(s+j-1)}{j!} & \text{if } j > 0 \end{cases}, \quad B_i = \begin{cases} 1 & \text{if } i = 0, \\ \frac{\prod_{k=0}^{i-1}(s+k)(s+j+k)}{i!\prod_{l=1}^{i}(j+l)} & \text{if } i > 0 \end{cases}$$

then

$$\frac{1}{2}b_s^{(j)}(\alpha) = A_j\alpha^j \sum_{i=0}^{\infty} B_i\alpha^{2i}. \tag{3}$$

Differentiating Eq. (3) with respect to α, the nth derivative of the Laplace coefficient can be written

$$\frac{d^n}{d\alpha^n}\frac{1}{2}b_s^{(j)}(\alpha) = A_j\alpha^{j-n} \sum_{i=0}^{\infty} F_i^{(j,n)} B_i\alpha^{2i} \tag{4}$$

where

$$F_i^{(j,l)} = \prod_{k=0}^{l-1}(l-k+2i). \tag{5}$$

The series in Eq. (4) has to be summed up to achive good precision. Unfortunatelly, the number of terms to be summed up can't be determined in advance to achive prefixed precision. Hereafter Eq. (4) will be referred as direct method.

An explicit formula for the numerical computation of the derivatives of the Laplace coefficients based on the integral representation is given in D'Eliseo (1989). Useful relations between the Laplace coefficients and their derivatives include:

$$Db_s^{(j)}(\alpha) = s\left(b_{s+1}^{(j-1)} - 2\alpha b_{s+1}^{(j)} + b_{s+1}^{(j+1)}\right), \tag{6}$$

$$\begin{aligned} D^n b_s^{(j)}(\alpha) = s\Big(&D^{n-1}b_{s+1}^{(j-1)} - 2\alpha D^{n-1}b_{s+1}^{(j)} \\ &+ D^{n-1}b_{s+1}^{(j+1)} - 2(n-1)D^{n-2}b_{s+1}^{(j)}\Big), \end{aligned} \tag{7}$$

where $n \geq 2$ in the last relation and $D = \frac{d}{d\alpha}$ is a differential operator. Using the recurrence formula of Eq. (7) to determine the nth derivative of $b_s^{(j)}(\alpha)$ one has to calculate all the derivatives up to n. It means that each time superfluous terms have to be determined.

2. The new method

The basics of the new method is to apply the D differential operator successively on $b_s^{(j)}(\alpha)$. The first three derivatives of the Laplace coefficients $b_s^{(j)}(\alpha)$ can be written as

$$
\begin{aligned}
D^0 b_s^{(j)} &= K_{0,0}\left\{ b_s^{(j)} \right\}_0, \\
D^1 b_s^{(j)} &= K_{1,0} s\left\{ -2\alpha b_{s+1}^{(j)} + b_{s+1}^{(j\pm 1)} \right\}_0, \\
D^2 b_s^{(j)} &= K_{2,0} s(s+1)\left\{ \left(4\alpha^2+2\right) b_{s+2}^{(j)} - 4\alpha b_{s+2}^{(j\pm 1)} + b_{s+2}^{(j\pm 2)} \right\}_0 + K_{2,1} s\left\{ b_{s+1}^{(j)} \right\}_1, \\
D^3 b_s^{(j)} &= K_{3,0} s(s+1)(s+2)\left\{ -\left(8\alpha^3+12\alpha\right) b_{s+3}^{(j)} + \left(12\alpha^2+3\right) b_{s+3}^{(j\pm 1)} + \right. \\
&\quad \left. - 6\alpha b_{s+3}^{(j\pm 2)} + b_{s+3}^{(j\pm 3)} \right\}_0 + K_{3,1} s(s+1)\left\{ -2\alpha b_{s+2}^{(j)} + b_{s+2}^{(j\pm 1)} \right\}_1,
\end{aligned}
$$

where

$$
\begin{aligned}
& b_s^{(j\pm k)} := b_s^{(j+k)} + b_s^{(j-k)} \quad k > 0, \\
& \left\{ \dots \right\}_0 \quad \text{contains } b_{s+n}^{(j\pm k)} \text{ terms}, \quad k = 0, 1, \dots n,
\end{aligned}
$$

Hereafter the coefficients of the Laplace coefficients in the above expressions will be referred as α-coefficients, since they depend only on α, and the $K_{i,l}$ factors will be referred as numeric factors.

Examining the α-coefficients in the $\{\dots\}_0$ brackets, it is easy to see that the following relation holds between them:

$$
\begin{aligned}
S_{i+1,0} &= -2\alpha S_{i,0} + 2S_{i,1}, \\
S_{i+1,l} &= S_{i,l-1} - 2\alpha S_{i,l} + S_{i,l+1}, \quad i = 0, \dots n, \quad l = 1, \dots n+1.
\end{aligned} \tag{8}
$$

Substituting the starting values $S_{0,0} = 1, S_{0,l} = 0, l = 1, 2, \dots n$ into the formulae (9) a table can be obtained (see Table 1).

The numeric factors $K_{i,l}$ can be calculated from the expression:

$$
\begin{aligned}
K_{i+1,0} &= K_{i,0} \quad i = 0, 1 \dots, n, \\
K_{i+1,l+1} &= -2(i-2l) K_{i,l} + K_{i,l+1} \quad i, l = 0, 1 \dots n.
\end{aligned} \tag{9}
$$

$i \setminus l$	0	1	2
0	1	0	0
1	$-(2\alpha)$	1	0
2	$+(4\alpha^2+2)$	$-(4\alpha)$	1
3	$-(8\alpha^3+12\alpha)$	$+(12\alpha^2+3)$	$-(6\alpha)$
4	$+(16\alpha^4+48\alpha^2+6)$	$-(32\alpha^3+24\alpha)$	$+(24\alpha^2+4)$
5	$-(32\alpha^5+160\alpha^3+60\alpha)$	$+(80\alpha^4+120\alpha^2+10)$	$-(80\alpha^3+40\alpha)$
$\vdots$	$\vdots$	$\vdots$	$\vdots$

Table 1: The α-coefficients of the Laplace coefficients in the $\{\dots\}_0$ bracket. In the text the elements of Table 1 are referred as $S_{i,l}$.

$i \setminus l$	0	1	2	3	4	...
0	1	0	0	0	0	0
1	1	0	0	0	0	0
2	1	-2	0	0	0	0
3	1	-6	0	0	0	0
4	1	-12	12	0	0	0
5	1	-20	60	0	0	0
6	1	-30	180	-120	0	0
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$

Table 2: The numeric factors $K_{i,l}$ of the derivatives (see the recurrence formula for $K_{i,l}$).

The starting values for this relation are $K_{0,0} = 1$ and $K_{0,l} = 0, l = 1, 2, \dots n$. Some values of $K_{i,l}$ are presented in Table 2.

The results can be summarized in the following formula:

$$D^n b_s^{(j)} = \sum_{m=0}^{\left[\frac{n}{2}\right]} \left(K_{n,m} \prod_{l=0}^{n-m-1} (s+l) \sum_{k=0}^{n-2m} S_{n-2m,k} b_{s+n-m}^{(j\pm k)} \right), \qquad (10)$$

where $n \geq 1$ is the order of the derivative.

The formula (10) has been tested via several relations between the derivatives of the Laplace coefficients.

