PADEU | PADEU **15**, 7 (2005)
ISBN 963 463 557

Heating of the Solar Corona: Review

R. Erdélyi

Solar Physics and upper-Atmosphere Research Group, Department of Applied Mathematics, University of Sheffield, S3 7RH, Sheffield, U.K.

E-mail: robertus@sheffield.ac.uk

Abstract
The heating of solar and stellar chromospheres and coronae are one of the key fundamental and yet unresolved questions of modern space and plasma physics. In spite of the multi-fold efforts spanning over half a century including the many superb technological advances and theoretical developments (both analytical and computational) the unveiling of the subtles of coronal heating still remained an exciting job for the 21st century! In the present paper I review the various popular heating mechanisms put forward in the existing extensive literature. The heating processes are, somewhat arbitrarily, classified as hydrodynamic (HD), magnetohydrodynamic (MHD) or kinetic based on the model medium. These mechanisms are further divided based on the time scales of the ultimate dissipation involved (i.e. AC and DC heating, turbulent heating). In particular, attention is paid to discuss shock dissipation, Landau damping, mode coupling, resonant absorption, phase mixing, and, reconnection. Finally, I briefly review the various observational consequences of the many proposed heating mechanisms and confront them with high-resolution ground-based and satellite data currently available.
Keywords: *Sun: atmosphere, Magnetohydrodynamics, Sun: coronal heating*

1 Introduction

The very high-temperature solar atmospheric plasma, in particular in the corona, is mainly confined in magnetic flux tubes. The actual operating heating process that generates and sustains this hot corona has so far defied a quantitative

understanding despite efforts spanning over half a century. In this review paper the most popular and viable heating mechanisms of the solar atmosphere are briefly summarised. We start our journey from the lower chromosphere where mainly a hydrodynamic approach is applicable, discuss the importance of the magnetic field at transition region level, and arrive at the strongly magnetised corona where the magnetohydrodynamic (MHD) descriptions seem to be a reasonable approximation. We address, by recalling the latest results of theoretical and observational studies, the source of plasma heating in the solar (and stellar) atmosphere and how do perturbations propagate from the source and dissipate efficiently resulting in a hot coronal plasmas.

Solar coronal observations go back a very long time, at least three millennia! The Babylonian astronomers reported during a solar eclipse in 1063 BC that, "...the day was turned to night, and fire in the midst of the heaven". The next great leap in coronal research was the spectroscopical discovery of the Sun in the late 19th century. The so-called coronal "green line" at 5303 Å was a real puzzle for astrophysicists for half a century. The observed wavelength of this mysterious spectral line did not match any known elements on Earth and it was concluded that a new element, called coronium, was discovered. However, Edlén in 1939 showed that the coronium line is emitted by highly ionised iron, i.e. Fe XIV, at temperature well over 1 MK. Probably this was the moment when the coronal heating problem was borned. Recent space observations, from Skylab in the 70th through SMM, Yohkoh and in very present times SoHO, TRACE and RHESSI, have investigated the solar atmosphere with unprecedented spatial and temporal resolutions covering wavelengths from (E)UV, through soft and hard X-ray to even gamma rays. These high-resolution imaging and spectroscopic observations contributed to many discoveries in the solar atmosphere. The solar atmospheric zoo, to the best of our knowledge today, consists of features from small-scale X-ray bright points to very large coronal loops (Figure 1a). For an excellent textbook on the corona see, e.g. Golub & Pasachoff (1997). Soon after the discovery of the approximately few MK hot plasma in the solar corona theoreticians came up with various physical models trying to explain the apparently controversial behaviour of the temperature in atmosphere. The key point is the observed distribution of temperature: the solar energy is produced by thermonuclear fusion in the very hot (approximately 14 MK) internal core of the Sun. This vast amount of energy then propagates outwards, initially in the form of radiation (radiation zone) up to about $0.72\mathrm{R}_{\odot}$ and later by convection (convective zone) right to the solar surface (photosphere) continuously cooling the solar plasma. Surprisingly, after reaching its minimum at the top of the photosphere, the temperature starts to rise slowly throughout the entire chro-

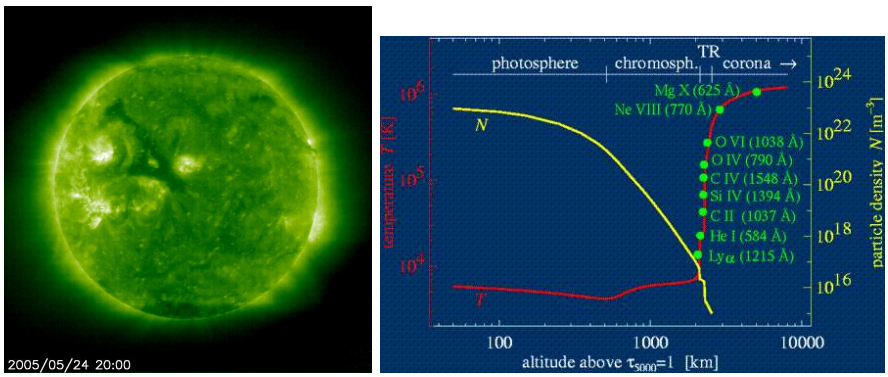


Figure 1: *Left: The very inhomogeneous and dynamic solar atmosphere. Right: Solar atmospheric temperature (red line) and density (yellow line) distributions as a function of the height measure in km. The formation of popular lines for observations is indicated by green. Note the logarithmic scales. Left figure is courtesy of H. Peter.*

mosphere (up to around 20,000 K), followed by a very steep and sharp increase in the narrow transition region (few 100,000 K) up to around 2 MK in the corona (Figure 1b). Although going continuously away from the energy producing solar core, instead of a temperature decrese, the tendency of temperature increase was found (Figure 1b). Maintaining this high temperature requires some sort of input of energy because, without it would cool down by thermodynamic relaxation on a minute-scale. Surprisingly, this non-thermal energy excess to sustain the solar corona is just a reasonably small fraction of the total solar output (see Table 1). It is relatively straightforward to estimates the entire energy budget needed for the solar corona: approximately just a tiny 10^{-4} fraction of the Sun's total energy output is needed giving, at least in theory, a fairly easy task for theoreticians to put forward various mechanisms that could divert 0.01% of the total solar output into heating the corona. The question is today not where does the coronal non-thermal energy come from, but how does it actually get to the corona and how does it dissipate efficiently there.

2 Importance of Atmospheric Magnetism

With increasing spatial and time resolution more and more structures and their dynamics were discovered at the solar surface and in the solar atmosphere.

Table 1: *Table 1: Average coronal energy losses (in erg cm^{-2} sec^{-1}).*

Loss mechanism	Quiet Sun	Active region	Coronal hole
Conductive flux	2×10^5	10^5 - 10^7	6×10^4
Radiative flux	10^5	5×10^6	10^4
Solar wind flux	$< 5 \times 10^4$	$< \times 10^5$	7×10^5
Total flux	3×10^5	10^7	8×10^5

Large-scale structures like sunspots, complex active regions, prominences, coronal loops, coronal holes are observed in great details. On the other hand, the improved resolution allowed to reveal fine structures like the magnetic pores, dark mottles, spicules, supergranular cells, filaments, X-ray and EUV bright points, etc. Since the discoveries of the solar cycle, the Hale's polarity law, the butterfly diagram for sunspots, the cyclic variations in sunspot numbers the role of solar magnetic fields became a central theme. Soon it turned out that these temporal phenomena are linked to the internal generation mechanism of the global solar magnetic field. Skylab observations made it clear for the first time that the x-ray emitting hot and bright coronal regions and the underlying surface magnetic field concentrations are strongly correlated suggesting that coronal heating and solar magnetism are intimately linked (Figure 2). Models of solar (and stellar) atmospheric heating have to comply with the observational facts (Cargill 1993, Zirker 1993). Today it is evident that the solar atmosphere is highly structured and is very likely that various heating mechanisms operate in different atmospheric magnetic structures. In closed structures, e.g. in active regions temperatures may reach up to $8 - 20 \times 10^6$ K, while in open magnetic regions like coronal holes maximum temperatures may only be around $1 - 1,5 \times 10^6$ K. Next, observations also show that temperature, density and magnetic field are highly inhomogeneous. Fine structures (e.g. filaments in loops) may have 3-5 times higher densities than densities in their environment. The fluctuating brightness and the associated fluctuating velocities as opposed to the quasi-static nature of the corona is a far-reaching observational constraint what is not yet modelled on a satisfactory level. There is also little known about how the heating depends on magnetic field strength, structure size (length, radius, expansion) and age and progress is accepted in the near future.

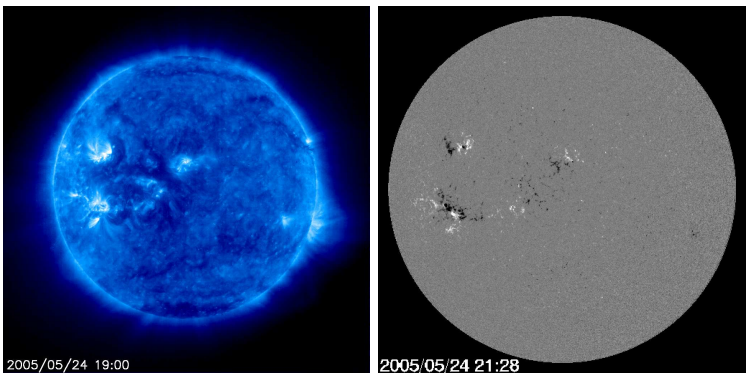


Figure 2: *Left: The solar corona in the 171 Å SoHO EIT spectral line. Right: Concurrently taken image of a SOHO MDI magnetogram at photospheric levels. Magnetic field concentrations coincide with bright patches in the SoHO EIT image indicating the role of magnetic field in the process of coronal heating.*

3 Atmospheric Heating Mechanisms

In order to explain solar (and stellar) atmospheric heating mechanism(s) models have to provide a mechanism or mechanisms that result(s) in a steady supply of energy not necessarily on a steady way. Random energy releases that produce a statistically averaged steady state are allowed for to balance the atmospheric (chromospheric and coronal) energy losses and these models became more viable (Mendoza-Briceno et al. 2004, 2005). Testing a specific heating mechanism observationally may be rather difficult because several mechanisms may operate at the same time. Ultimate dissipation occurs on very small spatial scales, sometimes of the order of a few hundred metres that even with current high spatial resolution satellite techniques cannot (and will not for a while!) be resolved. A distinguished signature of a specific heating mechanism could be obliterated during the thermalisation of the input energy. We should, instead, predict the macroscopic consequences of a specific favoured heating mechanism (Cargill 1993) and confirm these signatures by observations (Aschwanden 2003). For example one could predict the generated flows (see e.g. Ballai et al. 1998) or specific spectral line profiles or line broadenings (Erdélyi et al. 1998).

The heating process is usually split into three phases: (i) the generation of a carrier of energy; (ii) the transport of energy from the locii of generation into the solar atmospheric structures; (iii) and finally the actual dissipation of this energy

Table 2: *Table 2: Summary of the various popular heating mechanisms (see also Ulmschneider, 1998; Erdélyi 2004)*

Energy carrier	**Dissipation mechanism**
Hydrodynamic heating mechanisms	
Acoustic waves ($P < P_{acoustic\ cutoff}$)	Shock dissipation
Pulsational waves ($P > P_{acoustic\ cutoff}$)	Shock dissipation
Magnetic heating mechanisms	
1. Alternating current (AC) or wave mechanisms	
Slow waves Longitudinal MHD tube waves	Shock damping, resonant abs.
Fast MHD waves	Landau damping
Alfvén waves (transverse, torsional)	Mode coupling, res. heating, phase mixing, compressional viscous heating, turb. heating, Landau damping, res. absorption
2. Direct current (AC) mechanisms	
Current sheets	Reconnection (e.g. turbulent or wave heating)

in the various magnetic or non-magnetic structures of the atmosphere. Without contradicting observations it is usually not very hard to come up with a theory that generates and drives an energy carrier. The most obvious candidate is the magneto-convection right underneath the surface of the Sun. Neither seems the literature to be short of transport mechanisms. There is, however real hardship and difficulty in how the transported energy is dissipated efficiently on a time-scale such that the corona is not relaxed thermally. A brief and schematic summary of the most commonly accepted heating mechanisms is given in Table 2 (see also Narain & Ulmschneider 1996; Ulmschneider 1998; Erdélyi 2004). The operating heating mechanisms in the solar atmosphere can be classified whether they involve magnetisms or not. For magnetic-free regions (e.g. in the chromosphere of quiet Sun) one can suggest a heating mechanisms that yields within the framework of hydrodynamics. Such heating theories can be

classified as hydrodynamic heating. Examples of hydrodynamic heating are, among others, e.g. acoustic waves and pulsations. However, if the plasma is embedded in magnetic fields as it is in most parts of the solar atmosphere, the framework of magnetohydrodynamics (MHD) may be the appropriate approach. These coronal heating theories are called MHD heating mechanisms (for reviews see e.g. Browning 1991; Erdélyi 2004; Gomez 1990; Hollweg 1991; Pries & Forbes 2000; Roberts & Nakariakov 2003; and Walsh & Ireland 2003). The ultimate dissipation in MHD models invoke Joule heating or in a somewhat less extent viscosity. Examples of energy carrier of magnetic heating are the slow and fast MHD waves, Alfvn waves, magnetoacoustic-gravity waves, current sheets, etc. There is an interesting concept put forward by De Pontieu et al. (2005) where the direct energy coupling and transfer from the solar photosphere into the corona is demonstrated by simulations and TRACE observations. For a recent review on MHD waves and oscillations see e.g. Roberts (2004). Finally, a popular alternative MHD heating mechanism is the selective decay of a turbulent cascade of magnetic field (Gomez et al. 2000, Hollweg 2002, van Ballegooijen 1986).

Most of the hydrodynamic or MHD heating theories consider the plasma to be collisional. If however, the plasma, whether magnetised or not, is collisionless (and the plasma in the solar corona strictly speaking is!) one has to consider kinetic approaches (for a review see e.g. Scudder 1995). A proper description of the heating mechanisms is cumbersome and would require heavy computations and kinetic codes. Compromising ways to proceed may be the Chew-Goldberger-Low (CGL) closure or the semi-phenomenologic Abraham-Schrauner description of the plasma, where the latter formalism is based on a closure hypothesis of the kinetic equation that is not yet experimentally proven.

Based on the times-cales involved an alternative classification of the heating mechanism can be constructed. If the characteristic time-scale of the perturbations is less than the characteristic times of the back-reaction, in a non-magnetised plasma acoustic waves are good approximations describing the energy propagation; if, however, the plasma is magnetised and perturbation time-scales are small we talk about alternating current (AC-) heating mechanisms, e.g. MHD waves (Roberts 2000, Erdélyi 2001, Roberts & Nakariakov 2003). On the other hand, if perturbations have low frequencies hydrodynamic pulses may be appropriate in a non-magnetised plasma, while if the external driving forces (e.g. photospheric motions) operate on longer times-cales compared to dissipation and transit times very narrow current sheets are built up resulting in direct current (DC-) heating mechanisms in magnetised plasmas (Priest & Forbes 2000).

3.1 Hydrodynamic Framework

After the discovery of the hot solar atmosphere in the early 1940s the model of acoustic waves generated by solar granulation were put forward as energy carrier from the top of the convective zone into the corona. The very steep density decrease causes the sound waves to develop into shocks. These acoustic shocks dissipate their energy causing plasma heating in the solar corona.

After it was discovered that coronal plasma is heavily embedded into magnetic fields the relevance of the hydrodynamic heating mechanisms for the corona part of the atmosphere was re-evaluated. It is believed today that hydrodynamic heating mechanisms could still contribute to atmospheric heating of the Sun but only at lower layers, i.e. possibly in the chromosphere and up to the magnetic canopy (De Pontieu et al. 2004).

For late-type stars with spectral type of F to M acoustic shocks are important heating mechanisms. In early-type stars (O to A) with no convection zone the strong radiation plays the role of acoustic wave generator that steepens into shock waves.

Finally, pulsational waves are mainly prominent in Mira-stars and in other late-type giants where the wave generation is triggered by the κ-mechanism (that is related to the opacity increase of the stellar envelope).

3.2 MHD Framework

At least as a first approximation the plasma is considered frozen-in in the various magnetic structures in the hot solar atmosphere. The magnetic field plays a central and key role in the dynamics and energetics of the solar corona (see Figure 2). High-resolution satellite observations show the magnetic building blocks that seem to be in the form of magnetic flux tubes (Figure 3) in the solar atmosphere. These flux tubes expand rapidly in height because of the strong drop in density. Magnetic fields fill almost entirely the solar atmosphere at about 1,500 km above the photosphere. The flux tubes are shaken and twisted by photospheric motions (i.e. by both granular motion and global acoustic oscillations, the latter being called p-modes). This magnetic flux tubes are excellent waveguides. If the characteristic time of these photospheric footpoint motions is much less than the local Alfvénic transient time the photospheric perturbations propagate in the form of various MHD tube waves (e.g. slow and fast MHD waves; Alfvń waves). The dissipation of MHD waves is manifold: these waves couple with each other, interact non-linearly, resonantly interact with the closed waveguide (i.e. coronal loops) or develop non-linearly (e.g. solitons or shock

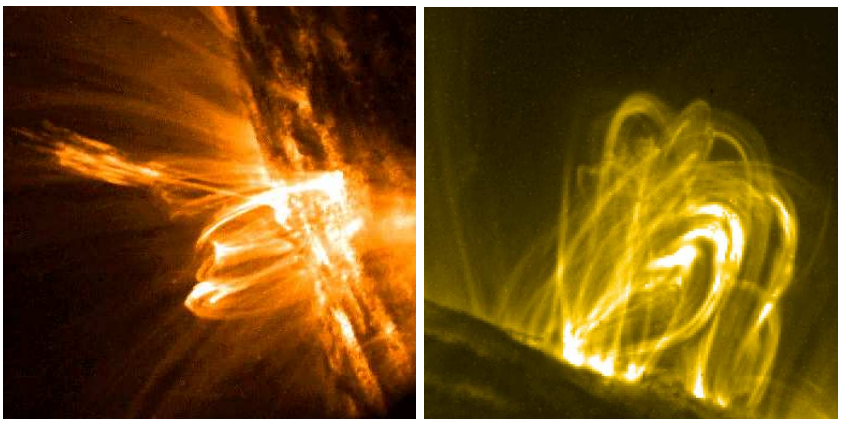


Figure 3: *TRACE images of the highly structured solar corona where the plasma is frozen in semi-circular shaped magnetic flux tubes. Left: The magnetic field in the solar atmosphere shapes the structures that we see, as the emitting gas can generally only move along the field. Sometimes, however, packets of gas are accelerated so much that they can shoot through the magnetic field almost in a straight line. Courtesy Charles Kankelborg. Right: The image shows the evolution of loop system: an increasing number of loops appears in the 1 MK range, probably as they cool from higher temperatures that they reached during the main X-ray flare phase. Courtesy TRACE (http://vestige.lmsal.com/TRACE/Public/Gallery/Images/TRACEpod.html)*

waves can form), etc. For an extensive review on the observations of MHD waves see e.g. Aschwanden (2003), while on theory see e.g. Roberts (2004).

In an inhomogeneous and magnetised plasma there are two particular dissipation mechanisms of MHD waves that received extensive attention in the past decades: resonant absorption and phase mixing. Although there are major theoretical advances on these two particular dissipation mechanisms unfortunatelly we still have only indirect evidences that they may actually operate under solar circumstances. Thanks to the fantastic imaging capabilities of TRACE, plenty of observations of MHD wave damping in coronal loops are available (Aschwanden 2003) and some of these cases may be an excellent candidate of resonant absorption. Further, it is less likely that phase mixing operates in closed magnetic structures, like solar coronal loops.

3.2.1 Mechanism of Resonant Absorption

Let us consider an ideal inhomogeneous vertical magnetic flux tube embedded in a magnetic free plasma such that the Alfvén speed has a maximum at the axis of the tube and the Alfvén speed is monotonically decreasing to zero as a function of the radial coordinate (Figure 4a). Let us suppose that there is a sound wave

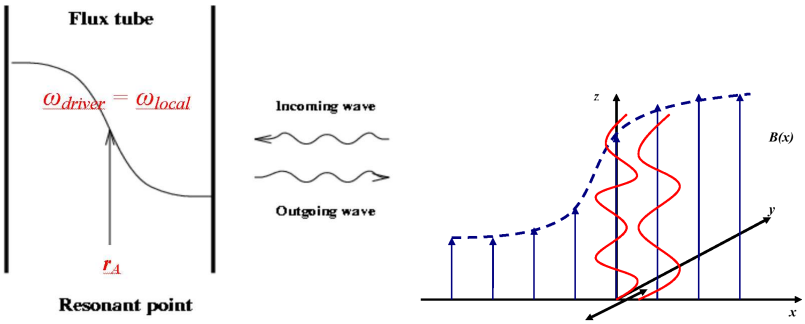


Figure 4: *Left: Schematic sketch of resonant absorption. The incoming driving wave with frequency ω_{driver} is in resonance with local oscillations at the resonant point r_A where the driver's frequency matches the frequency of the local eigenoscillations. Right: Phase-mixing of surface waves caused by gradients in the background magnetic field (or Alfvén speed) where the footpoints of the field lines are shaken in the y-direction.*

continuously impinging horizontally at the boundary of this flux tube. If the phase speed of this impinging (or driving) sound wave matches the local Alfvén speed at a given location of the radius, say at r_A, we say that the driving wave is in resonance with the local Alfvén waves at the magnetic surface at r_A. In ideal MHD this would result in infinite amplitudes of the perturbations resulting in large gradients. However, once the gradients of perturbations become large, one cannot assume any longer the plasma is ideal, i.e. dissipative effects (e.g. resistivity, viscosity) have to be considered at least within the vicinity of such resonant location leading to energy dissipation. Such dissipation, i.e. energy absorption of the driving wave, will result in heating of the plasma converting the energy of the driving wave into localised thermal heating. Resonant absorption, originally considered by plasma physicists as means of excess heating source for thermonuclear fusion, seems to work very well when modelling e.g. the interaction of solar global oscillations with sunspots; when applied to explain the damping of coronal loop oscillations (Ionson 1978, Erdélyi 2001), etc.

3.2.2 Process of Phase Mixing

Heyvaerts & Priest (1983) proposed another interesting mechanism that is in a way fairly similar to resonant absorption. There is a magnetised plasma that is inhomogeneous in the x-direction of the xz-plane where the magnetic field lines are parallel to the z-axis (Figure 4b). We perturb each field line in a coherent (e.g sinusoidal) way in the y-direction. Along each of the field lines an Alfvén wave will develop and will propagating in the z-direction with a speed characteristic to that field line. Since the plasma is inhomogenenous the Alfvén speed at two adjacent field lines is different and neighbouring oscillating field lines will be soon out of phase after some time resulting in large gradients of perturbations. At a given point when the gradients reach a critical value it is not correct anymore to assume that the plasma is ideal and dissipative effects have to be included in the analysis (just like in the case of resonant absorption) resulting in local heating. This dissipation of the initial perturbations is called phase mixing. Phase mixing is an excellent candidate for MHD wave energy dissipation in open magnetic regions like coronal funnels, plumes, solar wind.

3.2.3 Magnetic Reconnection

Finally, if the characteristic time scales of magnetic footpoint perturbations are much larger than the local Alfvénic transit times, magnetic tension is built up gradually involving highly localised current sheets that may release their energy through field line reconnection. This mechanism is called magnetic reconnection. There is plenty of evidence that reconnection works under solar atmospheric conditions at large scales releasing magnetic stresses at highly mixed polarity fields. However, whether this mechanism is viable to heat the solar atmosphere on micro- and nano-scales requires further detailed theoretical investigations and observations. An interesting attempt of solving this debate is to consider the power-law distribution of the various energy releases. It turned out that there is a critical value of the modulus of power-law distribution, approximately equal to 2, what could be measured by observing these small-scale energy releases. If the measured power-law index is greater than 2 that would indicate the solar atmosphere is heated by numerous localised events due to reconnection as a result of e.g. the continuous shuffling of the roots of coronal fields. However, if measurements would show a power index of less than two it is expected that a more global heating mechanism may be responsible for the observed temperature behaviour in the solar atmosphere. Unfortunately observations with the current accuracy could not allow drawing a final conclusion!

We have briefly listed a couple of popular heating mechanisms. We would like to emphasise that most probably different heating mechanisms operate in different solar and stellar structures. It is also likely that these mechanisms work simultaneously and their signatures are present in the high-resolution spectral and imaging data at the same time. Maybe the next-generation space missions like the much awaited Solar-B next year and Solar Dynamics Observatory (SDO) somewhat after or later on Solar Orbiter (probably around the mid of the next decade) will have the capability and capacity to answer the fundamental astrophysical question of: how solar and stellar atmosphere are heated?

4 Stellar Outlook

From the viewpoint of an astronomer the Sun is just a fairly ordinary main-sequence middle-aged low-mass star with a spectral type of G2V with an X-raying corona. Non-degenerate stars of nearly all spectral types show UV and X-ray emission and display evidence of chromospheric and coronal activities as was measured by the OSO-series, the IEU and Einstein satellites. F, G, K and M-stars have chromospheres and often coronae similar to the Sun where radiation is generally attributed to surface convection of these stars. Late giants and supergiants do not really seem to have coronae, while A-stars do not have either chromospheres or coronae. Since chromospheres and coronae of average stars do not receive energy from beyond the stellar atmosphere (except from the T-Tau stars where chromospheric emission originates from mass-infall from accretion disks) it means that stellar atmospheric emission depends solely on the structure of the underlying stellar interior structure. With increasing computer power one may expect that by carefully computing the energetics of surface convection one can predict the chromosphere and corona of a star.

Acknowledgement

Solar physics research at the Department of Applied Mathematics, University of Sheffield is supported by the Rolling Grant of the Particle Physics and Astronomy Research Council (PPARC) of the UK. The author also acknowledges M. Kéray for patient encouragement and is grateful to NSF Hungary (OTKA TO43741).

References

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

Ballai, I., Erdélyi, R. & Ruderman, M.S. *Phys. Plasmas*, **5**, 2264 (1998)

Browning, P.K. *Plasma Phys. and Controlled Fusion*, **33**, 539 (1991)

Cargill, P. in (eds.) J.L. Birch & J.H. Waite, Jr., *Solar System Plasma Physics: Resolution of Processes in Space and Time*, (1993)

De Pontieu, B., Erdélyi, R. & James, S.P. *Nature*, **430**, 536 (2004)

De Pontieu, B., Erdélyi, R. & De Moortel, I. *Astrophys. J.*, **624**, 61 (2005)

Erdélyi, R. in (eds.) J.L. Ballester & B. Roberts, *MHD Waves in Astrophysocal Plasmas*, UIB press, p.69 (2001)

Erdélyi, R. *Astron. & Geophys.*, **45**, p.4.34 (2004)

Erdélyi, R., Doyle, J.G., Perez, E.P. & Wilhelm, K. *Astron. Atrophys.*, **337**, 213 (1998)

Golub, L. & Pasachoff, J.M. *The Solar Corona*, CUP (1997)

Gomez, D.O. *Fund. Cosmic Phys.*, **14**, 131 (1990)

Gomez, D.O., Dmitruk, P.A. & Milano, L.J. *Solar Phys.*, **195**, 299 (2000)

Heyvaerts, J. & Priest, E.R. *Astron. Atrophys.*, **117**, 220 (1983)

Hollweg, J.V. in (eds.) P. Ulmschneider et al., *Mechanisms of Chromospheric and Coronal Heating*, Springer-Verlag, Berlin p.423 (1991)

Hollweg, J.V. *Adv. Space Res.*, **30** 469 (2002)

Ionson, J.A., *Astrophys. J.*, **226**, 650 (1978)

Mendoza-Briceno, C.A., Erdélyi, R. & Sigalotti, L.D. *Astrophys. J.*, **605**, 493 (2004)

Mendoza-Briceno, C.A., Sigalotti, L.D. & Erdélyi, R. *Astrophys. J.*, **624**, 1080 (2005)

Narain, U. & Ulmschneider, P. *Space Sci. Rev.*, **75**, 453 (1996)

Priest, E.R. & Forbes, T. *Magnetic Reconnection*, CUP (2000)

Roberts, B. *Solar Phys.*, **193**, 139 (2000)

Roberts, B. in (eds.) R. Erdélyi et al., *SOHO 13 - Waves, Oscillations and Small-Scale Transient Events in the Solar Atmosphere: A Joint View from SOHO and TRACE*, *ESA-SP*, **547**, 1 (2004)

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

Scudder, J.D. *Adv. Space Res.*, **15**, 181 (1995)

Ulmschneider, P. in (eds.) J.-C. Vial et al., *Lect. Notes in Phys.*, **507**, 77 (1998)

van Ballegooijen, A.A. *Astrophys. J.*, **311**, 1001 (1986)

Walsh, R.W. & Ireland, J. *Astron. Astrophys. Rev.*, **12**, 1 (2003)

Zirker, J.B., *Solar Phys.*, **148**, 43 (1993)

PADEU | PADEU **15**, 21 (2005)
ISBN 963 463 557

Comparison Between Explosive Events of Two Transition Region Lines of Similar Temperatures

B. Ishak, J. G. Doyle and I. Ugarte-Urra

Armagh Observatory, College Hill, Armagh BT61 9DG, N. Ireland

E-mail: bbk@arm.ac.uk, jgd@arm.ac.uk, iuu@arm.ac.uk

Abstract
We present the difference in behavior of two transition region lines at very close temperatures, observed with SUMER/SoHO. N V 1238.82 Å shows a series of explosive events with broadenings mostly shifted to the blue. This is not seen in O V 629.73 Å, the behavior of which remains "quiet" throughout most of the time series.
Keywords: *Sun: transition region - Sun: explosive events - Sun: line profiles*

1 Introduction

It has been recognized that small-scale quiet-Sun transient events, which occur over the entire solar disk, may provide direct evidence for magnetic reconnection, plasma acceleration and heating (Harrison et al., 2003). Using space observatories, many small-scale and globally distributed solar transient events have been reported in the literature. Their potential role in fundamental processes in the solar atmosphere, such as coronal heating, mass ejection, flare activity and wind acceleration, is currently under active investigation.

One of the most discussed and studied occurrences in the literature are explosive events, the term given to transient phenomena observed at temperatures around 1×10^5 K, first discovered and classified as turbulent events and jets by Brueckner & Bartoe (1983). Explosive events are the product of magnetic reconnection (Dere et al., 1984; Porter & Dere, 1991; Innes et al., 1997; Parker, 1998; Wilhelm et al., 1998; Roussev et al., 2001). They tend to occur throughout the quiet-Sun network

where mixed-polarity magnetic features are present (Chae et al., 1998), appearing as bi-directional jets (Innes et al., 1997). Recent observations have shown a birthrate of explosive events of approximately 2500 events per second over the entire Sun, with an average size of 1800 km (Teriaca et al., 2004).

The defining characteristic of an explosive event is a highly broadened red or blue shifted spectral line, affecting either or both wings of the profile. The majority of these non-Gaussian profiles are blue-shifted with velocities up to 150 km s^{-1}. A description on explosive events identification and their general characteristics can be found in Teriaca et al. (2004).

In this work, we have analyzed a series of explosive events in the transition region lines N V 1238.82 Å and O V 629.73 Å.

2 Observation

2.1 SUMER

The Solar Ultraviolet Measurements of Emitted Radiation (SUMER) spectrograph onboard the Solar Heliospheric Observatory (SoHO) has been designed to give measurements at high spatial and spectral resolutions over wide spectral coverage that ranges from less than 500 Å to 1610 Å. Within this wavelength range, spectral imaging of the Sun at short exposure times in the Extreme Ultra Violet (EUV) emission lines permits studies of the essential physical parameters of the solar atmosphere. With SUMER, the opportunity to study and analyze the density of plasma and its temperature, abundances of species, velocity fields, topologies of the plasma structures and their evolution at high temporal resolution of a few seconds has increased tremendously.

Full discussions on the instrumentation and performance of the SUMER spectrograph are given in Wilhelm et al. (1995, 1997) and Lemaire et al. (1997).

2.2 Observational Data

The present data set was observed on 1999 June 1 from 09:13 UT to 11:01 UT with the SUMER spectrograph, pointing at Solar_X and Solar_Y coordinates of (409″,400″). In order to remain at the same location on the Sun, a rotational compensation of 0.75″ has been applied to the X-pointing throughout the observation.

The observing sequence used slit 0.3″ × 120″, exposing for 25 s on detector B. This detector has wavelength ranges of 330 Å to 750 Å for second order lines and 660 Å to 1500 Å for first order. Spectral lines of first and second orders were obtained in the observation. They were N V 1238.82 Å, N V 1242.80 Å, C I 1249 Å and Si II 1251.16 Å for the first order, and O V 629.73 Å and Mg X 624.95 Å for the second. We only consider the lines N V 1238.82 Å and O V 629.73 Å for our analysis.

Table 1 shows the summary of the observation.

Table 1: *Summary of the SUMER spectral data taken on 1999 June 1.*

Date	1999 June 1
Start Time	09:13 UT
End Time	11:01 UT
Pointing (X,Y)	(409″,400″)
Detector	B
Slit	0.3″ × 120″
Exposure Time	25 s
First Order Lines	N V 1238.82 Å, N V 1242.80 Å, C I 1249 Å, Si II 1251.16 Å
Second Order Lines	Mg X 624.95 Å, O V 629.73 Å

2.3 Data Reduction

To reduce the SUMER raw data, several steps are to be followed. In general, the standard SUMER data reduction involves decompression, reversion, dead-time correction, local gain correction, flatfield correction, geometrical distortion correction and radiometric calibration.

However, we do not go through either the decompression or reversion procedure since the Flexible Image Transport System (FITS) files used in this analysis have been decompressed and reversed onboard the space craft. Also, we do not apply the dead-time correction procedure since it is only necessary when the total counts are above 50 000 per second, which is not the case in our data set.

In order to correct for non-uniformities in the sensitivity of the detector, a flatfield correction is applied to the data set by choosing the closest date of observation for the flatfield file to be used in this process. Since we want both the rest position of the line profiles to be on the correct spectral pixel and the slit images straightened, a geometrical distortion correction is also applied to the data set. Note that the radiometric calibration, used to convert the detected intensity unit, i.e. counts per pixel per sampling interval (in our case, second) to physical units, is also not applied, leaving the data set in its original unit of counts per pixel per second.

Due to instrumental broadening, the full width at half maximum (FWHM) of the line profiles needs to be corrected by applying the *con_width_funct_3.pro* routine. All routines used in this analysis are available from the SolarSoftWare (SSW) library. We have applied the Gaussian-fitting procedure to our data set to obtain the amplitude, central position, FWHM and χ^2 of the line profiles.

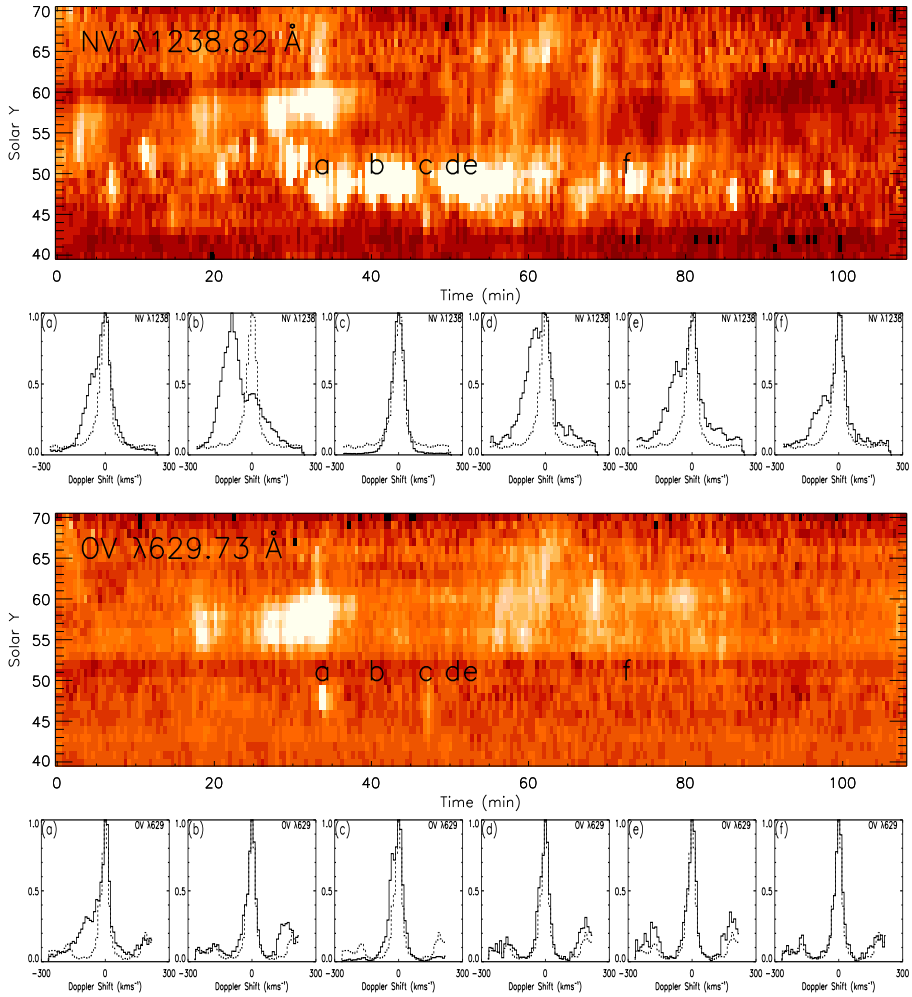


Figure 1: *Image plots of* N V *1238.82 Å and* O V *629.73 Å time series taken on 1999 June 1, starting at 09:13 UT and ending at 11:01 UT, showing the variation in the line width of a single Gaussian fit. Six different temporal locations are selected and labeled (a), (b), (c), (d), (e) and (f), corresponding to the times 09:47 UT, 09:53 UT, 10:00 UT, 10:03 UT, 10:05 UT and 10:26 UT, respectively. The selected profile plots, normalized to unity, are overplotted with the dash-lined quiet Sun profile for comparison.*

3 Results and Discussion

We have studied the two transition region lines, N v 1238.82 Å and O v 629.73 Å, and found that they behave differently so far as some explosive events are concerned, despite being formed at similar temperatures. The peak formation temperature of N v 1238.82 Å is 2×10^5 K while O v 629.73 Å is formed at 2.5×10^5 K. These temperatures are taken from CHIANTI (Dere et al., 1997; Young et al., 2003) data base, using Mazzotta et al. (1998) ionization balance calculations.

For our data analysis, we have performed a single Gaussian fitting for both N v 1238.82 Å and O v 629.73 Å to produce results in terms of amplitude, position, FWHM and χ^2 of the line profiles. χ^2 values are useful in determining the goodness of the fit.

Figure 1 shows an image plot of line width (FWHM) assuming a single Gaussian fit for each time series of N v 1238.82 Å and O v 629.73 Å. For clarity, the wider the width, the brighter it appears in both images. The "black" pixels in the images are due to bad fitting points. Plots of corresponding line profiles that have been normalized to unity for selected locations (a), (b), (c), (d), (e) and (f), at times 09:47 UT, 09:53 UT, 10:00 UT, 10:03 UT, 10:05 UT and 10:26 UT, respectively, are shown below each image. In each profile plot, the quiet Sun profile (dashed line) is overplotted for comparison.

As can be seen, except for profile (c), the line profiles of N v 1238.82 Å show prominent blue shift compared to those of O v 629.73 Å. For example, profile plot labeled (b) of N v 1238.82 Å shows an obvious shift to the blue part of the wavelength, in which a deviation from the quiet Sun profile is apparent, while the corresponding profile plot of O v 629.73 Å indicates little evidence of any such shift.

Although these two profiles are taken at the same location on the Sun and observed at the same time, the dissimilarity in their shape is conspicuous, demonstrating the absence of the explosive events phenomena in the higher temperature line, i.e. O v 629.73 Å, in this observation. The reasons for the discrepancy between the two transition region lines of N v 1238.82 Å and O v 629.73 Å are still under investigation, however, inspection at other locations show similar N v 1238.82 Å and O v 629.73 Å behavior. A possible explanation for the different behavior reported in Figure 1 may be where in the solar atmosphere these events occur.

In follow up work, we will look at the time variability of the chromospherical lines to test this idea. Furthermore, we plan to look at the formation process of the lines, in particular whether there is an effect due to electron density dependent ionization as in Doyle et al. (2005).

Acknowledgement

We wish to thank the SUMER team at Max Planck Institute for Solar System Research (MPS) (formerly Max Planck Institute for Aeronomy (MPAe)) in Lindau, Germany for their help in obtaining the data set. SUMER is part of the Solar and Heliospheric Observatory (SoHO) mission, an international collaboration between ESA and NASA.

This work was supported in part by a PRTLI research grant for Grid-enabled Computational Physics of Natural Phenomena (Cosmogrid). Research at Armagh Observatory is grant-aided by the N. Ireland Dept. of Culture, Arts and Leisure. BI would like to thank E. O'Shea for productive discussions, plus M. Madjarska and L. Xia for their help with the SUMER software.