2.1. Example of the new method

The derivative $D^3 b^{(1)}_{1/2}(\alpha_{JS})$ is calculated using formulae (10) as an example, and comparisons with the method based on the recurrence formula (7) with the direct method and with the method of D'Eliseo (1989) are given. In this case $n = 3$, $[n/2] = 1$, $j = 1$, $s = \frac{1}{2}$ and $\alpha_{JS} = \frac{a_J}{a_S} = 0.5444270016$, where a_J and a_S are the semi-major axes of the orbit of Jupiter and Saturn, respectively. The derivative consists of two terms, $m = 0$ and $m = 1$. The α-coefficients come from the third and first row of Table 1 ($S_{3-0,k}$ and $S_{3-2\cdot 1,k}$, $k = 0, 1, \ldots n+1$), and the numeric factors $K_{3,m}$ from the 3rd row of Table 2. The product in the case of $m = 0, n - m - 1 = 2$ is $\frac{1}{2}(\frac{1}{2}+1)(\frac{1}{2}+2) = \frac{15}{8}$, and in the case of $m = 1, n - m - 1 = 1$ is $\frac{1}{2}(\frac{1}{2}+1) = \frac{3}{4}$. The result is:

$$
\begin{aligned}
D^3 b^{(1)}_{\frac{1}{2}}(\alpha_{JS}) &= K_{3,0}\cdot\frac{15}{8}\left\{S_{3,0}b^{(1)}_{\frac{7}{2}} + S_{3,1}b^{(1\pm1)}_{\frac{7}{2}} + S_{3,2}b^{(1\pm2)}_{\frac{7}{2}} + S_{3,3}b^{(1\pm3)}_{\frac{7}{2}}\right\} + \\
&\quad + K_{3,1}\frac{3}{4}\left\{S_{1,0}b^{(1)}_{\frac{5}{2}} + S_{1,1}b^{(1\pm1)}_{\frac{5}{2}}\right\} = \\
&= \frac{15}{8}\left\{-\left(8\alpha^3_{JS}+12\alpha_{JS}\right)b^{(1)}_{\frac{7}{2}} + \left(12\alpha^2_{JS}+3\right)b^{(1\pm1)}_{\frac{7}{2}} + \right. \\
&\quad \left. - 6\alpha_{JS}b^{(1\pm2)}_{\frac{7}{2}} + b^{(1\pm3)}_{\frac{7}{2}}\right\} - \frac{9}{2}\left\{-2\alpha_{JS}b^{(1)}_{\frac{5}{2}} + b^{(1\pm1)}_{\frac{5}{2}}\right\} \\
&= 12.80441741,
\end{aligned}
$$

where we have substituted the numerical values of the Laplace coefficients computed from well-known recursive formulae.

3. Example and comparison

The calculation of $D^3 b^{(1)}_{1/2}$ with the recurrence formula is the following. Substituting $n = 3$, $j = 1$ and $s = \frac{1}{2}$ into the recurrence formula (7) one obtains

$$
D^3 A_1 = \frac{1}{2}\left(D^2 B_0 - 2\alpha_{JS} D^2 B_1 + D^2 B_2 - 4DB_1\right).
$$

For the determination of $D^3 A_1$ one has to calculate

$$
\begin{aligned}
D^2 B_j &= \frac{3}{2}\left(DC_{j-1} - 2\alpha_{JS} DC_j + DC_{j+1} - 2C_j\right), \quad j = 0, 1, 2, \\
DC_j &= \frac{5}{2}\left(D_{j-1} - 2\alpha_{JS} D_j + D_{j+1}\right), \quad j = -1, 0, \ldots 3,
\end{aligned}
$$

where we have used the denotations

$$A_j = b_{\frac{1}{2}}^{(j)}(\alpha_{JS}), \quad B_j = b_{\frac{3}{2}}^{(j)}(\alpha_{JS}), \quad C_j = b_{\frac{5}{2}}^{(j)}(\alpha_{JS}), \quad D_j = b_{\frac{7}{2}}^{(j)}(\alpha_{JS}).$$

Completing the calculations the result is $D^3 A_1 = 12.80441748$, and the difference between the two results is $\Delta = 7 \cdot 10^{-8}$.

The initial values needed by the two different methods are the same, and both methods use 10 Laplace coefficients. The powers of α_{JS} are present in each method, therefore the error in α_{JS} has the same effect on both results.

Comparing the number of operations the recurrence formula needs 24 additions and 24 multiplications, whereas the new method needs only 8 additions and 16 multiplications. At higher order derivatives the difference between the number of operations becomes even more significant. Generally the progression of error depends on the number of operations. Accordingly, the new method suffers less from this kind of error and so it is more accurate.

The computation of $D^3 b_{1/2}^{(1)}$ with the method of D'Eliseo (1989) was also carried out, the result is 12.80441828, and the differences between the recursive method and the new method are $-8 \cdot 10^{-7}$ and $8.7 \cdot 10^{-7}$, respectively. The equations of the calculations can be found in D'Eliseo (1989).

The numerical value of $D^3 b_{1/2}^{(1)}$ is 12.80441832 computed with the direct method. The difference between the direct method and the new method, the recursive method and the method of D'Eliseo are $9.1 \cdot 10^{-7}$, $8.4 \cdot 10^{-7}$ and $4 \cdot 10^{-8}$, respecively.

Finally, for the sake of a comparison a table is presented (Table 3), where the numerical results of $D^n b_{1/2}^{(3)}, n = 1, 2, 3$, obtained by the four different methods, are given. Examining the differencies in Table 4, the discussed methods can be classified into two separated groups:

- group 1: methods based on the hypergeometric (or integral) representation,
- group 2: methods based on the recurrence properties.

According to this classification it is only natural that the recursive and the new method deliver similiar numerical values and the direct method and the method of D'Eliseo (1989) give also similiar results. Of course the results of the two groups significantly differ from each other.

We have compared the CPU time, which the different methods require to deliver the result. Exact answers can not be given, because it strongly depends

Results:	Recursive Method	Method of D'Eliseo	New Method	Direct Method
$D^1 b_{1/2}^{(3)}$	0.7233938252	0.7233938483	0.7233938250	0.7233938413
$D^2 b_{1/2}^{(3)}$	3.520277603	3.520277431	3.520277597	3.520277503
$D^3 b_{1/2}^{(3)}$	15.37456178	15.37456299	15.37456183	15.37456284

Table 3: The numeric results of $D^n b_{1/2}^{(3)}, n = 1, 2, 3$.

Recursive - D'Eliseo	D'Eliseo - New	Recursive - New	Direct - New	Direct - D'Eliseo
-2.31e-8	2.33e-8	2e-10	1.63e-8	-7e-9
1.72e-7	-1.66e-7	6e-9	-9.4e-8	7.2e-8
-1.21e-6	1.16e-6	-5e-8	1.01e-6	-1.5e-7

Table 4: The numerical differencies between the different methods.

on the code. It turned out that the recursive formula is faster if it is explicitly written for the specific derivative term, whereas the general recursive-code is several times slower than the new method's formula. Taking into consideration that in most cases a general formula is utilized, the usage of the new method is more advantageous.

4. Summary

The literal expansion of the disturbing function is a fundamental task in dynamical systems. There are two ways to improve the accuracy of the development of the disturbing function, one way involves more terms into it and the other is to achieve better accuracy of the Laplace coefficients and their derivatives. As it has been shown, both the method based on the recurrence formula and the new method use the Laplace coefficients to deliver the derivatives of the Laplace coefficients, but the new method requires less operations. This has two advantages: better accuracy and efficiency. The application of the method has been shown via an example and it has been compared with the method of the recurrence formula, with the method of D'Eliseo (1989) and with the direct method. A classification of the methods was introduced based on the derivation of the method. The difference between the two groups is very well visible from Table 4.

References

Brouwer, D. and van Woerkom, A.J.J.:1950, *Astron. Papers Amer. Eph* **13**, Pt. 2, 81-107.

Hagihara, Y.: *Celestial Mechanics*, Vol II, Part I, p. 1-23.

Hill, G.W.:1897, *Astron. J.* **17**, 81-87.

Laskar, J.:1984, 'Theorie generale planetaire: Elements orbitaux des planetes sur 1 million d'annes.', Thesis, Paris.

Laskar, J.:1985, *Astron. Astrophys.* **144**, 133-146.

D'Eliseo, M. M.:1989, *Celest. Mech.& Dyn. Astron.* **46** 159-161.

Stumpf, K.: *Himmelsmechanik* III., VEB Deutscwer Verlag der Wissenschaften, p. 44-58.

Numerical investigation of the orbit of Interball-1

I. Nagy
Eötvös University, Department of Astronomy
H-1518 Budapest, P.O.Box 32., Hungary
E-mail: imre@ludens.elte.hu

Abstract

We investigated the orbit of the artificial satellite Interball-1 by using Zare's method (Zare et al., 1996). The satellite was perturbated the strongest by the Moon and it fell down as a result of this perturbation. The measured and computed orbital elements differ from each other within the measured error bound. We could predict the end of the lifetime of the satellite with large precision.
Keywords: *artificial satellites, Interball-1, atmospheric effects*

1. Introduction

In space research most often satellites orbiting in circular or small eccentricity orbits are used. The modelling of these orbits is relatively simple and can be done with great precision. For example, the members of the LAGEOS series move in such orbits, whose motions could be determined so precisely that even the effect of the infrared radiation of the Earth could be pointed out (Rubincam, 1987).

There are, however, satellites whose orbits are very elongated. To follow numerically the motion of these satellites is difficult, since when they approach the Earth their velocities become large and their orbits change quickly. A good example is the Russian Prognoz series, whose latest members are the two pairs of Interball or Prognoz-M (http://www/jki.rssi.ru/interball.html).

The orbital elements of the Interball-1 right after reaching the orbit were (3 August 1995): a=103 276.495 km, e=0.930 53, i=62.821 deg. The orbit was chosen so that it had a relatively small angle with the ecliptic. This made it possible that the satellite spent most of its orbital period in the magnetosphere of the Earth. The orbit is high above the atmosphere, thus the incalculable atmospheric perturbations do not influence the motion of the satellite, so a high precision orbit reconstruction is expected. A further important point is that the orbit was continuously tracked. Between its launch and

death much time elapsed (five years) so a large number of orbital date is available. The date were continuously displayed on the home page of the satellite (http://www.jki.rssi.ru/vprokhtor/tail.html) so they can be reached easily. The precision of the distances, computed from the published coordinates, is 30-40 km which can be judged as sufficiently good, considering that there could be distances of even 200 000 kms.

The computer program for the investigation of the orbit of the Interball-1 was written originally by Bálint Érdi. For the Earth's satellites both the Sun and the Moon have great influence. In order to compute to orbit of a space probe, first we have to determine the positions of the Sun and the Moon. The program uses the VSOP 87 theory (Bretagnon & Francou, 1988) to calculate the coordinates of the Sun with a time series (Simon et al., 1994) and the ELP 2000-85 theory (Chapront-Touzé & Chapront, 1988) to calculate the coordinates of the Moon. Beside the above effects, substantial perturbations are caused by the Earth itself. This comes from the fact that the Earth is not spherically symmetric. The potential of the Earth usually is given as a series expansion in spherical functions. The program takes into account the series expansion of the potential up to 12th order.

The equation of motion are regularized with Zare's method (Zare et al., 1996). This method uses a constant time element in place of time. The regularized equations form an 11th order system of differential equation.

2. The variations of the orbital elements

Figure 1 shows the orbital elements as the function of time. These results are similar to Zare's results (Zare et al., 1996) except the longer integration time. In the case of the semi-major axis it can be seen well, that it is not constant in time. Despite this, it changes almost in all times in a band of 200 km wide about the value 102 750 km.

The eccentricity changes very much during the lifetime of the satellite, this is an outstanding feature. From the initial 0.93 it decreases first to 0.7, then begins to increase and by the end of the interval reaches again above 0.93. Due to this, the perigee first goes higher and then after a maximum value it begins to decrease. The satellite enters into the denser regions of the atmosphere and gets destroyed although the semi-major axis is almost constant.

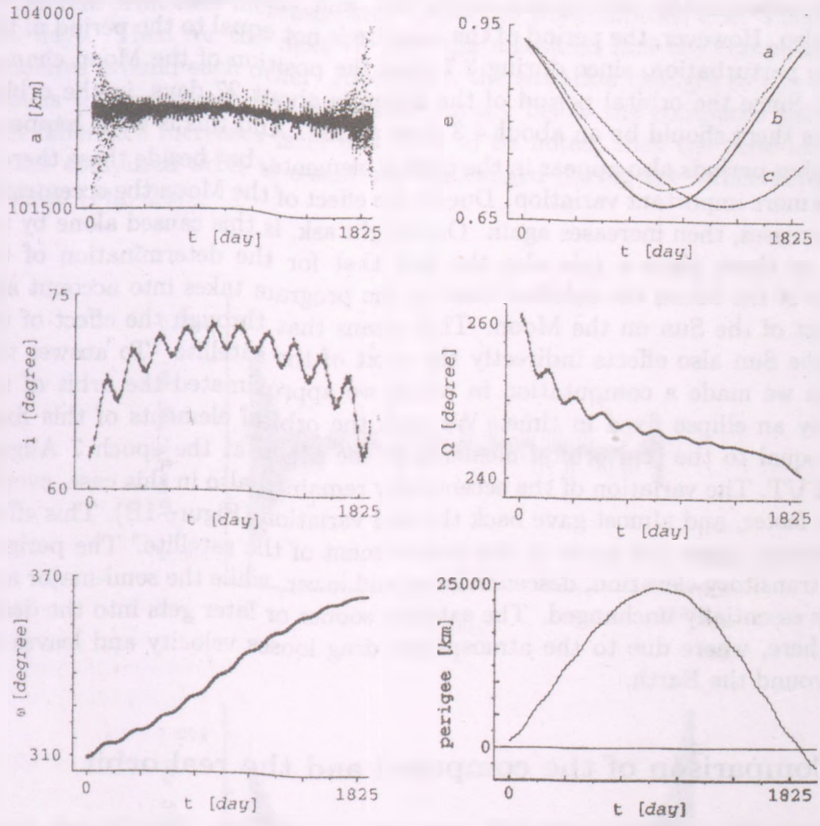

Figure 1: The orbital elements of Interball-1. In the eccentricity graph (B) shows three curve. The curve c shows the effect of the Moon (without other perturbation forces) where the position of the Moon was computed by the precise lunar theory. The curve b shows the effect of the Moon, when it is in a fixed elliptic orbit. The curve a shows the combined effect of all the important gravitational perturbations. In the graph of perigee (F), the curve intersects the horizontal line, corresponding to the Earth's surface, on 16 October 2002.

3. Influences of the Sun and the Moon

We studied the influence of the Moon and the Sun. The investigations show that the effect of the Moon is the most important. This effect results in the appearance of several periods. The distance between the satellite and the Moon

changes substantially during one revolution, and accordingly the orbital elements also. However, the period of the satellite is not equal to the period of this periodic perturbation, since during 3.7 days the position of the Moon changes largely. Since the orbital period of the Moon is about 27 days, in the orbital elements there should be an about 4.3 days period. And this is what happens.

Further periods also appear in the orbital elements, but beside these there is an even more important variation. Due to the effect of the Moon the eccentricity first decreases, then increases again. One might ask, is this caused alone by the Moon, or there plays a role also the fact that for the determination of the position of the Moon the solution used in the program takes into account also the effect of the Sun on the Moon. This means that through the effect of the Moon the Sun also effects indirectly the orbit of the satellite. To answer this question we made a computation in which we approximated the orbit of the Moon by an ellipse fixed in time. We took the orbital elements of this fixed ellipse equal to the real orbital elements of the Moon at the epoch 7 August 1995 12 UT. The variation of the eccentricity remained also in this case, even it became faster, and almost gave back the real variation (Figure 1B). This effect is important, since this leads to the destroyment of the satellite. The perigee, after a transitory elevation, descends lower and lower, while the semi-major axis remains essentially unchanged. The satellite sooner or later gets into the dense atmosphere, where due to the atmospheric drag looses velocity and leaves its orbit around the Earth.

4. Comparison of the computed and the real orbit

Comparing the computed and the measured coordinates, there is not much differences between them, except that the computed values (computed from the starting data) are late by 0.57 day with respect to the measured ones at the end of the integration interval. Due to this phase difference, the position of the satellite near the perigee can not be predicted (from the starting initial data), since the difference can be as large as 30 000 km. This difference is about 2-3000 km near the apogee.

The computed and the observed orbital elements differ from each other. This is most outstanding in the semi-major axis. In this orbital element the difference between the computed and the measured values is about 55 km at the end of the 1762.5 days interval for which measurements were available. Considering, that the semi-major axis is more than 100 000km, this does not seem much.