References

Brueckner, G. E. & Bartoe, J.-D. F., 1983, ApJ, 272, 329

Chae, J., Wang, H., Lee, C., Goode, P. R., Schühle, U., 1998, ApJ, 497, L109

Dere, K. P., Bartoe, J.-D. F. & Brueckner, G. E., 1984, ApJ, 281, 870

Dere, K. P, Landi, E, Mason, H. E., Monsignori Fossi, B. C., Young, P. R., 1997, A&AS, 125, 149

Doyle, J. G., Summers, H. P. & Bryans, P., 2005, A&A, 430, L29

Harrison, R. A., Harra, L. K., Brkovic, A., Parnell, C. E., 2003, A&A, 409, 755

Innes, D. E., Inhester, B., Axford, W. I., Willhelm, K., 1997, Nature, 386, 811

Lemaire, P., Wilhelm, K., Curdt, W., Schule, U., Marsch, E., Poland, A. I., Jordan, S. D., Thomas, R. J., Hassler, D. M., Vial, J. C., Kuhne, M., Huber, M. C. E., Siegmund, O. H. W., Gabriel, A., Timothy, J. G., Grewing, M., 1997, Sol. Phys., 170, 105

Mazzotta, P., Mazzitelli, G., Colafrancesco, S., Vittorio, N., 1998, A&AS, 133, 403

Parker, E. N., 1998, ApJ, 330, 474

Porter, J. G. & Dere, K. P., 1991, ApJ, 370, 775

Roussev, I., Galsgaard, K., Erdélyi, R., Doyle, J. G., 2001, A&A, 370, 298

Teriaca, L., Banerjee, D., Falchi, A., Doyle, J. G., Madjarska, M. S., 2004, A&A, 427, 1065

Wilhelm, K., Curdt, W., Marsch, E., Schuhle, U., Lemaire, P., Gabriel, A.,Vial, J.-C., Grewing, M., Huber, M. C. E., Jordan, S. D., Poland, A. I., Thomas, R. J., Kuhne, M., Timothy, J. G., Hassler, D. M., Siegmund, O. H. W., 1995, Sol. Phys., 162, 189

Wilhelm, K., Lemaire, P., Curdt, W., Schuhle, U., Marsch, E., Poland, A. I., Jordan, S. D, Thomas, R. J., Hassler, D. M., Huber, M. C. E., Vial, J.-C., Kuhne, M., Siegmund, O. H. W., Gabriel, A., Timothy, J. G., Grewing, M., Feldman, U., Hollandt, J., Brekke, P., 1997, Sol. Phys., 170, 75

Wilhelm, K., Innes, E. E., Curdt, W., Kliem, B. , Brekke, P., 1998, ESA SP-421: Solar Jets and Coronal Plumes, 103

Young, P. R., Del Zanna, G., Landi, E., Dere, K. P., Mason, H. E., Landini, M., 2003, ApJS, 144, 135

PADEU | PADEU **15**, 27 (2005)
ISBN 963 463 557

On the behaviour of sunspots in solar plasma: sunspot decay

B. Major

Eötvös University, Department of Astronomy, H-1518 Budapest, Pf. 32, Hungary

E-mail: [1]B.Major@astro.elte.hu

Abstract
There are two main fundamental achievements in connection with statistical studies of solar sunspots and sunspot groups. One of them is that the mean decay rates of sunspot groups are lognormally distributed, the other one is that the decay rate of sunpots is proportional to the relative radius of the spot. The preceding study is based on the Greenwich Photoheliographic Results (GPR), the latter one is based on the Debrecen Photoheliographic Results (DPR). Now the Debrecen Photoheloigraphic Data (DPD) will be used to verify the above achievements and discuss the usefulness of DPD for such statistical studies.
Keywords: *Sun, sunspots, sunspot groups, decay*

1 Introduction

Sunspots are very spectacular features. They appear in the solar photosphere as dark areas compared to their environment. It is commonly known that this is caused by their much larger magnetic field which originates from the bottom of the Sun's convective layer. In this layer a magnetic flux tube of some $10^5 G$ is situated. Some perturbation can lead to arising magnetic flux that can form a pair of sunspots in the photosphere after about one month. It has been shown that just before these appear, the flux tube takes a tree-shape: this is called a magnetic tree. Therefore, sunspots usually appear in groups. After its birth, a sunspot begins to grow, it reaches a maximum area, then it starts to decay and finally disappears. The latter phase is in the focus of our investigation. A review on sunspots can be found in Solanki (2003). Several papers studied the decay law of sunspots. There are two fundamental questions: what is the decay law and why is *that* the decay law.

The first paper on the form of decay curve was Bumba (1963). He concluded that there are two types of curve. One of them is a rapid, exponential curve for the decay phase and typical of non-recurrent groups. The other type is when the rapid decay phase is followed by a linear phase. Later Moreno-Insertis et al. (1988) showed that no such differences exist between recurrent and non-recurrent spot groups and a parabolic decay law is more likely. In these two studies the Greenwich Photoheliographic Results (GPR) was used. Most recently Petrovay and van Driel-Gesztelyi (1997) found evidence for a particular parabolic decay law, specifically the decay rate is proportional to the relative radius of the sunspot. They processed the very detailed Debrecen Photoheliographic Results (DPR) that contains data not only for spot groups but also for individual spots, but unfortunately only for the years 1977 and 1978.

The answer for the second question is an appropriate model that reproduces the fundamental requirements. The spot boundary has to be sharp during the decay phase; this is an obvious observational fact. The relation between the lifetime and the maximum area of a sunspot group is linear (Gnevyshev, 1938; Waldmeier, 1955). The central magnetic flux density has to be more or less constant in time. And finally, the model has to reproduce the decay law. In Petrovay and van Driel-Gesztelyi (1997) Table I. shows the predictions of different sunspot decay models. Among them the turbulent erosion model (Petrovay and Moreno-Insertis, 1997) can satisfy all of the above requirements.

Another interesting achievement is that the average decay rate of sunspot groups follows a lognormal distribution. Assuming a parabolic decay law, this distribution shape comes naturally (Martínez Pillet et al., 1993).

The goal of this paper is to detect the lognormal shape of decay rate distribution and the decay law, using the DPD catalogue. This is a preliminary work, hence we will invoke only some basic statistics in order to decide whether this catalogue is suitable for such an investigation or not. In Sec. 2 the catalogue will be introduced, and selection criteria will be presented. Then we will describe the methods applied to test the lognormality (Martínez Pillet et al., 1993) and to try the parabolic decay curve in Sec. 3. Finally, in Sec. 4 we will give a short discussion.

2 Data

DPD is a catalogue that will contain sunspot data from 1986 up to now. Now the processing of daily solar white-light plates is partly completed: data are available for the years 1986-1988 and 1993-1996 (Győri et al., 2003).[1] Incomplete data are also available for years 1989, 1997 and 1998. In the catalogue, areas and positions of sunspots can be found for every day. This study invokes areas and corresponding observational times in order to determine the decay rates. Because of the lack of day-by-day sunspot identification, we could only use sunspot group data. It has to

[1] ftp://fenyi.solarobs.unideb.hu/pub/DPD

be noted that the catalogue uses NOAA number that corresponds to an *active region* which is not necessarily the same as a *sunspot group*.

Some selection has been made on the database. It was demanded that at least 3 observations had to exist for a group. This is the minimum number of data that is required to calculate the instantaneuos and the average decay rate. Commonly, there is a requirement for the position, because near the solar limb the foreshortening effect causes a large error even if it is eliminated. In this study the absolute value of the distance in longitude from the central meridian (LCM) is less than 65°. Lastly, only those spot groups have been included whose areas reach the value of 10 MSH[2] at least once.

3 Methods

3.1 Lognormality

The density function of a lognormal distribution is

$$\frac{1}{\sqrt{2\pi}\sigma' D} e^{-\frac{(\log D-\mu')^2}{2\sigma'^2}}.$$

Here D is the average decay rate of a sunspot group, derived from a linear fit to area and time data. μ' and σ' are the mean and the standard deviation of $\log D$, respectively. Three methods have been used to test the hypotesis of lognormal distribution.

First, the skewness and the kurtosis of the distribution has been estimated:

$$g_1 = \frac{1}{n}\sum_{i=1}^{n}\left[\frac{(\log D)_i - \mu'}{\sigma'}\right]^3,$$

$$g_2 = \frac{1}{n}\sum_{i=1}^{n}\left[\frac{(\log D)_i - \mu'}{\sigma'}\right]^4 - 3.$$

In the precise form of g_1 and g_2, they are multiplied by a factor that depends on the sample size (n), but as we have $n = 886$, this factor is close to 1 with 10^{-3} error. If a random variable is lognormally distributed, it means that the logarithm of the variable is normally distributed. The skewness and the kurtosis of a normal distribution are equal to zero. The variance of g_1 and g_2 can be estimated from $\mathrm{var}(g_1) = \frac{6}{n}$, and $\mathrm{var}(g_2) = \frac{24}{n}$. Another way is to generate similar samples with Monte Carlo (MC) simulations and calculate the "real" variances of them. Here we have used 250 simulations with sample size $n = 900$. The correspondig values are in Tab. 1. Considering those values, $g_1 = -0.067 \pm 0.082$ and $g_2 = 0.201 \pm 0.16$, so we say the test is positive.

[2] 1 MSH = $10^{-6} A_{1/2\odot}$

Table 1: *Estimation of skewness and kurtosis.*

	value	var	var_{MC}
g_1	-0.067	0.007	0.007
g_2	0.201	0.027	0.023

However, higher moments of a distribution are not robust estimators, especially in the case when we have significant outlier points in the sample. Hence, another standard method has also been invoked. The χ^2-test for goodness of fit leads to $\chi^2 = 22$. For 95% significance level, the corresponding χ^2 value is 27.6 (for 17 degrees of freedom), therefore this test is positive, too. The observed and the estimated density functions are shown in Fig. 1.

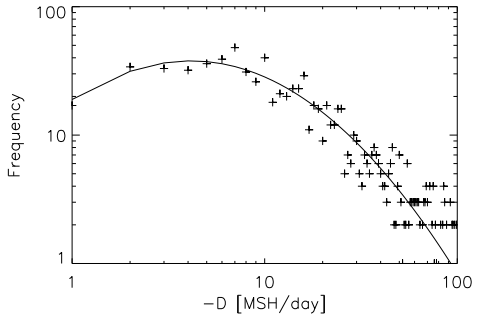


Figure 1: *The observed and the estimated density function: the crosses show the observed histogram of decay rates; the solid line shows the estimated density function scaled up with number of data. Both axis are logarithmic.*

3.2 Decay law

The form of the decay law according to Petrovay and van Driel-Gesztelyi (1997) is

$$D = C_D r / r_0,$$

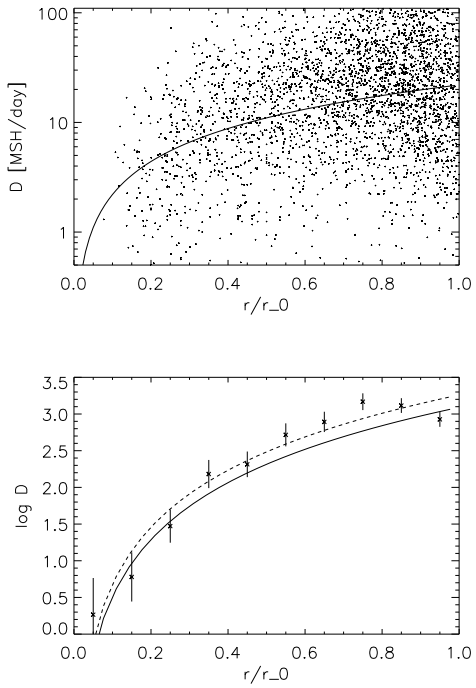


Figure 2: *Calculated instantaneous decay rate via relative equivalent radius are indicated with points for the original (left) and with crosses for the binned (right) data. The dashed line shows the linear fit to the data. The solid line comes from the turbulent erosion model (Petrovay and van Driel-Gesztelyi, 1997). 2σ error bars for the mean are shown.*

where r is the equivalent radius of the spot, r_0 is the maximum equivalent spot radius, and $C_D = 32.0 \pm 0.26$. Here, D means the instantaneous decay rate. This result was derived for individual sunspots using binned data from DPR. We have calculated the intantaneous decay rates with the same method as described in the latter paper but for spot groups using DPD. The results are depicted in Fig. 2. After binning the data, we can fit 3.2 to the data. This resulted in $C_D = 26.0 \pm 1.12$. The errors show that the relative error for our study is larger than for the previous study. However, our value for C_D is a bit closer to that value, which comes from the turbulent erosion model.

4 Discussion

Two statistical investigation has been made using DPD: the lognormal distribution of decay rates and the decay law of sunspot groups. The hypotesis of lognormal distribution was accepted, because both statistical studies led to positive results. This confirms a previous achievement of Martínez Pillet et al. (1993), which was made with GPR. The other investigation was to try the parabolic decay law. From this study, we can conclude that using sunspot group area data of DPD, the parabolic decay law - where the instantaneuos decay rate is proportional to the relative spot radius - can be verified.

We come to the conclusion that DPD is suitable for such statistical investigation. However, previously it has been shown that if we would like to get reliable information about the decay law, we have to use sunspot data. Hence, further effort will intent to identify sunspots day-by-day in the DPD catalogue. Another notable and relevant factor that sunspot groups and active regions (i.e., NOAA regions) are not necessarily the same and mixing them can lead to further errors in statistical studies.

Acknowledgement

The author is grateful for support from the European Solar Magnetism Network founded by the European Comission under contract HPRN-CT-2002-00313. We thank K. Petrovay and G. Mező for useful comments and discussions.

References

Bogdan, T. J., Gilman, P. A., Lerche, I., Howard, R. 1988, ApJ, 327, 451

Bumba, V. 1963, Bull. of the Astr. Inst. of Czech., 14, 91

Gnevyshev, M. N., 1938, Pulkovo Obs. Circ., 16, 36

Győri, L., Baranyi, T., Ludmány, A., Mező, G. 2003, in Solar Variability as an Input to the Earth's Environment (ESA Publ. SP-535), p. 707

Győri, L., Baranyi, T., Ludmany, A., Gerlei, O., Csepura, G., Mező G., 2004, Publ. Debrecen Obs. Heliogr. Ser., 17-19, 1

Martínez Pillet, V., Moreno-Insertis, F., Vázquez, M. 1993, A&A, 274, 521

Mező, G., Baranyi, T., Győri, L. 2003, in Proc. Solar Magnetic Phenomena Summer School and Workshop at Kanz. Sol. Obs., p. 320

Moreno-Insertis, F., Vásquez, M. 1988, A&A, 205, 289

Petrovay, K., Moreno-Insertis, F. 1997, ApJ, 485, 398

Petrovay, K., van Driel-Gesztelyi, L. 1997, Sol. Phys., 176, 249

Solanki, S. K. 2003, A&A Rev., 11, 153

Waldmeier, M. 1955, in Ergebnisse und Probleme der Sonnenforschung, 2. Aufl.

PADEU | PADEU **15**, 35 (2005)
ISBN 963 463 557

Impulsively heated solar coronal plasma

C. A. Mendoza-Briceño

Solar Physics & upper-Atmosphere Research Group, Department of Applied Mathematics, University of Sheffield, Hicks Building, Hounsfield Road, Sheffield, S3 7RH, England, UK

E-mail: c.mendoza@sheffield.ac.uk

Abstract
The dynamics response of the solar coronal plasma in a magnetic flux tube undergoing impulsive heating through the release of localized Gaussian energy pulses near the loop's footpoints is investigated. It was found that when a discrete number of randomly spaced pulses is released, loops heat up and stay at coronal temperatures for the whole duration of the impulsive heating stage provided that the elapsing time between pulses is less than a critical one. For elapsing times longer than this critical value, coronal temperatures can no longer be maintained and the loop apex cools down reaching chromospheric temperatures. For a large number of pulses having a fully random spatio-temporal distribution, the variation of the temperature along the loop is highly sensitive to the spatial distribution of the heating. As long as the heating concentrates more and more at the loop's footpoints, the temperature variation is seen to make a transition from that of a uniformly heated loop to a flat, isothermal profile along the loop length. Concentration of the heating at the footpoints also results in a more frequent appearance of rapid and significant depressions of the apex temperature during the loop evolution, most of them ranging from $\sim 1.5 \times 10^6$ to $\sim 10^4$ K and lasting from about 3 to 10 minutes. This behavior strongly resembles the intermittent variability of coronal loops inferred from SoHO observations in active regions of the solar atmosphere.
Keywords: *Sun: atmosphere, transition region, hydrodynamics, Sun: corona*

1 Introduction

The heating of the solar coronal plasma remains one of the most challenging problem in solar physics. Several heating mechanism have been proposed but recent interest has centered on the idea that flarelike discrete events heat the solar corona (Parker 1988). High resolution observations by space imaging telescope and spectrometers have revealed a variety of very small scale activities at transition region that may serve as building blocks of the heating mechanism (Innes et al. 1997; Harrison 1997; Pérez et al. 1999; Erdélyi et al. 2001). Numerical models aimed at studying the nature of localised energy deposition and their contribution to the coronal heating mechanisms have been proposed e.g. by Sarro et al. (1999), Teriaca et al. (1999), Roussev et al. (2001a,b), Bradshaw & Mason (2003), whereas Walsh & Galsgaard (2000) studied the response of the coronal plasma to dynamical heat input generated by the flux-braiding model. More recently, Mendoza-Briceño, Erdélyi, & Sigalotti (2002) and Mendoza-Briceño, Sigalotti, & Erdélyi (2003) investigated the hydrodynamical behavior of closed magnetic loops undergoing impulsive heating near the footpoints. They found that when a discrete number ($\lesssim 10$) of pulses are injected either periodically or randomly in space with constant elapsing times, the average plasma temperature stays over a million kelvins for the duration of the impulsive heating, with approximate *isothermal profiles* along the upper, hot loop segments. These temperature profiles are consistent with a heating function that decays *exponentially* from the loop's footpoints towards the apex in good agreement with observations of *TRACE* loops (Aschwanden et al. 2001).

In this paper, we report on the response of the coronal loop plasma to spatio-temporal microscale heating near the footpoints. First, we explore the effects of increasing the constant elapsing time between successive energy inputs on catastrophic cooling for a discrete number of randomly spaced pulses and varied total length of the loop. Second, we consider models with a large number of pulses where the energy releases are now applied randomly in space and time as well as in their occurrence at one or both footpoints. This gives rise to loop evolutions in which the heat injections are fully random and asymmetric. Finally, for these latter models we also consider the effects of varying the length of the bottom loop segments along which the localized pulses are randomly distributed. This allows for hypothetical loop models undergoing impulsive heating at the footpoints and across the transition region.

2 Hydrodynamic loop model

Coronal loops are modelled as semicircular magnetic flux tubes of constant cross section anchored in the photosphere. Plasma motion along the tube can be approximately described by solving numerically the standard set of transport equations for mass, momentum, and energy in one space dimension, including the effects of heat conduction, radiative cooling, and heating.

In this paper we assume that the heating deposition has a spatial and temporal dependence given by

$$H(s,t) = h_0 + H_0 \sum_{i=1}^{n} \exp[-\alpha(t-\tau_i)] \times$$

$$\times \left\{ \lambda_{l,i} \exp\left[-\frac{(s-s_{l,i})^2}{\beta^2} \right] + \lambda_{r,i} \exp\left[-\frac{(s-s_{r,i})^2}{\beta^2} \right] \right\} , \quad (1)$$

where t is time, s the curvilinear coordinate along the field line, $h_0 = 3.6 \times 10^{-4}$ ergs cm^{-3} s^{-1} is the uniform background heating rate, $H_0 = 30$ ergs cm^{-3} s^{-1} determines the maximum amplitude of the impulsive heating, $\beta \approx 3.6 \times 10^6$ cm is the spatial width of the heating, and $\alpha = -\ln(0.1)/\Delta t$ so that 90% of the total energy is deposited during a finite time taken to be $\Delta t = 150$ s. We use heat pulses of total energy $E_{tot} \sim 10^{25}$ ergs. The elapsing time τ_i between successive pulses is chosen randomly within the interval $20 \leq \tau_i \leq 190$ s. In addition, the pulses are centered at distances $s_{l,i} = \Delta L \cdot \mathrm{RAN}_{l,i}$ and $s_{r,i} = L(1 - \Delta L \cdot \mathrm{RAN}_{r,i})$ from the left and right footpoints, respectively, where RAN_l and RAN_r define different sequences of random numbers between 0 and 1 allowing for asymmetrical heat injections near the footpoints. Finally, the parameters $\lambda_{l,i}$ and $\lambda_{r,i}$ are randomly chosen to be either 0 or 1 so that the injections may arbitrarily happen at one or both footpoints. For each model calculation four distinct sequences of $n = 5000$ random numbers were employed to determine the parameters τ_i, $s_{l,i}$, $s_{r,i}$, and ($\lambda_{l,i}$, $\lambda_{r,i}$) in Eq. (1). In all models with constant τ_i, the first pulse is always released at the beginning ($t = 0$) of the calculation.

To solve the hydrodynamic loop equations, we use the 1D finite-difference code employed in previous models of impulsively heated loops (Mendoza-Briceño et al. 2002, 2003) and adopt many of the same parameters and assumptions that were made in those earlier simulations with this code. We refer the reader to Sigalotti & Mendoza-Briceño (2003) for a detailed account of the numerical methods and tests. All of our model calculations begin with an initial loop configuration in hydrostatic equilibrium. The initial cool atmosphere (≈ 0.55 MK) is such that the base pressure is always at 0.1 dyne cm^{-2} and is consistent with the value chosen for the background volumetric heating rate h_0. As in previous models, the total length of our basic loop model is $L = 1.0 \times 10^9$ cm (10 Mm), excluding the chromosphere. Additional calculations are presented for loops with total length $L = 5$, 20, and 30 Mm. Appropriate boundary conditions are applied at both footpoints by fixing the density and temperature there to their initial equilibrium values. Given that $p \propto \rho T$, this results in a constant pressure at $s = 0$ and L. In this way, the presence of a deep chromosphere is mimicked by evolving the velocity at the loop ends, thereby allowing for a flow of mass across the footpoints. All models are initialized the same way and the background heating is always maintained in which case we expect the loop model to return to the equilibrium density and temperature associated with this heating rate. Models with a discrete number (= 10) of randomly spaced pulses are all identical except for the constant

elapsing time τ_i between consecutive heat injections, which is varied from 60 to 240 s for the $L = 5$ and 10 Mm loops and to higher values (up to 300 s) for the $L = 20$ and 30 Mm cases. Finally, the model calculations with a large number (up to 1000 or more) of impulsive injections distributed randomly in space, time, and in their occurrence close to the footpoints all start with $L = 10$ Mm; the only variations in the simulations being the seeds generating the four distinct sequences of random numbers employed in Eq. (1) and the length $\Delta L/L$(=0.1, 0.2, 0.3, or 0.5) of the bottom loop segments along which the pulses are injected.

3 Results and Discussion

Mendoza-Briceño et al. (2003) investigated the evolution of a 10 Mm loop when 10 randomly spaced pulses were released near the footpoints over loop segments of length $0.1L$ and with constant elapsing times between successive injections of either 60 or 120 s. It was found that the instantaneous temperature profiles of the evolving loops were characterized by the appearance of localized thermal bumps along their hot coronal segments. Such bumps bear a strong resemblance with the intermittent behavior detected by Patsourakos & Vial (2002) from their analysis of light curves obtained in the O IV and Ne VIII transition region and low corona emission lines, as recorded by *SOHO*/SUMER in a quiet Sun region. The observed bursts exhibit a rather random temporal variation and are presumably due to intermittent energy release followed by its dissipation. In this way, Patsourakos & Vial (2002) concluded that the intermittency of the examined signals is related to well-known types of transition events in the corona such as explosive events, blinkers, and micro/nanoflares.

In this paper we further investigate the existence of a critical value of τ_i by performing further simulations of the 10 Mm loop model along a sequence of increasing constant elapsing times up to $\tau_i = 240$ s. In particular, Fig. 1a shows the evolution of the loop temperature when the time interval between the heat pulses is of 180 s. In this case, the impulsive heating can maintain the overall loop temperature at around 1.5 MK for about 1700 s. As expected the loop temperature achieved is lower compared to models having shorter elapsing times. After 1700 s, the loop cools down and returns to the initial equilibrium atmosphere, which is maintained until the termination of the calculation at 4000 s. The contour plot in Fig. 1a depicts the temperature variation during the impulsive heating phase. It is interesting to note the oscillatory behavior of the 0.5 MK contour line after about 2000 s, when the perturbations are relaxed and the loop returns to its equilibrium state. Similar qualitative trends are found for longer elapsing times up to the critical value of $\tau_i \approx 215$ s. For values higher than this, the loop is maintained at coronal temperatures only for the duration of the first two pulses. Soon after, the top temperature suddenly drops below the initial state towards typical chromospheric values.

When the elapsing time is further increased to $\tau_i = 240$ s, the evolution produced

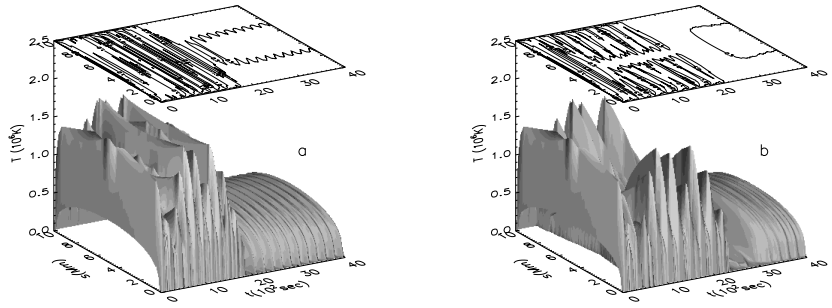


Figure 1: *Evolution of the temperature when* 10 *pairs of heat pulses are injected randomly over a segment of length* $0.1L$ *from each footpoint in a loop model of total length* $L = 10$ *Mm. The time interval between successive injections are of* 180 *s (a) and* 240 *s (b), with the first pair of pulses being released at the beginning of the evolution* ($t = 0$)*. A projected contour plot is shown for the temperature variations, where only contour lines for* 0.5*,* 1.0*, and* 1.5 *MK are shown.*

is qualitatively similar (see Fig. 1b). In this case, however, the cooling of the top loop segments lasts for a longer period ($\sim$ 1900 s) until the last pair of pulses is released (see Fig 2a). The spatial extent of the cool region around the loop apex is clearly evidenced by the temperature contour plot. We estimate that about 20% of the loop undergoes runaway cooling.

The critical elapsing time beyond which runaway cooling occurs is seen to depend on the total loop length. In particular, we find that loop models with the same heating parameters and increasing length experience catastrophic cooling at progressively higher critical elapsing times. For loops of length $L = 5$ Mm, the critical elapsing time is $\tau_i \approx 175$ s. This value rises to about 215, 240, and 263 s for loops of length 10, 20, and 30 Mm, respectively. Loops with the same subcritical value of τ_i and varied length evolve in a qualitatively similar fashion as we may see by comparing the time variations of the apex temperature depicted in Figs. 2*a* and *b* for $\tau_i = 120$ and 180 s. Since the timescale for heat conduction is $\tau_{cond} \propto \rho L^2$, conductive cooling is much more effective in the shorter than in the longer loops. This explains why during the heating stage the apex temperature in the 20 Mm loop oscillates about higher coronal values compared to the 10 Mm model. Similar trends are also seen when comparing the other model evolutions for $L = 5$ and 30 Mm. Larger values of the critical elapsing time are therefore required in the longer loops to allow conductive cooling have enough time to bring the loop apex past the equilibrium point, where radiative losses dominate and eventually induce the runaway cooling of the loop summit.

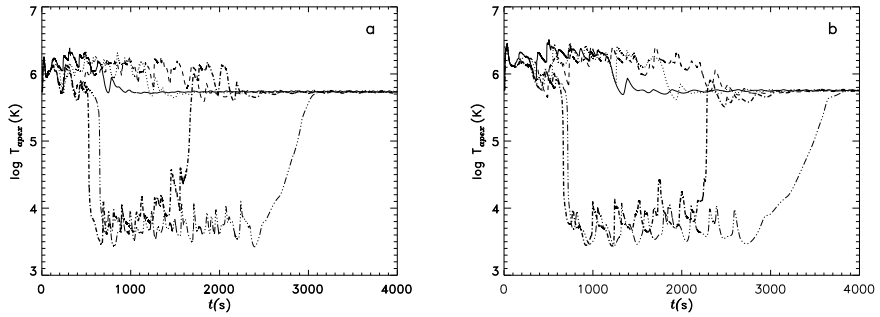


Figure 2: *Time variation of the apex temperature for loop models of total length (a)* $L = 10$ *Mm and (b)* $L = 20$ *Mm, undergoing impulsive heating through the release of* 10 *randomly spaced pulses near both footpoints. The elapsing time* τ_i *between successive pulses is of* 60 *s (solid line),* 120 *s (dotted line),* 180 *s (dashed line),* 220 *s (dot-dashed line), and* 240 *s (triple-dot-dashed line) in (a) and* 120 *s (solid line),* 180 *s (dotted line),* 240 *s (dashed line),* 245 *s (dot-dashed line), and* 260 *s (triple-dot-dashed line) in (b). Catastrophic cooling to temperatures* $\sim 10^4$ *K is evident for* $\tau_i = 220$ *and* 240 *s in (a) and* $\tau_i = 245$ *and* 260 *s in (b).*

We now consider the evolution of the 10 Mm loop model when it is heated by a large number of pulses distributed randomly in space and time near the footpoints. In contrast with the previous models, the pulses may not be simultaneous in the sense that they could appear randomly at one or both footpoints. The randomness in time is obtained by choosing arbitrarily the elapsing time between consecutive injections from the interval $20 \leq \tau_i \leq 190$ s. These values are all below the critical one for which we would expect the loop to undergo catastrophic cooling to chromospheric temperatures. Figure 3 shows the resulting temperature evolution for this loop model up to 15000 s, when more than about 1000 pulses have been released. We may see that the hottest segments of the loop reach temperatures higher than 1.5 MK, which are maintained until the termination of the calculation. When the randomness of the impulsive heating is varied, the evolution undergoes only quantitatively small changes. For instance, an obvious change involves differences in the spatio-temporal distribution of the thermal bumps which certainly modify the instantaneous shape of the temperature profiles when compared at identical evolutionary times. The variation of the apex density and temperature with time for the three model calculations is shown in Fig. 4. One important feature in these plots is the sporadic appearance of temperature drops accompanied by corresponding density rises. That is, during its impulsive heating the loop apex suddenly cools to temperatures $\gtrsim 10^4$ K and reheats to coronal temperatures

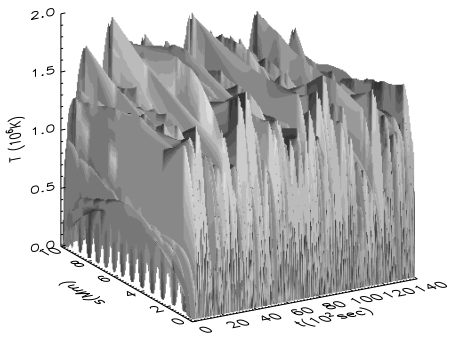


Figure 3: *Evolution of the loop temperature when a large number of heat pulses with a random distribution in space, time, and in their occurrence at one or both footpoints, is injected over a segment of length* $0.1L$ *from the feet in a loop model of total length* $L = 10$ *Mm.*

in a very short timescale. These variations in the local thermodynamic properties may be related to the observed rapid time variability of coronal loops detected in active regions of the solar atmosphere, which involves temperature variations in the interval from $\sim 10^4$ to $\sim 2.7\times10^6$ K (Kjeldseth-Moe & Brekke 1998). The predicted timescales for these rapid variations are of $\sim$ 3–10 min, which are towards the lower end of the range inferred observationally ($\sim$ 10–30 min). This implies that a loop at a given temperature that is missing in one location at a particular time may be present at another time. Also note that the number of the temperature depressions is related to the spatio-temporal dependence of the impulsive heating. We next consider the effects of varying the length ΔL of the bottom loop segments along which the pulses are randomly distributed for the same evolution model of Fig. 3. In particular, Fig. 5 depicts the integrated temperature profiles when $\Delta L/L = 0.1$ (solid line), 0.3 (dotted line), and 0.5 (dashed line). We may see that the form of the temperature variation is highly sensitive to the spatial distribution of the heating. As a consequence of distributing the pulses on a broader region, hotter loops are produced as confirmed by the $\Delta L/L = 0.3$ and 0.5 calculation models. We also notice that the upper, hot loop segments become progressively less flat when the impulsive heating is more broadly distributed along the loop. In particular, the form of the temperature variation for the $\Delta L/L = 0.5$ loop model is similar to that corresponding to a uniformly heated loop as described by Priest et al. (1998) from X-ray observations of the diffuse corona. Moreover, the results also imply that a quasi-isothermal temperature distribution along

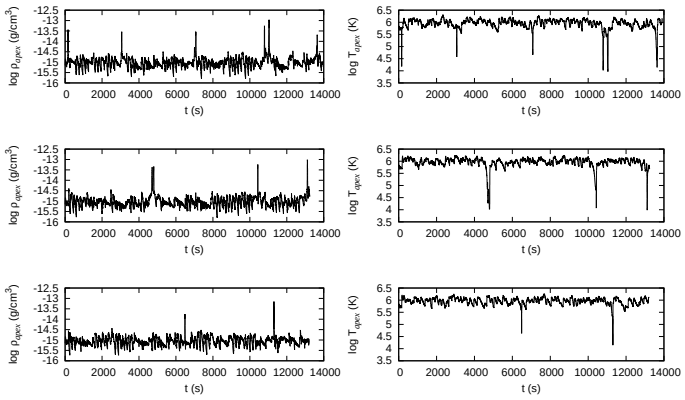


Figure 4: *Variation of the apex density (left panels) and apex temperature (right panels) with time for three loop models of total length* $L = 10$ *Mm each, undergoing impulsive heating through the release of a large number of pulses. The spatio-temporal random distribution of the pulses differ for each model. In all cases, the pulses were injected over segments of length* $0.1L$ *from the loop's footpoints. The upper two panels correspond to the model evolution shown in Fig. 3.*

the loop length is a clear signature of the heating being more strongly concentrated at the footpoints as deduced by Aschwanden et al. (2001) from observations of *TRACE* loops. Finally, we note that when the heating is less concentrated at the footpoints, the occurrence of the rapid temperature depressions strongly diminishes. This result is consistent with the observational lack of detected strong variability in the X-ray and EUV emission lines at higher coronal temperatures (Kjeldseth-Moe & Brekke 1998; Schrijver 2001).

4 Conclusions

In this paper we have described the evolution of a coronal loop model that has been heated impulsively near the footpoints.

It was found that successive microscale energy inputs are quite capable of heating up the loop plasma to typical coronal temperatures, which are, in general, maintained for the whole duration of the impulsive heating. As long as the elapsing time between successive pulses is increased in the 5 and 10 Mm models to ≈ 175 and ≈ 215 s, respectively, the loop heats up towards progressively lower temperatures. For elapsing

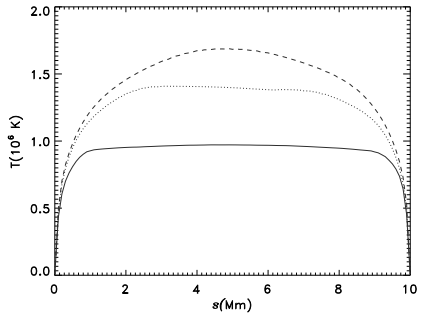


Figure 5: *Time integrated temperature profiles for the model evolution of Fig. 3, when the length of the bottom loop segments over which the pulses are randomly distributed is* 0.1 *(solid line),* 0.3 *(dotted line), and* $0.5L$ *(dashed line). The temperature is given in units of* 10^6 *K. In each case, the time integration was performed over the whole evolution (*$\approx$ 15000 *s).*

times longer than these critical values, coronal temperatures can no longer be maintained as radiative cooling proceeds faster than impulsive heating. As a consequence, the loop apex undergoes catastrophic cooling well below the initial state to typical chromospheric temperatures ($\sim 10^4$ K). The precise value of the critical elapsing time is seen to increase with increasing total loop length. In particular, catastrophic cooling in the 20 and 30 Mm loop models occurs at about 240 and 263 s, respectively.

Finally, loop simulations with a large number of pulses, having a fully random spatio-temporal distribution, confirm previous findings that the plasma would stay at coronal temperatures during the impulsive heating stage. Variations in the randomness of the heat releases produce qualitatively similar evolutions, differing mainly in the spatio-temporal distribution of the localized thermal bumps that appear randomly along the hottest loop segments. The model calculations also predict the occurrence of sporadic and very rapid temperature depressions near the loop apex, which are always accompanied by equally rapid rises of the apex density. These depressions may involve strong temperature variations, most of them from $\sim 1.5\times10^6$ to $\sim 10^4$ K, which may last from about 3 to 10 min, and their number may be sensitive to the details of the spatio-temporal distribution of the microscale heating. This behavior may be related to the observed rapid time variability of coronal loops inferred from *SOHO*-CDS observations in active regions of the solar atmosphere (Kjeldseth-Moe & Brekke 1998; Schrijver 2001). Moreover, when the pulses are less concentrated near the loop's footpoints, the evolution produces hotter loops and progressively less flat temperature profiles in the upper parts of the loop along with an appreciably reduced number of

the temperature depressions. This latter feature is consistent with the observational lack of strong variability at very high coronal temperatures (Kjeldseth-Moe & Brekke 1998; Schrijver 2001).

Acknowledgement

The author thanks to the British Council and the Particle Physics and Astronomy Research Council (PPARC), UK, for the financial support.

References

Aschwanden, M. J., Schrijver, C. J., & Alexander, D. 2001, ApJ, 550, 1036

Bradshaw, S. J., & Mason, H. E. 2003, A&A, 407, 1127

Erdélyi, R., De Pontieu, B., & Roussev, L. M. 2001, in ASP Conf. Ser. 223, Cool Stars, Stellar Systems, and the Sun: 11th Cambridge Workshop, ed. R. J. García López, R. S. Polidan, & R. W. Pogge (San Francisco: ASP), 619

Harrison, R. A., & Hood, A. W. 2002, A&A, 392, 319

Innes, D. E., Inhester, B., Axford, W. I., & Wilhelm, K. 1997, Nature, 386, 811

Kjeldseth-Moe, O., & Brekke, P. 1998, Sol. Phys., 182, 73

Mendoza-Briceño, C. A., Erdélyi, R., & Sigalotti, L. Di G. 2002, ApJ, 579, L49

Mendoza-Briceño, C. A., Sigalotti, L. Di G., & Erdélyi, R. 2003, Adv. Space Res., 32(6), 995

Parker, E. N. 1988, ApJ, 330, 474

Patsourakos, S., & Vial, J.-C. 2002, A&A, 385, 1073

Pérez, M. E., Doyle, J. G., Erdélyi, R., & Sarro, L. M. 1999, A&A, 342, 279

Priest, E. R., Foley, C. R., Heyvaerst, J., Arber, T. D., Culhane, J. L., & Acton, L. W. 1998, Nature, 393, 545

Roussev, I., Galsgaard, K., Erdélyi, R., & Doyle, J. G. 2001a, A&A, 370, 298

Roussev, I., Doyle, J. G., Galsgaard, K., & Erdélyi, R. 2001b, A&A, 380, 719

Sarro, L. M., Erdélyi, R., Doyle, J. G., & Pérez, M. E. 1999, A&A, 351, 721

Schrijver, C. J. 2001, Sol. Phys., 198, 325

Sigalotti, L. Di G., & Mendoza-Briceño, C. A. 2003, A&A, 397, 1083

Teriaca, L., Doyle, J. G., Erdélyi, R., & Sarro, L. M. 1999, A&A, 52, L99

Walsh, R. W., & Galsgaard, K. 2001, in Solar Encounter: Proc. First Solar Orbiter Workshop, ed. B. Battrick & H. Sawaya-Lacoste (ESA SP-493; Noordwijk: ESA), 427

PADEU | PADEU **15**, 45 (2005)
ISBN 963 463 557

MHD Simulation of Solar Coronal Mass Ejection

Cristiana Dumitrache

Astronomical Institute of Romanian Academy

E-mail: crisd@aira.astro.ro

Abstract
Time dependent magnetohydrodynamic computations on three solar radii, starting with a current sheet initial configuration, is presented. A prominence formation and eruption, trigger CMEs in 'raffales', with a coronal streamer formation as last result.
Keywords: *Sun, CME, MHD, numerical simulation*

1 Introduction

One of the amazing feature of the solar activity are the explosive phenomena, especially that which affect directly our planet life. Thecoronal mass ejections (CMEs) are huge magnetized bubbles of gas ejected from the Sun into the planetary space, which add to the solar wind. Large spatial scales are involved during a such process (average width of 47deg). The timescales of CMEs events range between several minutes to several hours. The CMEs are impressive phenomena: their velocities could rich from 10 to 2100 km/s, their masses range between $2 \times 10^{14} - 4 \times 10^{16}$ g, developing kinetic energies between 10^{29} to 10^{31} ergs.

Observationally, the CMEs are related to prominences eruptions or to coronal streamers disruptions, but also to big solar flares. CMEs are driven by the magnetic field. A CME appears in the active zones with closed magnetic field configurations stretched above the polarity inversion line. The material is coronal: once the CME produces, it observes a exhaustion of the mater back the bubble. Before the disruption of the material, strong movements are observed in zone. The ejection rate varies from a phenomenon to another, depending on the causes.

Many models and reviews were given within the past years and this topic is still on the top of the actuality.

Wu et al. (2000) distinct three CME initiation processes: streamer destabilization due to increase of currents via increase of axial fields, photospheric shears and plasma flow induced at the boundary region of a streamer and coronal hole.