In Figs. 2a and 2b one can see in a magnification, that the semi-major axis

of the real orbit first increases with respect to the computed one. This lasts for 530 days. Then for the next 70 days the measured and the computed values oscillates around each other. After that, the situation changes to the opposite, that is the measured data go systematically below the computed values, and this difference increases later on. It is to be noted, that the semi-major axis of the computed orbit is essentially unchanged during the whole investigated interval (Fig. 2b).

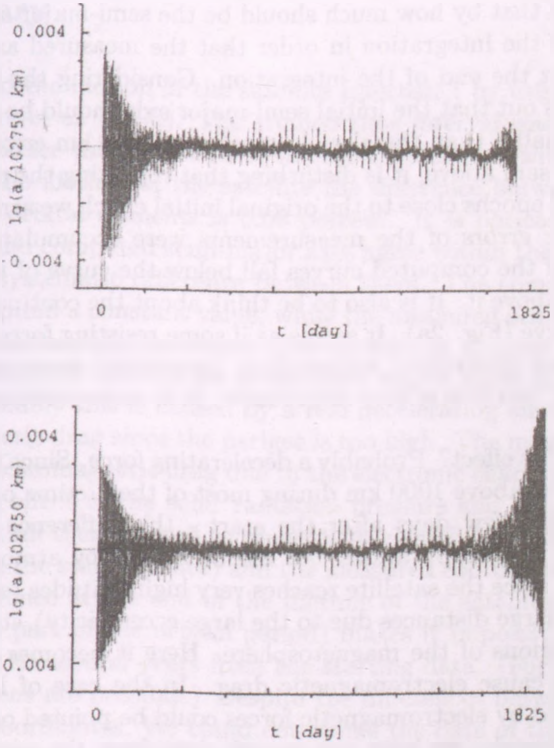

Figure 2: *Upper: (a)* The variation of the measured semi-major axis as a function of time. The continuous decrease can be seen well. *Lower: (b)* The computed semi-major axis. It can be seen, that it oscillates around a constant mean value, there is no indication for secular variation.

It is really an important difference? Considering, that the distance of the

satellite is known with 30-40 km precision, and assuming that the velocity measurements are without noticeable errors, the semi-major axis, computed from the measured data, has an error at least 22.5 km. Due to the errors in the initial data of the integration, the computed data should have at least such errors. This means, that the difference between the two curves can be as large as 55 km, that is the difference of the measured and the computed data are just in the error limit. If we take into consideration, that the gravitational many-body problem is chaotic (that is small initial differences grow very fast); then we can state with good heart that the two data series agree within the error bound. From the phase difference, found between the measured and computed coordinates, we computed that by how much should be the semi-major axis corrected at the beginning of the integration in order that the measured and computed coordinates agree at the end of the integration. Considering the laws of error propagation it turns out that the initial semi-major axis should be decreased by 22.1 km. This is smaller than the already mentioned 22.5 km error.

Despite what is said above, it is disturbing that repeating the computations starting from initial epochs close to the original initial epoch we arrive at similar results. If only the errors of the measurements were accumulated we would expect that some of the computed curves fall below the curve of the measured data, and some go above it. It is also to be think about the continuous decrease of the measured curve (Fig. 2a). It seems as if some resisting force would effect the satellite. As an alternative explanation, systematic measurement errors could be thought of, but this is very improbable. It is more probable, that there exists a real effect.

What could be this effect? Probably a decelerating force. Since the perigee of the Interball-1 is well above 1000 km during most of the lifetime of the satellite - apart from the first few days after the start - this difference between the measurements and the computations can not be caused by atmospheric drag. On the other hand, since the satellite reaches very high altitudes (even it spends most of its time at large distances due to the large eccentricity), it gets into the high ion density regions of the magnetosphere. Here it becomes substantially charged which can cause electromagnetic drag. In the case of LAGEOS the perturbations caused by electromagnetic forces could be pointed out (Barlier et al., 1986). Unfortunately, we have not enough data from the satellite and the magnetosphere to check this concept.

There is another interesting possibility to compare the measured and computed data. If the eccentricity increases while the semi-major axis does not change much, the perigee gets lower and lower. In the case of the Interball this descending was very fast, and it resulted in that the satellite entered into

the dense regions of the atmosphere on 16 October 2000 and ended its life (http://www.jki.rssi.ru/vprokhtor/tail/ib1bift.html). The computed data show that the perigee descends below 200 km on 14 October, and then the perigee height becomes negative on 16 October. Though we could not compute from the initial starting data the exact place, where the satellite would hit the Earth, due to the large phase shift in the coordinates, the date of the end of its lifetime we could determine with large precision (Fig. 1F).

5. Summary

We investigated the motion of the satellite Interball-1 by using the method developed by Zare et al. (1996). The studied time interval was about five years. We could reproduce the variations of the orbital elements with large precision, at the end of the lifetime of the satellite the difference between the measured and computed orbital elements is 0.06 percent. It is noticeable, that though the measured and computed semi-major axes agree within the error limit, there seems to be a systematic difference between them. The computed semi-major axis changes around a constant value, while the measured semi-major axis after the first year continuously decreases. Due to this, by the end of the life of the satellite the measured value of the semi-major axis is lower by 55 km than the computed. Probably this is caused by a real decelerating effect, which can not be the atmospheric drag since the perigee is too high. The most probable explanation is the electromagnetic drag due to the electronic charging of the satellite. We studied the effect of the solar radiation pressure and the perturbations of the planets, but all these proved to be negligibly small. Between the computed (computed from the starting data) and the measured coordinates we found a 0.6 day phase difference at the end of the lifetime of the satellite. This difference (more than 1/6 part of the orbital period) makes it impossible to predict the orbital position for several years from the starting data. (For shorter intervals precise predictions are possible.) Despite the differences between the observed and measured coordinates, We could determine the date of the end of the lifetime of the satellite. According to our computations the perigee height of the satellite decreases to zero on 16 October 2000, and this agrees with the real destroyment of the Interball. We have shown, that the large variations in the perigee distance and in the eccentricity are caused by the effect of the Moon. Because the satellite ended its life, there is no chance for more measurements to investigate the discovered decelerating effect.

Acknowledgments

I thank Bálint Érdi for the useful consultations and the numerical code.

References

Rubincam, D. P. 1987, NASA Technical Memorandum 87804
Zarc, K. et al. 1996, American Astronautical Society Publications 96
Bretagnon, P. & Francou, G. 1988, Astronomy and Astrophysics, 202, 309
Simon, J. L. et al. 1994, Astronomy and Astrophysics, 282, 663
Chapront-Touzé, M. & Chapront, J. 1988, Astronomy and Astrophysics, 190, 342
Barlier, F. C. et al. 1986, Ann. Geophys. 4, A3, 193

Distribution of Asteroids in the Solar System: The Trojans

A. Pál[1] and Á. Süli[2]
Eötvös University, Department of Astronomy
H-1518 Budapest, P.O.Box 32., Hungary
E-mail: [1]apal@szofi.elte.hu, [2]A.Suli@astro.elte.hu

Abstract

Although there is no difference between the stability properties between the Lagrangian points L_4 and L_5 both in the circular and elliptical restricted three-body problem, the number of the observed Trojan asteroids around the two Lagrangian points of Jupiter is not the same. We summarize our first numerical results concerning the difference of the phase space structure around the two Lagrangian points, taking into account the effect of Saturn as a perturber. These results show us that the size of the stable region around the leading point is even larger also for highly inclined and eccentric orbits.
KEYWORDS: *Solar System, Asteroids, Trojans, Asymmetry*

1. Introduction

Up to now, the orbital elements of more than 1380 Trojans are known with reasonably accuracy; 850 of them are around the leading Lagragian point, L_4 and 580 of them are around the trailing point, L_5. There is another prominent difference between the two Lagrangian points in the distribution of the (proper) eccentricities.

In Fig. 1 one can see the difference between the distribution of the Lagrangian eccentricities $(k, h) \equiv e \cdot (\cos\varpi, \sin\varpi)$ of the Trojans around L_4 and L_5. The center of distribution of these orbital elements around the leading point is $(\bar{k}, \bar{h})_{L4} = (0.01186, 0.04637)$ which corresponds to $\bar{e}_{L4} = 0.04794$ and $\bar{\varpi}_{L4} = 75.7°$, in accordance with the Lagrangian equilibrium solutions (the longitude of perihelion of Jupiter is $\varpi_J = 14.3°$). The mean value of the orbital elements (k, h) of the Trojans around the trailing point is $(\bar{k}, \bar{h})_{L5} = (0.03972, -0.02352)$ which corresponds to $\bar{e}_{L5} = 0.0462$ and $\bar{\varpi}_{L5} = -30.6°$. In the latter case, the difference between the mean values of the longitude of pericenters and the longitude of pericenter of Jupiter is not 60 degrees, but approximately 45 degrees, which makes the problem even more interesting.

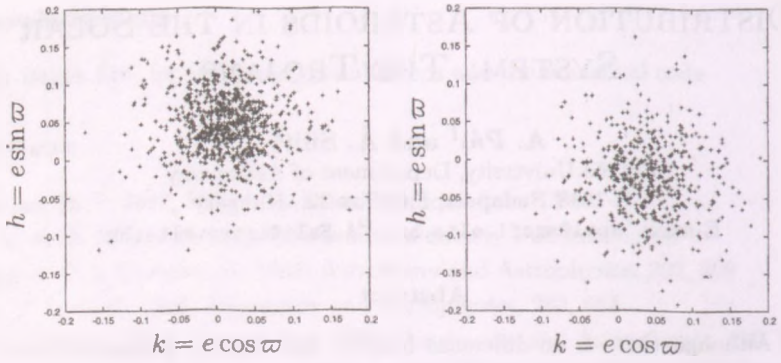

Figure 1: The distribution of the Lagrangian orbital elements $k = e \cdot \cos \varpi$ and $h = e \cdot \sin \varpi$ of the Trojans around L_4 (left) and L_5 (right).

Lag.point	$\bar{e}$	$\bar{\varpi}$
L_4	0.0480	$+75.7°$
L_5	0.0462	$-30.6°$

Table 1: The distribution of the eccentricites for the known Trojans.

One reason why the numbers of the Trojans around the leading point is larger than around the trailing point might be that the size of the stable region around the Lagrangian equilibrium solution is larger. To determine the size of the stable region and to size up the phase space around the Lagranigan points, *stability maps* can be created. While some of the initial orbital elements of a test particle are fixed, the other ones are varied and some indicators are calculated which correspond to the chaotic behaviour (and the stability) of such orbits. Obviously one can vary just two of the initial orbital elements if such stability maps are to be visualized easily, but for a complete study, all of the initial orbital elements should be varied and other kind of visualization might be used (such as a quadratic or Gaussian fit for the stability indicators or something like this).

In the next section, we describe the applied numerical methods which were used in our calculations. In the third section, the results concerning the difference of the phase space structure are presented.

2. Applied methods

In the investigations we used the well-known Lyapunov characteristic exponents (LCEs) to determine the long-term stability properties and the method of Fourier Spectra Numbers (FSNs) as short-term indicator. By the way, the latter was the only one short-time method which showed any difference between the two Lagrangian points. Our model was the Sun – Jupiter – Saturn – test particle restricted four-body system, with real orbital elements of Jupiter and Saturn corresponding to the epoch 1.0 Aug. 1993. For the numerical integrations and the data processing our package `chdet` (with a built-in 6th order RK integrator) and other pipelines have been used.

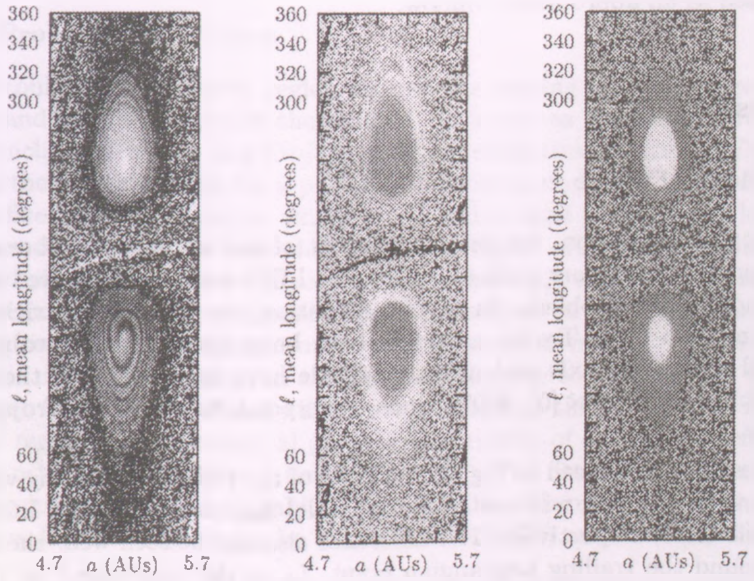

Figure 2: The $a - \ell$ plane of the Trojan belt with the calculated indicators: the plot of the FSNs (without any window), the plot of the FSNs (with Hanning window) and the plot of the LCEs, respectively.

2.1. The Fourier Spectra Number

The FSN is defined as follows. From the numerical integrations of the problem a time series of one of the coordinates $x_n \equiv x(t_n)$ of the test particle is given in discrete equidistant time instances $t_n = n \cdot \Delta t$ $(n = 0, \ldots, N)$. The FSN is defined as the number of the peaks in the power spectrum of x_n. To avoid aliases – between the sampling time Δt, the total length of integration $N \cdot \Delta t$, the orbital period of Jupiter and the libration period of the asteroids (which is approximately 140 years) – the series x_n should be multiplied by a window function before the evaluation of the DFT (discrete Fourier transformation, required by the calculation of the power spectrum). Usually, the Hanning function

$$W_n = \frac{1}{2} - \frac{1}{2}\cos\left(\frac{2\pi n}{N}\right)$$

can be used as an appropriate window.

3. Results

3.1. The $a - \ell$ plane

As a first approximation, the motion of a set of test particles has been integrated and the indicators such as FSNs and LCEs have been derived in the planar and circular problem. In this investigation, the initial eccentricity and inclination of the test Trojan asteroids have been fixed (set to zero), while the initial semimajor axis and mean longitude have been varied in the plane $(a, \ell) \in [4.7\text{AU}, 5.7\text{AU}] \times [0°, 360°]$ which corresponds to the whole Trojan belt of Jupiter.

The results can be seen in Fig. 2: the plots of the FSNs without any window function, with Hanning window (in linear scale) and the plot of the LCEs (in logarithmic scale) respectively. The difference also can be seen well: the stable region around the trailing Lagrangian point, L_5 at the longitude $\ell = 110°$ is even smaller than around the leading Lagrangian point. The length of the numerical integration was 5 000 yr for the FSNs and 1 Myr for the LCEs for all initial conditions. Note that the latter integration time was chosen to be larger than the longest secular period of the Jupiter – Saturn system (which is approximately 300 000 yr).

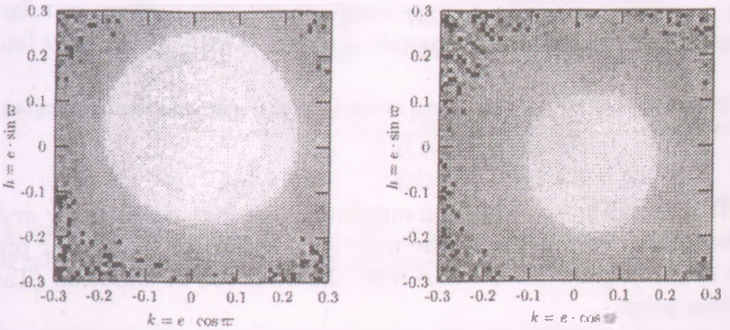

Figure 3: The stability of $k - h$ plane of the Trojans around the leading (left) and the trailing (right) Lagrangian point, calculated with the LCEs.