Klimchuk (cited by Poedts et al., 2002) reviewed the theoretical model for CME initialization as follow:

(1) Directly driven models

(a) Thermal blast - characterized by a sudden release of thermal energy (Dryer, 1982; Wu, 1982) - CME associated with flares

(b) Dynamo models - real-time stressing of the magnetic field involves the rapid generation of coronal magnetic flux (Chen, 1989)

(2) Storage and release models - three classes in which a slow buildup of magnetic stress precedes the eruption

(a) Mass loading models - the field is loaded with mass (Low, 1999)

(b) Tether release models - the strain increase on a decreasing number of tethers (Forbes & Isenberg, 1991)

(c) Tether straining models - total stress increases (Antiochos et al., 1999)

A 3-D numerical simulation of CMEs by coupled coronal and heliospheric model was performed by Odstrcil et al. (2002). The coronal model is based on the 3-D resistive MHD equations solved by a semi-implicit finite-difference scheme.. The output of this model consists of a temporal sequence of the MHD flow parameters, which are used as boundary conditions for the heliospheric solutions.

Forbes (2002) conclude that a CME is triggered by the disappearance of a stable equilibrium as a result of the slow evolution of the photospheric magnetic field. This disappearance may be due to a loss of ideal-MHD equilibrium or stability such as occurs in the kink mode, or to a loss of resistive-MHD equilibrium as a result of magnetic reconnection.

Our simulation is based on a prominence formation in a current sheet and its evolution. The current sheets naturally appear in the solar atmosphere. They are linked to the lines of magnetic polarity inversion line. They are places where prominences form and seldom their are at the base of a coronal streamer.

2 2-D numerical simulation

The MHD equations are solved with SHASTA method used by Alfven code, developed by Weber (1978). This code was described by Forbes & Priest (1982) and also by Dumitrache (1999).

$$\frac{\partial \rho}{\partial t} + \overrightarrow{\nabla}(\rho \cdot \overrightarrow{v}) = 0 \tag{1}$$

$$\rho[\frac{\partial \overrightarrow{v}}{\partial t} + (\overrightarrow{v} \cdot \overrightarrow{\nabla})\overrightarrow{v}] = -\overrightarrow{\nabla}(p) + (\overrightarrow{B} \cdot \overrightarrow{\nabla}) \cdot \overrightarrow{B} - \overrightarrow{\nabla} \cdot (\frac{B^2}{2}) + \rho \cdot \overrightarrow{g} \tag{2}$$

$$\frac{\partial \overrightarrow{B}}{\partial t} = \overrightarrow{\nabla} \times (\overrightarrow{v} \times \overrightarrow{B}) + \eta \overrightarrow{\nabla}^2 \cdot \overrightarrow{B} \tag{3}$$

$$\frac{\rho^\gamma}{\gamma - 1} \frac{d}{dt}(\frac{p}{\rho^\gamma}) = -\overrightarrow{\nabla}(k\overrightarrow{\nabla}T) - \rho^2 Q(T) + j^2/\sigma + h\rho \tag{4}$$

$$p = \rho T \tag{5}$$

The initial configuration is:

$$B_x = \begin{cases} \sin(\frac{\pi z}{2w}), \text{ for } |z| \leq w \\ 1, \text{ for } |z| > w \end{cases} \qquad B_z = 0 \tag{6}$$

with $v_x = 0$, $v_z = 0$, where $w(= 0,25$ from the computation grid) is the thickness of the sheet , $\rho = 1$ and $p = \rho T$.

The boundary conditions imposed on 49x97 mesh:

-at top (x=1)

$$\frac{\partial B_z}{\partial x} = \frac{\partial B_z}{\partial x} = 0 \tag{7}$$

$$\frac{\partial B_z}{\partial x} = -\frac{\partial B_x}{\partial z} \tag{8}$$

- at right (z=1)

$$\frac{\partial B_x}{\partial z} = \frac{\partial B_z}{\partial z} = 0 \tag{9}$$

$$\frac{\partial B_z}{\partial z} = -\frac{\partial B_x}{\partial x} \tag{10}$$

- at left - the symmetry axis (z=0)

$$\frac{\partial B_x}{\partial z} = B_z = 0 \tag{11}$$

- at bottom (x=0)

$$\frac{\partial B_x}{\partial x} = \frac{\partial B_z}{\partial x} = 0 \tag{12}$$

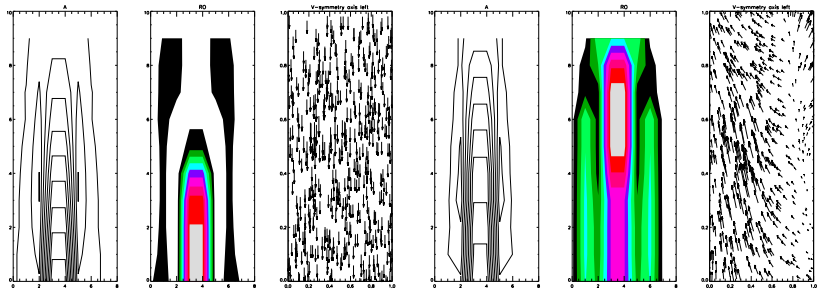


Figure 1: *Left: t=0.026; Right:t=0.031*

3 CMEs "en raffales"

Starting with a current sheet initial configuration and with $\beta = 0.5$ and $Rm = 10^3$, we performed this numerical experiment on 3 solar radii.

We obtain a prominences configuration, after a cooling process in the sheet, at the Alfven time $t = 0.026$. The temperature in the sheet is about 5000 K. The Figure 1 (left panel) displays the magnetic field lines at left, the density at middle and a half of grid for the velocity vector field at the right. The symmetry axis of the vector field is at left. At this stage, downward motions are registered.

The first CME starts at $t = 0.031$ (Fig.1, right), very impulsively with a $v = 1245$ km/s.

At $t = 0.036$ (Fig.2, left), the material is pushed to the lateral side of the current sheet and upward. Its velocity does dot exceed 507 km/s. The exhaustion of the matter is observed below the bubble, while the magnetic field picture displays plasma insulation. Lateral legs of the old loop could be seen steel at the bottom of the figure 2 (left), while at $t = 0.040$ (Fig.2, right) the magnetic loop is reformed, but plasma push strongly at the lateral side of the sheet. The legs of the old loop are ejected with about 566 km/s and this could be considered the second CME.

At $t = 0.042$ (Fig.3, left) gravitational instabilities appear again. On the feet of the old prominences, the matter is upward moved: as result a new CME consisting from the leg's material, appears again, but more impulsively. The velocity attain 715 km/s this time. From now, and continuing with the legs' elongation at $t = 0.043$ (Fig.3, right), to $t = 0.053$ (Fig.4) the matter is expulsed en raffales. This could be considered as the third CME.

The process of loop reforming replies at $t = 0.077$ (Fig.5, left). During this time, after each CME, the temperature in the sheet increased very much and attains now 10^6K, so we have a hot coronal loop. At $t = 0.081$ (Fig.5, right) we assist at a new magnetic reconnection, when the field lines open and a neutral O point from below

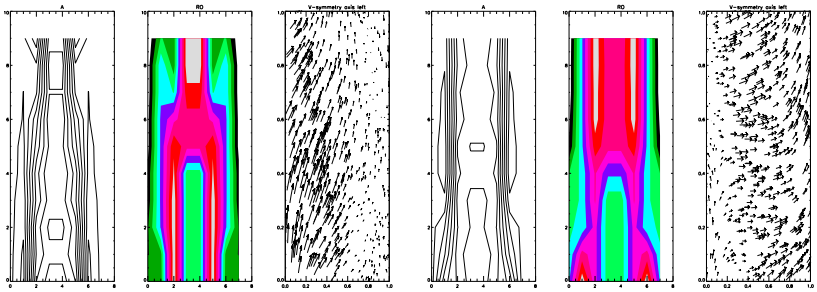


Figure 2: *Left: t=0.036; Right:t=0.040*

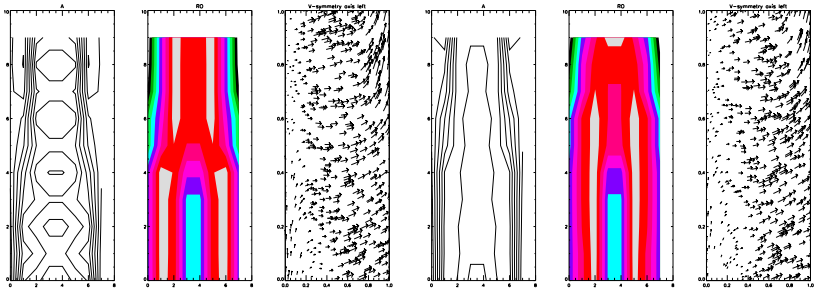


Figure 3: *Left: t=0.042; Right:t=0.043*

the bubble of plasma which elongated to start in a new CME. This new CME produce at $t = 0.085$ (Fig.6, left), but with low velocity (243 km/s). The temperature of the ejected bubble reaches 2×10^6 K. After this CME, a coronal helmet streamer installed, at $t = 0.096$ (Fig.6,right), with a temperature of 9×10^5 K at base and 1×10^6 K at top.

The first CME produced after 17.9 h after start of the simulated process, where the prominence formed in about 15 h. The second CME appeared after 23.2 h. New mass collected in the sheet and a hot loop reformed. The lateral material pushing has determined gravitational instabilities and new CMEs produced, at 24.36 h and 49.3 h. After 6.3 h from the last CME, the structure evolved in a helmet streamer configuration, which had its proper dynamics that our simulation follow till the streamer dissolution.

The regime of the velocities is displayed in Figure 7. The dot curve shows the modules of the velocities on the bottom of the computational grid, the doted curve represents the velocities at the middle of the grid and the solid one represents the top

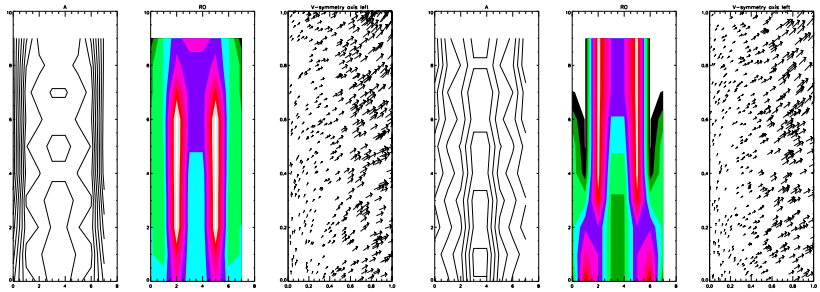


Figure 4: *Left: t=0.047; Right:t=0.053*

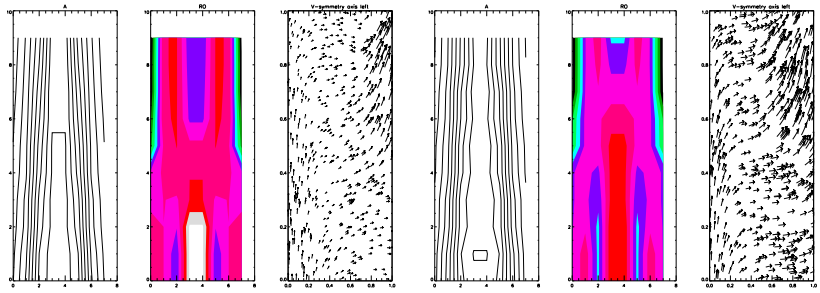


Figure 5: *Left: t=0.077; Right:t=0.081*

velocities, all of them considered on the symmetry axis.

The top-left side windows displays the modules of the velocities on all the computational grid. The axis with values from 0 to 15 represents moments of time (sixteen values). The top of the histograms are shaded at the CMEs start moments; also the correspondant points on the curves are shaded too. The solid curve with circle (corresponding to the top of the computational grid) is the most relevant for the velocity evolution during the "raffales". It seems the firsts two CMEs are impulsive ones, but the lasts are only a remanent process with slow movements of the matter.

4 Conclusions

Our simulation reveals a phenomenon reminding that observed by SOHO on 27 March 2001, the so-called "cannibal coronal mass ejections". In our simulation, a prominence structure, formed in a current sheet, evolved in CME disruptions "en raffales".

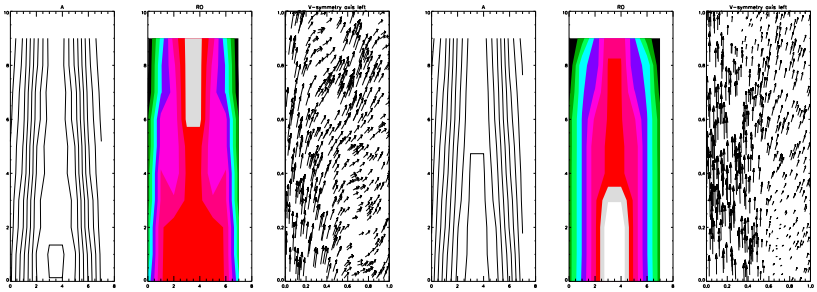


Figure 6: *Left: t=0.085; Right:t=0.096*

These transient phenomena "en rafalle" are expression of the small-scale reconnections in the current sheet. The reconnections produced between two open field lines from both sides of a streamer current sheet and created a new closed field line (which becomes part of the helmet with a prominence at the base) and a disconnected field line, which moves outward. The CMEs are formed by plasma collected in the sheet that is swept up in the trough of the outward-moving field line.

In this numerical experiment the first CMEs are impulsive, but the last ones moves slow, proving that the energy storage was exhausted and the sheet will accomplish a new equilibrium state. This new equilibrium permits that a new solar feature forms on this site: a new stage of evolution concerning in a coronal streamer. We account that the new feature is possible to form in our simulated case, with two shoks "raffales" and two slow CMEs later, and not in the "cannibal" case, where all the matter and energy is probable exhausted. On the other hand, the slowness is probable do to the second and third CMEs which consist in the legs of a loop or in the matter from lateral side of the sheet.

The streamer evolution will be treat in another paper.

References

Antiochos, S.K.; DeVore, C.R; Klimchuk, J.A. 1999, ApJ, 510, 485

Chen, J. 1989, ApJ 338, 453-470

Dryer, M. 1982, Space Sci.Rev., 33, 233

Dumitrache C. 1999, Romanian Astronomical Journal, 9,139

Forbes, T.G.; Priest, E.R. 1982, Solar Phys., 81, 303

Forbes, T.G., Isenberg, P.A. 1991, ApJ, 373, 294

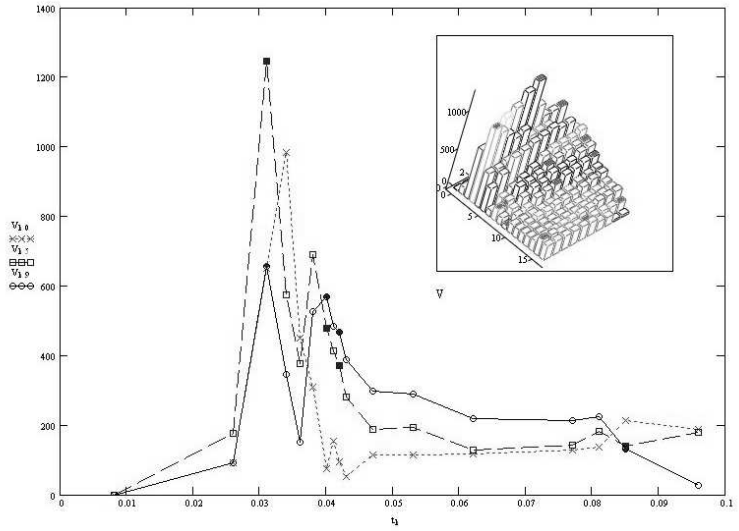


Figure 7: *Velocity evolution during all process (see text exolanations).*

Forbes, T.G. 2002, Bulletin of the American Astronomical Society, 34, 751

Low, B.C. 1999, in *Solar Wind Nine*, ed. Habbal S.R., Esser R., Hollweg J.V., Isenberg P.A., AIP, Woodburry, NY, 109

Odstrcil, D.; Linker, J. A.; Lionello, R.; Mikic, Z.; Riley, P.; Pizzo, V. J.; Luhmann, J. G. 2002, in *Solar variability: from core to outer frontiers. The 10th European Solar Physics Meeting, 9 - 14 September 2002, Prague, Czech Republic*, ed. A. Wilson. ESA SP-506, Vol. 1. Noordwijk: ESA Publications Division, 95

Poedts, S.; van der Holst, B.; de Sterck, H.; van Driel-Gesztelyi, L.; Csík, A.; Milesi, A.; Deconinck, H. 2002, in *Proceedings of the Second Solar Cycle and Space Weather Euroconference, 24 - 29 September 2001, Vico Equense, Italy*, ed. Huguette Sawaya-Lacoste. ESA SP-477, Noordwijk: ESA Publications Division, 263

Weber, W. 1978, PhD Thesis, Groningen

Wu, S.T: 1982, Space Sci. Rev. 32, 1150

Wu, S.T; Guo, W.P.; Plunkett, S.P.; Schmieder, B.; Simnett, G.M. 2000, Jurnal of Atmospheric and Solar-Terrestrial Physics, 1489

PADEU | PADEU **15**, 53 (2005)
ISBN 963 463 557

The Sun as a laboratory for turbulence theory: the problem of anomalous diffusion

K. Petrovay

Eötvös University, Department of Astronomy, H-1518 Budapest, Pf. 32, Hungary

E-mail: K.Petrovay@astro.elte.hu

Abstract
The solar atmosphere offers a unique possibility to study tuubulent motions under conditions presently unattainable in laboratory experiment or even numerical simulations. This short review will focus on one controversial issue in turbulence theory, on which some light can be shed by solar observations: anomalous turbulent diffusion.
Keywords: *Sun: interior, MHD, Sun: rotation*

1 Introduction: Random walk and diffusion on the solar surface

Granular and supergranular flows on the solar surface lead to a random motion of tracers, the most important of which are magnetic flux tubes. A simple random walk of stepsize Δx and timestep Δt over a plane is known to lead to an increase of the rms separation r of a tracer from its starting point (or of two tracers from each other) according to the law

$$r^2 = 4Dt \tag{1}$$

where $D = \Delta x^2/2\Delta t$. The time development of a continuous distribution of tracers is then described by a diffusion equation with diffusivity D.

As a first approximation, the advection of tracers by (super)granules may be represented by such a simple random walk/diffusion, identifying Δx with the spatial scale l of the cells and Δt with their lifetime τ.

This is the approach used in Babcock–Leighton-type models of the solar cycle where poloidal fields are brought to the surface in a concentrated form in active regions, and thereafter they are passively transported to the poles by transport processes (diffusion and meridional circulation). The diffusivity in these models is a free parameter: a best fit to the observations yields $D = 600\,\mathrm{km}^2/\mathrm{s}$. Despite the vectorial character of the magnetic field, these 1D models have been remarkably successful in reproducing the observed temporal evolution of the flux distribution. A possible explanation was proposed in their model by Wang et al. (1991): they assume that field lines are vertically oriented throughout much of the convective zone and this essentially reduces the problem to one dimension. Some support for this conjecture has come from the 2D flux transport models of Petrovay & Szakály (1999). Thus, in a first approximation, 1D models may be used for the description of meridional transport, as these fields pervade the convective zone and are continuously reprocessed through it.

The empirically determined value of the diffusivity, $600\,\mathrm{km}^2/\mathrm{s}$, seems to agree with the primitive random walk model if the steps are identified with granular sizes/lifetimes. Supergranulation, however, should lead to a diffusivity that is by an order of magnitude higher than this calibration. The continuous reprocessing of large-scale fields throughout the convective zone offers a plausible explanation for this inconsistency: the empirical value of the diffusivity reflects the turbulent diffusivity in the lower convective zone where the pressure scale height is $\sim$50 Mm and the turnover time $\sim$1 month.

An alternative explanation was put forward by Ruzmaikin & Molchanov (1997) who pointed out that, owing to the cellular nature of photospheric flows, identifying cell size l and lifetime τ with random walk steps is an oversimplification. The fact that a tracer cannot leave a cell during the cell's lifetime, even if it was originally placed next to its border, reduces the effective stepsize significantly. The resulting reduction in D is very sensitive to the value of the Strouhal number $\mathrm{St} = \tau v/l$ and is especially strong for $\mathrm{St} \gg 1$. This effect may be sufficient to reduce supergranular diffusivity to the observed value.

2 Anomalous diffusion

The cellular and turbulent nature of the flow implies that a simple random walk cannot account for the motion of magnetic elements. As a consequence, the actual flux redistribution may differ from a simple (Fickian) diffusion process (Isichenko, 1992) and, instead of equation (1) in general one has

$$r = 2Kt^{\zeta}. \tag{2}$$

If $\zeta \neq 1/2$ it is customary to speak of *anomalous* or *non-Fickian diffusion,* $\zeta > 1/2$ corresponding to *superdiffusion,* $\zeta < 1/2$ to *subdiffusion.* As equation (2) means a unique relation between r and t one might formally still write $r = 2K'(r)t^{1/2}$, leading

to the concept of a "scale-dependent diffusivity"

$$D(r) = K'^2 = K^{1/\zeta} r^{2-1/\zeta} \tag{3}$$

It is, however, clear that such a concept is in general useless for the description of the evolution of a continuous field where no preferred scale exists. Anomalous diffusion thus cannot be described by a diffusion equation or, indeed, by any partial differential equation.

How can anomalous diffusion come about? One possibility was suggested by Schrijver & Martin (1990). Magnetic flux tubes are located at junctions of a fractal lattice between supergranules, mesogranules and granules. Assuming that limitations exist for the motion of individual flux elements along this lattice, for certain lattice properties subdiffusion may result. They made an attempt to detect subdiffusion by the analysis of observed flux redistribution in the photosphere; however, ζ was not found to differ from 0.5 within the observational uncertanties.

Being a multiscale phenomenon, turbulence can also naturally lead to a "scale-dependent diffusivity". In order to understand the nature of the diffusion process in a turbulent medium let us consider the question how a random continuous velocity field of a given characteristic scale λ (i.e. one level in the turbulent hierarchy) can be best represented by a random walk with steps Δx and Δt. For the best representation one should set $\Delta t = \tau_L$, the Lagrangian correlation time of the flow, as this is just the time after which the advected particle experiences a significant change in its velocity. The distance the particle travels in this time is $\Delta x = v\tau_L = \min(v\tau_E, \lambda)$ where τ_E is the Eulerian correlation time, λ the correlation length, and v the rms velocity. The diffusivity for this random walk will thus depend on the Strouhal number $\mathrm{St} = \tau_E v/\lambda$; assuming a non-cellular flow

$$D = \begin{cases} \tau_E v^2 & \text{if St} < 1 \\ \lambda v & \text{if St} > 1 \end{cases} \tag{4}$$

In a multiscale flow both τ_E and v scale with λ:

$$\tau_E \sim \lambda^z \qquad v^2 \sim \lambda^{\alpha-1} \tag{5}$$

During the random walk, motions on scales exceeding the separation r of two tracers do not contribute to their further separation while all other scales contribute to it. Of these scales, according to equation (3), the smallest one will dominate in the diffusion process for $2 - 1/\zeta < 1$ i.e. $\zeta < 1/2$. In this case, then, the diffusivity will not significantly depend on the separation for all scales above the viscous scale: turbulence can never lead to subdiffusion.

In the case when the relatively largest scale $\lambda \sim r$ dominates, for given values of z and α the anomalous diffusion exponent ζ can be determined by substituting (5) into (4) and equating the resulting scaling exponent of D to $2 - 1/\zeta$, as given by (3). The Strouhal number scales as $\mathrm{St} \sim \lambda^{z+\alpha/2-3/2}$. For high Reynolds numbers, then, the

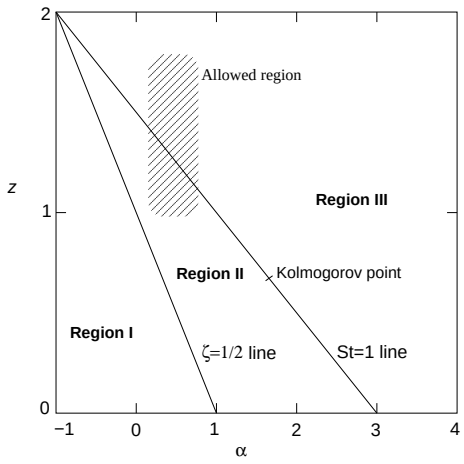


Figure 1: *Regimes of anomalous diffusion on the α–z plane. K is the Kolmogorov point; the shaded area indicates the approximate position of the photospheric flow field in the 1–30 Mm size range. (Non-cellular case)*

sign of $(\mathrm{St}-1)$ at the larger scales depends on the sign of the scaling exponent of St, i.e. the line

$$2z + \alpha - 3 = 0 \tag{6}$$

defines two regimes in the α–z plane (Fig. 1). Above the line, in what is called Region III (Avellaneda & Majda, 1992), one finds $\zeta = 2/(3-\alpha)$ (except in the case of a cellular flow when $\zeta = 1/z$ —cf. the discussion at the end of Sect. 4.1). This Region is clearly superdiffusive for all values of $\alpha > -1$ (or $z < 2$). Below the line, in Region II we have $\zeta = 1/(3-z-\alpha)$, independent of cellularity, as here we have low Strouhal numbers at the large scales. It is then clear that a second dividing line will also exist at

$$z + \alpha = 1 \tag{7}$$

as below that line (Region I) $\zeta < 1/2$ would result, in which case, as we have seen, the smallest scales dominate the diffusion process. Region I is thus characterized by a Fickian diffusion, while Region II is again superdiffusive. Point K in Figure 1 denotes the case of a Kolmogorov spectrum, $\alpha = 5/3$, $z = 2/3$, $\zeta = 3/2$.

3 Observational constraints on anomalous diffusion in the Sun

In order to determine the place of photospheric velocity fields on the α–z diagram, Ruzmaikin et al. (1996) fitted power laws to the spatio-temporal power spectra of photospheric velocity fields with the result $\alpha \sim 1.5$–1.8, $z \sim 0.15$–0.85. This would localize solar turbulence to the neighbourhood of the Kolmogorov point K. However, in Section 2.2.1 above we already stressed the perils of power-law fits to power spectra of solar velocity fields. There is simply no theoretical reason or observational evidence to suggest that these fields should follow a power-law spectrum from supergranular scales down to the resolution limit. Indeed, the well known fact that meso- and supergranular motions have a lower velocity amplitude than granulation, tells us that $\alpha < 1$ in the regime $\lambda > 1\,\mathrm{Mm}$! Using observational estimates for these velocity amplitudes and for correlation times one arrives at much more robust limits that are in plain contradiction to the ones quoted above: $\alpha \sim 0.0$–0.7, $z \sim 0.9$–1.8, leading to $\zeta \sim 0.48$–1.2. These limits in themselves would indicate superdiffusion (shaded area in Fig. 1).

Turbulent erosion models of sunspot decay can also be used to constrain anomalous diffusion in the photosphere (Petrovay, 1998). The size of sunspots spans the granular-supergranular size range that is of interest in this respect, and the fortuitous property of the erosion models that they *do* show a characteristic scale, the radius of the spontaneously formed current sheet, makes it possible to test for ζ by using a scale-dependent diffusion coefficient with the current sheet radius as defining scale. In this way, ζ is found to lie in the range 0.44–0.59, i.e. any deviation from a Fickian diffusion seems to be modest, if present at all. A possible explanation for why the diffusion exponent is lower than suggested by velocity power spectra may be that superdiffusion due to turbulence is offset by subdiffusion effects due to diffusion on a bond lattice.

Acknowledgement

This work was supported in part by the OTKA under grant no. T043741 and by the European Commission through the RTN programme (European Solar Magnetism Network), contract no. HPRN-CT-2002-00313.

References

Avellaneda, M., Majda, A. J. 1992, Phys. Fluids A, 4, 41

Isichenko, M. B. 1992, Rev. Mod. Phys., 64, 961

Petrovay, K. 1998, in E. R. Priest, F. Moreno-Insertis, R. A. Harris (eds.), A crossroads for european solar and heliospheric physics: recent achievements and future mission possibilities, ESA, Publ. SP-417, p. 273

Petrovay, K., Szakály, G. 1999, Sol. Phys., 185, 1

Ruzmaikin, A. A., Cadavid, A. C., Chapman, G. A., Lawrence, J. K., Walton, S. R. 1996, ApJ, 471, 1022

Ruzmaikin, A. A., Molchanov, S. A. 1997, Sol. Phys., 173, 223

Schrijver, C. J., Martin, S. F. 1990, Sol. Phys., 129, 95

Wang, Y.-M., Sheeley, N. R., Nash, A. G. 1991, ApJ, 383, 431

PADEU | PADEU **15**, 59 (2005)
ISBN 963 463 557

Dichotomic model for describing the solar granulation dynamics

Mircea Rusu

University of Bucharest, Atomistilor 405, RO-76900 Bucharest-Magurele, Romania

E-mail: mrusu@dnt.ro

Abstract
In the present communication, we describe a simple model for the observed solar granulation. The proposed model is a phenomenological one, and describes the dynamics of granulation as complex changing patterns observed on the Sun's surface. Based on some observed facts, we describe the variability of the granules form and the statistics observed, as a continuous process of "splitting" each granule in two pieces, in a roughly constant area ratio. This process is named a dichotomic process, and produces in time, granules of different forms and dimensions. Simulating such kind of dichotomic "fragmentation" of the granules, starting with a given initial distribution, we can obtain the distribution of the granules area at a given moment of time that can be compared with the observed statistics.
Keywords: *Solar granulation, dihotomic model, fractal interpretation*

1 Introduction

Solar granulation is a common feature observed inspecting the Sun surface. Its mosaic, grainy structure is easy to observe in images obtained with a high resolution. The origin of granulation is complex but mainly is due to the convection of the fluids from the depth to the vicinity of the observed surface of the Sun. Many aspects contribute to the feature observed: the temperature gradients in the convection zone, the (differential) rotation of the Sun, the magnetic field from the vicinity, the composition of the fluids (plasmas) and the mixing processes, and so on.

Extremely complex simulations of magneto-hydrodynamics near the Sun surface to describe the convection and to understand the granulation are far to be satisfying (Berger & Scharmer , 1999).

In the present communication, we describe a much simpler version of a model for the observed granulation.

The proposed model is a phenomenological one, and describes the dynamics of granulation as complex changing patterns observed on the Sun's surface. Based on some observed facts, we describe the variability of the granules form and the statistics observed, as a continuous process of "splitting" each granule in two pieces, in a roughly constant area ratio. Various hypotheses for the fragmentation mechanisms and a given initial distribution can be used, in order to choose the best one. We will show that such a simple model can reasonable describe some of the observed facts, giving some hints for understanding solar granulation.

Convection is the dominant mechanism of energy transport in the envelopes of the Sun. The flow in the free convection is driven by buoyancy forces, which are induced by a temperature gradient between the lower and upper boundaries of the plasma in a gravitational field. The convection zone is an extremely dynamic layer form of the Sun, that shows itself in high resolution images, as a collection of relatively small, variable form, patches named granules (Fig.1), one division equal 1000 km). At the centers of granules hot solar gas rises and radiates its heat rapidly into space; the gas is then diverted horizontally, and sinks back into the Sun in the darker intergranular lanes. The sizes of the granules range from approximately 250 km (the limit set by the telescope and the Earth's atmosphere) to more than 2000 km with an average diameter of 1300 km. Lifetimes of granules typically range from 8 to 15 minutes. Horizontal and vertical velocities of the gas motion are 1 to 2 km/s.

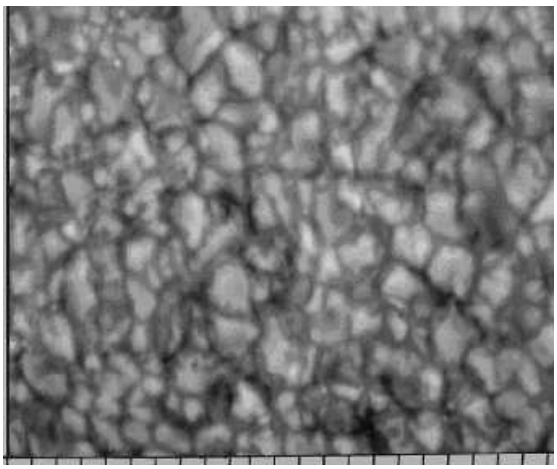


Figure 1: *A picture of solar granules*

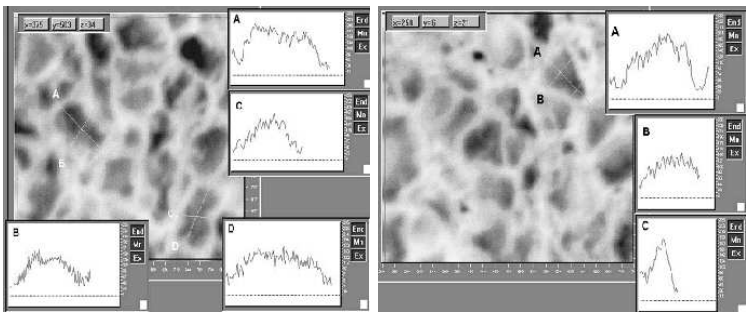


Figure 2: *Luminosity profiles on negative images of the surface*

Examples of such convection pattern in gravitational field can be observed also on Earth, for example in huge fire in the forests.

It is possible to try to describe as precise as it is possible, the dynamics of the fluid circulation in the convection zone, but the huge amount of data needed, and the complexity of equations and conditions are formidable. At least, simulations of this dynamics can be done with limited accuracy, and the results are promising (Brummel et al , 1995), giving us a much confident image of the convection. The result obtained from simulation could be described as a "forest" of jets of fluids that comes from inside and ends at the surface, cooling the fluid, which after is sinking back in the depth of Sun.

2 The model

2.1 The morphology of granules

The variation in luminosity and the form of the end part of the columns (the granule) shows variation in space and time, and defines (in a loosely speaking way) the boundaries of granules. Our computer image analysis of granule luminosity could be seen in figure 2 (luminosity profiles on negative images of the surface). From this analysis we can conclude that a very sharp and well defined boundary of a granule is difficult to find. However, having high-resolution images it is possible to describe some statistical characteristics of the granules.

Very high-resolution pictures (0".25 - Pic du Midi Observatory) used for analyzing solar granulation using computer-processed images found that the distribution of the number of granules increases continuously towards smaller scales. This means that the solar granulation has no characteristic or mean scale. Nevertheless, the granules appear to have a critical scale of 1".37, at which dramatic changes in properties of granules

occur; in particular, the fractal dimension changes at the critical scale (Roudier et al , 1991).

2.2 Fractal analysis of solar granulation

The study of turbulent phenomena has been a demanding task in astrophysics. The photosphere of the Sun is one of the few places in astrophysics where turbulent motion could be, in principle, observed directly. The high spatial resolution, which can now be attained, and the development of new ways to describe and analyze chaotic systems, i.e. the concept of fractals changed our possibilities, today.

We made fractal analysis of some images in order determine the fractal dimension (D) of the granulation filed (Munteanu et al , 1994). The method used was the box-counting one. Determination of the fractal dimensions of the solar granules, using luminosity analysis has some difficulties because of the uncertainties of the granules boundaries. The results of our work reveal that there are variations of the fractal dimension, if we analyze pictures of granules in regions in which exists sunspots, (D = 1.37) or pictures in regions without such features (D = 1.81). The same, large distribution of fractal dimension for granulation was obtained by Brundt et. al (Brundt et al , 1991), but without a correlation to the regions on the Sun.

2.3 Granule statistics

Being interested in the granule structure and their geometry and dynamics, we focused on the image analysis and recognition of patterns, and on the statistics of the granules population. We used images obtained by The Swedish Solar Vacuum Telescope. Almost all the conclusions described here were drawn using this source.

For the statistics of inter granular distances, we made extensive measurements on a large surface of the Sun. Radial distribution function (number of granules versus inter granular distance) was computed from the list of all distances found on the image, measured and processed using a special design computer code. The "position" of a granule used for computing the distances, was defined using different algorithms: the geometrical mean of the roughly ellipsoidal granules, the point of highest luminosity inside the granule boundary, or a simple visual center of the granule. The results obtained for the distribution are practically the same for all methods of finding the position of the granules.

The whole investigated region was divided in 10 rectangular fields, and the radial distribution function of the inter granular distances was computed. The results are presented in figure 3, for nine regions. A striking and unexpected result appeared, without exception: all the distributions show asymmetry. The distribution suggested that a deconvolution of the asymmetric distributions in two symmetric Gaussians could be possible. Figure 4 shows the experimental points (dots), the two Gaussian distributions (green line) and the reconstructed - convoluted distribution (red line).

It is without doubt that the deconvolution works well. If this deconvolution has any physical meaning is another question.

Analyzing the parameters of the Gaussian distributions, it was also evident that the positions of the two maxima are correlated in each investigated region. Such a correlation could be seen in figure . The chart of the region studied is also presented here. The two groups of distributions have distinct average distances between granules. We denoted "line 1" and "line 2" the two distinct distribution found in each region. We can consider that the analysis reveals two distinct "populations" of granules.

Results and conclusions:

The slope of graphs: position of line 1 versus position of line 2 is 1.64 considering all the points (the red line),

The line of slope 2 is shown for comparison (the black line),

The points 1, 2, 3, 5, and 7 are from regions free of sun spots;

The points 6, 8, and 9 are from regions near sunspots; they are the most distant points from the best fit line,

Part of the dispersion of the fitted parameters of the deconvolution could be attributed to the difficulties of assessing the center of the granule from the images.

The whole trend suggests that the two populations could be a result of a splitting mechanism applied to the "mature" granules, that suddenly reduces approximately to half the average distances between granules. This is a statistical inference which we can explain simply using a phenomenological model.

2.4 The dihotomic model

We tried a phenomenological model, namely a dihotomic model. We consider the following scenario:

The grains are in continuous change of form and size. There are moments when grain splits in two smaller grains. The population statistics at one moment contains in fact at least two different populations. The distribution shows after several generations of splitting process an increases of the diversity of grains size.

New grains are added to the distribution, and the small ones disappear. The result is a statistically stable structure of two distinct populations of grains. A simple example of a dihotomic fragmentation and self-similarity could be used as a model for the dynamics of the two populations in equilibrium. The main hypothesises are: constant ratio of division, K, and a conservation of the initial area (or other measure, denoted here by V) at the moment of division. These two hypothesises could be written as:

$$V_0 = V_1 + V_2 \ ; \ K = \frac{V_2}{V_1}$$

To summarize the results, after one division, the parts are:

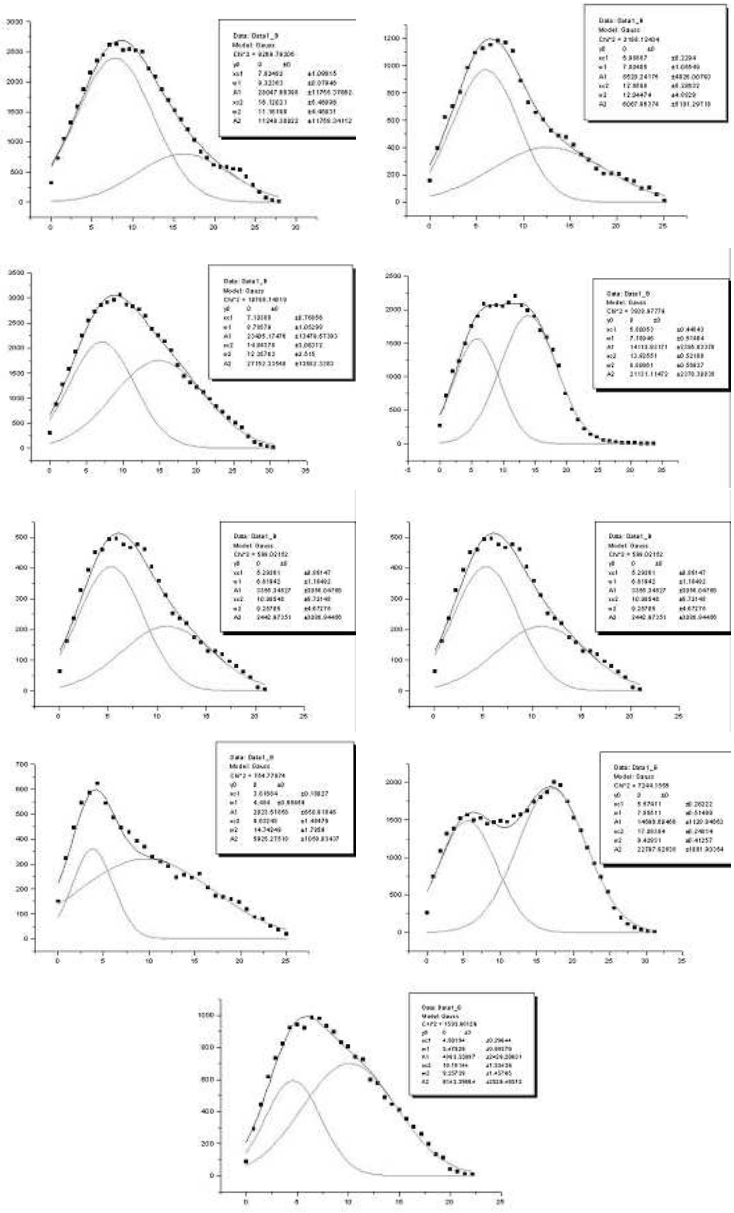


Figure 3: *The computed radial distribution function of the inter granular distances*

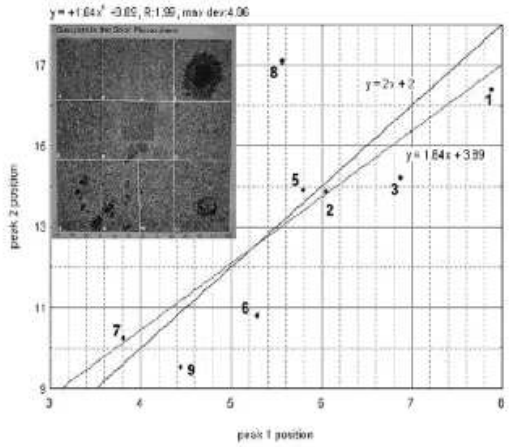


Figure 4: *Representation of experimental points, Gaussian distribution and reconstructed-convoluted distribution*

$$V_1 = \frac{V_0}{1+K} \; ; \; V_K = \frac{V_0 K}{1+K}$$

After n-divisions (generations) the size of p-fragment has:

$$V_{p;n,K} = V_0 \frac{K^p}{(1+K)^n} \; , \; p = 0, 1, 2, 3, ..., n$$

The lowest size are: $V_{min} = \frac{V_0}{(1+K)^n}$ The largest size are: $V_{max} = \frac{V_0 K}{(1+K)^n}$ The difference and ratio are: $d = V_{max} - V_{min} = V_0 \frac{K^n - 1}{(1+K)^n}$ and $r = \frac{V_{max}}{V_{min}} = K^n$; The mean values are: $\bar{V} = \frac{V_0}{2^n}$ and $\bar{N} = 2^n$

Figure 5 shows the distribution after N = 11 generations, with a fragmentation ratio K = 1.6

We can make this process more "realistic" if we add a noise in the fragmentation ratio: $K = K_0 + K_{noise}(t)$, that can be used in computation. The result of such a noise, is a spreading (and smoothing) of the distribution groups (Fig.6).

If the time between generation, splitting and disappearance of the grains is short, just two generations could be seen in the distribution, and the distribution is "smoothed" by the random size of generated grain, and of the moment of splitting. In addition, this could account for the limitation of the maximum size of a grain, explained by this mechanism of splitting. The moment of splitting is probable triggered in the column of the convection tube, by some instability in the flow.

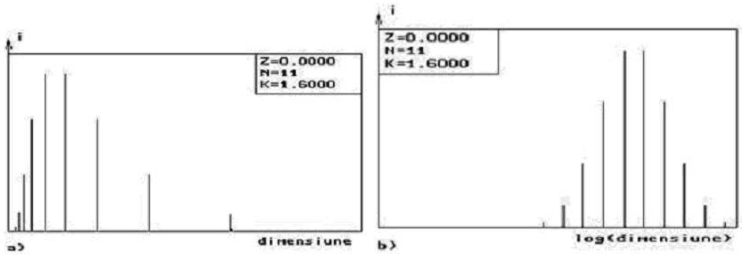


Figure 5: *Distribution after N=11 generations*

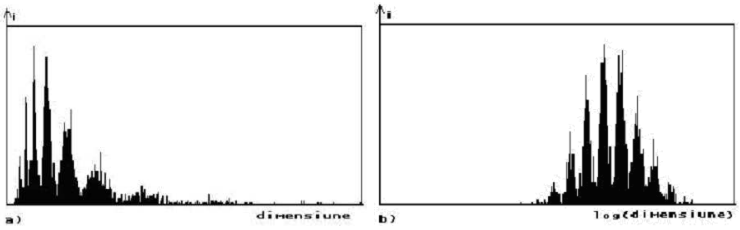


Figure 6: *Distribution groups*

2.5 The dynamics of granules evolution

In order to see if this mechanism could be real, we examined the movie pictures of the granulation dynamics. The series of the images was observed with a fast frame selection system on June 5, 1993, at the SVST (La Palma) in cooperation with G. Scharmer (Stockholm) and G. W. Simon (Sunspot); N. Hoekzema (Utrecht), W. Mhlmann (Graz), and R. Shine (Palo Alto) were involved in the data analysis. Technical data: wavelength 468 ± 5 nm; exposure time 0.014 s; rms contrast (uncorrected) between 7 and 10.6 %. The images were registered, destretched, corrected for the telescope's point spread function, and sub sonically filtered after interpolation to equal time steps. For each frame, both area and total time are indicated.

A sequence of these time-lapse series of the evolution of the solar granulation is represented in figure 7. Qualitatively (this analyze is at the moment under computation, using a special code for pattern recognition and granule characteristic extraction) the above phenomenological model seems to be correct.

3 Conclusions

The fractal structure of the granulation field suggests some self-similar mechanism that acts in the dynamics of the convection.

The simple model of dihotomic self-similar fragmentation could reasonable explain the presence of the two distinct populations observed in the statistics of radial distribution of distances between granules The dynamics of fragmentation, which is revealed in the motion pictures of the Sun's surface, qualitatively confirms the idea of continuously dihotomic fragmentation of large granules.

This model gives a simple phenomenological mechanism that could describe the facts found in the statistics of granules and gives us a hint about the phenomena inside the convection zone. We can make the hypothesis that the convection tube in his upward moving can exhibit instabilities that split the tube, most probable in two adjacent tubes, and a hierarchy of splitting could follow this process.

The assemble of the tubes will exhibit self-similarities revealed in the measured fractal dimension.

Acknowledgement

I we would like to thank prof. Don Wentzel for his suggestions and help in completing the image and to the work of some students involved in measurements: Adriana Durbala, Oana Stere, Constantin Oprea and to Alex Curutiu for producing the computer codes used in automatic processing of the images

References

Berger T., Scharmer G., Swedish Vacuum Solar Telescope, on 12 May 1998, SVST Telescope 50 cm f/45 vacuum refractor, Instituto Astrofisica de Canarias, Isla de La Palma, Spain

Brummel, N., Cattaneo, F., Toomre, J., 1995, Science V., 269, 5229, 1370

Brummel, N.H., Cattaneo, F., Tobias, S.M., 1999, "Stelar Dynamos: Nonliniarity and Chaotic Flows", ASP Conference Series, 178, Eds. M. Nunez and A. Ferriz-Mas, Astronomical Society of the Pacific, San Francisco, p.23

Brun A.S., 2000, "L'origine des relations Soleil-Terre: Modlisation en convection solaire, MHD et Cycle solaire", *http* : *//dphs*10.*saclay.cea.fr/Sap/Activites/Science/Soleil/Simulations*$_M$*HD /page.shtml*

Brundt, P.N., et all, 1991, "Turbulence, Fractals, and Solar Granulation, in Applying Fractals in Astronomy", A.Heck, J.M.Perdang Eds., Springer-Verlag

Munteanu F., Zugravescu D., Rusu M., Suteanu C., 1994, "On the synergy of ruptures", Revue Roumaine de Gophysique, 38

Roudier, T. et al, 1991, A&A, 248, 245

Rusu, M.V, Durbala, A., 2000, "Fractal Analysis of Solar Granulation", National Physics Conference, Bucharest

Sadovskii M.A., Golubeva T.V., Pisarenko V.F., et al., 1984, "Characteristic dimensions of rock and hierarchical properties of seismicity", Izv. Acad. Sci. USSR Phys. Solid Earth, 20, 87-96

Suteanu C., Munteanu F., Zugravescu D., 1993, "Recursive models for solids fragmentation"; Bul. Inst. Geodin. 4, 2

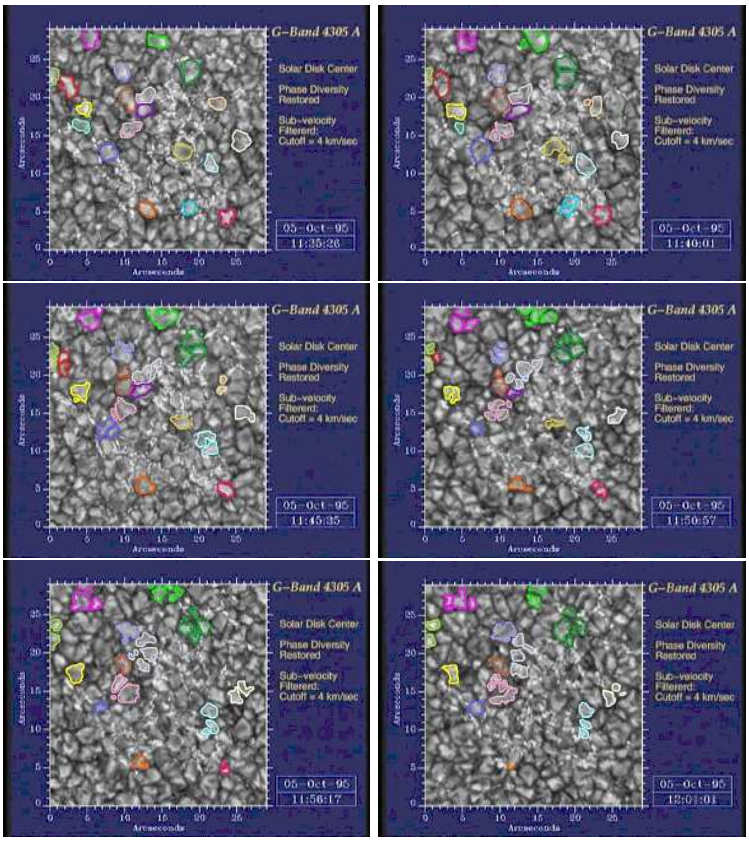


Figure 7: *Movie pictures of the granulation dynamics*

PADEU | PADEU **15**, 73 (2005)
ISBN 963 463 557

Nonlinear waves in solar plasmas

I. Ballai

SPARG, Dept. of Applied Mathematics, University of Sheffield, Hounsfield Road, Hicks Building, Sheffield, S3 7RH, UK

E-mail: i.ballai@sheffield.ac.uk

Abstract
Nonlinearity is a direct consequence of large scale dynamics in the solar atmosphere. Here, the nonlinear steepening of waves balanced by dispersion generates solitary waves. Nonlinear waves can also appear in the vicinity of resonances, influencing the efficiency of energy deposition. Here we review recent theoretical breakthroughs that have lead to a greater understanding of many aspects of nonlinear waves arising in homogeneous and inhomogeneous solar plasmas.
Keywords: *Sun: waves, MHD, Sun: wave heating*

1 Introduction

One of the most interesting processes in solar and astrophysical plasmas is the complicated interaction of plasma motions with magnetic fields. These media are highly non-uniform and as a consequence are a natural environment for magnetohydrodynamic (MHD) waves. Waves can transport energy and momentum. When part of their energy or momentum is transferred to the plasma they can heat and accelerate the plasma (e.g. resonant absorption). Waves can carry information about the medium in which they propagate, therefore they can provide a unique tool for plasma diagnostics.

In the present contribution we review two important nonlinear waves arising in inhomogeneous solar plasmas. Firstly, solitary waves arising in structured plasmas (i.e. waveguides) are discussed in different structures and for different dispersions. Secondly, nonlinear waves generated in the vicinity of resonant positions (slow resonance) are revisited and we show how nonlinearity will influence the efficiency of heat deposition.

2 Nonlinear waves in waveguides

One of the basic properties of solar plasma is that is structured, the magnetic field is not distributed smoothly over the surface of the Sun, but it tends to accumulate in entities called *magnetic loops*, the building blocks of the solar corona. These structures can support, e.g. longitudinal wave propagation over long distances. The effect of the structuring is that it introduces dispersion, i.e. a modification in the propagation characteristic of the wave.

Solitons are finite-amplitude waves of permanent shape which owe their existence to the balance between nonlinear wave-steepening and wave dispersion. Nonlinearity appears for waves of finite amplitude and generally is a consequence of large scale dynamics. Dispersion could arise due to two different effects. *Geometrical dispersion* appears for waves propagating in a magnetic guide (flux tube or sheet). This dispersion does not depend on the reaction of the external media and its value is defined by the geometrical scale of the duct (the tube diameter or the thickness of the sheet). Alternatively, waves in open ducts could have dispersion due to the reaction of the external media. It is not always simple to separate these two sources of dispersion in spite of their different behavior. Furthemore, *physical dispersion* appears due to plasma (magnetic) effects (generalized Ohm's law or Finite Larmor Radius (FLR) effects). In general, these two dispersive effects give rise to different dispersive behavior but they have the same result: creation of a new length scale in addition to the natural length scale of the waves, i.e. their wavelength.

Guided waves in solar and space plasmas are investigated in two cases: magnetic slab (Cartesian geometry) and magnetic tube (cylindrical geometry). The dynamics of solitary waves are best described in the so-called *thin flux tube approximation*. For a motion $v(z,t)$ along a tube (slab or cylinder) of cross-sectional area $A(z,t)$, the one-dimensional equations of continuity, longitudinal momentum, isentropic energy and flux conservation are

$$\frac{\partial}{\partial t}(\rho A)+\frac{\partial}{\partial z}\rho v A=0,$$

$$\frac{\partial v}{\partial t}+v\frac{\partial v}{\partial z}=-\frac{1}{\rho}\frac{\partial p}{\partial z},$$

$$\frac{\partial}{\partial t}\left(\frac{p}{\rho^\gamma}\right)+v\frac{\partial}{\partial z}\left(\frac{p}{\rho^\gamma}\right)=0,\quad BA=const,$$

where the quantities $p(z,t)$, $\rho(z,t)$, $B(z,t)$ and $v(z,t)$ are supposed uniform across the tube.

In a magnetic slab of width $2a$ with the magnetic field along the structure, the dispersion relation of slow sausage modes with their wavelength (k^{-1}) much larger than the width of the slab is (Roberts, 1981)

$$\omega/k=c_T-\alpha_1|k|,\quad \alpha_1=\frac{1}{2}\frac{\rho_e}{\rho_e}\left(\frac{c_T}{v_A}\right)^3 ac_T, \tag{1}$$

where c_T is the tube speed (the propagation speed of slow magnetoacoustic waves in an unbounded medium). The $\alpha_1|k|$ term in Eq. (1) arises due to dispersion and in the long wavelength limit is a small quantity. If the amplitude of slow waves becomes large enough, the nonlinear evolution of these waves is described by the Benjamin-Ono (BO) equation written for the z-component of the velocity perturbation (Roberts & Mangeney, 1982; Edwin & Roberts, 1986; Ballai et al., 2002)

$$\frac{\partial v}{\partial t} + c_T \frac{\partial v}{\partial z} + \beta v \frac{\partial v}{\partial z} + \frac{\alpha_1}{\pi} \frac{\partial^2}{\partial z^2} \int \frac{v(z',t)}{z'-z}\, dz' = 0, \tag{2}$$

where β is a coefficient which depends on the characteristic speeds (sound, Alfvén and cusp speeds). The single-soliton solution of this equation is the algebraic soliton,

$$v(z,t) = \frac{A}{1+[(z-st)/L]^2}, \tag{3}$$

where A is the velocity amplitude of the soliton, s and L are the speed and scale of the soliton related by

$$s = c_T + \frac{\beta A}{4}, \quad L = \frac{4\alpha_1}{L\beta}. \tag{4}$$

In a magnetic cylinder with radius R, the dispersion relation of slow surface sausage modes in the long wavelength limit is

$$\omega/k = c_T - \alpha_2 k^2 K_0(|k|\lambda), \tag{5}$$

where $K_0(x)$ is the modified Bessel function of the zeroth-order. The quantities α_2 and λ depend on characteristic speeds and the radius of the tube and λ^2 can be both negative or positive quantity. If these waves steepen into nonlinear waves, their evolution is described by the Leibovich-Roberts (LR) equation (Roberts, 1985),

$$\frac{\partial v}{\partial t} + c_T \frac{\partial v}{\partial z} + \beta v \frac{\partial v}{\partial z} + \frac{\alpha_2}{\pi} \frac{\partial^3}{\partial z^3} \int_{-\infty}^{\infty} \frac{v(z',t)dz'}{[\lambda^2+(z'-z)^2]^{1/2}} = 0. \tag{6}$$

Although this equation was derived 20 years ago, there is no known analytical solution, however, numerical investigations showed that it has a solitary-like solution (Weisshar, 1989). If the propagation speed of the slow waves inside and outside the tube are approaching each other, the LR equation reduces to a nonlinear wave equation without dispersion which describes shock waves with zero-width. If the internal cusp speed approaches either the external sound or Alfvén speed (supposing a magnetized environment), the LR equation reduces to the Leibovich equation describing nonlinear waves on a cylindrical vortex core. The LR equation is valid provided $\lambda^2 > 0$. If $\lambda^2 < 0$, then slow leaky sausage modes will propagate in the tube draining energy away from the structure. In this case, the LR equation can be modified to describe slow leaky sausage modes as (Ballai & Zhughzda, 2002)

$$\frac{\partial v}{\partial t} + c_T \frac{\partial v}{\partial z} + \beta v \frac{\partial v}{\partial z} + \alpha_1 \frac{\partial^3}{\partial z^3} \left[\int_{-\infty}^{-|\lambda|+z} + \int_{|\lambda|+z}^{\infty} \frac{v(s,t)\, ds}{\sqrt{\lambda^2+(z-s)^2}} \right] = 0. \tag{7}$$

One of the limitations of these equations is that the solitary wave solution appears only up to some critical amplitude. This amplitude threshold appears because the dispersion relation has a maximum, i.e the maximum value of the dispersion is not enough to smooth out the front of the waves if the amplitude of the waves exceeds a critical threshold value. Thus, the (LR) equation describes the nonlinear behavior of weakly nonlinear slow sausage modes whose phase velocity in the linear limit has an extremum. For solitons with negative dispersion this limitation does not occur, instead they are subject to an aperiodic instability.

There is one aspect which so far has been neglected, and this is related to the dissipative character of the plasma. In fact, the right choice for a dissipative mechanism depends on the location where physical processes are to be studied and also on the physical mechanism itself. For instance, Ohmic dissipation of wave propagation in the solar corona does not result in significant damping (unlike viscosity or thermal conduction) but this dissipative effect must be taken into account when studying effects which require small length scales, e.g. coronal heating.

When dissipation is taken into account, solitary waves will exhibit a slow damping, which means that the energy and momentum of solitary waves are not conserved quantities any longer. The most important dissipative mechanisms are viscosity, thermal and electrical conduction and radiation. If we take into account the first three mechanisms, the solitary wave equations must be supplemented by an extra term proportional to $\partial^2 v/\partial z^2$ which results in an algebraic decay of the soliton. If radiation is considered, the nonlinear equations will have an extra term proportional to v which leads to a slow exponential decay of the solution. Illustrations of when these dissipative terms are added to a nonlinear evolutionary equation are the Leibovich-Roberts-Burgers or the Korteweg-de Vries-Burgers equation.

Dispersion can arise not only due to a geometrical structuring, but also due to the presence of the magnetic field, through, e.g. the Hall term in the generalised Ohm's law. Strictly speaking, Hall MHD is relevant to plasma dynamics occurring on length scales shorter than the ion inertial length, c/ω_i, where c is the speed of light and ω_i is the ion plasma frequency. Inclusion of the Hall term in the magnetohydrodynamic induction equation is known to affect the polarization of waves because it includes the dispersion of Alfvén waves near the ion cyclotron frequency.

The nonlinear wave evolution in the presence of a Hall effect in a viscous plasma has been studied in connection to the acceleration of the solar wind. When the nonlinear steepening of compressional waves is balanced by the broadening of the wavefront caused by the Hall effect, we obtain that the dynamics of solitary waves propagating in a super-radial magnetic field is described by the Korteweg-de Vries-Burgers (KdV-B) equation

$$\frac{\partial v}{\partial t} + c_f \frac{\partial v}{\partial z} + \alpha_1 v \frac{\partial v}{\partial z} - \alpha_2 \frac{\partial^3 v}{\partial z^3} - \alpha_3 \frac{\partial^2 v}{\partial z^2} = 0, \tag{8}$$

where c_f is the phase speed of linear waves and the coefficients α_i depend on characteristic speeds and the angle of propagation with respect to the magnetic field.

Choosing a nearly-parallel propagation, we obtain that solitons arising from the nonlinear steepening of compressional slow waves are able to accelerate the plasma, while solitons which are generated by the nonlinear steepening of fast waves will decelerate the plasma. The speed at which the solar wind is accelerated by means of solitons agrees very well with the observed speeds by UVCS-SOHO at $1.3R_{\odot}$.

Solitary waves have unique properties which make them special for mathematics and their applications to other fields: *(i) Integrability*: Before the discovery of solitons, mathematicians were under the impression that nonlinear PDEs could not be solved, at least analytically. However, solitons showed us that it is possible to solve PDEs (at least the solitary wave equations) exactly, which gives us a tremendous "window" into what is possible in nonlinearity. *(ii) Nonlinear superposition*: In linear theory, there is a simple way to generate a new solution from known ones, just by multiplying them with a scalar and adding them together. This is known as superposition. Before the discovery of solitons, there was no analogue of this construction for nonlinear equations, but the way that a 2-soliton solution can be viewed as a combination, although not a simple linear combination, of two 1-soliton solution leads us to the recognition that (at least for solitons) there is a nonlinear superposition principle, as well. *(iii) The particle-like behaviour* of solitons leads to a large number of applications. This is true to some extent: there are soliton models for nuclei and the technique known as bosonization allows us to view fermions as being solitons in appropriate situations. Recently, the transport of energy and information along DNA chains was described by the so-called Davydov-solitons. Solitons have also a series of other applications in fields like oceanography, fiber optics, telecommunications and geophysics. Solitary waves carry a large amount of energy, therefore if they are dissipated over short length scales they could provide, e.g. the energy required to heat the coronal plasma (resonant solitary waves).

3 Nonlinear resonant waves

Resonances are ubiquitous every time MHD (magnetohydrodynamic) waves are driven in inhomogeneous plasmas. However in weakly dissipative plasmas (as in the case of solar plasma) driven MHD waves show nearly resonant behaviour, which deviates from the resonant behaviour in ideal plasmas only in thin *dissipative layers* surrounding the ideal resonant positions.

A very important property of these nearly resonant waves is that their damping rate is almost independent of the values of dissipative coefficients. As a result, the damping rate of nearly resonant MHD waves can be many orders of magnitudes larger than the damping rate of MHD waves with the same frequencies in homogeneous plasmas. This property of resonant waves being strongly damped in weakly dissipative plasmas has attracted ample attention from plasma physicists since the transferred energy can be converted into heat (Sakurai et al., 1991; Ruderman et al., 1997a; Ballai et al.,

1998a) or it might give valuable information about the density of the plasma and the characteristic scale of inhomogeneity.

Resonant absorption can be considered as an effective process of generating small length scales comparable to the dissipation length scales. The local oscillation modes of an inhomogeneous plasma are represented by continuous spectra for slow MHD and Alfvén waves and a discrete spectrum for fast MHD waves. The resonant absorption occurs when the frequency of a laterally driven oscillation matches the local slow and/or Alfvén wave frequency and a resonant field line is created which transfers energy from the surface disturbance to its environment.

Usually, the importance of the dissipation is characterized by the viscous and magnetic Reynolds numbers (if viscosity and magnetic diffusion are considered as dissipative effects) and we denote by R the total Reynolds number which under solar conditions is a very large number (10^6 in the photosphere and up to 10^{12} in the corona).

Linear theory of resonant absorption has shown that in the vicinity of a resonant position the perturbations have steep gradients and large amplitudes and therefore the linear theory in this region can break down and nonlinear theory has to be considered. Nonlinearity in the dissipative layer was first taken into account in the theory of resonant absorption by Ruderman et al. (1997a) and Ballai et al. (1998a) where they studied the nonlinear evolution of slow resonant MHD waves in the isotropic and anisotropic dissipative layer using a Cartesian geometry. These theories were applied to study the resonant absorption of sound and fast magneto-acoustic waves in solar structures (Ruderman et al., 1997b; Ballai et al., 1998b; Erdélyi & Ballai, 1999). One of their main results was that in contrast to the linear theory, the coefficient of wave energy absorption was dependent on the particular type of dissipation. They have also found that the general tendency of nonlinearity is to decrease the absolute value of the coefficient of wave energy absorption when the wavelength of the incoming wave is much larger than the characteristic scale of the inhomogeneity and nonlinearity is considered weak.

Characteristic quantities used to scale the problem are ϵ (the dimensionless amplitude of perturbations away from the dissipative layer) and the total Reynolds number. One way to determine the importance of nonlinerarity is to calculate the ratio

$$\delta = f\frac{\partial f}{\partial \theta}/\nu\frac{\partial^2 f}{\partial x^2} = \epsilon R^{2/3}, \tag{9}$$

where f is any large variable, i.e. the most singular perturbations (e.g. for slow wave resonance, the most singular are the parallel component of the velocity and magnetic field perturbation). Linear theory works as long as $\delta \ll 1$, i.e. $\epsilon R^{2/3} \ll 1$. For a typical value of $\epsilon \approx 10^{-2}$ to have resonant absorption described by linear theory, we need $R \ll 10^3$ which is in contrast to previously accepted values. Based on these scalings, it is obvious that resonant absorption is a *nonlinear phenomenon*.

In nonlinear theories perturbations cannot be Fourier analysed. However, to be as close as possible to the linear results, we suppose that waves are plane periodic

propagating modes with permanent shape, i.e. all perturbations depend only on $\theta = z - Vt$ so they are periodic with respect to θ.

Outside the dissipative layer, the plasma dynamics can be described by the linear ideal MHD system of equations which can be reduced to two coupled first order PDE

$$\frac{\partial u}{\partial x} = \frac{V}{D}\frac{\partial P}{\partial \theta}, \quad \frac{\partial P}{\partial x} = \frac{\rho_0 D_A}{V}\frac{\partial u}{\partial \theta}, \tag{10}$$

where

$$D = \frac{\rho_0 D_A D_C}{V^4 - V^2(v_A^2 + c_S^2) + v_A^2 c_S^2 \cos^2\alpha},$$
$$D_A = V^2 - v_A^2\cos^2\alpha, \quad D_C = (v_A^2 + c_S^2)(V^2 - c_T^2\cos^2\alpha). \tag{11}$$

In the case of cylindrical tube when the equilibrium magnetic field is such that $\mathbf{B}_0 = (0, B_{0\varphi}(r), B_{0z}(r))$ and the wave-vector now has a helical component, therefore the running variable is $\theta = m\varphi + kz - \omega t$. The governing equations outside the dissipative layer are

$$D\frac{\partial ur}{\partial r} = C_1 ur + \omega C_2 r\frac{\partial P}{\partial r},$$
$$\omega r D\frac{\partial^2 P}{\partial r \partial \theta} = c_3 ur - \omega r C_2, \tag{12}$$

where the coefficient functions are given by

$$C_1 = 2\omega^4\frac{B_{0\varphi}}{\nu r} - 2\frac{m f_B B_{0\varphi}}{\nu r^2}D_C, \quad C_2 = \omega^4 - \left(\frac{m^2}{r^2} + k^2\right)D_C,$$

$$C_3 = D\left[\rho_0 D_A\frac{\partial}{\partial\theta^2} + \frac{2B_{0\varphi}}{r}\frac{d}{dr}\left(\frac{B_{0\varphi}}{r}\right)\right] - 4\omega^4\left(\frac{B_{0\varphi}^2}{\nu r}\right)^2 + \frac{4\rho_0 D_C\omega_A^2}{\nu r^2}B_{0\varphi}^2,$$
$$f_B = \frac{m}{r}B_{0\varphi} + kB_{0z},$$

and the coefficients D, D_A and D_C are similar to the equations given by Eq. (11) with V replaced by ω.

In the present study we only focus on the slow resonance given by the condition $V^2 = c_T^2(x)$ or $\omega^2 = \omega_C^2(r)$. The resonant position ($x = x_C$ in Cartesian geometry and $r = r_C$ in cylindrical geometry) is a regular singular point of the system of Eqs. (11)-(12) and as a consequence, the solutions are obtained in form of Fröbenius series. The equilibrium quantities have a slight change across the dissipative layer and they are approximated by the first non-vanishing term in their Taylor expansion. These expansions are valid in a layer wider than the dissipative layer since the characteristic scale of the inhomogeneity is larger than the scale of dissipation.

Inside the dissipative layer, the solutions are obtained in form of asymptotic expansions. In order to connect the solutions in the two regions (inside and outside the dissipative layer) we use the so-called *matched asymptotic expansions* developed

by Nayfeh (1981). Both the internal and external solutions have to coincide in the overlap regions.

The dynamics of resonant slow waves in the vicinity of the resonance propagating along the magnetic field is given in cartesian geometry (with isotropic anisotropy) by

$$(x-x_c)\frac{\partial v_\parallel}{\partial \theta}+\Phi_1 v_\parallel \frac{\partial v_\parallel}{\partial \theta}+k\frac{\partial^2 v_\parallel}{\partial x^2}=\Phi_2 P(\theta), \tag{13}$$

and in cylindrical geometry

$$(r-r_c)\frac{\partial v_\parallel}{\partial \theta}+\Psi_1 v_\parallel \frac{\partial v_\parallel}{\partial \theta}+k\frac{\partial^2 v_\parallel}{\partial r^2}=\Psi_2 \mathcal{C}(\theta), \tag{14}$$

where the function $\mathcal{C}(\theta)$ is a sum of the θ-derivative of the total pressure and a function containing the φ-component of the magnetic field.

There are two interesting points to be mentioned. The coefficient Φ_1 and Ψ_1 in Eqs. (13)-(14) are similar to the coefficient of the nonlinear terms found for solitons and it provides a measure of nonlinearity in compressional modes. Secondly, in the coronal case, where all transport coefficients are anisotropic, the nonlinear governing equation is modified in the dissipative term (the third term in the LHS) and instead of having a 2^{nd} order derivative with respect to the transversal coordinate, we have a 2^{nd} order derivative with respect to θ. Eqs. (13) and (14) should be understood in the sense that the nonlinear behaviour of slow waves in the vicinity of the resonance is driven by the variation of the total pressure.

When solving the MHD equations for the entire domain, the resonances are considered as singularities, therefore the evolution of physical quantities in the vicinity of resonances are given as jumps (connection formulae), exactly as the Rankine-Hugeniot relations for shock waves. The jump in a quantity Q across the dissipative layer can be calculated with the aid of

$$[Q]=\lim_{x\to x_C}\left\{Q(x)-Q(-x)\right\}.$$

When connecting the solutions, the jump conditions serve as boundary conditions. In the case of Cartesian geometry, the jumps in the total pressure and the normal component of the velocity are given by

$$[P]=0, \quad [u]=\Omega_1 \mathcal{P}\int_{-\infty}^{\infty}\frac{\partial v_\parallel}{\partial \theta}\,dx, \tag{15}$$

and in cylindrical geometry by

$$[P]=\Lambda_1 \mathcal{P}\int_{-\infty}^{\infty} v_\parallel\,dr, \quad [u]=\Lambda_2 \mathcal{P}\int_{-\infty}^{\infty}\frac{\partial v_\parallel}{\partial \theta}\,dr. \tag{16}$$

Here $\mathcal{P}$ is used for the Cauchy principal part because the integrals are divergent at infinity.

When calculating the efficiency of the resonant absorption (coefficient of wave energy absorption) it is found that the effect of nonlinearity is to decrease the net coefficient of wave absorption. This means that the largest amount of energy stored in nonlinear waves does not go into increasing the absorption rate but into generating a mean flow outside the dissipative layer. This turbulent flow is generated by the absorption of wave momentum in the dissipative layer and its amplitude is determined by the balance of forces created by resonant absorption and shear viscosity. The mean shear flow is a piecewise continuous function of r (e.g. in cylindrical geometry) but its vorticity has a jump given by

$$[v_\varphi = A_1 \int_{-\infty}^{\infty} \left\langle \left(\frac{\partial v_\parallel}{\partial r} \right)^2 \right\rangle dr, \quad [v_z] = A_2 \int_{-\infty}^{\infty} \left\langle \left(\frac{\partial v_\parallel}{\partial r} \right)^2 \right\rangle dr, \tag{17}$$

where the coefficients A_1 and A_2 depend on characteristic speeds, the location, r_C, of the resonance and the dissipative coefficients and $<,>$ is the mean value of a quantity over a period. Estimates of this mean shear flow give us speeds of the order of $0.1 km/s$ in the solar photosphere and a few km/s in the solar corona. Observation of this flow might be a first indirect evidence for resonant absorption in solar plasmas. The properties of generated mean turbulent flow are not fully understood and they are an important topic for further investigations.

The results presented here considered that the equilibrium is static; in reality the plasma is very dynamic, showing motion on all time and space scales. Including an equilibrium steady flow, Ballai & Erdélyi (1998c) obtained the governing equations inside and outside the dissipative layer, as well as the jump conditions across the singularity.

The model described here considered a simplified atmosphere. Possible further investigations could be performed for a more realistic equilibrium (e.g. equilibrium quantities vary not only across the field but also along the field, inclusion of gravity, etc.). The governing equations were obtained in the limit of weak nonlinearity and long wavelength approximation. Recently, Ruderman (2000) considered the analysis of resonantly interacting waves in the limit of strong nonlinearity. He has obtained that the decreasing tendency of the coefficient of wave energy absorption by nonlinearity does not persist in this limit for intermediate values of wave vector. In the long wavelength limit, however, he found that the difference between strong nonlinear and linear limit does not exceed 20%.

Acknowledgement

The author acknowledges the financial support by the British Council, Nuffield Foundation and the great effort of the local organizers to make the meeting a time to remember.

References

Ballai, I., Ruderman, M.S. & Erdélyi, R. 1998, *Phys. Plasmas*, 5, 252.

Ballai, I., Erdélyi, R. & Ruderman, M.S. 1998, *Phys. Plasmas*, 5, 2264.

Ballai, I. & Erdélyi, R. 1998 Sol. Phys., 180, 65

Ballai, I., Erdélyi R., Voitenko, Y. & Goossens, M. 2002, *Phys. Plasmas*, 9, 2593.

Ballai, I. & Zhugzhda, Y.D. 2002, *Phys. Plasmas*, 9, 4280.

Edwin, P & Roberts, B. 1986, *Wave Motion*, 8, 151.

Erdélyi, R. & Ballai, I. 1999, Sol. Phys., 186, 67

Nayfeh, A.H. 1981, *Introduction to Pertubation Techniques*, Wiley-Interscience, New York.

Roberts, B. & Mangeney, A. 1982, *MNRAS*, 198, 7P-11P

Roberts, B. 1981, *Sol. Phys*, 69, 27

Roberts, B. 1985, *Phys. Fluids*, 28, 3280

Ruderman, M.S., Goossens, M. & Hollweg, J.V. 1997a, *Phys. Plasmas*, 4, 75.

Ruderman, M.S., Hollweg, J.V. & Goossens, M. 1997b *Phys. Plasmas*, 4, 92.

Ruderman, M.S. 2000 *J. Plasma Phys.*, 63, 43.

Sakurai, T., Goossens, M. & Hollweg, J.V. 1991a, Sol. Phys., 133, 227.

E. Weisshaar, *Phys. Fluids* **1**, 1406 (1989)

PADEU | PADEU **15**, 83 (2005)
ISBN 963 463 557

Wave Propagation In Stratified One Dimensional Waveguides

J. Hargreaves

Solar Physics and upper-Atmosphere Research Group, Department of Applied Mathematics, University of Sheffield, S3 7RH, Sheffield, U.K.

E-mail: j.hargreaves@shef.ac.uk

Abstract
We investigate the effect of stratification by gravity on the propagation of linear magnetohydrodynamic (MHD) waves along a magnetic flux tube. For a quiescent environment linear wave propagation is governed by a Klein-Gordon equation. We consider the presence of a background flow in the tube and show the response to various atmospheric footpoint drivers. We solve for three drivers corresponding to a monochromatic source, a delta-function pulse and a sinusoidal pulse. This work is motivated by the vast amount of recent observational evidence supporting the existence of waves and flows in solar MHD waveguides.

1 Introduction

The effect of stratification on the propagation of MHD waves is important, especially in the lower atmosphere. These effects, of course, have been considered previously (e.g Rae and Roberts, 1982). In this paper we consider take a look at the effect of flows upon wave propagation in such atmospheres.

Recent observational evidence gives strong support to the long discussed possibility of flows in the solar atmosphgere. Waves will be influenced and possibly even generated by their presence. A detailed derivation of steady flow effects on MHD waveguides can be found e.g. in Terra-Homem, Erdélyi & Ballai (2003). The identification of a fast wind originating from the polar coronal holes has been established for many years (e.g., Watanabe, 1975; Gloeckner and Geiss, 1998). Observations from SOHO and TRACE have indicated the presence of steady flows in the south polar coronal hole

and the equatorial quiet Sun-region (Buchlin and Hassler, 2000). Bi-flows, associated with tens of thousands of small scale explosions, have been found in the chromosphere (e.g., Innes et al., 1997; Perez et al. 1999; Sarro et al., 1999; Teriaca et al. 1999; Roussev et al. 2001a,b,c, etc.). Background flows have been noted in arched isolated magnetic flux tubes, steady flows have been observed in slender magnetic flux tubes and in return flows from spicules. Thus there is a clear need to study the effect of such flows on the wave characteristics in the solar atmosphere, as flows may influence not just the dynamic properties of wave propagation, but also wave dissipation (Erdélyi 1996; Erdélyi & Goossens 1996; etc.).

2 Wave Propagation in an Elastic Tube

The Sun's photospheric region exhibits enormous complexity, being both highly dynamic and highly structured, with magnetic structuring over many scales. Further, stratification due to gravity is important and must be taken into account. The use of slender flux tube theory (wavelengths much greater than tube width) allows an analytical description (see Roberts and Webb, 1978). Consider a general slender elastic tube in a stratified atmosphere (e.g. Rae and Roberts 1982; Erdélyi & Fedun 2004). The atmosphere is stratified such that

$$p_0' = -g\rho_0(z), \tag{1}$$

where the prime ($'$) shows a derivative with respect to the z direction. The governing equations of the system are given by the linearised equations for continuity, momentum and energy

$$\frac{\partial}{\partial t}\left(\rho_1 A_0 + \rho_0 A_1\right) + \left(\rho_0 A_0 u_1\right)' = 0, \tag{2}$$

$$\rho_0 \frac{\partial u_1}{\partial t} = -p' - \rho g, \tag{3}$$

$$\frac{\partial p}{\partial t} + p_0' u_1 = c_0^2 \left(\frac{\partial \rho}{\partial t} + \rho_0' u_1\right), \tag{4}$$

where perturbed quantities are functions of z and t and equilibrium quantities are functions of z only. It is assumed that vertical motions are dominate and thus are considering sausage modes only.

With the introduction of $c(z)$ (Lighthill, 1978)

$$\frac{1}{c^2} = \frac{1}{c_0^2} + \frac{\rho_0}{A_0}\left(\frac{\partial A_1}{\partial p_1}\right)_{p_1=0}, \tag{5}$$

then $Q(t,z)$ (a scaled velocity) can be seen to satisfy

$$\frac{\partial^2 Q}{\partial t^2} - c^2(z)\frac{\partial^2 Q}{\partial z^2} + \Omega_0^2(z)Q = 0, \tag{6}$$

where

$$\begin{aligned}\Omega_0^2 &= N_0^2 + c^2\left[\frac{1}{2}\left(\frac{\rho_0'}{\rho_0} + \frac{A_0'}{A_0} + \frac{c_0^{2\prime}}{c_0^2}\right)' + \frac{1}{4}\left(\frac{\rho_0'}{\rho_0} + \frac{A_0'}{A_0} + \frac{c_0^{2\prime}}{c_0^2}\right)^2 + \right.\\ &\left. + \left(\frac{g}{c_0^2} - \frac{A_0'}{A_0}\right)' + \left(\frac{g}{c_0^2} - \frac{A_0'}{A_0}\right)\left(\frac{\rho_0'}{\rho_0} + \frac{c_0^{2\prime}}{c_0^2} + \frac{g}{c_0^2}\right)\right],\end{aligned} \tag{7}$$

and N_0^2, given by

$$N_0^2 = -g\left(\frac{\rho_0'}{\rho_0} + \frac{g}{c_0^2}\right), \tag{8}$$

is the square of the Brunt-Väisälä frequency. Notice that (6) is a Klein-Gordon type differential equation, with all its attendant features (see Rae and Roberts 1982, Roberts 2004).

2.1 Addition of a Background Flow

We extend the earlier work of Rae and Roberts by the addition of a background flow inside the tube. A full treatment, with the flow, U_0, varying with height, is as yet unavailable since the problem becomes rapidly intractable. Here we consider the simpler, but physically less realistic, case of steady flow inside the tube, such that $U_0'(z) = 0$. The addition of the flow modifies the governing equations and following a similar method, $Q(t,z)$ can be shown to satisfy

$$\frac{D^2 Q}{Dt^2} - U_0\left(\frac{A_0'}{A_0} + \frac{c_0^{2\prime}}{c_0^2}\right)\frac{DQ}{Dt} - c^2(z)\frac{\partial^2 Q}{\partial z^2} + \Omega_0^2(z)Q + S = 0, \tag{9}$$

where

$$\begin{aligned}\Omega_0^2 &= N_0^2 + (c^2 - U_0^2)\left[\frac{1}{2}\left(\frac{\rho_0'}{\rho_0} + \frac{A_0'}{A_0} + \frac{c_0^{2\prime}}{c_0^2}\right)' \frac{1}{4}\left(\frac{\rho_0'}{\rho_0} + \frac{A_0'}{A_0} + \frac{c_0^{2\prime}}{c_0^2}\right)^2 + \right.\\ &\left. + \left(\frac{g}{c_0^2} - \frac{A_0'}{A_0}\right)' + \left(\frac{g}{c_0^2} - \frac{A_0'}{A_0}\right)\left(\frac{\rho_0'}{\rho_0} + \frac{c_0^{2\prime}}{c_0^2} + \frac{g}{c_0^2}\right)\right],\end{aligned} \tag{10}$$

and

$$S = \frac{-1}{\rho_0 R^{\frac{1}{2}}} \left[\left(\frac{c^2 U_0}{A_0} (\rho_0' A_1 + A_0' \rho_1) \right)' + \frac{g U_0 c^2}{c_0^2} A_0 (\rho_0' A_1 + A_0' \rho_1) \right], \tag{11}$$

and where

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + U_0 \frac{\partial}{\partial z}. \tag{12}$$

Equation (9) describes the wave motion in a general elastic tube with the presence of a steady background flow. It should be noted that Eq. (9) is not of the Klein-Gordon type and also note the term S, given by (11), is a function of the perturbations ρ_1 and A_1. We are seeking a result that is a function of the scaled velocity Q only and thus to proceed we take S as being negligible.