3.2. Proper eccentricities

As we could see, the stable region around the leading point is larger in the planar and circular model of the motion. It is known that there are several highly inclined Trojans, and Trojans with quite elliptical orbits. In Fig. 3 one can see the stability maps for a planar test body around the L_4 and L_5 points with a fixed inital semimajor axis $a = a_J$ and a fixed initial mean longitude ($\ell = \ell_J \pm 60°$). It can be seen that the stable region around the leading point is larger (similarly in the $a - \ell$ plane). The center (ϵ_c, ϖ_c) and the size of the stable region can be derived easily by fitting a circle around the boundary of such regions. To get an estimation how these boundaries differ from a circle, a quadratic curve (an ellipse) can also be fitted and the eccentricity (or the ratio of the two axes) of this ellipse will be proportional to this difference. The sizes of these regions are the maximal proper eccentricities of a Trojan asteroid with a stable orbit (so hereafter it is denoted as $e^{(P)}$). The fitted parameters of these stable regions can be seen in Table 2. The analitical theory of the motion of the perihelion for the Trojans can be found in Érdi (1978) and Érdi (1979).

3.3. Inclined orbits

To get an almost complete view of the stable regions around the two Lagrangian points, all of the initial orbital elements should be varied simultaneously. This kind of variation is also a requirement for such an investigation because there

Lag.point	e_c	ϖ_c	$e^{(P)}$
L_4	0.044937	+68.202°	0.21192
L_5	0.046216	−45.721°	0.14936

Table 2: The parameters of the stable regions in the $k-h$ plane around the Lagrangian points in the planar model.

are several Trojans with high libration amplitude *and* high eccentricity *and* with quite inclined orbit. Furthermore, the numerical results in the planar problem (Table 2) is not in a good accordance with the observed distribution (Table 1) for the trailing point.

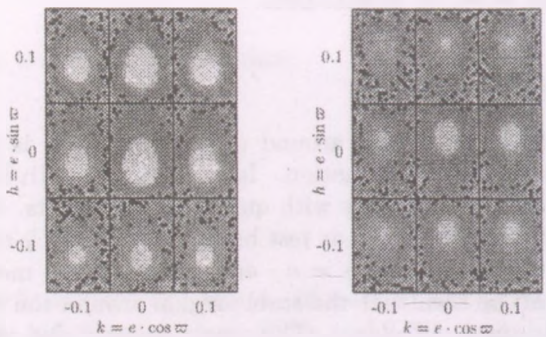

Figure 4: The stability of the $k-h-a-\ell$ section of the phase space around the leading (left) and the trailing (right) Lagrangian point in the planar model, calculated with the LCEs.

First, we varied the initial semimajor axis, mean longitude and Lagrangian eccentricities simultaneously while the inclination was fixed (to join somehow Fig. 2 and Fig. 3). The plots of this investigation can be seen in Fig. 4. As one can see, these results are consistent with the previous ones.

Furthermore, we varied all of the initial orbital elements, such as the axis, longitude, eccentricities (k, h) and Lagrangian inclinations ($q = \sin i \cos \Omega$, $p = \sin i \sin \Omega$) in a grid of $(21 \times 31) \times 5 \times 5 \times 5 \times 5$ points. (21×31) denotes that we made cross-sections in the $a - \ell$ plane through the Lagrangian points, the complete $a-\ell$ map was not calculated as in the previous calculations (see Fig. 2 or Fig. 3). It seems to be a good approximation: the size of the stable regions in both of the cross-sections are correlated.

The maximal investigated inclination and eccentricity were 34° and 0.28. Altough this kind of stability maps cannot be visualized, one can estimate the size and the center of it by fitting a Gaussian or quadratic distribution.

Lag.point	$e_{\rm tot}$	$\varpi_{\rm tot}$
L_4	0.047	+62°
L_5	0.043	−44°

Table 3: The center of the stable region around the two Lagrangian points derived from the stability maps while all of the initial orbital elements were varied.

The center of this distribution can be seen in Table 3. As one can see, the magnitudes of the eccentricities are in a good agreement but there is a difference (about 15°) between the observed and numerically derived mean longitude of the perihelion.

4. Conclusions

Our numerical results show clearly that the size of the stable region around the two Lagrangian points is smaller around the trailing point. It is true for the planar, the circular model and also for larger inclinations. The smaller area of the stable region derived with the LCEs can be an explanation why there are less Trojans observed around the trailing point.

On the other hand, the observed mean value of the longitude of the perihelion differs a bit from the calculated center of the stable region for the trailing point which shows us that further investigations should be done or the fitting or the determination of the stable regions should be revised (the raw data of the LCEs on the grid of $(21 \times 31) \times 5 \times 5 \times 5 \times 5$ points can also be interpreted in another way).

For a further investigation, such stability maps can be created also for the Lagrangian points of Saturn. As it is well-known, the equilibrium solutions are not stable due to the perturbation of Jupiter, but orbits with a high libration amplitude can be stable. Recent works (e.g. Mazari et al., 2002) showed that stable orbits can exist with a libration amplitude of 80°, these calculations are based on the method of the FMA (Frequency Map Analysis). Our first results derived with LCEs and long-term integrations show that the stable orbits have smaller libration amplitudes, approximately $55 - 60°$ (see Fig. 5).

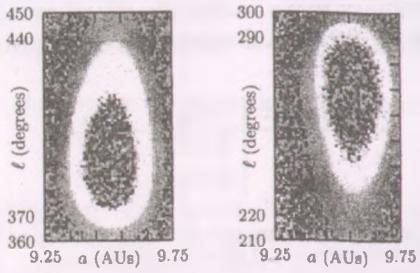

Figure 5: The stable area around the Lagrangian points of the Saturn in the planar model with zero initial proper eccentricities, calculated with the LCEs.

Acknowledgments

The authors thank B. Érdi for his useful comments and ideas related to this topic. The first author (A. Pál) would also thank R. Dvorak for the great time in Potsdam where this work has been started.

References

Érdi, B., 1978, Cel. Mech., **18**, p. 141

Érdi, B., 1979, Cel. Mech., **20**, p. 59

Mazari, F., Tricarico, P. and Scholl, H., 2002, ApJ., **579**, p. 905

On the stability of exoplanets in binary systems

T. Kovács
Eötvös University, Department of Astronomy
H-1518 Budapest, P.O.Box 32., Hungary
E-mail: t.kovacs@astro.elte.hu

Abstract

In this paper we examine the dynamical stability of exoplanets around binary stars depending on the inclination. To explore the structure of the phase space we have applied the method of the relative Lyapunov indicators. We have used the models of the elliptic restricted three- and four-body problem to study the stability of exoplanets in fictitious and real binary systems.
KEYWORDS: *Exoplanets, stability, chaos, binary stars*

1. Introduction

To date 119 extrasolar planets have become known, and the discovery of these objects is still going on. In the last decade one of the most important problem of celestial mechanics was whether exoplanets could exist in stable orbits around their host star. It was not by chance that researchers turned their attention to long-term orbital stability of extrasolar planets. The suitability of life on terrestrial planets is referred to as habitability, and the habitable zone is the region in which a terrestrial planet's atmosphere is able to sustain liquid water (Kasting et al., 1993). The extrasolar planets discovered until now have many different properties comparing with the planets of our Solar System. There are many massive objects in very close orbits to their host stars. About half of the planets move in highly eccentric orbits. However, by the present day observing techniques only giant planets can be detected, and the discovery of Earth-like planets is the task of future space missions (Kepler, TPF).

Most of the stars in our Galaxy belong to double or multiple systems. In the future probably we observe more and more exoplanets around double stars. There are studies on planetary orbits in binary systems (Pilat-Lohinger et al., 2003). These orbits can be classified to three categories (Dvorak et al., 1986). The P-type is when the planet revolves around the center of mass of the stars. It is called planetary motion. The second type is the S-type or satellite motion. In this case the planet orbits one component of the binary. The third group

is the motion around the triangular Lagrangian points L_4 or L_5 of the binary system. In this case for stable motion the mass ratio of the stars of the binary must satisfy the criterion $0 < \mu = \frac{m_1}{m_1+m_2} < 0.0385\ldots$ (with the exception of a few μ values corresponding to resonances between the short and long period librations).