2.1.1 Basic Formation of Solution

In order to demonstrate some of the features of (9) we consider further simplifications to a straight ($A_0' = 0$) and rigid tube ($c = c_0$). Further, if the gas inside the tube is considered isothermal (c_0 is a constant) then (9) reduces to that describing isothermal acoustic-gravity waves with a background flow

$$\frac{D^2 Q}{Dt^2} - c_0^2(z) \frac{\partial^2 Q}{\partial z^2} + \Omega_a^2(z) Q = 0, \tag{13}$$

where

$$\Omega_a^2 = \frac{c_0^2 - U_0^2}{4\Lambda_0^2}, \tag{14}$$

where Ω_a^2 is a modified acoustic cut-off frequency, and Λ_0 is the isothermal scale height, given by $\Lambda_0 = p_0/\rho_0 g$. If we consider Fourier forms of solution, $e^{i(\omega t - kt)}$, we obtain the following dispersion relation

$$\omega_D^2 = k^2 c_0^2 + \Omega_a^2, \tag{15}$$

where $\omega_D = \omega + U_0 k$ is the *Doppler-shifted* frequency.

To solve (13) analytically we consider Laplace transformations in time. Since we consider an isothermal medium, c_0 and Ω_a are constant. A wave source is introduced into the atmosphere by applying a driver, $A(t)$, at $z = 0$. Thus the atmosphere is bounded and it is assumed that the atmosphere is initially at rest with no wave motions or derivatives of wave motions present. Further it is assumed there are no wave motions at $z = \infty$. The problem is a homogeneous Klein-Gordon type equation with inhomogeneous boundary conditions. We can thus obtain the general solution, given by

$$Q(t,z) = A_0((t-(t_c-t_u))H(t-(t_c-t_u)) + \int_0^t A_0(t+t_u-\tau)\cdot H(t+t_u-\tau)\cdot P(\tau,z)d\tau, \tag{16}$$

where

$$P(\tau,z) = -\frac{z}{2\Lambda_0}\frac{J_1\left(a\sqrt{\tau^2-t_c^2}\right)}{\sqrt{\tau^2-t_c^2}} \times H(\tau-t_c), \tag{17}$$

and

$$a = \sqrt{\frac{\Omega_a^2}{c_0^2}(c_0^2-U_0^2)} = \frac{c_0^2-U_0^2}{2\Lambda_0 c_0}, \qquad t_c = \frac{c_0 z}{c_0^2-U_0^2}, \qquad t_u = \frac{U_0 z}{c_0^2-U_0^2}, \tag{18}$$

where $\Omega_a^2 = (c_0^2-U_0^2)/4\Lambda_0^2$, a can thus be thought of as an acoustic cut-off frequency analogous to $\omega_a(=c_0/2\Lambda_0)$, but having been modified by the presence of the flow, from this point on referred to as the *flow acoustic cut-off*. Note the fact the as $U_0 \to 0$, $\Omega_a \to \omega_a, a \to \omega_a$, $t_c \to z/c_0$ and $t_u \to 0$, thus we return exactly to the Sutmann et al case. We consider only the case for $U_0 > 0$ in this analysis.

We now solve the general solution for variuos prescribed drivers, namely a monochromatic driver, an impulsive driver and finally a sinusoidal driver, each applied at $z = 0$.

3 Monochromatic Source

As per Sutmann et al. (1998) we consider first the boundary conditions to be satisfied by a source of monochromatic acoustic waves, generated continuously with the frequency ω and amplitude Q_0. Thus we have

$$A_0(t) = Q_0 e^{-i\omega t}. \tag{19}$$

Substituting this boundary condition into (16) we obtain

$$Q(t,z) = Q_0\left[e^{-i\omega(t-(t_c-t_u))} \times H(t-(t_c-t_u)) + \int_0^t e^{-i\omega(t+t_u-\tau)}P(\tau,z)d\tau\right]. \tag{20}$$

Notice the removal of the Heaviside function from (16), this is due to the fact that in the interval $(0,t)$, $H(t+t_u-\tau) = 1$ since we consider t_u to be positive. Let the integral in this expression be given by $I = I_1 - I_2$ where

$$I_1 = -e^{-i\omega(t+t_u)}\frac{z}{2\Lambda_0}\int_0^\infty \frac{J_1\left(a\sqrt{\tau^2-t_c^2}\right)}{\sqrt{\tau^2-t_c^2}}H(\tau-t_c)e^{i\omega\tau}d\tau, \tag{21}$$

and I_2 being the same, save for the limits being now (t, ∞). The integral I_1 may be evaluated directly, however the integral I_2 cannot be evaluated analytically unless we apply the condition $\tau >> t_c$ and thus $t >> t_c$. Thus we take

$$I_2 \approx -e^{-i\omega(t+t_u)} \frac{z}{2\Lambda_0} \int_t^\infty \frac{J_1(a\tau)}{\tau} e^{i\omega\tau} d\tau. \tag{22}$$

We have applied the condition that τ is large and so $J_1(a\tau)$ can be expanded asymptotically, thus

$$J_1(a\tau) \approx \sqrt{\frac{2}{\pi a\tau}} \cos\left(a\tau - \frac{3\pi}{4}\right), \tag{23}$$

and by applying the exponential form of the cosine function we obtain

$$I_2 = -e^{-i\omega(t+t_u)} \frac{z}{4\Lambda_0} \sqrt{\frac{2}{\pi a}} \left[\int_t^\infty \frac{e^{i(a\tau - 3\pi/4)} e^{i\omega t}}{\tau^{3/4}} d\tau + \int_t^\infty \frac{e^{-i(a\tau - 3\pi/4)} e^{i\omega t}}{\tau^{3/4}} d\tau \right]. \tag{24}$$

Which may be evaluated (see Sutmann et al. (1998) Appendix B for a detailed discussion) to give

$$I_2 = \frac{z}{2\Lambda_0} \sqrt{\frac{2}{\pi a}} \frac{1}{t^{3/2}} \frac{1}{a^2 - \omega^2} e^{-i\omega t_u} \left[a \sin\left(at - \frac{3\pi}{4}\right) + i\omega \cos\left(at - \frac{3\pi}{4}\right) \right]. \tag{25}$$

Substituting $I_{1,2}$ into (20), we obtain

$$\begin{aligned} Q(t,z) &= Q_0 e^{-i\omega t_u} \left[e^{-i(\omega t + t_c\sqrt{\omega^2 - a^2})} - \right. \\ &\quad \left. - \frac{z}{2\Lambda_0} \sqrt{\frac{2}{\pi a}} \frac{1}{t^{3/2}} \frac{1}{a^2 - \omega^2} \left(a \sin\theta + i\omega \cos\theta\right) \right], \end{aligned} \tag{26}$$

where $\theta = (at - 3\pi/4)$, being an expression fully describing the atmosphere responding to a source of monochromatic acoustic waves, valid for $w \neq a$. Following the discussion in Sutmann et al. (1998) we see the velocity $Q(t, z)$ is again comprised of two ocscillations superimposed upon each other. Firstly, the *forced atmospheric osciallations* at the driving frequency ω, representing well-known prpoagating acoustic waves for $\omega > a$ and evanescent waves for $\omega < a$. Secondly, we have the *free atmospheric oscillations* at the frequency a, given by the second term on the RHS, decaying with $t^{-3/2}$ and increasing linearly with z.

4 Delta Function Pulse

Here we consider the source to be a single pulse of δ- function form. Thus the boundary condition is

$$A_0(t) = Q_0'\delta(t). \tag{27}$$

The δ-function has a non dimensional argument, hence Q_0' has different dimensions to Q_0. Thus we introduce a non-dimensional time, t', by $t = Rt'$, where $R = 2\pi/a$. Letting $Q_0 = Q_0'/R$ and by assuming $t' >> t_c/R > t_u/R$, the fist term on the RHS of (16) disappears and we obtain

$$Q(t', z) = \int_0^{t'R} Q_0 R\delta(t' - \tau')P(\tau' R, z)d\tau', \tag{28}$$

thus by employing a fundamental property of the δ-function and rewriting for $Q(t, z)$ we obtain

$$\begin{aligned} Q(t, z) &= Q_0 RP(t, z), \\ &= Q_0 R\frac{z}{2\Lambda_0}\frac{J_1\left(a\sqrt{t^2 - t_c^2}\right)}{\sqrt{t^2 - t_c^2}}, \end{aligned} \tag{29}$$

and since $t' >> t_c/R$ the Bessel function reduces to $J_1(at)$ which can be evaluated asymptotically as before, giving

$$Q(t, z) = -Q_0 R\frac{z}{2\Lambda_0}\sqrt{\frac{2}{\pi a}}\frac{1}{t^{3/2}}\cos\theta, \tag{30}$$

where $\theta = (at - 3\pi/4)$, being very similar to the no-flow case in from apart from the change in frequencies.

5 Sinusoidal Pulse

The final case we consider is a sinusoidal pulse generated at the lower boundary and lasting for one wave period P, where $P = 2\pi/\omega$. Thus the boundary condition is given by

$$A_0(t) = Q_0\left[H(t) - H(t - P)\right]e^{-i\omega t}. \tag{31}$$

Considering the first part of (16) we see that since we are mainly concerned with $t >> t_c$ and our analysis is for $c_0 > U_0$ we can see that the arguments of both Heaviside functions are positive and so they cancel each other out. Thus we have

$$Q(t,z) = Q_0 \int_0^t P(\tau,z) H(t+t_u-\tau) e^{-i\omega(t+t_u-\tau)} d\tau - \\ -Q_0 \int_0^t P(\tau,z) H(t+t_u-P-\tau) e^{-i\omega(t+t_u-\tau)} d\tau. \quad (32)$$

In the interval $(0,t)$ the first Heaviside function is unity since we assume t_u positive, however two cases can be considered for the second, depending on the relative values of P and t_u. For the case of $P < t_u$ then in the interval $(0,t)$ the Heaviside function is unity and so the two terms cancel resulting in $Q(t,z) = 0$, implying there is no atmospheric response to the sinusoidal pulse. For $P > t_u$ then we see that the limits of integration may be changed to be between $(0, t-b)$, where $b = P - t_u$. Thus we can write (32) as

$$Q(t,z) = Q_0 \int_0^t P(\tau,z) e^{-i\omega(t+t_u-\tau)} d\tau - Q_0 \int_0^{t-b} P(\tau,z) e^{-i\omega(t+t_u-\tau)} d\tau, \quad (33)$$

and thus combining the two integrals together by changing the limits and rewriting it in the form two integrals, I_1 and I_2 so $Q(t,z) = I_1 - I_2$, giving

$$Q(t,z) = Q_0 \int_{t-b}^{\infty} P(\tau,z) e^{-i\omega(t+t_u-\tau)} d\tau - Q_0 \int_t^{\infty} P(\tau,z) e^{-i\omega(t+t_u-\tau)} d\tau, \quad (34)$$

and solving asymptotically and then by parts and taking the first term only, as in Section 3, giving

$$I_1 = \frac{z}{2\Lambda_0} \sqrt{\frac{2}{\pi a}} \frac{1}{(t-b)^{3/2}} \frac{1}{a^2-\omega^2} e^{-i\omega t_u} \times \\ \times \left[a \sin\left(a(t-b) - \frac{3\pi}{4} \right) + i\omega \cos\left(a(t-b) - \frac{3\pi}{4} \right) \right], \quad (35)$$

and

$$I_2 = \frac{z}{2\Lambda_0} \sqrt{\frac{2}{\pi a}} \frac{1}{t^{3/2}} \frac{1}{a^2-\omega^2} e^{-i\omega t_u} \left[a \sin\left(at - \frac{3\pi}{4} \right) + i\omega \cos\left(at - \frac{3\pi}{4} \right) \right]. \quad (36)$$

Since we are dealing with large times we can expand the $1/(t-b)^{3/2}$ in equation (35) binomially and taking the first term only the combined result is

$$Q(t,z) = \frac{z}{2\Lambda_0}\sqrt{\frac{2}{\pi a}}\frac{1}{t^{3/2}}\frac{1}{a^2-\omega^2}\left\{a\left[\sin\left(a(t-b)-\frac{3\pi}{4}\right)-e^{-iwt_u}\sin\theta\right]+ \right.$$
$$\left. +i\omega\left[\cos\left(a(t-b)-\frac{3\pi}{4}\right)-e^{-iwt_u}\cos\theta\right]\right\}, \quad (37)$$

where $\theta = (at - 3\pi/4)$.

References

Buchlin, E. & Hassler, D.M. in *AAS/Solar Physics Division Meeting*, Vol. 32, 201 (2000)

Erdélyi, R. in Y. Uchida, T. Kosugi, & H. S. Hudson (eds.) *Magnetohydrodynamic Phenomena in the Solar Atmosphere'*, Kluwer, pp.47-48 (1996)

Erdélyi, R. & Fedun, V. *Solitary wave propagation from the photopshere into the Solar Corona*, in (ed.) H. Lacoste, SOHO 13 *Waves, Oscillations and Small Scale Transient Events in the Solar Atmosphere*, *ESA-SP*, **547**, 63 (2004)

Erdélyi, R. & Goossens, M. *A&A*, **313**, 664 (1996)

Gloeckner, G. & Geiss, J. *Space Sci. Rev.*, **86**, 127 (1998)

Innes, D.E., Inhester, B., Axford, W.I. & Wilhelm, K. *Nature*, **386**, 811 (1997)

Lighthill, J. *Waves in Fluids* (Cambridge University Press) (1978) p.93

Perez, M.E., Doyle, J.G., Erdélyi, R. & Sarro, L.M., *A&A*, **342**, 279 (1999)

Rae, I. & Roberts, B. *ApJ*, **265**, 761 (1982)

Roberts, B. *MHD Waves in the Solar Atmosphere*, in (ed.) H. Lacoste, SOHO 13 *Waves, Oscillations and Small Scale Transient Events in the Solar Atmosphere*, *ESA-SP*, **547**, 1 (2004)

Roberts, B. & Webb, A.R. *Sol. Phys.*, **56**, 5 (1978)

Roussev, I., Galsgaard, K., Erdélyi, R. & Doyle, J.G. *A&A*, **370**, 298 (2001a)

Roussev, I., Galsgaard, K., Erdélyi, R. & Doyle, J.G. *A&A*, **375**, 228 (2001b)

Roussev, I., Doyle, J.G., Galsgaard, K., Erdélyi, R. & *A&*, **380**, 719 (2001c)

Sarro, L.M., Erdélyi, R., Doyle, J.G. & Perez, E.P. *A&A*, **351**, 721 (1999)

Sutmann, G., Musielak, Z.E. & Ulmschneider, P. *A&A.*, **340**, 556 (1998)

Teriaca, L., Doyle, J.G., Erdélyi, R. & Sarro, L.M. *A&A*, **352**, L99 (1999)

Terra-Homem, M., Erdélyi, R. & Ballai, I. *Sol. Phys.*, **217**, 199 (2003)

Watanabe, T. *Publ. Astron. Soc. Japan*, **27**, 385 (1975)

PADEU | PADEU **15**, 93 (2005)
ISBN 963 463 557

Identification and Statistical Analysis of Hot Flow Anomalies Using Cluster Multi-Spacecraft Measurements

G. Facskó, K. Kecskeméty, M. Tátrallyay and G. Erdős

KFKI Research Institute for Particle and Nuclear Physics, H-1525 Budapest, Pf. 49, Hungary

E-mail: gfacsko@rmki.kfki.hu

Abstract
Discovered nearly 20 years ago near the Earth's bow shock, the identification and separation of Hot Flow Anomalies (HFAs) from other events is still under debate. We have used the observations of instruments FGM, CIS, and RAPID aboard the four Cluster spacecraft to detect and study these phenomena. The definition and basic features of HFAs (size, direction of tangential discontinuity, electric field, speed of propagation) have been refined, several series of events identified, and a preliminary statistical analysis carried out. After combining data from RAPID and FGM the pitch angle distributions of protons have been calculated. The measured and calculated features of HFA events are confronted with the results of hybrid simulations.
Keywords: *hot flow anomaly, Earth's bow shock, tangential discontinuity*

1 Introduction

ESA's cornerstone mission, CLUSTER has been the most successful magnetosphere project since the satellites were launched in 2000 (Escoubet et al., 1997). The orbits of the four satellites cross the most important parts of the cosmic neighborhood of our planet. Their apogee are upstream of the Earth's bow-shock during spring on the

Northern Hemisphere. This fact, using FGM[1] and CIS[2] instruments allows to detect the relatively rare phenomena called HFA[3] effectively whereas the RAPID energetic particle instrument provides additional information. We processed the measurements of CLUSTER from February 1 through April 8, 2003 and identified nearly 50 HFA events, performed a statistical analysis and refined their basic features. After combining data from the RAPID and FGM instruments pitch angle distributions were calculated using a transformation into the plasma frame (Facskó, 2004). Although no perfect theoretical explanation of HFAs has been made so far, the two hybrid simulations developed seem to reproduce the basic features (Thomas et al., 1991; Lin, 1997, 2002, 2003). The observations of the HFA events found are confronted with the results of these hybrid simulations.

The outline of the paper is as follows: identification, analysis, measured and calculated features of hot flow anomalies are described in Section 2. The comparison of hybrid simulations with observations is presented in Section 3. A summary is given in Section 4

2 Hot Flow Anomalies

No detailed theory has been presented yet which could reproduce all properties of HFAs since they were discovered (Schwartz et al., 1985; Thomsen et al., 1986). Reconnection and ion-beam instability are assumed as energy source of HFAs (Thomas et al., 1991; Lin, 1997). The formation and development of HFAs have been modeled by hybrid simulations (Thomas et al., 1991; Lin, 1997, 2002, 2003). Several features of HFAs were given and their sketch was constructed based upon single spacecraft observations (See: Fig. 1, left panel based on Sibeck et al. (1999); Fig. 1, right panel). The main problem of detecting HFAs is the relatively small size of the volume affected. Satellites must have been at the right place at the right time. The Cluster project has changed this situation revolutionary: the four satellites cover huge space and all satellites have suitable instruments detecting HFA events. Lucek et al. (2004) made the first attempt to use 4-point measurements and examined 3 HFAs within a short time interval when the s/c separation was small (≈100 km).

2.1 Detecting HFAs

Figure 2 depicts a typical hot flow anomaly event and represents criteria for searching. HFAs are the result of the interaction of a tangential discontinuity with the bow-shock of the Earth (Schwartz et al., 2000) or of other object (Øieroset, 2001). We set up the following conditions, which might indicate the presence of a HFA:

[1] **F**lux **G**ate **M**agnetometer
[2] **C**luster **I**on **S**pectrometer
[3] **H**ot **F**low **A**nomaly

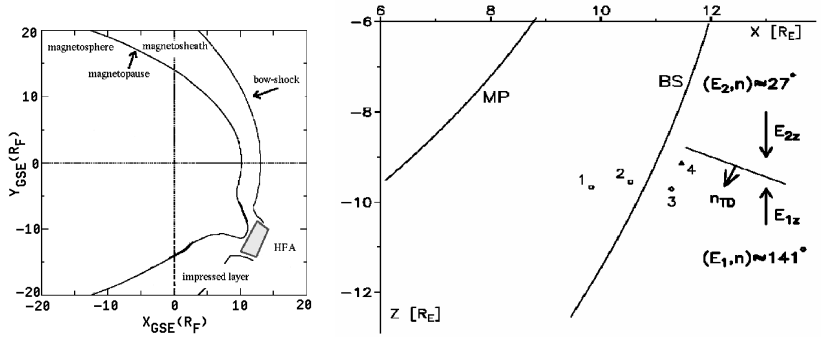


Figure 1: *Left: The supposed structure of HFAs: a tangential discontinuity crosses the bow-shock and they form a hot, tenuous diamagnetic cavity. HFAs seem to be a bulge on both bow-shock and magnetopause. Right: Sketch of the HFA event on February 16, 2003. The normal vector of the tangential discontinuity, the direction of the electric field and the location of the bow-shock and the magnetopause are calculated and plotted.*

- Behavior of the magnetic field, measurements of FGM (Fig. 2, 2nd panel):
  - HFAs appear as a bulge on the bow-shock so one should search them upstream of the shock. That means that the magnitude of the magnetic field has to be near the average interplanetary value.
  - The event begins when the magnitude of the magnetic field drops.
  - FGM observes fast fluctuations in the magnitude of the field and its direction turns around.
  - After the HFA B reaches its value prior to the event.
- Behavior of the solar wind, measurements of CIS HIA[4]:
  - The solar wind slows down, its flow direction might turn back (Fig. 2, 3rd panel).
  - The plasma temperature increases up to several 10 MK (Fig. 2, 4th panel).
  - The plasma density decreases (Fig. 2, 5th panel).
- The fluxes of energetic particles usually increase in the four lower energy channels of RAPID but not always. One can obtain an angular resolution of the

[4] **H**ot **I**on **A**nalyser

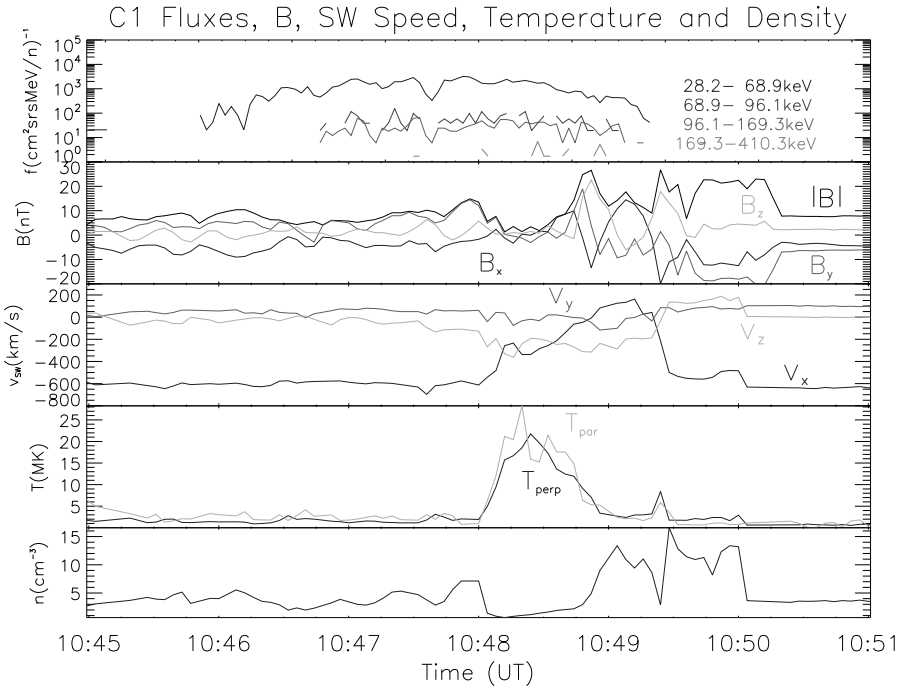


Figure 2: *The first HFA event, detected at 10:45-10:50 UT, on February 16, 2003. Top panel: proton fluxes in four energy channels measured by RAPID aboard Cluster-1. 2nd panel: components and absolute value of magnetic field. 3rd, 4th and 5th panels: components of the solar wind speed, parallel and perpendicular temperature, and the particle density of the solar wind.*

particle fluxes (16×12pixel) and calculate pitch angle distributions by combining data from the RAPID and FGM instruments (Fig. 2, 1st panel).

- The right direction of electric field seems to be the most important condition. The E field focuses particles towards the tangential discontinuity so that its direction should point towards it on both sides.

We found about fifty candidates of HFA events after processing Cluster measurements from February to April, 2003 when the separation between the 4 s/c was of the order of 10,000 km.

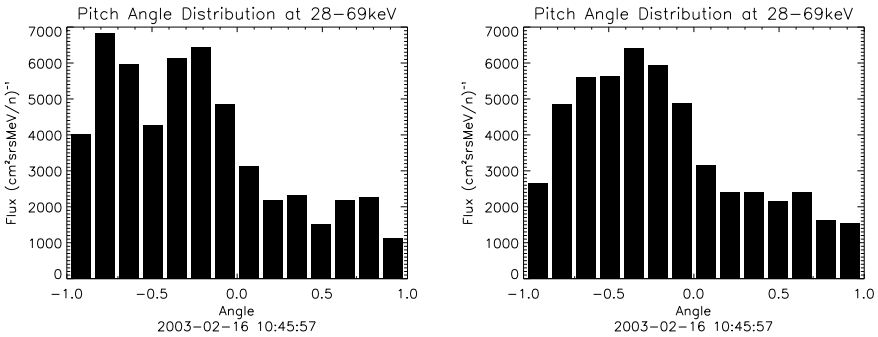


Figure 3: *Pitch angle distribution of 28-69 keV protons during the HFA event at 10:45-10:50 UT on February 16, 2003. Left: PAD calculated with 128s averaged magnetic field. Right: Average of four PADs calculated with 32s averaged magnetic field.*

2.2 Pitch Angle Distributions

After combining data from the RAPID and FGM instruments pitch angle distributions (PADs) of the lowest energy (28-69 keV) protons were calculated for the HFA event 10:45-50 UT on February 16, 2003. The sensitivities of the IIMS[5] detector heads have significantly decreased since launch, we have corrected for this, then the fluxes were transformed to the frame of solar wind using actual CIS solar wind speed measurements and spectral slope from RAPID (Compton-Getting effect). Unfortunately, the cycle of collecting a full directional distribution was about 128 s, 32 times the spin period of Cluster. The simplest choice is to average the magnetic field direction over this period and then compute the pitch angles accordingly. However, the direction of B usually changes much more rapidly than that (Fig. 3, left panel). In lack of higher resolution information one can try to average the B directions over shorter periods (say 32 s instead of 128 s) and assume that the measured directional distribution does not change significantly over that time. Then one has 4 different PADs for the 128 s period which can be averaged. We carried out the two procedures, divided the $[-1, 1]$ interval of $\cos\alpha$ into subintervals, added the fluxes and counted the number of points of those points which have suitable pitch angle. We calculated the average flux of subintervals and then we plotted both pitch angle distributions (Fig. 3, right panel).

2.3 Statistical Aspects

As a result of a survey of the period between February 15 and April 20 about 50 events were found that fulfilled out requirements for HFA events. Most of them were seen by

[5]**I**maging **I**on **M**ass **S**pectrometer

all Cluster satellites while some of them only by one or two spacecraft. Summarizing their general parameters we found that x component of the solar wind decreased usually by 200-400 km/s (in 3 extreme cases the plasma speed became antisolar but in some other events the drop was not more than 50 km/s). V_z changed significantly in all events because they were observed at higher latitudes. The proton density as measured by the HIA sensor of CIS dropped to or below $1/cm^3$ in most cases (about 80% of all), and the parallel proton temperature increased to more than 10 MK in nearly all events. The magnetic field in the cavity was usually below 3 nT. The differences observed in parameters at different spacecraft are generally small. The energetic proton signatures associated with these events were highly variable: in about 20 cases the 28-69 keV proton flux peak exceeded 1000 p/(cm^2 s sr keV) for at least one spacecraft. The particle events usually exhibited smooth profiles starting before and ending after the plasma and magnetic field signatures.

The events were not randomly distributed in time, many of them appeared in sequences within about 1-2 hours (4 on 16 February, 4 on 17 February, 3 on 7 March, 7 on 17 March, 5 on 19 March, 7 on 21 March, and 4 on 24 March). This may indicate preferable conditions rather than grouping of discontinuities. In most cases the solar wind velocity and dynamic pressure was very high compressing the magnetopause and the bow shock.

3 Hybrid Simulations

The hybrid simulation of the plasma is a combination of a full particle approach and a fluid approach. It models the plasma dynamics by treating the ions as particles and the electrons as a charge-neutralizing fluid. Many kinds of hybrid code exist; ions can be treated as macroparticles, the electron fluid might be massless or can have finite mass, the network of the simulation might be two or three dimensional, Cartesian or curvilinear, adaptive or not adaptive. Thomas et al. (1991) and Lin (1997, 2002, 2003) have developed hybrid HFA simulation codes. We studied, checked and compared their result with our observations.

3.1 Results

The hybrid code developed by Thomas et al. (1991) studied the close surroundings of HFAs. The tangential discontinuity intersected the bow shock perpendicularly and they sliced the simulation space and made magnetic field, temperature and particle density diagrams. The form of both magnetic field and particle density diagrams is similar: two small peaks appear at the beginning and the end of the event and both quantities decrease in the middle of the event. The temperature increases in his model.

The other hybrid code developed by Lin (1997, 2002, 2003) uses larger simulation space. A solid target is inserted into superalvénic flow and bow shock form. The flow is

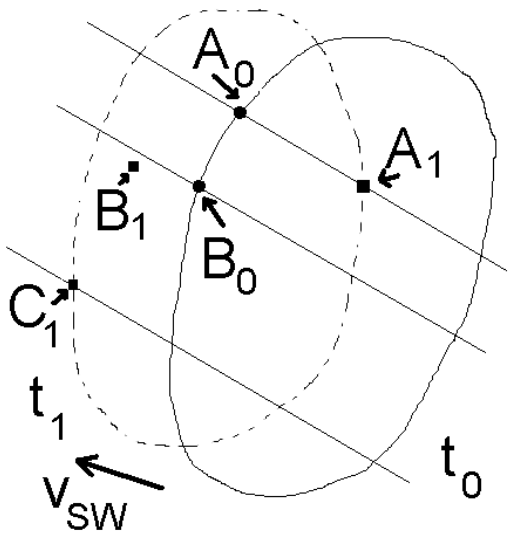


Figure 4: *Two methods for HFA size estimation. A_0, A_1: estimation with one spacecraft. B_0, B_1 and C_1: estimation method with two spacecraft.*

parallel with the x axis of the simulation box. A tangential discontinuity is generated in the flow after forming bow shock. The angle between the normal vector of the tangential discontinuity and the direction of the flow might be variable. Simulations with different angles were performed. The results predict that the size of HFAs are of the order of 1-3 Earth radii. The form of the profiles look alike in the first simulation, however, the temperature increases to a value 100 times higher, the solar wind speed decreases to 50-80%. The density of the solar wind decreases to 55-75% in the middle of the event but increases to 140% at the rim.

3.2 Comparison with observations

Analyzing fifty HFA events we can say that the shape of the temperature, the density and the magnitude of the magnetic field profiles qualitatively agree with the results of simulations. The quantitative result is not so close, however, these simulations were performed for idealistic and very simple cases thus we must not expect a better accordance. The simulations also describe well the attendance and development of HFAs.

We used two different methods to estimate the size of HFAs (See Fig. 4). The

spacecraft enters into the HFA in A_0 at t_0 and leaves it in A_1 at t_1. The HFA is traveling with the solar wind till the satellite is flying inside of the structure thus we get the distance from the $d = v_{SW} \cdot (t_1 - t_0)$ equation. We might neglect the speed of the satellite because $v_{s/c} \ll v_{SW}$. This is not very accurate because the solar wind speed might change and the HFA is not frozen into the plasma of the solar wind. So we have used another multispacecraft method, too (See Fig. 4). The B spacecraft intersects the border of HFA in B_0 at t_0. The C spacecraft do the same in C_1 at t_1. However, the position of the HFA changes until the second (C) satellite intersects its border. We can calculate the new position of the B_0 point from the position of the B_0 and the speed of the solar wind using the $\overrightarrow{r}_{B_1} = \overrightarrow{r}_{B_0} + \overrightarrow{v}_{SW} \cdot (t_1 - t_0)$. The HFA size must be larger than the distance between B_1 and C_1. We calculate this distance from the coordinates. If more then two spacecraft detect the same HFA then we can calculate all distances and use the longest.

We obtained very exciting result after determining the size of the HFA events. The size of the affected region seems to be 2-10 Earth radii with the first method based upon the time of pass. If at least two spacecraft cross HFA we could estimate the minimal diameter of the event. This second method gave values of 0.65-2.2 Earth radii. Both results are in agreement with the forecast of the simulations. Actually the first method based upon the time of cross doesn't seem to be accurate. Its error larger then the second method thus it is more practical to use this method if it is possible.

4 Summary

We set up conditions for HFA events and processed the mesurements of CLUSTER RAPID, FGM and CIS instruments between February and April, 2003. We found about fifty new events and developed a method for calculating transformed pitch angle distributions taking the Compton-Getting effect into account. We demonstrated our method by calculating of two kinds of PADs for the same HFA event. We also estimated the size of HFAs and compared with observations and found agreement with predicted values.

The second simulation (Lin, 2002) with different angles between the tangential discontinuity and the direction of the flow indicated that the size depends on this angle (γ). Higher γ values result in larger HFA up to about 80^o. If γ is greater than 80^o, the size decreases. The size of HFA also depends on the directional change of the magnetic field *before* and *after* the tangential discontinuity ($\Delta\Phi$). Unfortunately it seems to be very difficult to determine the values of these γ and $\Delta\Phi$ angles, so this problem needs to be solved in the future.

A new research area was created when Mars Global Surveyor discovered HFA events around Mars (Øieroset, 2001). This fact might be the first proof of that HFAs are general phenomena. HFAs might appear where bow shock and tangential discon-

tinuity interacts: around Jupiter, Saturn, ahead of CME[6]s, at the heliopause, or in interstellar clouds as well. The solution of this problem deserves further experimental and theoretial research.

Acknowledgement

The authors thank the OTKA grant T037844 of the Hungarian Scientific Research Fund for support.

References

Compton, H., Getting, I. 1935, Phys. Rev. E, 47(11), 817

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

Facsko, G., 2004, Publications of the Astronomy Department of the Eötvös Lóránd University, Vol. 14, p. 25-33

Lin, Y., 1997, J. Geophys. Res., Vol. 102, NO. A11, p. 24265-24281

Lin, Y., 2002, Planet. Space Sci., 50, p. 577-591

Lin, Y., 2003, J. Geophys. Res., Vol. 108, NO. A11, 1390, doi:10.1029/2003JA009991

Lucek, E. A., Horbury, T. S., Balogh, A., Dandouras, I., Rme, H., 2004, J. Geophys. Res., Vol. 109, A06207, doi:10.1029/2003JA010016.

Øieroset, M., Mitchell, D. L., Phan, T. D., Lin, R. P., Acuña, M. H., Geophys. Res. Lett., 2001, Vol. 28, NO. 5., p. 887-890

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, Vol. 318, p. 269-271

Schwartz, S. J., Paschmann, G., Sckopke, N., Bauer, T. M., Dunlop, M., Fazakerley, A. N., Thomsen, M. F., 2000, J. Geophys. Res., Vol. 105, NO. A6, p. 12639-12650

Sibeck, D. G., Borodkova, N. L., Schwartz, S. J., Owen, C. J., Kessel, R., Kokubun, S., Lepping, R. P., Lin, R., Liou, K., Lühr, H., McEntire, R. W., Meng, C.-I., Mukai, T., Nemecek, Z., Parks, G., Phan, T. D., Romanov, S. A., Safrankova, J., Sauvaud, J.-A., Singer, H. J., Solovyev, S. I., Szabo, A., Takahashi, K., Williams, D. J., Yumoto, K., Zastenker, G. N., 1999, J. Geophys. Res., Vol. 104, NO. A3, p4577-4594

Thomas, V. A., Wiske, D., Thomsen, M. F., Onsager, T. G., 1991, J. Geophys. Res., Vol. 96., NO. A7, p. 11625-11632

[6]**C**oronal **M**ass **E**jection

Thomsen, M. F., Gosling, J. T., Fuselier, S. A., Bame, S. J., Russell, C. T., 1986, J. Geophys. Res., Vol. 91, NO 10., p11311-11325

PADEU | PADEU **15**, 103 (2005)
ISBN 963 463 557

Linear MHD waves in penumbral structures

A. Marcu[1] and I. Ballai[2]

[1] Centre for Research in Space Plasmas, Dept. of Theoretical and Computational Physics, Babes-Bolyai University, 1, M. Kogalniceanu, 3400 Cluj-Napoca, Romania
[2] SPARG, Dept. of Applied Mathematics, University of Sheffield, Hicks Building, Hounsfield Road, Sheffield, S3 7RH, England (UK)

E-mail: [1]amarc@phys.ubbcluj.ro, [2]i.ballai.sheffield.ac.uk

Abstract
In this paper we study the oscillations of a magnetic medium periodic in the x-direction with B parallel to z in the presence of a steady equilibrium field aligned flow. The case with no gravity and stepwise profile for B(x), allowing a normal mode analysis, is studied and the dispersion relation for linear compressional waves is derived. The propagation of waves is studied in particular case modelling the propagation of waves in the penumbral filamentary structures in the photosphere.
Keywords: *Sun: magnetohydrodynamics(MHD) waves, Sun: magnetic field*

1 Introduction

The recent observational evidence for the existence of waves and oscillations in the solar atmosphere has invigorated theoretical developments in this field. Waves are responsible for carrying energy and momentum, creating instabilities, generating phenomena like magnetic reconnection, phase mixing, etc. They can serve as a unique tool for plasma diagnostics due to their capability of carrying information about the medium in which they propagate.

Solar and space plasmas have a very dynamic character, showing steady flow on all sorts of time and space scales. This property is confirmed by recent ground and space-born satellite observations Steady flows have been observed in the photosphere, chromosphere (November et al. 1979, Athay and Dere 1991), corona (Winebarger et al. 2002) and beyond, in the solar wind (Gabriel et al. 2003)). Therefore, theoretical

models should include the presence of an equilibrium steady flow. Equilibrium flows are known to introduce a series of new effects such as Kelvin-Helmholtz and resonant instabilities, negative energy waves, etc.

Observations have also revealed that the solar plasma is structured and this is determined and controlled by the magnetic field. In the photosphere, the magnetic field is concentrated in thin flux tubes with field strengths exceeding 1000 G (Spruit 1981). The flux tubes extend well into the chromosphere and corona contributing to the net heating of the solar upper layers, creating solar spiculae (Roberts 1979, James and Erdélyi 2002) and even contributing to the solar wind acceleration.

Many solar features (e.g. the granular pattern in the photosphere, the penumbra, plume/interplume, etc.) show a transversal periodicity (or quasi-periodicity). Waves propagating along magnetic structures with a transversal periodicity is likely to 'feel' the effect of this structure provided their transversal wavelength is of the same order as the transversal spatial organisation scale. Effects of periodicity, though it is not perfect on the Sun's surface, could possibly be observationally detected on diagnostic ω-k diagram. The possibility of wave propagation in a periodic structure has been discussed by Hollweg (1982), Berton and Heyvaerts (1987) and Uralov (2003) in the linear regime, while the steepening of linear waves into nonlinear waves (solitons) in periodic structures has been studied by Hollweg and Roberts (1984) and Ruderman et al. (2001).

Here we present some preliminary results on the propagation of compressional MHD waves in a magnetically periodical structure with field aligned flow. Weak field regions alternate with strong field regions modelling the filamentary structure of the penumbra.

2 Derivation of Dispersion Relations

We consider an ideal, perfectly conducting fluid permeated by a magnetic field of constant direction along the z-axis, and periodic along the x-axis. We suppose that the wavelengths are smaller than the gravitational scale-height, i.e. gravitational effects are neglected.

A field-aligned equilibrium steady flow is present in the system. In the present paper we suppose that the medium consists of alternating magnetic slabs (with width L_i and L_e) with homogeneous magnetic fields inside them (B_i and B_e, with $B_i > B_e$) and sharp discontinuities at the boundaries. Let us denote by $L = L_i + L_e$ the periodicity of the medium. The continuity of the total pressure at each boundary requires an equation of the type

$$\frac{d}{dx}\left(p_0 + \frac{B_0^2}{2\mu}\right) = 0, \tag{1}$$

which, in particular, leads to a density contrast in two adjacent regions

$$\frac{\rho_i}{\rho_e} = \frac{2c_{0e}^2 + \gamma c_{Ae}^2}{2c_{0i}^2 + \gamma c_{Ai}^2}, \tag{2}$$

where $c_{0i,e}$ and $c_{Ai,e}$ denote the sound and Alfvén speeds in the two layers, γ is the adiabatic index and μ is the magnetic permeability.

We perturb the system of linearized ideal MHD equations and write all physical quantities in the form $f_0 + f$ where f_0 are the equilibrium values and f their Eulerian perturbations. Since the equilibrium quantities depend on x only, we write all perturbations as

$$f = \widehat{f}(x)e^{i(\omega t - k_z z)}.$$

The dispersion relation of linear compressional MHD waves can be written as (Roberts 1981)

$$\frac{d^2\hat{v}_x}{dx^2} - q^2\hat{v}_x = 0, \tag{3}$$

where the magnetoacoustic parameter, q^2 is defined as

$$q(x)^2 = \frac{\left(k_z^2 c_A^2 - \Omega^2\right)\left(k_z^2 c_0^2 - \Omega^2\right)}{\left(c_0^2 + c_A^2\right)\left(k_z^2 c_T^2 - \Omega^2\right)}, \tag{4}$$

with $\Omega = \omega - k_z v_0$ being the Doppler-shifted frequency and $c_T = c_0 c_A/(c_0^2 + c_A^2)^{1/2}$ the tube (cusp) velocity. The quantity q^2 is constant in each region and it can take both, positive and negative values.

The solutions of Eq. (3) inside and outside the slab can be written as

$$\begin{cases} \widehat{v}_{xi} = \alpha_i e^{q_i x} + \beta_i e^{-q_i x} \\ \widehat{v}_{xe} = \alpha_e e^{q_e x} + \beta_e e^{-q_e x} \end{cases} \tag{5}$$

where the coefficients α_i, β_i, α_e and β_e are four constants which can be determined by matching conditions at the boundaries of two adjacent regions (at $x = L_i/2$ and $x = L_e + L_i/2$). Due to the periodicity of the medium we employ a method widely used in solid state physics. We introduce the quantity K_0 (called the *Bloch wavenumber*) which plays the same role as the wavenumber k_x in a homogeneous medium. According to the the Bloch's theorem, a solution of the Eq. (3) which is bounded at infinity can be written as

$$\widehat{v}_x = aF(x)exp[iK_0x] + bF(-x)exp[-iK_0x], \tag{6}$$

where $F(x)$ is a periodic function with the period L. Then

$$\frac{\widehat{v}(L)}{\widehat{v}(0)} = \cos K_0 L.$$

In the case of surface ($q_i^2 > 0$) and body ($q_i^2 < 0$) modes, the dispersion relations can be written as

$$\cos(K_0 L) = \cosh(\theta_i)\cosh(\theta_e) + \frac{1}{2}\left(S + \frac{1}{S}\right)\sinh(\theta_i)\sinh(\theta_e), \tag{7}$$

$$\cos(K_0 L) = \cos(\theta_i)\cosh(\theta_e) + \frac{1}{2}\left(S - \frac{1}{S}\right)\sin(\theta_i)\sinh(\theta_e), \tag{8}$$

$$S = \frac{\rho_i}{\rho_e}\frac{q_e}{q_i}\frac{k_z^2 c_{Ai}^2 - \Omega_i^2}{k_z^2 c_{Ae}^2 - \Omega_e^2}. \tag{9}$$

The dispersion curves $\omega(k_z)$ will depend upon the Bloch's wave number, K_0, as a parameter.

According to the relative magnitude of the characteristic speeds c_0, c_A,and c_T in the internal and external regions, various situations are possible depending on the signs of the quantities q_i^2 and q_e^2. For a simpler representation it is convenient to use the following dimensionless quantities

$$R_\rho = \frac{\rho_i}{\rho_e}, \qquad R_T = \frac{T_i}{T_e}, \qquad R_B = \frac{B_i}{B_e}, \qquad R_L = \frac{L_i}{L_e}, \tag{10}$$

where T_i and T_e denote the temperatures in the two regions. In these new notations, Eq. (2) can be rewritten as

$$\frac{c_{0i}^2}{c_{Ai}^2} = \frac{\gamma}{2}\frac{R_\rho R_T}{R_B^2}\frac{R_B^2 - 1}{1 - R_\rho R_T}. \tag{11}$$

In what follows, the dispersion relations are solved for incompressible and compressible plasmas.

2.1 Incompressible Modes

A first insight into the properties of the possible modes propagating in periodic plasma structures can be acquired by considering the incompressible plasma limit which yields a relatively simple analytical solution. This solution enables us to gain a better understanding of the behaviour of the different modes with respect to the equilibrium steady flow. For an incompressible plasma ($\gamma \to \infty$) $q^2(x) = |k_z| > 0$, i.e. only surface trapped modes are allowed to propagate. The dispersion relation of surface waves (Eq. (7)) becomes

$$\cos\Phi = \cosh(\theta)\cosh(\frac{\theta}{R_L}) + \frac{1}{2}\left(S + \frac{1}{S}\right)\sinh(\theta)\sinh(\frac{\theta}{R_L}) \tag{12}$$

where

$$S = R_\rho \frac{c_{Ai}^2 - (c - v_{0i})^2}{c_{Ae}^2 - (c - v_{0e})^2}, \qquad \theta = k_z L_i, \qquad \Phi = K_0 L. \tag{13}$$

and $c = \omega/k_z$ is the phase speed of the waves. The numerical investigation of these modes requires the function c(θ) satisfies Eqs. (12) and (13), where Φ is a parameter. Inverting Eq. (12) yields

$$S^2 - 2\frac{\cos\Phi - \cosh(\theta)\cosh(\frac{\theta}{R_L})}{\sinh(\theta)\sinh(\frac{\theta}{R_L})}S + 1 = 0 = S^2 - 2US + 1, \tag{14}$$

with

$$\frac{\cos\Phi - \cosh(\theta)\cosh(\frac{\theta}{R_L})}{\sinh(\theta)\sinh(\frac{\theta}{R_L})} = U. \tag{15}$$

Using this notation, Eq. (14) can have two roots

$$S_\pm = U \pm \sqrt{U^2 - 1}, \tag{16}$$

where the $\pm$ sign corresponds to sausage and kink modes, respectively. In Eq. (16) we have to impose that $U^2 \geq 1$, i.e. the quantity S is real. Since

$$S = \frac{c_{Ai}^2 - (c - v_{0i})^2}{c_{Ae}^2 - (c - v_{0e})^2}R_\rho, \tag{17}$$

the dispersion relation of fast magnetoacoustic modes propagating in periodic structures in incompressible limit is given by the roots of

$$c^2(R_\rho - S_\pm) + 2c(S_\pm v_{0e} - R_\rho v_{0i}) - R_\rho(c_{Ai}^2 - v_{0i}^2) + S_\pm(c_{Ae}^2 - v_{0e}^2) = 0. \tag{18}$$

Eq. (18) posses two real roots corresponding to two possible sausage or kink modes provided

$$\frac{R_\rho S}{R_\rho - S} > \frac{S c_{Ae}^2 - R_\rho c_{Ai}^2}{(v_{0i} - v_{0e})^2}$$

If the two equilibrium flow speeds are set to be zero we recover the results by Berton and Heyvaerts (1987).

Before turning to the general case, let us first discuss two particular cases, i.e $R_L \to 0$ (the isolated slab) and $R_L \to \infty$ (homogeneous plasma). In the case of an isolated tube, $U = -\coth\theta$, so the phase speed of fast surface sausage modes is given by

$$c_+^2 - 2c_+\frac{v_{0e}\tanh\frac{\theta}{2} + v_{0i}R_\rho}{\tanh\frac{\theta}{2} + R_\rho} - \frac{(c_{Ae}^2 - v_{0e}^2)\tanh\frac{\theta}{2} + (c_{Ai}^2 - v_{0i}^2)R_\rho}{\tanh\frac{\theta}{2} + R_\rho} = 0, \tag{19}$$

while the propagation speed of fast surface kink modes is described by

$$c_+^2 - 2c_+\frac{v_{0e}\coth\frac{\theta}{2} + v_{0i}R_\rho}{\coth\frac{\theta}{2} + R_\rho} - \frac{(c_{Ae}^2 - v_{0e}^2)\coth\frac{\theta}{2} + (c_{Ai}^2 - v_{0i}^2)R_\rho}{\coth\frac{\theta}{2} + R_\rho} = 0, \tag{20}$$

identical to the results obtained by Nakariakov and Roberts (1995). This form of Eqs. (19) and (20) allows us to interpret the effect of the equilibrium flow. In both equations, the second terms owe their existence to the equilibrium internal and external flows.

In the limit of homogeneous plasma, we let $R_L \to \infty$ and we obtain the well-known propagation speeds for sausage and kink modes

$$c_+ = \pm c_{Ai} + v_{0i}, \quad c_- = \pm c_{Ae} + v_{0e} \tag{21}$$

where here the $\pm$ indicates forward (parallel to the magnetic field) and backward (anti-parallel to the magnetic field) propagation.

In the general case we can derive analytical solution for the phase speed of sausage and kink modes propagating in the incompressible periodic plasma in two extreme limits corresponding to thin ($\theta \ll 1$) and thick ($\theta \gg 1$) slab.

In a thin slab, U can be approximated as

$$U \sim -\frac{2R_L}{\theta^2} \sin^2 \frac{\Phi}{2}, \tag{22}$$

which in the case of $\theta \to 0$ leads to the approximate expressions

$$S_+ \sim \frac{1}{2U}, \qquad S_- \sim 2U. \tag{23}$$

These expressions can now be used to obtain the phase speed of fast sausage and kink modes as

$$c_+^2 - 2c_+ \frac{4\frac{R_L R_\rho}{\theta^2} \sin^2 \frac{\Phi}{2} v_{0i} + v_{0e}}{1 + 4\frac{R_L R_\rho}{\theta^2} \sin^2 \frac{\Phi}{2}} - \frac{4\frac{R_L R_\rho}{\theta^2} \sin^2 \frac{\Phi}{2}(c_{Ai}^2 - v_{0i}^2) + c_{Ae}^2 - v_{0e}^2}{1 + 4\frac{R_L R_\rho}{\theta^2} \sin^2 \frac{\Phi}{2}} \tag{24}$$

and

$$c_-^2 - 2c_- \frac{4\frac{R_L}{\theta^2} \sin^2 \frac{\Phi}{2} v_{0e} + v_{0i} R_\rho}{R_\rho + 4\frac{R_L}{\theta^2} \sin^2 \frac{\Phi}{2}} - \frac{4\frac{R_L}{\theta^2} \sin^2 \frac{\Phi}{2}(c_{Ae}^2 - v_{0e}^2) + R_\rho(c_{Ai}^2 - v_{0i}^2)}{R_\rho + 4\frac{R_L}{\theta^2} \sin^2 \frac{\Phi}{2}} \tag{25}$$

In the wide slab limit, $S_\pm \to -1$ and the dispersion relation is given by

$$c_\infty^2 - 2c_\infty \frac{v_{0e} + v_{0i} R_\rho}{R_\rho + 1} - \frac{R_\rho(c_{Ai}^2 - v_{0i}^2) + c_{Ae}^2 - v_{0e}^2}{R_\rho + 1} = 0 \tag{26}$$

The possible modes propagating in a incompressible periodic structure for an arbitrary θ and R_L are shown in Figure 1. All characteristic speeds are expressed in units of the internal Alfvén speeds and the following notations have been used

$$a = \frac{c_{Ae}}{c_{Ai}}, \quad M_i = \frac{v_{0i}}{c_{Ai}}, \quad M_e = \frac{v_{0e}}{c_{Ai}}.$$

Figure 1 depicts the possible modes propagating in a periodic structure when no plasma flow is present in the system.

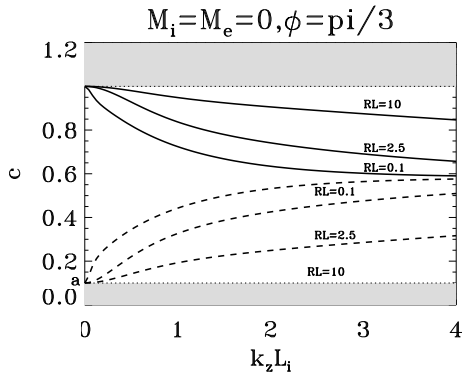


Figure 1: *The phase speed of sausage (solid line) and kink modes (dashed lines) as a function of the dimensionless wavelength, θ for an incompressible periodic plasma with $a = 10$ and $a = 1$ and $M_i = M_e = 0$.*

As expected, these modes are identical to the results obtained by Berton and Heyvaerts (1987). Since the equilibrium state is static, the modes are symmetric with respect to the horizontal axis, therefore in Figure 1 we plot only the forward propagating modes. The possible modes are allowed to propagate between the two Alfvén speeds, the grey regions are the intervals where waves become leaky. The phase speed given by the dispersion relation has been plotted for $\Phi = \pi/3$ and for three different values of $R_L(0.1, 2.5, 10)$. A simple visual inspection shows that waves are more dispersive for small values of R_L, i.e. for transversal dimensions of the adjacent slabs close to each other. The phase speed of both types of waves is increasing with increasing the R_L number. In general, sausage waves (with positive dispersion) propagate faster that kink modes (with negative dispersion). For large values of θ both, sausage and kink modes will tend towards the positive value of c_∞ given by Eq. (26).

When an equilibrium flow is taken into account (Figures 2-3) the symmetry of the modes is broken although the propagation windows of the forward and backward propagating modes are shifted by an amount proportional to the flow speed. Figures 2 and 3 were obtained for two values of $\Phi(0, \pi/2)$ and for three different values of $R_L(0.1, 2.5, 10)$. We have considered the case $R_B > 1$ with $a = 0.1$ and $R_\rho = 0.5$, $M_i = 0.04$ and $M_e = -0.02$.

Repeating the plotting for other values of Φ, we can observe that the dependence of the phase speed on θ is very similar to the case obtained for $\Phi = \pi/2$, only small changes are observed for intermediate values of θ. When $\Phi = 0$, sausage and kink

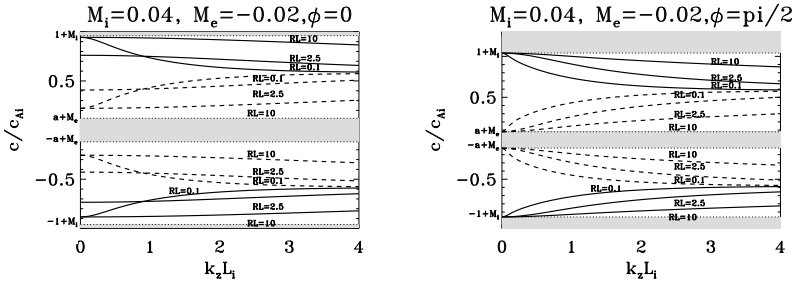


Figure 2: *The same as in Figure 1 but now* $M_i = 0.04$, $M_e = -0.02$ *and* $\Phi = 0$ *(left panel) and* $\Phi = \pi/2$ *(right panel)*

modes show a significant dispersion for small values of R_L, and for values greater than 1, waves are weakly dispersive. When $\Phi \neq 0$ the behaviour of oscillating modes is similar to the case obtained in a static equilibrium.

2.2 Compressible modes

The dispersion relations for surface and body modes have a rich variety of solutions. In what follows we only the slender-slab ($\theta \ll 1$) approximation.

Surface modes propagate with phase speeds given by the dispersion relation (7), which in the $\theta \ll 1$ limit reduces to

$$-\frac{4\sin^2\frac{\Phi}{2}}{\theta^2} \approx q_i^2 + \frac{q_e^2}{R_L^2} + \left(S + \frac{1}{S}\right)\frac{q_i q_e}{R_L}. \tag{27}$$

If $\Phi \neq 0$, the two possible modes propagate with phase speeds given by

$$c \approx v_{0i} \pm c_{Ti}\left\{1 + \frac{c_{Si}^2}{c_{Ai}^2 c_i^2}\left[R_L R_\rho \frac{c_{Ai}^4}{c_i^2} + c_{Ae}^2 - (v_{0i} \pm c_{Ti} - v_{0e})^2\right]\frac{\theta^2}{8sin^2\frac{\Phi}{2}R_L R_\rho}\right\} \tag{28}$$

and

$$c \approx v_{0e} \pm c_{Te}\left\{1 + \frac{c_{Se}^2}{c_{Ae}^2}\left[\frac{c_{Ae}^4}{c_e^2}\frac{1}{R_L R_\rho} + c_{Ai}^2 - (v_{0e} \pm c_{Te} - v_{0i})^2\right]\frac{\theta^2 R_\rho}{8sin^2\frac{\Phi}{2}R_L c_e^2}\right\} \tag{29}$$

In a compressible plasma in the thin slab limit, body modes can also propagate and their dispersion relation is given by

$$-\frac{4\sin^2\frac{\Phi}{2}}{\theta^2} \approx -q_i^2 + \frac{q_e^2}{R_L^2} + \left(S - \frac{1}{S}\right)\frac{q_i q_e}{R_L}, \tag{30}$$

Using the same method as in the case of surface waves, we obtain a dispersion relation of body waves in the thin slab approximation which is similar to the equation describing surface modes to the second order terms in Eqs. (28)–(29). Since the second order terms in these equations contain a term proportional to the very small θ^2, changing the sign from plus (surface modes) to minus (body modes) will not affect our results significantly.

3 Waves in photospheric structures

A possible photospheric application would be the study of wave propagation in sunspots' penumbra. Viewed in white light, sunspots consist of a dark, central umbra surrounded by brighter, radially striated penumbra made up from alternating bright and dark filaments. Martinez Pillet (2000) suggested that these horizontal tubes might have diameters around 100 km.

The magnetic field in the umbra is vertical, but as soon as we approach the penumbra, the field lines tend to be horizontal, relative to the solar surface. Observations by Beckers and Schröter (1969) suggested that, while the magnetic field in the dark filaments is nearly horizontal, the field in the bright filaments is inclined to the horizontal with a mean inclination of 10°-15°. However, here we will suppose that in both types of filaments, the magnetic field is horizontal. Observations also showed that the magnetic field strength is weaker in dark regions than in the bright ones.

The plasma dynamics show a systematic inward (towards the umbra) plasma flow in the bright elements of the umbra and outward Evershed flow occurring mostly in the dark filaments. The most important motion in the penumbra is the radial, nearly horizontal (field aligned) outflow of plasma responsible for the Evershed effect seen in spectral lines formed in the penumbral photosphere. This motion is thought to be a siphon flow driven by a pressure difference between the footpoints of arched magnetic flux tubes. The siphon flow is probably also responsible for the reversed, inward Evershed flow seen higher up in the penumbral atmosphere, in chromospheric spectral lines such as H_α (Schlichenmaier, 2002).

Radially outward propagating running penumbral waves have speeds of 10-20 km s^{-1}, repeating with a period in the range of 200-300 s with a horizontal wavelength of 2.3-3.8 Mm (Giovanelli, 1974). These propagation speeds are larger than the local sound speed, therefore they might be fast magnetoacoustic waves. Penumbral structures also support waves with periods around 5 min, which are related to the interaction of the sunspot with the resonant acoustic (p-modes) in the surrounding quiet Sun.

The possible modes arising in penumbral fine structures modelled as a periodic structure in the long wavelength limit (slender slab) are shown in Figures 3. Plots are obtained for three values of Φ ($\pi/10, \pi/2, \pi$) and three value of R_L (0.1, 2.5, 10). All characteristic speeds are normalized to the internal Alfvén speed.

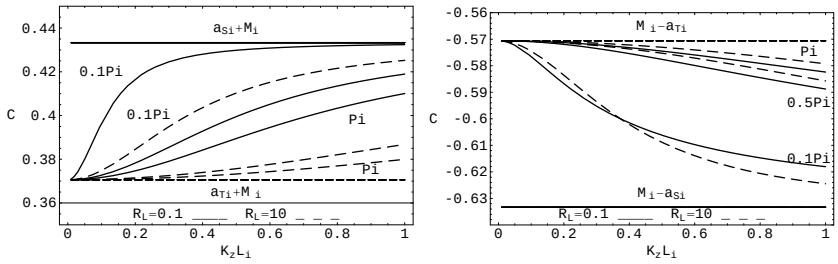


Figure 3: *The phase speed of surface modes as a function of the dimensionless wavelength, θ arising in an compressible periodic plasma (sunspot's penumbra). The solid lines correspond to $R_L = 0.1$ and dashed lines to $R_L = 10$. ($a = 0.8$, $a_{Si} = 0.533$, $a_{Ti} = 0.47$, $M_i = -0.1$ and $M_e = 0.2$)*

With an Alfvén speed in bright regions of 15 km s^{-1}, we choose $c_{Si} = 9$ km s^{-1}, $c_{Se} = 8$ km s^{-1} and v_{Ae}=12 km s^{-1}. For the equilibrium flow we choose Mach numbers such as $M_i = -0.1$ and $M_e = 0.2$.

Since the ratio of the widths of two adjacent slabs appears in the dispersion relation, the dependence of the phase speed (in Alfvén speed units) with respect to R_L for a thin slab is shown in Figure 4.

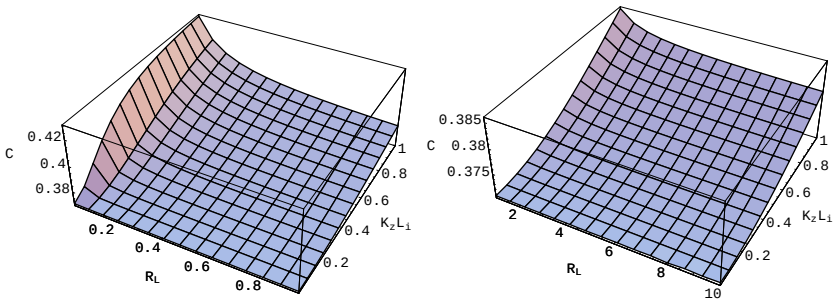


Figure 4: *The phase speed of surface modes as a function of the dimensionless wavelength, θ, and $R_L < 1$ (left panel) and $R_L > 1$ (right panel) in photospheric penumbrae.*

The presented findings are preliminary results and a full investigation on the effect of equilibrium flows on the propagation of linear compressional MHD waves in periodic structures is underway. Part of this investigation will be the calculation of

the threshold for negative energy waves which can appear in these structures. Further models will be made for the plume/interplume system and the spaghetti structures in the solar wind.

Acknowledgement

The authors acknowledge the financial support by the British Council and the Royal Society. One of the authors (IB) acknowledges the financial support by the Nuffield Foundation.

References

Athay, R.G. and Dere, D.P. 1991, ApJ, 381, 323

Beckers, J.M. and Schröter, E.H. 1969, Sol. Phys., 10, 384

Berton, R. and Heyvaerts, J. 1987, Sol. Phys., 109, 201

Gabriel, A.H., Bely-Dubau, F. & Lemaire, P. 2003, ApJ, 589, 623

Giovanelli, R.G. 1974, in R. Grant Athay (ed.) *Chromospheric Fine Structures*, IAU Symp., 56, 137

Hollweg, J.V. 1982, J. Geophys. Res., 87, 8065

Hollweg, J.V. and Roberts, B. 1984, J. Geophys. Res., 89, 9703

James, S.P. and Erdélyi, R. 2002, A&A, 393, L11

Roberts, B. 1979, Sol. Phys., 61, 23

Roberts, B. 1981, Sol. Phys., 69, 27

Ruderman, M.S., Roberts, B., Pelinovsky, E.N. and Petrukhin, N.S. 2001, *Phys. Plasmas*, 8, 2628

Martinez Pillet, V. 2000, A&A, 361, 734

Nakariakov, V.M and Roberts, B. 1995, Sol. Phys., 159, 213

November, L.J., Toomre, J., Gebbie, K.B. and Simon, G.W. 1979, ApJ, 227, 600

Schlichenmaier, R. 2002, *Astron. Nachrichten*, 323, 303

Spruit, H.C. 1981, in *The Sun as a star*, ed. Jordan, S.D. *NASA Spec. Publ.*, 450, 385

Uralov, A.M. 2003, Sol. Phys., 218, 17

Winebarger, A. R., Warren, H., van Ballegooijen, A., DeLuca, E.E. & Golub, L. 2002, ApJ, 567, L89

PADEU | PADEU **15**, 115 (2005)
ISBN 963 463 557

MHD Waveguide Modes in Structured Magnetic Flux Tubes

B. Carter

Solar Physics and upper-Atmosphere Research Group (SPARG), University of Sheffield, Department of Applied Maths, Hounsfield Road, Sheffield, S3 7RH, England

E-mail: b.carter@sheffield.ac.uk

Abstract
The propagation of MHD waves in a structured magnetic flux tube embedded within a straight vertical magnetic environment are studied analytically. The motivation behind this study comes from the observations of damped loop oscillations showing that only part of the loop is homogeneous. The magnetic tube considered contains two characteristically distinct parts, namely an internal straight core, and a surrounding uniformly twisted magnetic annulus envelope. The general dispersion relation is derived. Modes of oscillations are examined from this general dispersion relation that is suitable for obtaining information on not just oscillations but some instability properties of this complex tube structure. Both short and long wavelength approximations are considered analytically for the symmetrical mode and with small twist.

1 Introduction

The recent launch of sophisticated satellites such as SOHO and TRACE, that have imaging telescopes on-board, have lead to very detailed data and a surge in the development of the theory of oscillations of MHD waveguides. Numerous oscillatory periods within the Sun's surface have been observed and, in an attempt to account for these periods, numerous theoretical models developed (see reviews e.g. Aschwanden (2003); Roberts (2004); Roberts & Nakariakov (2003)). We consider one specific area of the many theoretical models: the effects of magnetic structuring. The role of magnetic structuring on wave existence and propagation has been studied in some detail (see e.g. De Pontieu et al. (2004); Erdélyi et al. (2004), James et al. (2003), Mendoza-Briceno

et al. (2004), Ruderman & Erdélyi (2003), Taroyan et al. (2004)). Here I discuss previous literature on this specific aspect of coronal seismology. Gravity effects are ignored, the emphasis being on the role of the magnetic structuring and, from the basic MHD equations, dispersion relations are found and specific analysis carried out.

2 Waves in a Strongly Inhomogeneous Medium

When the wavelength is greater than or approximately equal to the length-scale the inhomogeneous nature of the medium determines the behavior of any disturbances. Within the Sun and the solar atmosphere the principal causes of inhomogeneity are gravity and the magnetic field. Gravity creates a vertical stratification in plasma pressure and the magnetic field can cause the plasma pressure to increase in a direction normal to the field. These stratifications introduce significant effects, namely amplification (of the wave amplitude (increase or decrease) as it propagates), evanescence (regions in which waves, otherwise oscillating spatially, may decay exponentially) and surface modes (a discontinuity in the medium (magnetic, pressure,...) may give rise to decaying surface waves and will modify the 'body' waves).

2.1 A Magnetic Interface

Ignoring effects due to gravity, we are now interested in the effect of magnetic structuring. Assume that in the basic state the plasma is permeated by a magnetic field $B_0(x)\hat{\mathbf{z}}$, working in Cartesian coordinates x, y, z. Then the pressure and density are structured by the x-dependence of the magnetic field and the basic states are found to be

$$p_0 = p_0(x), \quad \rho_0 = \rho_0(x), \quad \frac{d}{dx}\left(p_0 + \frac{B_0^2}{2\mu}\right) = 0, \tag{1}$$

for the pressure, density and total (gas + magnetic) pressure. Linearised perturbations from this state are taken and the continuity, momentum, induction and energy equations, after Fourier analysis, give a single ODE for $\hat{v}_x(x)$. Consider the magnetic interface:

$$B_0(x) = \begin{cases} B_e, & x > 0, \\ B_0, & x < 0, \end{cases} \tag{2}$$

with B_0 and B_e both constants. Considering pressure continuity at $x = 0$ leads to

$$\frac{d^2\hat{v_x}}{dx^2} - m_0^2\hat{v_x}(x) = 0 \quad \text{for} \quad x < 0, \tag{3}$$

and

$$\frac{d^2\hat{v_x}}{dx^2} - m_e^2\hat{v_x}(x) = 0 \quad \text{for} \quad x > 0. \tag{4}$$

where

$$m_0^2(x) = \frac{(k^2c_0^2(x) - \omega^2)(k^2v_A^2(x) - \omega^2)}{(c_0^2(x) + v_A^2(x))(k^2c_T^2(x) - \omega^2)}, \quad c_T^2(x) = \frac{c_0^2(x)v_A^2(x)}{c_0^2(x) + v_A^2(x)}, \tag{5}$$

and m_e^2 is defined in a similar way to m_0^2 except that the Alfvén and sound speeds appropriate to $x > 0$ are taken.
It is the presence of the discontinuity in $B_0(x)$ that is responsible for the existence of surface waves which may arise if m_0^2 and m_e^2 are both real and positive. Solving (3) and (4) for $\hat{v}_x(x)$ gives

$$\hat{v}_x(x) = \begin{cases} \alpha_e\, e^{-m_e x}, & x > 0, \\ \alpha_0\, e^{m_0 x}, & x < 0. \end{cases} \tag{6}$$

In writing this solution we are excluding laterally propagating waves, so only the surface modes arise. Then, using the continuity of both $\hat{v_x}(x)$ and of the total pressure perturbation $\hat{p_T}(x)$ across the interface $x = 0$ the general dispersion relation for the magnetic interface is

$$\rho_0(k^2v_A^2 - \omega^2)m_e + \rho_e(k^2v_{Ae}^2 - \omega^2)m_0 = 0, \tag{7}$$

valid for m_0 and m_e both positive and for $l = 0$. From this one is able to expand, simplify and conclude that the existence of the magnetic interface supports the propagation of surface waves (see e.g. Roberts (1981a)).

2.2 Waves in a Magnetic Slab

First consider a magnetic slab (Edwin & Roberts (1982)) with zero field surrounding it so that

$$B_0(x) = \begin{cases} B_0, & |x| < x_0, \\ 0, & |x| > x_0, \end{cases} \tag{8}$$

with pressure p_0 and density ρ_0 inside the slab, p_e and ρ_e outside. The two regions are related by

$$p_e = p_0 + \frac{B_0^2}{2\mu}, \quad \rho_e = \left(\frac{c_0^2 + \frac{1}{2}\gamma v_A^2}{c_e^2})\right)\rho_0 \tag{9}$$

where c_o and c_e are the sound speeds inside and outside the slab and v_A is the Alfvén speed in the slab. Again, attention is confined to two-dimensional disturbances so velocity perturbation component v_y and wavenumber l are supposed zero. Again, the boundary conditions, $\hat{v}_x(x)$ and $\hat{p}_T(x)$ continuous across the boundary $x = \pm x_0$ are used and the general dispersion relation recovered as:

$$(k^2v_A^2 - \omega^2)m_e = \left(\frac{\rho_e}{\rho_0}\right)\omega^2 m_0 \begin{pmatrix} tanh \\ coth \end{pmatrix} m_0 x_0, \tag{10}$$

valid for $\omega^2 < k^2 c_e^2$. Further analysis discovers the existence of slow magnetoacoustic waves, both as a body wave and as a surface wave and fast magnetoacoustic waves only if the slab is cooler than the surrounding plasma also as body and surface waves. When the slab is considered thin in comparison to the wavelength (long wavelength approximation that is of interest for photospheric and coronal conditions) the kink mode vibrates as a single thin string and sausage mode vibrates as both surface and body waves.
Expanding the above case to one of a slab embedded in a magnetic environment certain differences were discovered. The dispersion relation for this case is found to be:

$$\rho_e(k^2 v_{Ae}^2 - \omega^2) m_0 \begin{pmatrix} tanh \\ coth \end{pmatrix} m_0 x_0 + \rho_0(k^2 v_A^2 - \omega^2) m_e = 0, \tag{11}$$

for $m_e > 0$. First, considering incompressible motions, m_0 and m_e both tend to k and thus the modes are Alfvén surface waves. The sausage and kink modes both exist but their general behaviour regarding phase speeds exchanges for v_A greater or less than v_{Ae}. For compressible motions, the transcendental nature of (11) makes the analysis more difficult but when taking the slender slab (long wavelength) approximation the situation become slightly easier. It is supposed that $m_0 x_0 \to 0$ as $k x_0 \to 0$ so that $\tanh(m_0 x_0) \simeq m_0 x_0$ for $k x_0 \ll 1$. Equation (11) then reduces to

$$\rho_0(k^2 v_A^2 - \omega^2) m_e + \rho_e(k^2 v_{Ae}^2 - \omega^2) m_0^2 x_0 = 0, \tag{12}$$

which, for the tanh function (corresponding to the sausage mode) gives solutions indicating that the sausage modes are only weakly affected by the external field. For the kink mode (coth) the dispersion relation becomes

$$\rho_e(k^2 v_{Ae}^2 - \omega^2) + \rho_0(k^2 v_A^2 - \omega^2) m_e x_0 = 0 \tag{13}$$

which, when solved, shows the external field to dictate the behaviour for a sufficiently thin slab. Only one mode is shown to exist.

For a finite slab, applicable to coronal structures which have been observed as about one tenth wide as they are long, some other simplifying assumptions are made. In the limit of large $k x_0$ both kink and sausage modes result in

$$m_0 \rho_e(k^2 v_{Ae}^2 - \omega^2) + m_e \rho_0(k^2 v_A^2 - \omega^2) = 0 \tag{14}$$

for the dispersion relation which can be reduced to a cubic and solved. Two cases are considered, firstly a small non-zero external field then a large external field with small plasma beta. Sausage and kink modes are both found to exist and behaviour for various situations are discussed.

2.3 Wave Propagation in a Magnetic Cylinder

After observations in $H\alpha$ and soft and hard X-rays, it became quite clear that cylinders would be a better approximation to the tubular structures seen in the corona and photosphere (Edwin & Roberts (1983)).

With the equilibrium configuration, in cylindrical coordinates $(r,\ \theta,\ z)$, taken as

$$\rho,\ p,\ \mathbf{B}_0 = \begin{cases} \rho_0,\ p_0,\ (0,0,B_0), & r < a \\ \rho_e,\ p_e,\ (0,0,B_e), & r > a \end{cases} \tag{15}$$

using Fourier analysis and continuity across $r = a$ and needing bounded solutions at $r = 0$ the dispersion relation was found. For surface waves ($m_0^2 > 0$)

$$\rho_0(k^2 v_A^2 - \omega^2) m_e \frac{K_n^{'}(m_e a)}{K_n(m_e a)} = \rho_e(k^2 v_{Ae}^2 - \omega^2) m_0 \frac{I_n^{'}(m_0 a)}{I_n(m_0 a)}, \tag{16}$$

and for body waves ($m_0 = -n_0 < 0$)

$$\rho_0(k^2 v_A^2 - \omega^2) m_e \frac{K_n^{'}(m_e a)}{K_n(m_e a)} = \rho_e(k^2 v_{Ae}^2 - \omega^2) n_0 \frac{J_n^{'}(n_0 a)}{J_n(n_0 a)}, \tag{17}$$

where K_n, I_n, J_n, are Bessel functions and the dash denotes the derivative with respect to the argument. Three cases of study were chosen for the complex array of modes given by equations (16) and (17).
Incompressible modes. In the incompressible limit $(c_0^2 \to \infty, c_e^2 \to \infty), m_0$ and m_e become $|k|$. The kink and sausage modes are then given explicitly and it is noted that the phase speed for the kink mode is not monotonic as a function of k but has a maximum/minimum and the sausage mode is monotonically increasing/decreasing. This max/min feature of the kink mode is absent in the slab case so can be deduced to be a reflection of the geometry of the magnetic field.
Photospheric tubes. The dispersion relations were solved for photospheric values, paying particular attention to the slender tube case and not including stratification. The kink mode with phase-speed close to c_k is considered and it is noticed that the corresponding equation is of the general form previously discussed. This suggests this mode may propagate nonlinearly as a solitary wave. By similar argument it is concluded that two kink (surface) modes exist, a slow and a fast mode, the slow mode having phase-speed close to c_T.
Coronal loops. The coronal conditions, i.e. $v_{Ae},\ v_A > c_e,\ c_0$, imply that there are no longer surface waves present but two classes of body waves can occur - fast, which is of particular interest, and slow body waves, which i shall not mention in detail. Fast kink modes, when $v_{Ae} > v_A$, are sustained in dense loops with periods on an Alfvénic timescale and it is found that these body waves have a low wavenumber cut off implying that only wavelengths shorter than the diameter of a loop can propagate

freely. The sausage mode, however, has a much shorter period, approximately one tenth of that of the kink mode. Sausage and kink fast modes exist only in high density loops. However, the slow modes appear in both high and low density cylinders.

2.4 Wave Propagation in Twisted Magnetic Cylinder

Granular shear motions in the photosphere can introduce a twist to the flux tube and prominences often appear to have twisted field lines so the study of this modification to the straight flux tube is of some importance (Bennett et al. (1999)). Twisted tubes have been studied before but only in terms of stability. Bennett et al. (1999) investigates the details of different modes in a uniformly twisted flux tube embedded in a straight magnetic field given by:

$$\mathbf{B} = \begin{cases} (0, Ar, B_0), & r < a, \\ (0\ ,\ 0\ ,\ B_e), & r > a. \end{cases} \tag{18}$$

The plasma is taken as incompressible, with the field and plasma pressure being structured in the radial direction. Again, using continuity of total pressure p_T and perturbation velocity v_r across $r = a$ and searching for a bounded solution at $r = 0$ and as $r \to \infty$ leads to the recovery of dispersion relation

$$\frac{(\omega^2 - \omega_{A0}^2)\dfrac{xI_m^{'}(x)}{I_m(x)} - 2m\omega_{A0}\dfrac{A}{\sqrt{\mu\rho_0}}}{(\omega^2 - \omega_{A0}^2)^2 - 4\omega_{A0}^2\dfrac{A^2}{\mu\rho_0}} = \frac{\dfrac{x_1 K_m^{'}(x_1)}{K_m(x_1)}}{\dfrac{\rho_e}{\rho_0}(\omega^2 - \omega_{Ae}^2) + \dfrac{A^2}{\mu\rho_0}\dfrac{x_1 K_m^{'}(x_1)}{K_m(x_1)}}. \tag{19}$$

In this equation the dash, as before, denotes derivative with respect to the argument of the bessel function, ω_{A0} and ω_{Ae} are the internal and external alfvén speeds, $x = m_0 a$, $x_1 = |k_z|a$ and

$$m_0^2 = k_z^2 \left(1 - \frac{4A^2\omega_{A0}^2}{\mu\rho_0(\omega^2 - \omega_{A0}^2)^2}\right). \tag{20}$$

This is the dispersion relation for waves in an incompressible magnetic tube with uniform twist embedded in a straight magnetic environment. For an incompressible tube with no twist there are no body waves but when twist is introduced body waves appear. It was found that as twist is increased the body modes cover a wider range of the phase speeds and become more distinct. As the value of $k_z a$ increases the distribution of the body waves decreases showing that long wavelength modes display body mode features. A dual nature is discovered where body waves exist for long wavelengths but surface wave characteristics are displayed for shorter wavelengths. When $m = 1$, the kink modes, it is noted that the phase speed of the body modes tends to infinity as $k_z a \to 0$, so waves with larger wavelengths propagate with larger phase speeds. Bennett et al. (1999) explored further the cases of large and small $k_z a$. These approximations, when plotted and compared to the full dispersion relation

solved numerically, provide a useful check and cast additional light on the behaviour of the various modes.

2.5 Wave Propagation in Twisted Magnetic Annulus

The interpretation of damping of loop oscillations by resonant absorption given by eg. Robbrecht et al. (2001); Ruderman & Roberts (2002); Goossens et al. (2002) indicates that the coronal flux tubes are homogeneous for only a percentage of their radius ($\sim$5–45%) being inhomogeneous, in the first approximation, both in the centre and in the surrounding plasma. To model this we consider, not only the structuring of a filament surrounded by a coaxial shell as already considered by, among others, Mikhalyaev & Solov'ev (2004), but also added twist that makes the annulus inhomogeneous. We consider a uniformly twisted magnetic annulus embedded in a vertical straight magnetic field.

$$\mathbf{B} = \begin{cases} \mathbf{B}_i = (0\,,\,0\,,\,B_i), & r < a, \\ \mathbf{B}_a = (0, Br, B_0), & a < r < R, \\ \mathbf{B}_e = (0\,,\,0\,,\,B_e), & r > R. \end{cases} \tag{21}$$

For equilibrium state there is no background flow and gravity effects are neglected. Careful analysis is used to find the dispersion relation in terms of the Bessel functions $I_m(z), K_m(z)$ and their derivatives:

$$\frac{\Xi_{aK} - \Xi_i + \Xi_{aK}\Xi_i \frac{A_0^2}{\mu}}{\Xi_{aI} - \Xi_i + \Xi_{aI}\Xi_i \frac{A_0^2}{\mu}} \frac{K_m(m_o a)}{I_m(m_o a)} = \frac{K_m(m_o R)}{I_m(m_o R)} \frac{\Xi_{RK} - \Xi_e + \Xi_{RK}\Xi_e \frac{A_0^2}{\mu}}{\Xi_{RI} - \Xi_e + \Xi_{RI}\Xi_e \frac{A_0^2}{\mu}}. \tag{22}$$

where

$$\Xi_i = \frac{|k|a\, I_m^{'}(|k|a)}{\rho_i(\omega^2 - \omega_{Ai}^2) I_m(|k|a)}, \qquad \Xi_e = \frac{|k|R\, K_m^{'}(|k|R)}{\rho_e(\omega^2 - \omega_{Ae}^2) K_m(|k|R)}$$

and

$$\Xi_{\alpha X} = \frac{(\omega^2 - \omega_{A0}^2)\frac{m_0 \alpha X_m^{'}(m_0\alpha)}{X_m(m_0\alpha)} - \frac{2mB\omega_{A0}}{\sqrt{\mu\rho_0}}}{\rho_0(\omega^2 - \omega_{A0}^2)^2 - \frac{4B^2\omega_{A0}^2}{\mu}} \qquad \text{for} \quad \begin{array}{l} \alpha = a,\, R, \\ X = \text{Bessel fncs } I,\, K. \end{array}$$

From this equation, setting first $a = 0$ then setting twist to zero and $a = 0$ one can corroborate this result by recovering the general dispersion relations from Bennett et al. (1999) and from Edwin & Roberts (1983).
Setting $m = 0$, for the sausage modes, let us consider both the long/short wavelength approximations for slender annulus ($|k_z|(R-a) \ll 1$).