In this paper we examine the dynamical stability of fictitious and real planets in some binary systems, and investigate the structure of the phase space by using fast chaos detection methods.

2. The model

In our computations we used the model of the general three-body problem, but the mass of the third body - the test planet - was thousandth or less than the masses of the stars. All through the investigations the emphases was on the elliptic problem as a special case.

Our aim is to map the phase space structure of the test planet by determining the regular and chaotic regions. From this we can decide whether and where can the planet revolve around the binary for a long time.

The chaoticity of orbits is given by the Lyapunov Characteristic Exponents. The LCEs express that observation how trajectories starting from two close initial points behave in time. When the trajectories are in a regular domain of the phase space, their separation grows linearly in time, when they are in a chaotic region, the separation grows exponentially.

The equations of motion of the problem can be written in the form

$$\dot{x}_i = f_i(x_1, x_2, \ldots, x_n), \qquad (i = 1, 2, \ldots, n).$$

The compute the LCEs we integrate the equations of motion together with their linearized equations

$$\dot{\xi}_i = \sum_{j=1}^{n} \frac{\partial f_i}{\partial x_j} \xi_j \qquad (\text{i=1,2,...,n}),$$

where $\frac{\partial f_i}{\partial x_j}$ is the Jacobian-matrix. Most often it is enough to compute the largest LCE

$$L^1(x_0) = \lim_{t\to\infty} \frac{1}{t} ln \frac{\|\xi(t)\|}{\|\xi_0\|},$$

where $\boldsymbol{x}_0$ is the initial condition vector of the studied orbit, $\boldsymbol{\xi}_0$ is an infinitesimal small initial deviation vector between this orbit and a very close orbit, and $\boldsymbol{\xi}(t)$ is its image after time t. If $L^1(\boldsymbol{x}_0) > 0$, the orbit is chaotic, if $L^1(\boldsymbol{x}_0) = 0$ it is regular (stable). Since the $L^1(\boldsymbol{x}_0)$ can be obtained only as a limit, in practice they compute its finite-time approximation, the finite-time Lyapunov indicator

$$L(\boldsymbol{x}_0, t) = \frac{1}{t} ln \frac{\|\boldsymbol{\xi}(t)\|}{\|\boldsymbol{\xi}_0\|}.$$

It has been shown by Sándor et al. (2000) that the difference between the finite-time Lyapunov indicator of two neighbouring orbits is an efficient indicator of the regular or chaotic behaviour.

This difference is called the relative Lyapunov indicator (RLI) and computed as

$$\Delta L(\boldsymbol{x}_0, t) = |L(\boldsymbol{x}_0 + \Delta\boldsymbol{x}, t) - L(\boldsymbol{x}_0, t)|,$$

where $|\Delta\boldsymbol{x}|$ is the difference vector between the two initial condition vectors. This should be taken small enough in order to reflect local properties of the phase space. During our simulations we used $\|\Delta\boldsymbol{x}\| = 10^{-8}$. It has been shown by Sándor et al. (2004) that selecting $\Delta\boldsymbol{x}$ from a quite large interval $\|\Delta\boldsymbol{x}\| \in [10^{-14}, 10^{-8}]$ does not change essentially the behaviour of the RLI as a function of the time. The RLI shows characteristically different behaviour in different regions of the phase space. In a regular region it oscillates around a small constant value, while in a regular region it increases rapidly as the function of time. This different behaviour becomes evident during a very short integration time, corresponding to 200 periods of the primaries, therefore the RLI is a very efficient chaos indicator.

3. Results

In this section we present our results which show the phase space structure of fictitious and real binary systems. We want to show the dependence of the stability of P-type motion of exoplanets on the inclination. In our study we used the general three- (binary + planet) and the restricted four-body problem (binary + planet + test particle).

The stability computations are based on the method of the relative Lyapunov indicators, which in short time characterize the different types of motions. We investigated the stability of orbits depending on the initial conditions: the semi-major axis and the inclination.

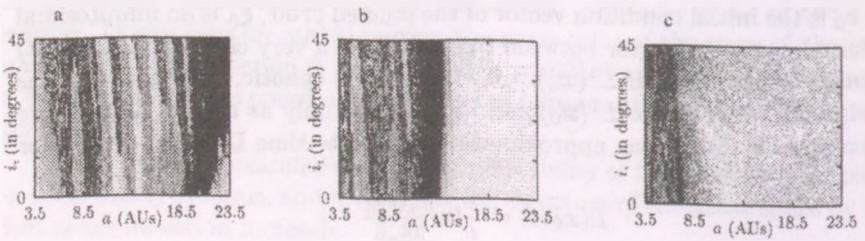

Figure 1: The dynamical structure of the $a - i$ plane of the fictitious elliptic restricted three-body system with $\mu = 0.5$ ($m = 2M_{\odot}$). The mass of the planet is $m_p = 5M_{Jup}$ and the $e_p = 0.05$. The separation of the binaries is $a = 3AU$. Figure *(a)* shows the case $e_b = 0.0$, figure *(b)* $e_b = 0.4$, and figure *(c)* $e_b = 0.8$. The integration time was $t_{max} = 500000day$, the stepsize $\Delta t = 50day$. We can see that when the eccentricity of the binarys grows, the (white) stable region draws near to the primaries.

The other orbital elements of the test particles were zero, that is $\omega = 0, \Omega = 0, \lambda = 0$, (where ω is the argument of periapsis, Ω is the longitude of the ascending node, and λ is the mean longitude) with the exception of the eccentricity, it was in all case 0.05. In our study the two stars had equal masses, so the mass ratio (μ) was 0.5. We examined only this case, because the stability of P-type motion was almost independent of the mass-ratio of the primaries (Dvorak et al., 1989).

3.1. Fictitious systems

Figure 1. shows the results in the (a, i) plane obtained in the model of the elliptic restricted three-body problem. This is not a real system, the separation of the two stars is $a_b = 3AU$. We examined the region of $a \in [3.5, 23.5]AU$. In all computations of our investigation we varied the inclination from 0° to 45° with a stepsize $\Delta i = 0.25°$.

For each pair (a, i) of the initial orbital elements we calculated the avarage value of $\Delta L(a, i)$ for the investigation time, and for a good visualization we used a logarithmic gray scale to show the structure of the phase space. In this map the light regions correspond to regular, while darker regions correspond to chaotic motions.

In the first case (Fig. 1a) the eccentricity of the binary (e_b) was zero, in the second case (Fig. 1b) $e_b = 0.4$, and in the third (Fig. 1c) $e_b = 0.8$. In the circular problem, $e_b = 0.0$, we can see perpendicular (dark) stripes in the (white) stable

region. These unstable segments correspond to mean motion resonances. The two most prominent stripes represent the 2:7 and 2:11 resonances. In Fig. 1b the main difference from Fig. 1a is that from $a = 14AU$ we can find only stable orbits. Figure 1c shows the most eccentric case. Here the stable domain begins from $a = 8.5AU$. We found that the unstable region migrates inwards as the eccentricity grows from $e_b = 0.0$ to $e_b = 0.8$.

Fig. 2 shows again in the $a - i$ plane for different eccentricities of the binary the stable (white) and unstable (dark) regions. In these cases there is a planet with semi-major axis $a_p = 15AU$ and $e_p = 0.1$ around the two stars. We investigate the stability of the test particles around this planet. The circular problem (Fig. 2a) is very similar to the previous circular case. In the restricted four-body problem (Fig. 2.) we can see the effect of the planet ($m_p = 5M_J$) which orbits around the two stars. The difference from Fig. 1. is the little unstable island around $a = 15AU$. This is significant at low inclination, namely to $i = 12°$.