Long wavelength. Let us denote $m_0 a = x,\ \ m_0 R = y$. Letting $ka < kR \ll 1$, the dispersion relation reduces to:

$$\frac{K_0(x)}{I_0(x)} \frac{1 + \dfrac{c_{a\theta}}{2ic_{Az}} \dfrac{K_0^{'}(x)}{K_0(x)}}{1 + \dfrac{c_{a\theta}}{2ic_{Az}} \dfrac{I_1(x)}{I_0(x)}} = \frac{1 + \dfrac{c_{R\theta}}{2ic_{Az}} \dfrac{K_0^{'}(y)}{K_0(y)}}{1 + \dfrac{c_{R\theta}}{2ic_{Az}} \dfrac{I_1(y)}{I_0(y)}} \frac{K_0(y)}{I_0(y)}. \tag{23}$$

If twist is small it follows from equation(23), after expansion of the Bessel functions and noting that $x,\ y\ \ll 1$:

$$1 + \frac{1}{2ic_{Az}} \left(\frac{c_{a\theta}}{xK_0(x)} - \frac{c_{R\theta}}{yK_0(y)} \right) = \frac{K_0(x)}{K_0(y)}, \tag{24}$$

and, since $y = \frac{R}{a}x$ and $c_{R\theta} = \frac{R}{a}c_{a\theta}$ we obtain

$$xK_0(x) = -\frac{c_{a\theta}}{2ic_{Az}}, \tag{25}$$

a transcendental equation determining x. We are now left to determine c_{ph} from

$$c_{ph}^2 = c_{Az}^2 \left(1 \pm \frac{2c_{a\theta}/c_{Az}}{\sqrt{x^2 - k^2 a^2}} \right).$$

Short wavelength ($1 \ll ka < kR$). We introduce the following notation.

$$c_0 = \rho_0(c_{ph}^2 - c_{Az}^2), \quad c_i = \rho_i(c_{ph}^2 - v_{Ai}^2), \quad c_e = \rho_e(cph^2 - v_{Ae}^2),$$
$$c_a = \rho_0 c_\theta^2, \quad c_R = \rho_0 c_{R\theta}^2 \quad \text{and} \quad \varphi_0 = \tfrac{K_0(ka)}{I_0(ka)} \tfrac{I_0(kR)}{K_0(kR)}$$

After some algebra we find that:

$$\varphi_0 = \frac{c_e - c_0}{c_e + c_0} \frac{c_i - c_0}{c_i + c_0} \left(1 - \frac{1}{x} \frac{c_i^2 - c_i c_0 - 2c_0 c_a}{c_i^2 - c_0^2} \ + \ \frac{1}{y} \frac{c_e^2 - c_e c_0 - 2c_0 c_R}{c_e^2 - c_0^2} \right). \tag{26}$$

This is of too high order in c_{ph} to be solvable. However, it is possible to further simplify for the slender annulus limit. So, for $k(R-a) \ll 1$, we find

$$\varphi_0 \approx 1 \ + \ 2(R-a)m_0, \tag{27}$$

and using the fact that $x, y \gg 1$ (short wavelength) and $k(R-a) \ll 1$ (slender annulus) equation (26) can be reduced to give

$$(c_e + c_0)(c_i + c_0) = (c_e - c_0)(c_i - c_0) \tag{28}$$

which yields the solutions

$$c_{ph}^2 = c_{Az}^2, \qquad c_{ph}^2 = \frac{\rho_i v_{Ai}^2 + \rho_e v_{Ae}^2}{\rho_i + \rho_e} = c_k^2, \tag{29}$$

a similar result to that found by Bennett et al. (1999).

It remains, in the near future, to study, for these dispersion relations, c_{ph} as a function of wavenumber $k_z a$ and to investigate the existence and behaviour of the various modes. Further study will hopefully lead to examining the case of a twisted annulus without the discontinuity in magnetic field at the inner annulus boundary in the hope to extend understanding of heating processes within the solar atmosphere.

References

Aschwanden, M.J., 2003, in (eds.) R. Erdélyi et al., *Turbulence, Waves, and Instabilities in the Solar Plasma*, NATO ASI **124**, p.215, Kluwer

Bennett, K., Roberts, B. & Narain, U., 1999, *Solar Phys.*, **185**, 41

De Pontieu, B., Erdélyi, R. & James, S.J., 2004, *Nature*, **430**, 536

Edwin, P.M. & Roberts, B., 1982, *Solar Phys.*, **76**, 239

Edwin, P.M. & Roberts, B., 1983, *Solar Phys.*, **88**, 179

Erdélyi, R., Mendoza-Briceno, C.A. & Sigalotti, L.D.G., 2004, in R.W. Walsh et al. (eds.), *ESA-SP*, **575**, 300

Goossens, M., Andries, J. & Aschwanden, M.J., 2002, *A&A*, **394** 39

James, S.P., Erdélyi, R. & De Pontieu, B., 2003, *A&A*, **406**, 715

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

Mikhalyaev, B.B. & Solov'ev, A.A., 2004, *Astron. Lett.*, **30**, 268

Robbrecht, E., Verwichte, E., Berghmans, D., Hochedez, J.F., Poedts, S., et al. 2001, *A&A*, **370**, 591

Roberts, B., 1981a, *Solar Phys.*, **69**, 27

Roberts, B., 1981b, *Solar Phys.*, **69**, 39.

Roberts B., 2004, in (eds.) H. Lacoste, *ESA-SP*, **547**, 1

Roberts, B., Edwin, P.M. & Benz, A.O., 1984, *ApJ.*, **279**, 857

Roberts, B. & Nakariakov, V.M., 2003, in (eds.) R. Erdélyi et al., *Turbulence, Waves, and Instabilities in the Solar Plasma*, NATO ASI 124, p.167, Kluwer

Ruderman, M.S. & Erdélyi, R., in (eds.) H. Lacoste, *ESA-SP*, **507**, 1 (2004)

Ruderman, M.S. and Roberts, B., 2002, *ApJ.* **577,** 475.

Taroyan, Y.T., Erdélyi, R. & Doyle, J.G., 2004, in R.W. Walsh et al. (eds.), *ESA-SP*, **575**, 443

Wang, T.J., Solanki, S.K., Curdt, W., Innes, D.E. & Dammasch, I.E, 2002, *ApJ*, **574**, L101

PADEU | PADEU **15**, 125 (2005)
ISBN 963 463 557

Acoustic Oscillations in Solar and Stellar Flaring Loops

A. Kelly, V. M. Nakariakov and T. D. Arber

Space and Astrophysics Group, University of Warwick, Coventry, CV4 7AL, UK

E-mail: kellya@astro.warwick.ac.uk

Abstract
Observations of flaring loops in radio, visible and x-ray bands show quasi-periodic pulsations with periods from a few seconds to several minutes. Recent numerical studies have shown that some of these oscillations can be interpreted as standing slow magnetoacoustic waves. Energy deposition from the flare excites the second standing harmonic, with a period determined by the temperature and loop length. The excited longitudinal oscillations can be practically dissipationless and can, possibly, be considered MHD autowaves. Numerical simulations with a wide range of flare durations and choices of heat deposition location show that the second harmonic is a common feature of flaring loops.
Keywords: *Sun: corona, Sun: oscillations, MHD.*

1 Introduction

Observations of the solar corona provide us with many observations of oscillations and waves in a variety of wavelength bands and with a huge range of periodicities. We are interested in quasi-periodic pulsations (QPP) observed in coronal loops during solar flares. The flare causes a temporary increase in the intensity of the signal from the loop, followed by a cooling phase in which the intensity returns to its normal level. Quasi-periodic variations in intensity are seen during this cooling phase. These pulsations have periods from a few seconds up to thirty minutes. The two main mechanisms proposed for these oscillations are the sausage mode and slow magnetoacoustic mode. For the sausage mode there exists a cut-off which imposes a limit of around 20 s on the period. (Nakariakov et al., 2003) However, post-flare oscillations often show periods

much longer than this and thus the sausage mode cannot explain all observations. In this study, we develop further an alternative suggested in Nakariakov et al. (2004): flare-generated acoustic oscillations.

There are numerous observations of quasi-periodic compressible pulsations in coronal loops. For example, Wang et al. (2003) present a review of hot coronal loop oscillations as observed by SUMER with periods between 7 and 31 minutes. Harrison (1987) presents solar X-ray pulsations observed by the Hard X-Ray Imagining Spectrometer on the Solar Maximum Mission and reports a 24 minute periodicity. Similar oscillations but of shorter periodicity are presented by McKenzie et al. (1997) and Terekov et al. (2002). Wang and Xie (2000) observed flare associated pulsations in the microwave band with a period of about 50 s. Pulsations have been observed simultaneously in X-ray and microwave bands, for example by Fu et al. (1996) and Tian et al. (1999). Similar oscillations, have been observed in the stellar case such as the oscillation with a period of 220 s observed by Mathioudakis et al. (2003) on the star Peg II.

2 Numerical Model

We model coronal oscillations using a 1D radiative hydrodynamic code. The numerical code is a 1D version of the Lagrangian Re-map code (Arber et al., 2001). This code includes effects such as thermal conductivity, gravitational stratification and radiative losses. For the radiative loss function we use form given by Rosner et al. (1978) extended to a wider temperature range. (Peres et al., 1982; Priest, 1982) The simulation domain consists of a hot (initially 1 MK) coronal loop with a dense, cool (10^4 K) plasma region at each footpoint. This region mimics the chromosphere and acts as a source of plasma to fill the loop during the flare.

The flare is modelled by the application of a Gaussian heat pulse given by

$$E_H = E_0 \exp\left(-\frac{(s-s_0)^2}{2\sigma_s^2}\right)\left[1 + Q_P \exp\left(-\frac{(t-t_p)^2}{2\sigma_t^2}\right)\right], \quad (1)$$

where the energy deposition width is σ_s and its duration is σ_t. The flare amplitude is Q_p. This function includes time independent background heating which is applied in order to maintain the equilibrium. The magnitude of this background heating is given by E_0. For simplicity this background heating is applied at the same position as the flare. The default values for the various parameters are given in Tab. 1. In our analysis of the results times are quoted relative to the flare peak time, ie. $t - t_p$.

In the first phase of the simulation, the loop is allowed to settle into an equilibrium between radiative losses and the time independent heating function. Next the flare causes the temperature to rise and heat travels down the loop by thermal conductivity. This in turn causes evaporation of material from the model chromosphere to fill the loop. After the flare peak the loop cools and its temperature and density return to

Table 1: *Default simulation parameters.*

Parameter		Value
Loop Length	L	55 Mm
Initial Temperature	T	1 MK
Background Heating	E_0	0.004 erg cm^{-3} s^{-1}
Heating Width	σ_s	7 Mm
Flare Peak Time	t_p	7200 s
Flare Amplitude	Q_P	2×10^4
Flare Location	s_0	0 Mm (Apex)

their initial values. We study the oscillations which appear in the density, temperature and velocity profiles during this cooling process.

3 Data Analysis

We consider acoustic waves set up in a coronal loop due to an impulsive energy deposition. We observe these waves by analysing temperature, density and velocity time-series from the simulations. Temperature oscillations are of smaller amplitude than the velocity and density oscillations and therefore we don't examine them. The time series that we use are the variations at the apex. The structure of standing modes is such that odd numbered modes have a node in density at this point while odd numbered modes have a node in velocity. The standing wave has a frequency given by

$$f_n(T) = 152\frac{\sqrt{T}n}{L}. \tag{2}$$

The wave frequency changes with temperature and the temperature changes throughout the flare development. Thus, we can identify not the standing not just by the frequency, but by the appropriate modulation of the frequency with time. We use a Morelet wavelet transform (Torrence and Compo, 1998) to observe the distribution of wave power over both time and frequency.

Fig 1 shows time-series for the density at the loop apex and the average loop temperature for a typical flare. ($\sigma_t = 100$) As described in the previous section the density and temperature are seen to rise until the flare peak and then return slowly to their initial values. The density reaches its peak value after the temperature. Quasi-periodic variations can be clearly seen in the density, and less clearly in velocity.

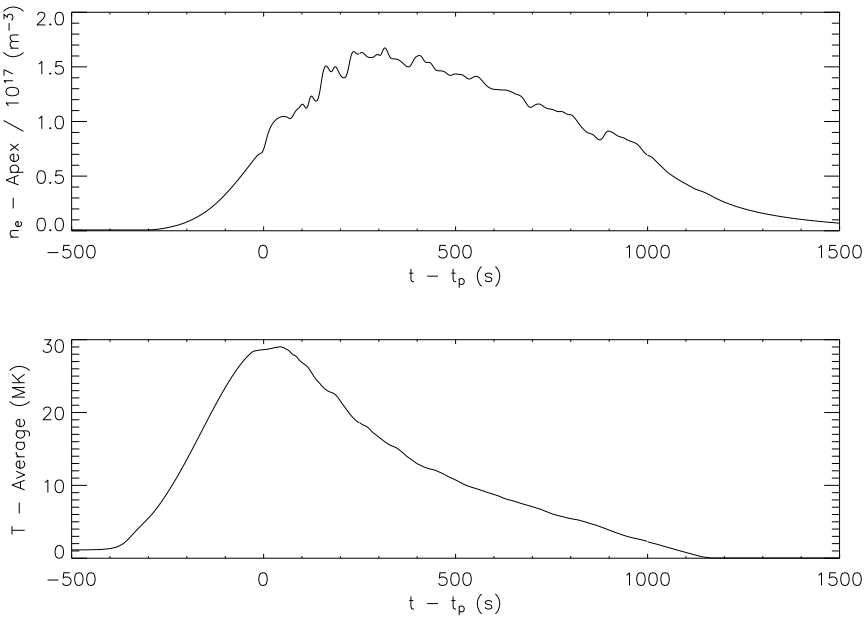


Figure 1: *Time evolution of the number density at the apex (upper panel) and average temperature (lower panel) for the $\sigma_t = 100$ s flare. ($Q_p = 2 \times 10^4$ and $s_0 = 0$)*

Before taking the wavelet transform we use a low pass filter to remove the long time scale variations from the signal. Fig 2 shows the filtered density from the 100 s flare and the wavelet the transform of this signal. This plot shows peaks in the wave power firstly around 200 seconds after the flare with a period of approximately 80 seconds and then later around 900 seconds after the flare with a period of approximately 180 seconds. The solid curved line across this plot shows the period of the harmonic. This oscillation with an amplitude of around 5–10% is the second harmonic standing mode acoustic wave.

As odd numbered modes have a node at the apex, it is unsurprising that the fundamental mode is not seen in these density plots. Fig 3 shows the velocity signal and its wavelet transform. There is no significant oscillation in the immediate aftermath of the flare with some oscillations appearing later. The wavelet transform shows that this later oscillation has power in the fundamental, second and fourth harmonics. We therefore conclude that the fundamental mode is not strongly excited by the energy

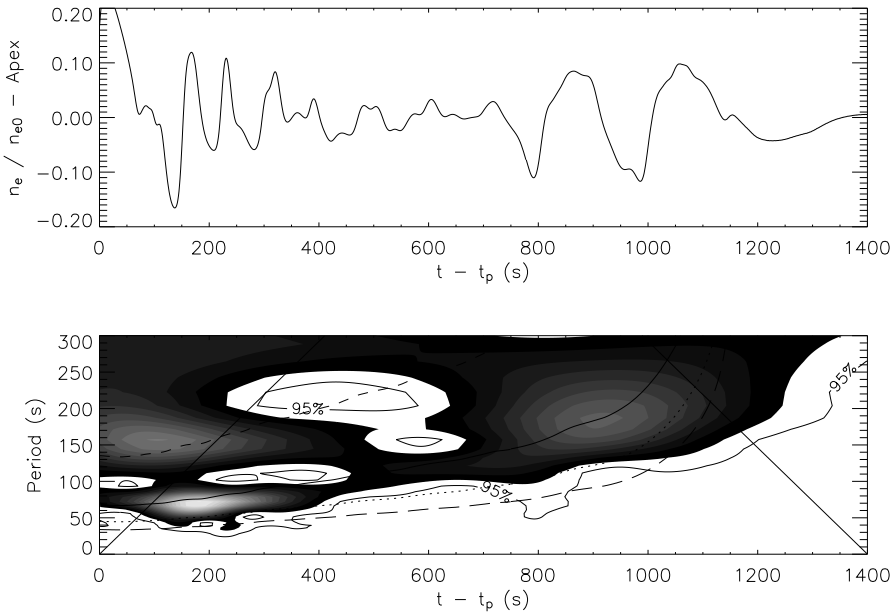


Figure 2: *The relative density oscillation n_e/n_{e0} at the apex for the 100 s flare (shown in Fig. 1) and a wavelet transform of the same signal. The curved lines across the wavelet transform show fundamental (upper dashed), second (solid), third (dotted) and forth (lower dashed) harmonic periods.*

deposition.

4 Discussion

The example shown in the previous section is a fairly typical simulation result. The second harmonic is seen to be the dominant excited mode for a wide range of flare parameters, while the fundamental mode is rarely seen. The second harmonic is a symmetric mode, and it would be natural to assume that the application of the flare at the apex leads to a symmetric oscillation. This, however, is not the case. The second harmonic continues to dominate, even when the flare energy is deposited asymmetrically at one footpoint. (Tsiklauri et al., 2004)

By examining the parameters of solar and stellar oscillations we can consider

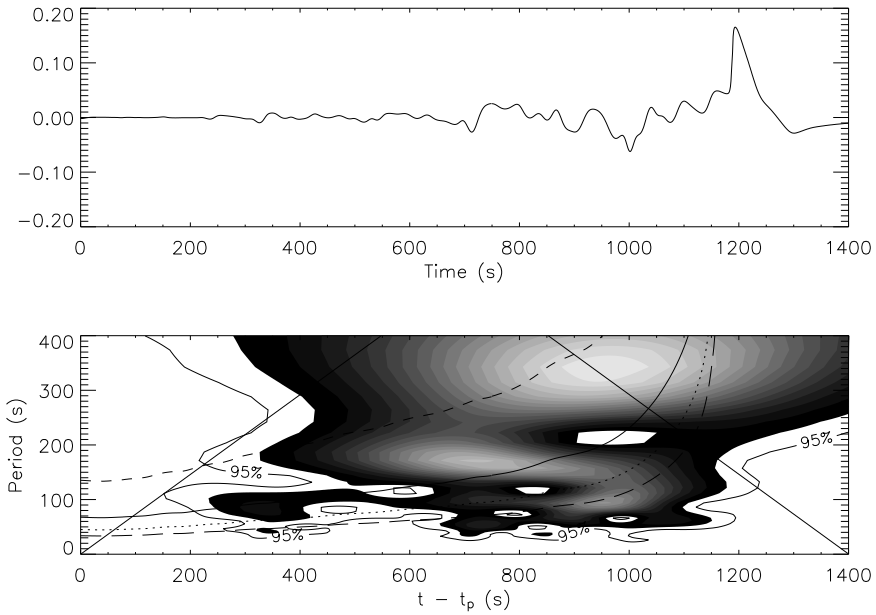


Figure 3: *The relative velocity oscillation V/C_S at the apex for the 100 s flare (shown in Fig. 1) and a wavelet transform of the same signal. The curved lines across the wavelet transform show fundamental (upper dashed), second (solid), third (dotted) and forth (lower dashed) harmonic periods.*

whether or not they are likely to be second harmonic standing acoustic waves as predicted by the simulations. Mathioudakis et al. (2003) studied white light oscillations during a flare on II-Peg and reported a period of 220 seconds. That paper gives estimates for the temperature and loop length of 200 MK and 500 Mm respectively. Using these values and Eq. 2, we can derive a value for the period of 233 seconds. This is consistent with the observed periodicity.

Considering solar observations Wang et al. (2003) examine a number of loop oscillations seen by SUMER. Many of these oscillations are not flare associated, but one example flare associated oscillation is that seen on the 29th of September 2000. An oscillation with a period of 28 minutes is observed in a loop with an estimated length of 515 Mm. This time, we can use Eq 2 to make a estimate of the temperature. This gives us a value of 4 MK. The observations are made using spectral lines at a

temperature of 6.4 MK.

The oscillations have two interesting properties. Firstly, they can persist for a number of oscillations with no significant damping, and sometimes even amplification. Secondly, they are often seen to disappear quite suddenly. The main dissipation mechanism for slow waves is well known to be thermal conductivity. From simple linear theory, the damping times for our example flare is 1.2 wave periods. Despite, this very strong damping several clear oscillations can be seen in Fig 2. As the oscillations appear despite the strong damping, then there must exist some instability which is amplifying the waves. One possible mechanism is the thermal instability. The simulation contains a heat loss function ($\mathcal{L}$) which gives the radiative losses from the plasma as a function of temperature. In the temperature regime at which oscillations occur $\frac{d\mathcal{L}}{dT} < 0$, meaning that cooler plasma loses energy more quickly. This instability can amplify waves. Waves may therefore exist in a balance between amplification due to thermal instability and high frequency dissipation. In order to test this hypothesis we run the flare simulations with radiative losses switched off. No oscillations are then seen in our example 100 s flare. If the waves are sustained by thermal instability then the waves would disappear suddenly when the temperature falls into the stable regime. This explains the sudden disappearance of the oscillations.

5 Conclusions

Observed quasi-periodic oscillations in coronal loops can be interpreted as standing magnetoacoustic waves. Hydrodynamic simulations of these waves show that energy deposition from a flare can excite standing modes and that the second harmonic is the most easily excited mode. This second harmonic seems to be the natural response of the loop to energy input and occurs for a wide range of energy deposition parameters such as position, duration and amplitude. The second harmonic is excited even for asymmetric energy deposition. The period of the oscillations can be calculated from the loop length and plasma temperature.

The acoustic waves interpretation is often excluded on the basis of strong thermal conductivity. Calculations based on the thermal conductivity suggest that the oscillations in our simulations should not occur. In the absence of thermal instability, these oscillations disappear and we therefore suggest that the waves exist as a result of competition between thermal over-stability and thermal dissipation.

References

Arber, T. D., Longbottom, A. W., Gerrard, C. L., Milne, A. M., 2001, J. Comput. Phys., 171 151.

Fu, Q.-J., Liu, Y.-Y., Li, C.-S., 1996, Chin. Astron. Astrophys, 20, 487.

Harrison, R. A., 1987, A&A, 182, 337.

Mathioudakis, M., Seiradakis, J. H., Williams, D. R., Avgoloupis, S., Bloomfield, D. S., McAteer, R. T. J., 2003, A&A, 403, 1101.

McKenzie, D. E., Mullan, D. J., 1997, Sol. Phys., 176, 127.

Nakariakov, V. M., Melnikov, V. F., Reznikova, V. E., 2003, A&A, 362, 1151.

Nakariakov, V. M., Tsiklauri, D., Kelly, A., Arber, T. D., Aschwanden, M. J., 2004, A&A, 414, L25.

Peres, G., Serio, S., Vaiana, G. S., Rosner, R., 1982, ApJ, 252, 791.

Priest, E. R., 1982, Solar Magnetohydrodynamics, Dordrecht,: D. Reidel Publ. Comp.

Rosner, R., Tucker, W. H., Vaiana, G. S., 1978, ApJ, 220, 643.

Terekov, O. V., Shevchenko, A. V., Kuz'min, A. G., et al., 2002, Astron. Lett., 28, 397.

Tian, D.-W., Gao, Z.-M., Fu, Q.-J., 1999, Chin. Astron. Astrophys., 23, 208.

Torrence, C., Compo, G. P., 1998, Bull. Amer. Meteor. Soc., 79, 61.

Tsiklauri, D., Nakariakov, V. M., Arber, T. D., Aschwanden, M. J., 2004, A&A, 422, 351.

Wang, M., Xie, R. X., 2000, Chin. Astron. Astrophys., 23, 208.

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

PADEU | PADEU **15**, 135 (2005)
ISBN 963 463 557

Local plasma diagnostics research on Electron Cyclotron Resonance Ion Source

L. Kenéz[1], J. Karácsony[2], A. Kitagawa[3], M. Muramatsu[3], S. Biri[4], A. Valek[4]

[1] Sapientia Hungarian University of Transylvania, Dept. of Electrical Engineering, Targu-Mures, RO-540053, 61, P-ta. Trandafirilor, Romania
[2] Babes-Bolyai University, Cluj-Napoca, RO-400084, 1, M. Kogalniceanu, Romania
[3] National Institute for Radiological Sciences, 4-9-1 Anagawa, Inage, Chiba 265-8555, Japan
[4] Institute of Nuclear Research (ATOMKI), H-4026 Debrecen, 18/c, Bem tér, Hungary

E-mail: l_kenez@ms.sapientia.ro

Abstract
Electron Cyclotron Resonance Ion Sources (ECRIS) are important tools of the atomic physics research, which find applications in the field of physics, medicine etc. The most important part of the ion sources is the complex plasma, which determines their performance but is still less known. Therefore we need to make efforts to find out more about them to improve ion sources. For this reason we started local plasma diagnostics research project, at the ATOMKI-ECRIS in Hungary and continued our work at the NIRS-ECRIS in Japan. In the framework of the project we studied the cold plasma of the source, developed a new method to evaluate the ion current flowing to the probe that is useful to calculate local electron density.We designed a mechanism to position the probe and determined various electron density distributions. Using emitting probes we determined the plasma potential and its distribution.
Keywords: *ECRIS, ECR plasma, Langmuir-probe, plasma diagnostics*

1 Motivation

The Electron Cyclotron Resonance Ion Source is one of the most successful machines built for Highly Charged Ion production (HCI). ¿From their late ′70 invention, the ion sources have been developed from complex prototypes to high-performance, highly sophisticated experimental facilities. ECR ion sources can be used in diverse research fields as stand-alone machines or injectors of high-energy accelerators. They were found useful in basic research and application either (e.g. tumour therapy, production of new materials etc.). The core of the source is the ECR plasma, which is decisive factor regarding the performance of the ion source. However, the theoretical knowledge we have on ECR plasma is incomplete, therefore diagnostic research of the plasma is important. Diagnostic methods canbe divided in two groups, global and local methods. Global methods provide overall information on the plasma parameters using the intense electromagnetic radiation coming out of the source, so there is no possibility to determine the local values of the plasma parameters. However, knowledge of the local plasma parameters and their evolution during (e.g. beam optimization process, biased-disk on/off etc.) external interventions would be of great interest in better understanding of the source and explore hidden possibilities. In case of ECRIS's mainly global diagnostic methods have been applied, while physicists just nowadays started to show interest in using local diagnostic methods.

2 Short presentation of the ECR ion source

Detailed description of ECR ion sources are described by many authors (e.g. Geller 1996; Brown 1989; Wolf 1995). This paper presents the block diagram (Fig.1.) of the ECR ion source.

¿From left to right one can see the microwave system responsible for plasma ignition and support, the magnetic system responsible for the confining of the plasma, the extraction system which allow us to extract ions from the plasma using electrostatic fields, the gas feed system which allow us to introduce different gases to obtain different plasmas (eventually other mechanisms to create plasmas from solid state materials), the vacuum system which is crucial for highly-charged ion production, and the radiation protection and safety systems which are important from the user and apparatus protection point of view. There are many possibilities to make experiments in the complex ECR plasma. The plasma is created by high-frequency microwaves and confined by complex magnetic trap. The magnetic trap is formed by the superposition of two different magnetic filed configurations, an axial field created by $DC4$ solenoid coils and a radial field created by permanent magnets. It is important that in the central region of the plasma chamber, the resulted field can be imagined as constant induction egg-shaped surfaces embedded in each other. One of these surfaces play crucial role in plasma ignition and maintenance. This surface is the so-called Resonant Zone, the inner region of this surface is the so-called hot plasma, while the outer region is the cold

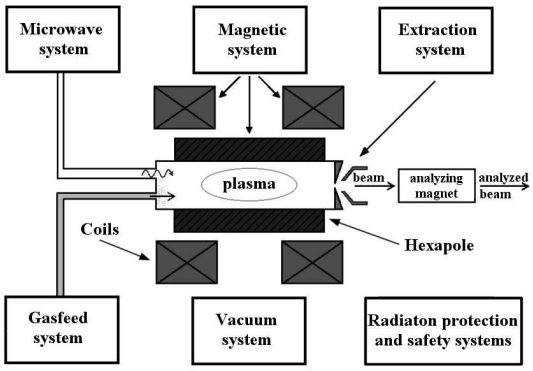


Figure 1: *Block diagram of an ECR ion source.*

plasma. The magnetic induction of the resonant zone is such that the frequency of the circular motion of the electrons equals the frequency of the microwaves introduced in the plasma chamber. When an electron goes through the resonant surface, stochastic resonant energy exchange occur between the electron and the microwaves. This is the so-called ECR heating. However, only the velocity component perpendicular to the magnetic field can be increased in this way, while the parallel remain unchanged. This way large anisotropy occur. Due to this fact, energetic electrons are trapped in the resonant zone. They collide with neutral atoms or low charged ions, which loose electrons step-by-step. These electrons form a negative cloud while ions are trapped due to electrostatic interaction between the electron cloud and ions. The final result of these processes is plasma ignition. As another result of a collision between electron and ion the anisotropy can completely or partially disappear, the electron is no more trapped and can get out of the resonant zone and get into the cold plasma. In this region electron cannot get energized by the microwaves, they cannot ionize, and most likely they escape from the magnetic trap, or mirrored back toward the resonant zone (Geller 1996).

3 Langmuir probe and ECR ion source

Depending on the particular plasma region we have to deal with and the plasma parameter we want to determine we have to choose properly the probe configuration. There are many choices, e.g. simple cylindrical probe, double cylindrical probe, emitting probe etc. Because the ECR plasma is complex there are many difficulties we have to deal with. Here we mention some of them, detailed description can be found

in (Kenéz 2002a).

- the plasma is confined in a strong and inhomogeneous $B - min$ magnetic trap,
- small plasma and chamber dimensions at high frequency microwaves (> 5 GHz),
- presence of multiply charged ions of different atomic species,
- secondary electron emission of the probe,
- sputtering of the probe metal caused by energetic particles of the plasma.

By careful design of the probe and careful definition of the experimental conditions many of these difficulties can be handled or their influence can be minimized. The experimental data we obtain is the voltage-current (U-I) curve of the probe. Fig.2. shows typical probe $U - I$ curves.

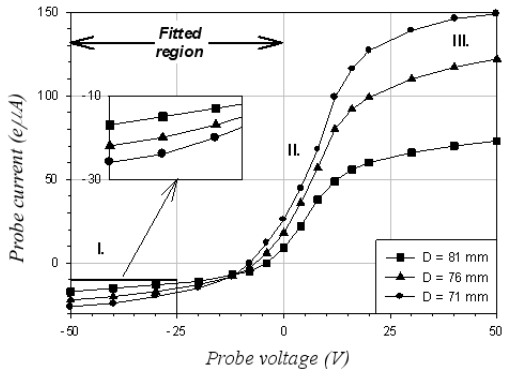


Figure 2: *Langmuir probe $U - I$ curves. D-distance between the probe and $B_M IN$ plane of the source*

These were measured in the cold region of the ECR plasma (oxygen plasma, source tuned for O^{1+} production). For the measurements we used cylindrical probe with 0.4 mm base diameter and 3 mm height. To calculate plasma parameters we fitted the [-50,0] voltage region assuming Maxwell-Boltzmann statistics for the collected electron component (Kenéz 2002a).

4 Theoretical model

The ECR plasma is confined in a magnetic mirror trap, which affects mostly the electrons of the plasma and makes the probe theory complicated. Plasmas confined in

magnetic fields cannot reach a complete equilibrium state; in consequence they can be described using different temperatures for the motions perpendicular and parallel to the magnetic field. In such cases one can expect that the slope of the probe characteristics give the parallel electron temperature (Chen 1965). The particle distribution of an ECRIS is well described by a "loss-cone" distribution (Dory 1965), which presents Maxwellian character for the particle motion parallel to the magnetic field. Thus we consider Maxwell-Boltzmann electron distribution function for the parallel component. Description of a theoretical model dealing with the case of singly charged ions could be found in (Chen 1965).

4.1 Multiply charged ions

Theoretical models concerning Langmuir probes do not take into account the fact that ECR plasmas contain multiply charged ions of different atomic species, so we have to include them into the calculations to evaluate the probe data correctly. An exact ion distribution function is not available, so it has to be introduced empirically. We assume that the plasma conditions are well reflected by the extracted beam when the plasma contains mostly lower charge states (Dou 1998). The electron density can be calculated as (Kenéz 2002a):

$$n_e = \frac{I_{sat}^{ion}\overline{j^+}}{0.61Ae}\left(\frac{k_B T_e}{M}\right)^{-1/2}\left(\sum_j \alpha_j j^{3/2}\right)^{-1}, \tag{1}$$

where n_e is the electron density, M the ion mass, A the normal area of the probe with respect to the magnetic field lines, e the elementary charge, k_B the Boltzmann constant, j the charge state; I_{sat}^{ion} the saturation ion current and T_e the parallel electron temperature, determined from the mean square fit of the characteristics; $\overline{j^+}$ the average charge and α_j the percentage composition of the beam, determined experimentally from the beam spectra.

4.2 Multi-component plasma

The vacuum of the source always contain residual gas ($N, C, H...$), while highly charged ion production needs gas mixing. All these atoms getionized in the plasma and eventually collected by the probe so they should be considered when more accurate electron density calculations are the objectives. Generalization of Eq. (1) is straightforward (Kenéz 2002a). We only have to make a summation for the different components of the plasma.

$$n_e = \frac{I_{sat}^{ion}\overline{j^+}}{0.61Ae}\left(k_B T_e\right)^{-1/2}\left(\sum_{k,j} \frac{\alpha_{k,j} j^{3/2}}{M_k^{1/2}}\right)^{-1}, \tag{2}$$

where now $\overline{j^+}$ is the estimated average charge of the multi-component plasma and $\alpha_{k,j}$ is the percentage composition of the beam (k denotes the different plasma components), determined experimentally from the beam spectra. The experiments have shown that differences between the electron densities yielded by the two models are significant when great amount of residual gas atoms are present in the plasma or two gases are used.

5 Experimental results

5.1 Axial electron density distribution. Biased-Disk effect

The current collector surface of the probe was set on the symmetry axis of the source. Probe $U-I$ curve series were taken in different plasmas tuned for different charge state production, e.g. O^{1+}, O^{3+}, O^{5+}, Ar^{4+}, Ar^{1+}, Ar^{11+} etc. During data acquisition the probe was moved step by step in the cold plasma region. Data processing was performed using all theoretical models presented in section 4. Three purposes were followed: electron density distribution along the source axis, effect of the multiply charged ions on the electron density calculated with the different models and study of the so-called Biased-Disk effect. The Biased-Disk is an internal electrode, which is negatively biased with respect to the ion source potential. When is turned on and its voltage is optimized, the extracted ion current can increase with a factor of $4-10$ depending on the optimized charge state. However, it is not completely understood the mechanism of this effect. Fig.3. presents the results of an experiment performed in Ar plasma, ion source tuned for Ar^{8+} production. Two series of $U-I$ curves were taken in Biased-Disk ON/OFF cases, respectively.

Conclusions. Moving the probe from the edge of the plasma chamber toward the resonant zone, the electron density increases. As it can be seen on the figure, withinin a relatively short distance ($20cm$) the electron density increases by one order of magnitude (true for every experiment we performed). Taking into account the different charge states of every atomic species, we calculated differences between the electron densities values calculated using SCSC and MCMC models up to a factor of 3 (Kenéz 2002b). Performing experiments increasing the charge state of the optimized ion we observed that that the calculated values of the electron density also increased. This is in agreement with the assumption that increasing the charge state of the optimized ion, the average charge state of the plasma ion component must also increase which naturally must be followed by an increase of the electron density. This result shows that for realistic description of the ECR plasma, multiply charged ions must be taken into account. We also showed that the Biased-Disk does not affect the cold plasma regions. The difference between the electron density values calculated for the Biased-Disk ON/OFF cases, are below the error limit (Kenéz 2002b).

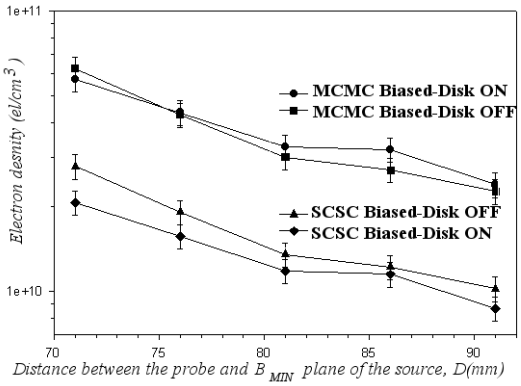


Figure 3: *Axial electron density distributions; Biased-Disk ON/OFF cases. Effect of the multiply charged ions*

5.2 Azimuthal electron density distribution

The intersection points of the end plane of the plasma chamber and those magnetic field lines which cross the resonant zone is a star-shaped configuration. Due to the cylindrical symmetry of the coils and the hexapole permanent magnet, the stars have three identical branches at each end of the chamber, but 60 degrees rotated. The purpose of the next experiment was to determine the electron density distribution in the branches of the stars. The holder of the probe was introduced on the axis of the source, so its current collector surface could be rotated on a given radius circle (e.g. $14mm$). Measurements were performed in the cold plasma region, in different axial planes, in a 100-degree angle region of only one branch of the injection side star. For this experiment low ionized O plasma was generated. We calculated the electron density using the models presented in section 4, but due to the low average charge of the plasma ion component (close to 1) no relevant differences we observed. But it is important to note, that using MCMC model the differences were larger than the error limit. It can be seen on Fig.4., that the electron density has maximum value in the middle plane of the star, than falls rapidly.

Outside the star-shape no electron current could be measured, only a little ion current, which proves that ions are electrically confined. Moving toward the resonant zone the electron density increase as we showed at the axial measurements, while the angular region where the plasma is located widens. The result of this experiment is the first three-dimensional electron density distribution of the ECR plasma in the literature (Kenéz 2004).

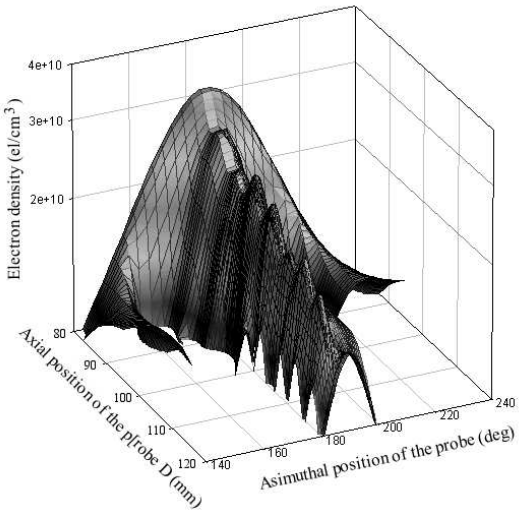


Figure 4: *3D - electron density distribution*

5.3 Plasma potential measurements using emitting probe

The experiments presented in this section were performed at the all-permanent NIRS-ECRIS, Japan. For well-known reasons, e.g. electron current suppression caused by the confining magnetic field, in case of ECR plasmas simple Langmuir-probes are not suitable for plasma potential measurement (Kenéz 2002a, Chen 1965). However, another type of probe, the emitting probe is available for this purpose. Let's see what makes possible to use emitting probes for plasma potential measurement. Curve 1 in Fig.5. is a typical cold probe characteristic curve.

There are two important points in this curve, one is the so-called floating potential (V_f) where the total probe current is zero and the other one is the plasma potential (V_p) where the plasma and the probe are on the same electric potential, there are no electric fields, the number of the collected charged carriers is determined by their energies. Due to the larger mobility of the electrons, V_f is negative for simple Langmuir-probes used in the cold regions of the ECR plasma. When a probe is heated (Fig.1. curves 2,3 and 4) it starts emitting secondary electrons. Due to the secondary electron emission, the shape of the probe voltage-current ($U - I$) curve changes in some regions. When the probe potential is above the plasma potential, the low energy secondary electrons are trapped by the positive potential barrier of the probe so they cannot leave and do not contribute to the total current. Decreasing the probe voltage below the plasma

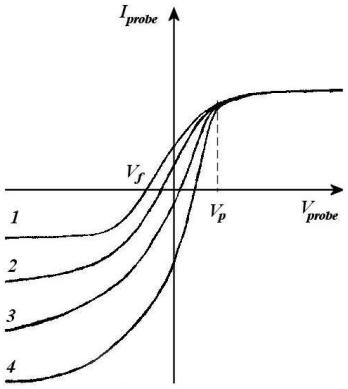


Figure 5: *Probe voltage-current curves; curve 1 - cold probe, curves 2,3 and 4 - heated, emitting probe (increasing heating current)*

potential, the secondary electrons are repelled by the probe and enter into the plasma contributing to the total current. It can be seen on the figure, that proportionally to the heating current, the steepness of the transition region of the curves is emphasized. Of course, the number of the secondary electrons increases proportionally with the probe temperature. This means, heating continuously up the probe, the floating potential gets higher and higher closing to the plasma potential, which is the saturation value of this process. Taking into account these considerations, ensuring sufficiently high secondary electron emission, the emitting probe can be used to determine the local plasma potential and consequently its distribution in ECR the plasma. Emitting probes can be simply realized making a loop of tungsten (or other material) of little dimensions and insulating it properly from other conducting parts of the ion source. The probe must have a heater electrical circuit and another circuit to bias it to different voltages with respect to the plasma chamber (which is on high voltage) and to measure the current. The emitting probe used at the present experiments was made of $0.1mm$ diameter tungsten wire and approx. 1 mm length. The emitting probe was made splicing one 0.1 mm diameter tungsten wire with two 0.1 mm diameter copper wires (Kenéz 2002b) and introduced inside a two-bored (2×0.4 mm) ceramic insulator tube. For test experiments different kinds of plasmas were generated at different microwave power using Ar, N and O as working gas. The ECR plasma contains highly charged heavy ions, which collide with the probe sputtering it. Due to this effect, the lifetime of the probe is limited. However, carefully controlling the experimental conditions we reached lifetimes as long as 10 hours. The behavior of the emitting probe and ion source was tested. During heating when the probe was kept in fixed position the

plasma was only slightly disturbed. The disturbance was slightly higher when the heated probe was moved inside the plasma chamber. The heated probe was used only in the cold regions of the ECR plasma. To determine the local plasma potential, $U - I$ curves were taken while the heater current was step-by-step increased to get higher and higher emission. Using these curves the behavior of the floating potential was analyzed. It was observed that corresponding to the theory, the floating potential increased as the probe temperature increased and finally saturated (Fig.6.).