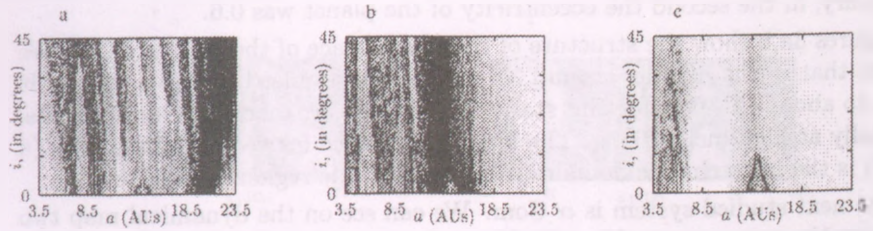

Figure 2: The $a - i$ plane of the restricted four-body system. The mass parameter is again $\mu = 0.5$ ($m = 2M_{\odot}$, and $m_p = 5M_{Jup}$). The distance of the two stars is $a_b = 3AU$. The planet revolves around the binary with semi-major axis $a_p = 15AU$, and with eccenticity $e_p = 0.1$. The integration time was $t_{max} = 500000day$, the stepsize $\Delta t = 50day$. *(a):* $e_b = 0.0$ *(b):* $e_b = 0.4$ *(c):* $e_b = 0.8$. We can see in each figure the small stable island around $a = 15AU$ corresponding to the planet.

3.2. Real binary systems

In this section we present the results of our stability investigation of real binary systems having an exoplanet. The parameters of the studied systems are in Table 1 (Holm & Wiegert, 1999):

Name	$a_b(AU)$	e_b	$m_1(M_\odot)$	$m_2(M_\odot)$	$a_p(AU)$
ϵ Cet	1.57	0.27	1.3	1.3	4.9
α Com	12.49	0.5	1.43	1.37	45
α Cen	23.57	0.516	1.12	0.95	87

Table 1. Data of the investigated systems. Columns 2 and 3 show the semi-major axis of the binary (a_b) in AU and the eccentricity of the binary (e_b). Columns 4 and 5 list the masses of the binaries (m_1, m_2) in solar mass. Column 6 shows the semi-major axis of the planet (a_p) around the binary in AU. Since the mass of the planet is unknown we assumed a planet which have 5 Jupiter masses.

In Figures 3a,b,c and d we plotted the dynamical structure of the $(a - i)$ plane of a test particle in the above systems. In each case we investigated the phase space of the region between the binary and the planet. In these cases we made two simulations. In the first the planet revolved in circular orbit around the binary, in the second the eccentricity of the planet was 0.6.

Figures 3a,b show the structure of the phase space of the system ϵ Cet. We can see that in the middle, around $3.4AU$ there is a quiet big stable region. It grows to about 40°. Outside this stable island there are many regular segments, especially at high inclinations. The largest difference between the two maps (a and b) is the little chaotic domain in the large stable region in Fig. 3b.

The next studied system is α Com. We can see on the dynamical map two larges stable regions around $a = 20AU$, and around $a = 34AU$, the latter is far from the binary and ends about $i = 20°$. In the second half of the investigated region there are many stable orbits at high inclinations, but they are separated from each other.

In Figures 3e,f we present the system α Cen. We found in this map more stable regions in the phase space than in the system ϵ Cet. In the figures the dark stripes correspond to the 1/2, 4/15, 16/69, 3/16 and 11/64 mean motion resonances ($a = 36.41, 56.85, 62.43, 71.99, 76.23AU$, respectively). Looking carefully at Figure 3f, we can see a little unstable region around $37AU$. On the other hand, there is a chaotic island in the middle of the largest stability region. Figures 3c and 3d, and also 3e and 3f are very similar. Hovewer, we found slight differences between them. For example, there is a little unstable island in the stable region (Fig. 3d - around $a = 19AU$, Fig. 3f - around $a = 37AU$). These chaotic orbits correlate with the growth of the eccentricity of the planet which revolve around the binary.

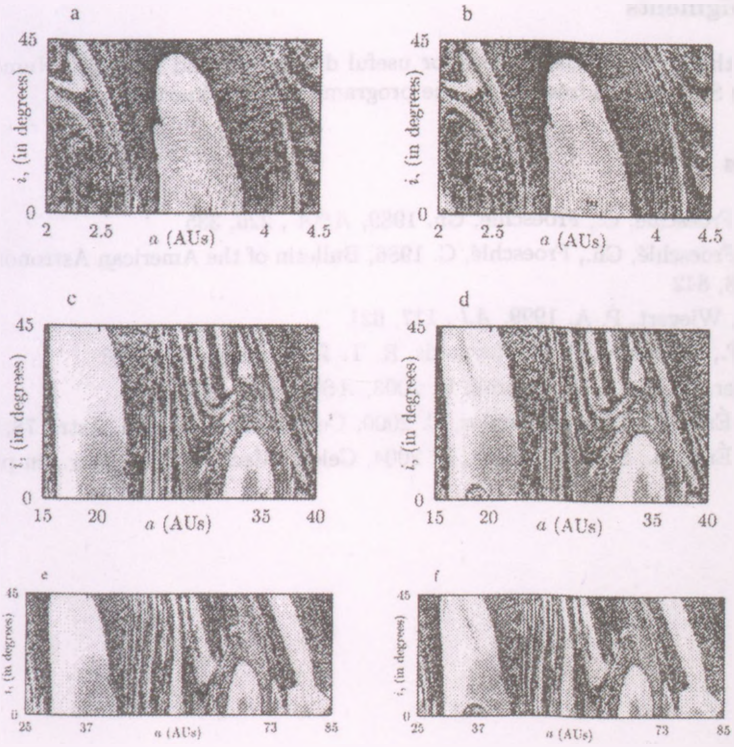

Figure 3: The dynamical structure of the $a-i$ plane of the above-described three systems. The eccentricity of the test particles in all cases was 0.05. The integration time varied depending on the period of the systems. It was in 3a,b (ϵ Cet) 200000 day, in 3c,d (α Com) 1400000 day and in 3e,f (α Cen) 4200000 day. The mass of the planet was $5M_{Jup}$.

In this study we examined the dynamical stability of P-type motion of exoplanets around binaries. The numerical computations were based on the method of relative Lyapunov indicators. We showed that in the fictitious systems the stable region migrates inward when the eccentricity of the binaries increases. We showed also that in the real systems the stability is independent from the eccentricity of the exoplanets. We note that if we search for earth-like planets around these binary systems, we have to do that in the founded stable regions. We have seen that mean motion resonances play an important role in the stability of the terrestrial planets in the studied binary systems.

Acknowledgments

The author thanks Prof. Bálint Érdi for useful discussions and remarks. Numerical help of Áron Süli and András Pál for the programs is also acknowledged.

References

Dvorak, R., Froeschlé, C., Froeschlé, Ch. 1989, *A&A* , 226, 335

Dvorak, R., Froeschlé, Ch., Froeschlé, C. 1986, Bulletin of the American Astronomical Society, 18, 842

Holm, M. J., Wiegert, P. A. 1999, *AJ* , 117, 621

Kasting, J. F., Whitmire, D. P., Reynolds, R. T. 1993, Icarus, 101, 108

Pilat-Lohinger, E., Funk, B., Dvorak, R. 2003, *A&A* , 400, 1085

Sándor, Zs., Érdi, B., Efthymiopoulos, C. 2000, Celest. Mech. & Dyn. Astr., 78, 113

Sándor, Zs., Érdi, B., Széll, A., Funk, B. 2004, Celest. Mech. & Dyn. Astr., in press

Published by the Department of Astronomy of the Eötvös University
Budapest 2004

The N+N Young Researchers' Workshop scheme, initiated and funded by the British Council, aims to provide an opportunity for young researchers to exchange ideas, knowledge and information by coming together in the form of N+N workshops and meetings. Here the term N+N workshop refers to a workshop involving a number of researchers from the UK and an equal number of local researchers. The workshops are followed by real and virtual networking to sustain the contacts made with a view to produce a proposal for longer term externally funded collaboration or applications for further funding.

As at the Department of Astronomy of the Eötvös University we have had a tradition of national workshops of young researchers in astronomy and astrophysics with similar characteristics, the call for meeting proposals by the British Council seemed an ideal way to extend the geographical scope of our meeting, while still maintaining its general format and spirit. Thus, this year's Hungarian Young Researchers' Workshop in Astronomy and Astrophysics was held in tandem with the British-Hungarian N+N workshop, and the contributions are also presented together in this volume.