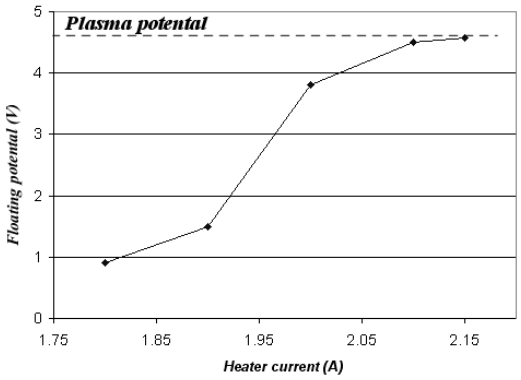


Figure 6: *Evolution of the floating potential during continuous heating of the probe; the saturation value is the local potential*

Under these circumstances we consider that the saturation value of the floating potential is the local plasma potential.

6 Conclusions

We adapted a method well known in the filed of plasma physics and successfully applied to calculated local plasma parameters and parameter distributions in the complex ECR plasma even though many difficulties are encountered. Both theoretical and experimental work has been done. We developed new theoretical model to take into account the multiply charged nature of the ECR plasma. We also measured the first three-dimensional local electron density distribution. Using the emitting probe method, we measured local plasma potential. Further experiments are needed to build the three-dimensional plasma potential map of the cold ECR plasma.

7 Acknowledgments

One of the authors (L.K.) wishes to thank A. Kitagawa for making possible the research in NIRS. In addition, L.K. and J.K. wish to thank Domus Hungarica Foundation for the awarded fellowship.

References

Geller, R. 1996, *Electron Cyclotron Resonance Ion Sources and ECR plasmas*, IoP Publishing Ltd.

Brown, I.G. 1989, *The Physics and Technology of Ion Sources* John Wiley $ Sons, Inc. 207-229

Wolf, B. 1995, *Handbook of Ion Sources*, CRC Pres, Inc. 109-121

Kenéz, L., Biri, S., Karácsony, J., Valek, A., 2002, *Nucl. Instrum. Methods Phys. Res. B* 187/2, 249-258

Chen, F.F. 1965, *Plasma Diagnostics Techniques*, Academic, New York, Hapter 4.

Douysset, G., Khodja, H., Briandt, J.P. 1998, *IXth International Conference on the Physics of Highly Charged Ions* 514, Bensheim

Kenéz, L., Biri, S., Karácsony, J., Valek, A., Nakagawa, T., Stiebing, K.E., Mironov, V., 2002, *Rev. Sci. Instrum.* 73/2, 617-619

Kenéz, L., 2004, *Fizikai Szemle* 2004/12, 411-415

PADEU | PADEU **15**, 147 (2005)
ISBN 963 463 557

Self-consistent modelling of glow discharges

K. Kutasi[1], P. Hartmann[2] and Z Donkó[3]

Research Institute for Solid State Physics and Optics of the Hungarian Academy of Sciences, POB 49, H-1525 Budapest, Hungary

E-mail: [1]kutasi@sunserv.kfki.hu, [2]hartmann@sunserv.kfki.hu, [3]donko@sunserv.kfki.hu

Abstract
The paper presents a self-consistent discharge model of gas discharges, the hybrid model. The applicability of the model is illustrated on a helium gas discharge. With the hybrid model the role of molecular ions in helium glow discharges is investigated. Because of the uncertainties in the determination of the electron temperature the effect of kT_e (used as an input parameter of the model) on the calculated discharge characteristics is investigated.
Keywords: *dc glow discharges, hybrid models*

1 Introduction

Low temperature, cold-cathode glow discharges are used in various application fields: in the semiconductor industry for plasma etching and deposition, for lighting and laser purposes, for plasma display panels, in analytical chemistry as spectroscopic sources for the analysis of solid materials, etc. In order to optimize these applications, a good insight into the glow discharge processes is desirable. Numerical modelling proved to be a powerful technique for this purpose.

Glow discharges are composed of many regions, which have very different emission light intensity, electric field and charge density distributions. The most important part of the glow discharges is the cathode region with the cathode sheath and the negative glow. In this region are created the charges which are necessary for the self-sustainment of the discharge. Many applications use the cathode region because of the strong light emission characteristic to the negative glow and also because of the

presence of high energy electrons and consequently of active radicals. Further in our work we talk about glow discharges which contains only the cathode sheath and the negative glow region.

In this paper, the basics of hybrid model are described and the applicability of hybrid modelling is illustrated on a helium gas discharge.

2 Hybrid model

Hybrid models consist of a fluid model and a Monte Carlo model. The fluid model makes it possible to describe the motion of charges which are in hydrodynamic equilibrium with the electric field in the low electric field region (negative glow). With the Monte Carlo model the electrons which move in the high electric field present in the cathode vicinity can be traced. The fluid models are based on a two-component fluid description of the plasma. The self-consistency is achieved by solving the Poisson equation together with the continuity (Eq. 1) and momentum transfer equations (Eq. 2) for fluid species (positive ions and electrons):

$$\frac{\partial n_e}{\partial t} + \nabla(n_e \mathbf{v}_e) = S_e,\ \frac{\partial n_i}{\partial t} + \nabla(n_i \mathbf{v}_i) = S_i,\ \triangle V = -\frac{e}{\epsilon_0}(n_i - n_e), \tag{1}$$

where $\mathbf{v}_e$ and $\mathbf{v}_i$ are the mean velocities, S_e and S_i are the source functions, n_e and n_i are the densities of electrons and ions, respectively, e is the elementary charge, V is the electric potential and ϵ_0 is the permittivity of free space. The mean velocities $\mathbf{v}_e$ and $\mathbf{v}_i$ are calculated from the momentum transfer equations for electrons and ions:

$$\Phi_e = n_e \mathbf{v}_e = -n_e \mu_e \mathbf{E} - \nabla(n_e D_e),\ \Phi_i = n_i \mathbf{v}_i = n_i \mu_i \mathbf{E} - \nabla(n_i D_i), \tag{2}$$

where $\mu_{e(i)}$ and $D_{e(i)}$ are the mobility and diffusion coefficients of electrons (ions) and $\Phi_{e(i)}$ are the corresponding fluxes. The set of fluid equations is closed by the equation describing the production of primary electrons at the cathode $\mathrm{j}_c^- = \gamma \mathrm{j}_c^+$, where γ is the secondary electron emission coefficient and j_c^+ and j_c^- are the ion and electron current at the cathode, respectively. The source functions of the fluid species are calculated in the Monte Carlo routine for fast electrons, which are traced in a potential distribution obtained from the fluid model (Surendra et al., 1990; Fiala et al., 1994; Donkó et al., 1998).

In the MC algorithm random numbers are used to determine the positions and the types of the collisions. The random numbers (R_{01}) have a uniform distribution in the [0,1) interval. The primary electrons emitted from the cathode and their secondaries produced in ionizing collisions are traced until they are absorbed by the anode or, due to their energy losses in inelastic collisions, they are no longer capable of producing any additional ions. The $\mathbf{r}(t)$ trajectory of electrons between successive collisions is followed by direct integration of their equation of motion:

$$m\frac{\mathrm{d}^2\mathbf{r}}{\mathrm{d}t^2} = e\mathbf{E}, \tag{3}$$

where e and m are the electron charge and mass, respectively and $\mathbf{E}$ is the electric field. The free path of electrons is assigned randomly and the position of the collision is calculated from (Boeuf et al., 1982):

$$\int_{s_0}^{s_1} n\sigma[\varepsilon(s)]\mathrm{d}s = -\ln(1 - R_{01}), \tag{4}$$

where s_0 is the position of the last collision and s_1 is the position of the next collision measured on the curvilinear abscissa s, n is the background gas density, σ is the sum of cross sections of all possible elementary processes, ε is the kinetic energy of the electron.

The type of the collision which occurs after the free flight is chosen randomly, taking into account the values of cross sections of different processes at the energy of the colliding electron. The source function of ions S_i is accumulated from the individual ionization processes. The electrons are transferred to the slow electron group when their (kinetic+potential) energy falls below the ionization potential of the gas atoms. Here the potential energy is considered to be the difference between the maximum value of the potential in the discharge and the potential at the actual position of the electron. In the hybrid model the Monte Carlo and fluid models are solved iteratively until the stationary state of the discharges is reached.

3 Modelling of helium glow discharges

3.1 Description of the self-consistent model

It is well known that in helium glow discharges at high pressures $\geq$ 100 mbar high density of molecular ions are present in the discharge and are responsible for the strong UV and VUV radiation. At low pressures (several mbar) the molecular ions have been investigated only in positive column discharges. At pressures higher than 10 mbar in positive column discharges molecular ions become the dominant ions (Ichikawa et al., 1980). The aim of our work is to investigate low pressure negative glow helium discharges and to answer the two main questions: (i)Are the molecular ions present and do they play an important role in the self sustainment of low pressure negative glow helium discharges? (ii) How does the molecular ion to atomic ion density ratio change in the discharge with pressure? In order to answer these questions we investigate similar glow discharges (pL = const., j/p^2=const.) in the 2-60 mbar pressure range. Four different discharges are studied at the conditions which correspond to the same pL = 6 mbar cm and reduced current density j/p^2 = 0.027 mA cm^{-2} mbar^{-2}: (i) p = 2 mbar, j = 0.108 mA cm^{-2}, L = 3 cm (ii) p = 6 mbar, j = 0.972 mA cm^{-2}, L = 1 cm, (iii) p = 20 mbar, j = 10.8 mA cm^{-2}, L = 0.3 cm and (iv) p = 60 mbar, j = 97.2 mA cm^{-2}, L = 0.1 cm.

The discharges are described by a one-dimensional hybrid model which combines a fluid model for atomic and molecular ions and slow electrons, the Monte Carlo (MC)

simulation of fast electrons, and a diffusion-reaction model of the metastable species (it has been found from the literature that the metastables play on important role in the formation of the atomic and molecular ions). In the simulation the fluid, Monte Carlo and metastable models are solved in an iterative way until the stationary state of the discharge is reached. The structure of the hybrid model and the transfer of physical quantities between the three submodels are presented on the flowchart shown in Fig. 1. The elementary processes taken into account in the model are summarized in Table 1 (for a more detailed discussion see Ref. (Kutasi et al., 2001).) For fast electrons we take into account elastic scattering (p1), excitation to metastable and several higher excited states (up to $n = 5$) (p2-p3) as well as ionization (p4). The excited atoms (including the $n = 3$ to $n = 5$ states) can participate in associative ionization process (p5) in which molecular ions are created. The singlet and triplet atomic metastables may convert into triplet atomic and molecular metastables, respectively (p6-p7). The singlet atomic metastables also may convert to ground state atoms due to collision with the gas atoms (p8). The atomic and molecular ions are partly created in metastable-metastable associative ionization processes which result in the loss of metastables (p9). The metastables are also lost in deexcitation processes (p10-p11). The atomic ions convert into molecular ions through the ion conversion process (p12). The atomic ions are lost through collisional radiative (p13) and radiative recombination (p14). The molecular ions are lost through dissociative recombination (p15) (this process was not taken into account in the model of (Kutasi et al., 2001)), collisional radiative recombination (p16) and three body recombination (p17) processes.

Table 1: *Elementary processes considered in the model. He(S), He(T) and He(M) denote the singlet atomic, triplet atomic and molecular metastables, respectively.*

Proc. id.	Reaction Process	Proc. id.	Reaction Process
p1	$He+e^-\rightarrow He+e^-$	p10	$He(S)+e^-\rightarrow He+e^-$
p2	$He+e^-\rightarrow He(S,T)+e^-$	p11	$He(T)+e^-\rightarrow He+e^-$
p3	$He+e^-\rightarrow He^*+e^-$	p12	$He^++2He\rightarrow He_2^++He$
p4	$He+e^-\rightarrow He^++2e^-$	p13	$He^++2e^-\rightarrow He^*+e^-$
p5	$He^* + He \rightarrow He_2^+ + e^-$	p14	$He^++e^-\rightarrow He+h\nu$
p6	$He(S)+e^-\rightarrow He(T)+e^-$	p15	$He_2^++e\rightarrow He(1^1S)+He(2^3S)$
p7	$He(T)+2He\rightarrow He(M)+He$	p16	$He_2^++2e^-\rightarrow He_2^*+e^-$
p8	$He(S)+He\rightarrow 2He$	p17	$He_2^++e^-+He\rightarrow He_2^*+He$
p9	$He(S,T)+He(S,T)\rightarrow He^++He+e^-$ $He_2^++e^-$		

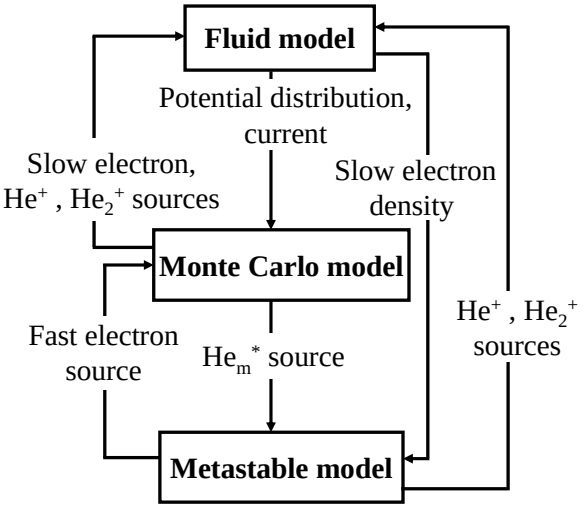


Figure 1: *Flowchart of the model.*

The input parameters of such a hybrid model are the discharge voltage, gas pressure, rate coefficients of different processes, the electron collisional cross sections, diffusion and mobility coefficients, the temperature of bulk (slow) electrons kT_e, and the secondary electron emission coefficient γ. One part of these parameters can be determined experimentally, the other part can be found in the literature, however there is a lack of data for the (i) electron temperature and (ii) secondary electron emission coefficient. Hybrid models of negative glow discharges conventionally use a constant characteristic energy for the slow electrons, which is chosen to be $kT_\mathrm{e} = 1$ eV in almost all studies. In some previous investigations on low pressure negative glow discharges cold electron temperatures significantly lower than 1 eV have been found. These studies include laser based diagnostics of Den Hartog et al. (1989), theoretical calculations of Arslanbekov et al. (1998), Langmuir probe measurement of Bogaerts et al. (1995), Angstadt et al. (1993) and Ohsawa et al. (1991), Thomson scattering measurements of Gamez et al. (2004). In all of these works cold electron temperatures ranging between 0.08 and 0.5 eV have been found. Considering these data, the 1 eV value, used in most hybrid model-based simulations, may be too high. As the determination of the electron temperature in a wide pressure range is difficult in our work we also investigate the effect of the slow electron temperature on the discharge characteristics.

3.2 Results of the model

First the discharge characteristics calculated for 2 mbar pressure are presented. The model predicts the formation of cathode sheath - negative glow structure, see Fig. 2(a); the electric field falls nearly linearly from the cathode, and it is closely zero in the negative glow region. The charge density distributions - shown in Fig. 2(b) - indicate the presence of a quasi-neutral plasma in the negative glow and the dominance of the positive ions in the cathode sheath. According to the calculated charge fluxes Fig. 2(c) a part of the ions flows to the anode, which is the consequence of the small negative field present in the anode vicinity, due to the reversal of the electric field in the negative glow, see Fig. 2(d). The modelling results show that the position of the field reversal coincides with the position of the maximum density.

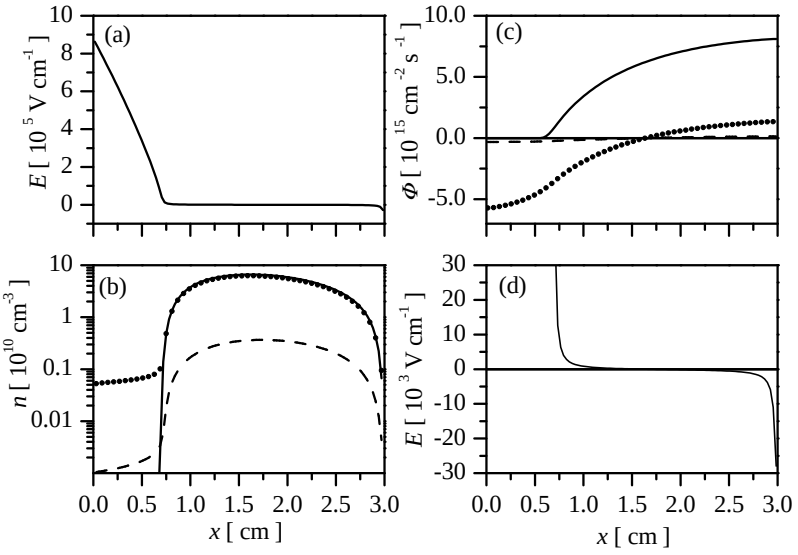


Figure 2: *(a) Axial distribution of the electric field; (b) Density distribution of the slow electrons (—), atomic (•) and molecular (– – –) ions; (b) Flux of the slow electrons (—), atomic (•) and molecular (– – –) ions; (d) Axial distribution of the electric field enlarged for the negative glow.*

The dependence of calculated discharge characteristics on the assumed value of the electron temperature is studied in details in the forthcoming part of the paper. First the dependence of the current density on kT_e at different pressures is studied. The results are illustrated in Fig. 3: the current density increases with increasing electron temperature. The results show that in order to obtain by modelling the experimental

current density the electron temperature in the discharges of 2-60 mbar pressure have to be chosen in the 0.1 - 0.3 eV range. Assuming higher kT_e values in the hybrid models the current density can be strongly overestimated. At low pressures assuming an electron temperature of 1 eV introduces a 20% error in the current density, however with increasing pressure this error raises and at 60 mbar reaches 60%. This strong dependence of the current densities on the electron temperature let us conclude that for correct modelling the accurate determination of the electron temperature is required.

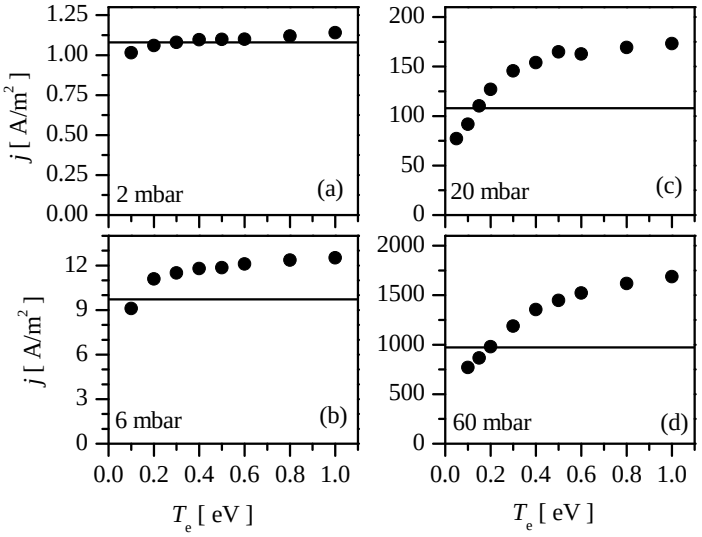


Figure 3: *Calculated current density (•) as a function of the assumed electron temperature at different gas pressures. The heavy horizontal lines in the panels represent the experimental current density at $U = 350$ V.*

In the following the percentage of molecular ions in the negative glow as a function of the assumed kT_e is studied and presented in Fig. 4(a). The percentage of molecular ions decreases with increasing electron temperature, in comparison with the 0.1 eV case, at 1 eV the percentage of molecular ions decreases by about 10% at $p = 60$ mbar. The results show that in the 0.1-0.2 eV range, the percentage of molecular ions decreases by about 5% in the case of 2, 6 and 20 mbar, while at 60 mbar increases by about 2%. In the case of $kT_e = 0.2$ eV - which is supposed by us to be a realistic value for the electron temperature - at 2 mbar 6% of ions are molecular ions, at 6 mbar 16%, at 20 mbar 30% and at the highest pressure investigated 60 mbar 42%. In Fig. 4(b) the percentage of the ion current carried by molecular ions at the cathode as a func-

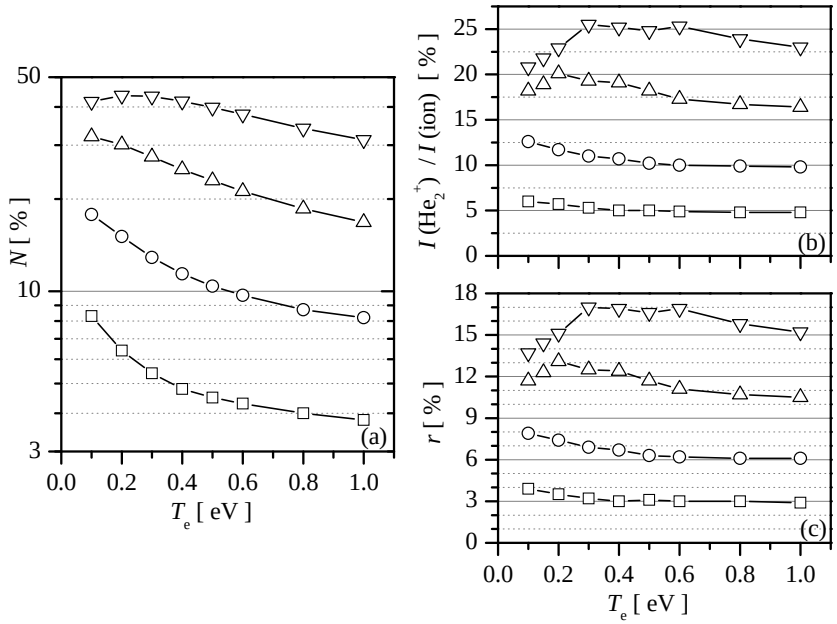


Figure 4: *(a) Percentage of molecular ions in the negative glow as a function of the assumed electron temperature at 2 mbar (□), 6 mbar (○), 20 mbar (△), 60 mbar (▽). (b) Percentage of the ion current carried by the molecular ions at the cathode and (c) percentage of secondary electrons released by molecular ions at the cathode as a function of the assumed electron temperature at 2 mbar (□), 6 mbar (○), 20 mbar (△), 60 mbar (▽).*

tion of the assumed electron temperature is presented. The ratio of the molecular ion current $I_{\mathrm{He}_2^+}$ to the total ion current shows a slight dependence on the electron temperature. In the case of $kT_\mathrm{e} = 0.2$ eV at 2 mbar 5% of the ion current is carried by the molecular ions, at 6 mbar this percentage increases up to 12% and at 60 mbar reaches 22%. The self sustained mode of operation of the discharge is assured by the ions arriving at the cathode surface which induce the emission of secondary electrons. Fig. 4(c) shows the percentage of secondary electrons released by the molecular ions, $r = 100 \times \gamma_{\mathrm{He}_2^+} I_{\mathrm{He}_2^+} / (\gamma_{\mathrm{He}_2^+} I_{\mathrm{He}_2^+} + \gamma_{\mathrm{He}^+} I_{\mathrm{He}^+})$. At 2 mbar 3% of electrons are released by the molecular ions, at 6 mbar 7% and at 60 mbar 15%. These results give us an insight about the increasing importance of molecular ions in the self-sustainment of the discharge.

4 Summary

In the paper the hybrid model and the investigations carried out on a helium glow discharge in the 2-60 mbar pressure range have been presented. The calculations have been carried out for similar discharges (pL and j/p^2 const.) at a constant discharge voltage $V = 350$ V. Besides the role of molecular ions in the discharge the effect of the electron temperature on discharge characteristics have been investigated. The model have shown that even at low pressures like 6 mbar 16% of ions are molecular ions and they play on important role in the secondary electron emission. The comparison of the measured and calculated current densities indicated that the electron temperature in the 2-60 mbar pressure domain is in the 0.1-0.3 eV range, which is significantly lower than the $kT_e = 1$ eV value conventionally used in hybrid models. We have shown that by assuming $kT_e = 1$ eV – as the majority of hybrid models do – the current density of the discharges can be significantly overestimated.

From our investigations we conclude that for correct modelling the accurate experimental determination of electron temperature is necessary and in the case of helium discharges even at low pressures the presence of molecular ions should be taken into account.

Acknowledgement

This work was supported by the Hungarian Science Foundation (Grants OTKA T-34156 and T-48389).

References

Surendra, M., D. B. Graves, G. M. Jellum 1990, Phys. Rev. A, 41, 1112

Fiala, A., Pitchford, L. C., Boeuf, J. P. 1994, Phys. Rev. E, 49, 5607

Donkó, Z. 1998, Phys. Rev. E, 57, 7126

Boeuf, J. P., Marode, E. 1982, J.Phys. D: Appl. Phys., 15, 2169

Ichikawa, Y., Teii, S. 1980, J. Phys. D: Appl. Phys., 13, 2031

Kutasi, K., Hartmann, P., Donkó, Z. 2001, J. Phys. D: Appl. Phys., 34, 3368

Den Hartog, E. A., O'Brian, T. R., Lawler, J. E. 1989, Phys. Rev. Lett., 62, 1500

Arslanbekov, R. R., Kudryavtsev, A. A. 1998, Phys. Rev. E, 58, 6539

Bogaerts, A., Quentmeier, A., Jakubowski, N., Gijbels, R. 1995, Spectrochim. Acta B, 50, 1337

Angstadt, A. D., Whelan, J., Hess, K. R. 1993, Microchem. J., 47, 206

Ohsawa, A., Ohuchi, M., Kubota, T. 1991, Meas. Sci. Technol., 2, 801

Gamez, G., Bogaerts, A., Andrade, F., Hieftje 2004, Spectrochim. Acta B, 59, 435

PADEU | PADEU **15**, 159 (2005)
ISBN 963 463 557

Do type Ia supernovae prove a positive cosmological constant?

Zsolt F. Hetesi[1] and Lajos G. Balázs[2]

[1] Eötvös University, Dept. of Astronomy, H-1518 Budapest, P.O.Box 32., Hungary
[2]Konkoly Observatory of HAS, H-1525 Budapest, P.O. Box 67., Hungary

E-mail: [1]Zs.Hetesi@astro.elte.hu, [2]balazs@konkoly.hu

Abstract
Nowadays the only direct evidence for a model Universe with non-zero cosmological constant is the Hubble diagram of the distant Ia type supernovae. Other observations, e.g. the WMAP measurements, do not support the existence of a non-zero cosmological constant without any doubt (see Blanchard et al. (2003)). Therefore there is a growing interest in studying the redshift distance diagram of SN Ia supernovae to verify the existence of $\Lambda \neq 0$ cosmological constant. In this paper we show that (i) there is a correlation between the statistical residuals of SN Ia distance moduli in the Hubble-diagram and the calculated internal extinction values of the host galaxy. It suggests that there is something wrong with the previous estimations of internal extinction. Furthermore, we show that (ii) the correction for these correlation results a Hubble-diagram which does not support models with $\Omega_\Lambda \simeq 0.7$, rather an Einstein-de Sitter Universe.
Keywords: *SN Ia:general, cosmology:miscellaneous*

1 Introduction

It seems evident that after an alternative explanation of the WMAP results (Blanchard et al., 2003) the Hubble-diagram of the distant Ia type supernovae will become a more fundamental (and maybe the only direct) evidence for a cosmology with *non-zero* cosmological constant. Therefore there is an increasing interest in studying the fine details of this diagram to exclude the possible alternative explanations.

Type Ia supernovae (SNe Ia) have an important role in the chemical evolution of the Universe and the determination of cosmological distance scale. This later application

is based on their well defined $M_V = -19.4 \pm .5$ maximum luminosity (Richardson et al., 2002).

It is worth noting that the mean absolute magnitude of SNe Ia is assumed to be a universal standard in the whole Universe. Therefore they are used as standard candles. The paper of Tonry et al. (Tonry et al. (2003) and references therein; "TONRY" in what follows) concluded that $\Omega_\Lambda \simeq 0.7$ and $\Omega_M \simeq 0.3$, where Ω_M is the ratio of the density of the non-relativistic matter in Universe to the critical density and $\Omega_\Lambda = \lambda c^2/(3H_o^2)$, where λ is the cosmological constant, c is the velocity of light, and H_o is the Hubble constant. Previously, both the High-z Supernova Search Team (Schmidt et al. (1998), Riess et al. (2000)) and the Supernova Cosmology Project (Perlmutter et al. (1999)) came to the conclusion that supernovae (SNe) between redshifts of $z \simeq (0.3 - 1.0)$ had in the average $\simeq 0.28$ mag higher distance moduli than expected assuming $\Omega_M \simeq 0.3$ and $\Omega_\Lambda = 0$ (Riess, 2000).

Recently, two independent studies queried the reality of the nonzero cosmological constant deduced from the studies of high-redshift type Ia supernovae in the last years. Rowan-Robinson (2002) argued that the internal extinctions in the host spiral galaxies are underestimated, and he obtained inconclusive evidence for the positive cosmological constant. Mészáros (2002) used pure statistical arguments, and also showed that the introduction of the positive cosmological constant is premature yet.

The paper is organized as follows: After defining some important cosmological quantities and equations related to SNe Ia the statistical studies mentioned above are further extended in this paper. Out of 230 supernovae in TONRY dataset 188 have given extinction values. We separated it into subsamples with respect to their extinction values. Then the whole sample of 188 supernovae and its different subsamples are tested whether the statistical structure of the redshift distance relation depends on the internal extinction. These tests suggested that distances obtained for the high z part of the dataset were not independent from the internal extinction. Therefore any cosmological conclusion drawn from the supernova data should be taken with care. A more extensive sample of supernovae containing much more objects is highly required for a more reliable analysis.

We give the results of these tests. We show that former investigations, for example Choudhoury et al. (2004) failed to recognize the inadequacy of the removal of internal extinction from the data because of the improper definition of the subsamples investigated.

Using the result of statistical tests connected with the internal extinction values of SNe Ia we conclude that the corrected Hubble-diagram does not support models with $\Omega_\Lambda \simeq 0.7$, rather an open Universe with $\Omega_m \simeq 0.1$ matter.

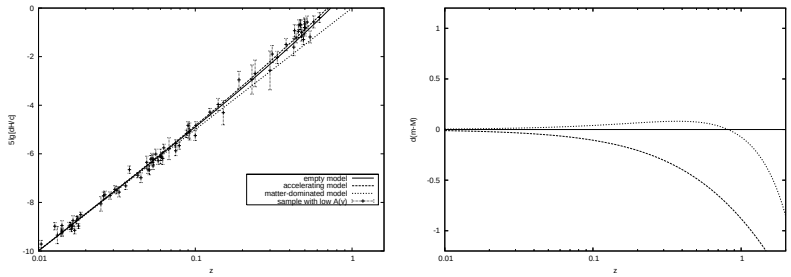


Figure 1: *Left: The distance modulus-redshift relation of SNe Ia. The data-set is form Tonry et al. 2003. There are three cosmological models marked in the figure. The "empty model" is dominated by only the cosmological constant, there is no matter density:* $\Omega_m = 0, \Omega_\Lambda = 1$. *The "accelerating model" is the today mostly accepted flat model, with cosmological constant, or dark energy:* $\Omega_m = 0.3, \Omega_\Lambda = 0.7$. *The "matter-dominated model" is a flat model without cosmological constant:* $\Omega_m = 1, \Omega_\Lambda = 0$. *Right: The deviation of the measured distance moduli from the calculated ones of the empty Universe. The uppermost curve represents the acceelrating model, the zero line belongs to the empty model, the lower curve is the Einstein-de Sitter Universe.*

2 Observational cosmology with SNe Ia

The luminosity distance (D_L) of cosmological objects depends on some cosmological parameters. These parameters determine the structure of the space-time. In the practice we measure some appropriate quantities of celestial objects from which we can calculate their luminosity distances. If we have the value of luminosity distance, we are in the position to compare these data with the predictions of theoretical models. SNe Ia are especially good objects for measurements like this because they are standard candles, i.e. their absolute luminosities (after corrections) are the same everywhere in the Universe. We can measure the redshift, and the apparent brightness of an individual supernova and one can calculate from these data the luminosity distance. The latest extensive data-set (Tonry et al. 2003) lists the logarithm of the redshift, multiplied by the speed of light c, and the logarithm of the luminosity distance multiplied by the present value of the Hubble constant, H_0. In this case there is no need for the accurate value of H_0. If we draw the distance modulus-redshift relations of different models we obtain the left side of Fig. 1.

The following equation specifies the relationship between the apparent magnitude and the redshift of an object:

$$m = M + 5 \log_{10} Q(z, \Omega) \tag{1}$$

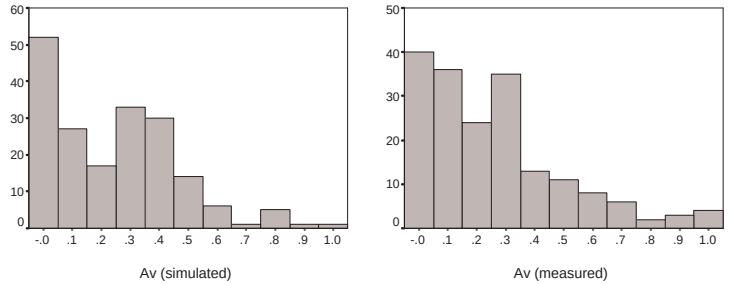


Figure 2: *Histogram of simulated SNe extinctions (left), and the real sample (right). The extinction is in magnitude on the horizontal axis. There is a well pronounced peak at $A_V = 0$ due to the objects at the front side of the host galaxy. The second smaller peak can be accounted for objects at the opposite side.*

where

$$Q(z, \Omega) = \frac{D_L H_0}{c} \tag{2}$$

As it was noticed at the beginning of this section the D_L luminosity distance strongly depends on some cosmological parameters. By definition the distance modulus is the difference between the apparent and the absolute magnitude: $\mu = m - M$ so Eq. 1 can be written in the following form using this definition:

$$\mu = 5 \log_{10} Q(z, \Omega) \tag{3}$$

In the redshift range investigated there are only small differences between the different models: we displayed in the left panel of Fig. 1 the deviation of the distance moduli obtained from different models and the distance modulus calculated from the empty Universe $\Delta\mu = \mu_i - \mu_{empty}$ (Fig. 1, right)

3 Data and statistical methods

Before putting the supernova data into a scatter plot, let us cast a glance on the sample! The extinction values scatter between 0 and 4.1 magnitudes with a median of 0.2 magnitude. Except a few outliers the data are concentrated in the 0-1 magnitude range. The distribution of the extinction values can be well understood if we compare it with a simple model. Assuming that the Ia type supernovae belong to the old disc population with a scale height of 340 pc and the interstellar dust concentrate to the plane of the host galaxy within a layer of 100 pc FWHM the observed distribution

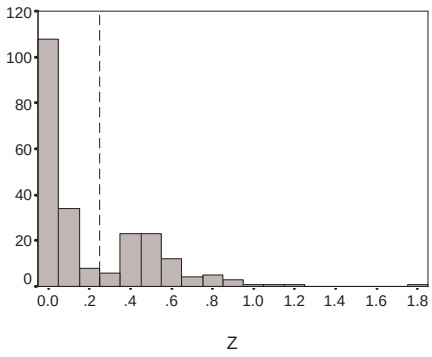


Figure 3: *Histogram of z distribution in the TONRY sample. The vertical dashed line indicates a cut between the low and high redshift part of the sample.*

of internal extinction is simply a projection effect. The peak near to zero extinction comes from supernovae in front of the host galaxy. Fig. 2 shows a comparison of the simulated distribution of the observed internal extinction along with that of TONRY data. In both samples there is a peak at $A_V = 0$ corresponding to the objects at the front side of the host galaxy. Both samples have another peak at $A_V \simeq 0.3$, which can be accounted for objects at the opposite side of the host. Despite of these similarities the peaks of the real sample are less pronounced probably due to observational errors.

The distribution of redshifts in the TONRY sample is bimodal. There is a dip at around $z = 0.25$. Having a cut at this point we defined two different subsamples. The high z subsample is interesting from at least two point of views. First, the calculated distance moduli from different cosmological models depart from each other exceeding the error of direct measurements obtained from SN Ia supernovae. Second, as we demonstrate below the estimated internal extinction in the high redshift part has some interrelation with the distance modulus which is not the case at $z < 0.25$.

The calculated distance moduli obtained from the observation of SN Ia events assumed to be independent from the internal extincion listed in the TONRY sample. It is easy to infer from the low redshift subsample that this is really the case. In the $z > 0.25$ range, however, at the low extinction part the majority of the SN Ia distance moduli exceeds those obtained from an empty Universe. Distance moduli from an empty Universe can be treated as an upper bound for those obtained from models without cosmological constant. Only in models of non zero cosmological constant can be exceeded this bound. The excess of SN Ia distances exceeding the empty case is considered as a firm evidence favoring the non zero cosmological constant.

The difference between distance moduli obtained from the observation of SN Ia

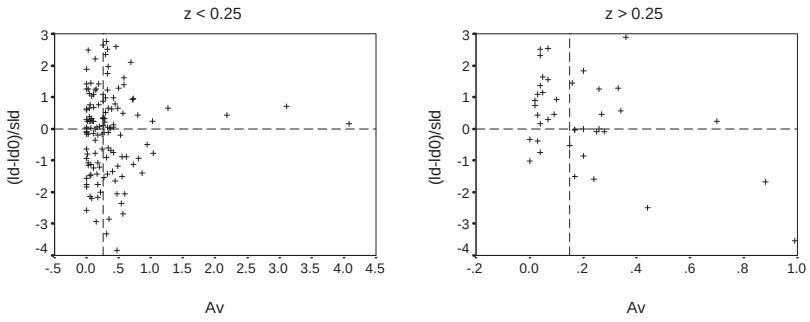


Figure 4: *The absorption-luminosity distance relation of the low (left) and high-z (right) subset (data source: TONRY). On the vertical axis the standardized deviation from distance moduli calculated in an empty Universe is given. Vertical dashed line mark the median of the extinction data. Horizontal dashed line marks the reference level of an empty Universe. Note the difference between the left and the right panel. In the* $z < 0.25$ *case the distribution of residuals is symmetric to the reference level of an empty Universe, independently from the extinction. On the contrary, the low exinction part (left from the median line) of the* $z > 0.25$ *panel clearly has an excees of the points above the reference line but it disappears at higher extinction values displaying a pronounced negative trend.*

events and calculated from an empty cosmological model assumed to be independent from the internal extinction of the host galaxy. As we demonstrated in the right panel of Figure 4 it is not the case in the $z > 0.25$ range. While the absorption values right from the median are symmetrical to the line representing the empty model, left from it, however, there is a remarkable asymmetry because there are more points above the line of the empty model than below it. If there was no interrelation between the distance moduli residuals and the internal extinction this trend would not be present. (The detailed calculation of the statistical significance of this trend will be published in a forthcoming paper). [1]

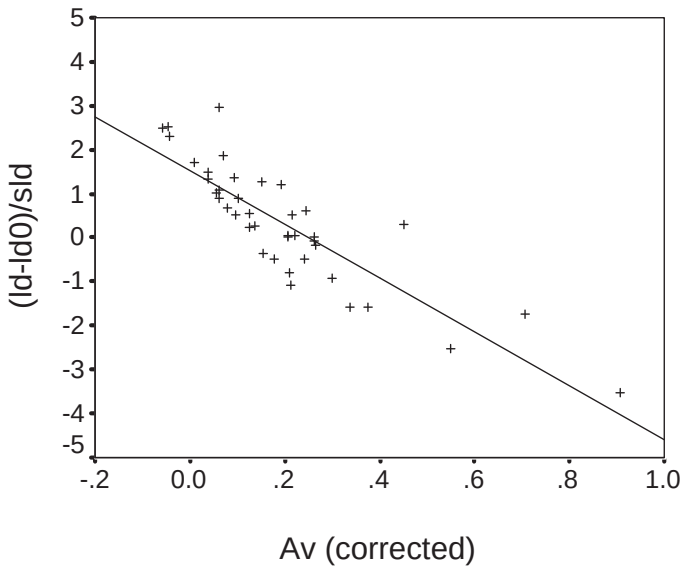


Figure 5: *Dependence of the uncorrected standardized residual on the corrected extinction represented by the f background variable. Full line displays the obtained relationship between the uncorrected residual and internal extinction.*

It is obvious, the distance moduli obtained from SN Ia observations have to be freed from the interrelation with the internal extinction before using it for testing

[1]It is interesting to note that some recent papers (for example Choudhoury et al. (2004)) divided the sample into two parts as follows: They cut off points with high absorption values and fitted some model on the remainder. As one can see on Fig. 4 this remainder subset really shows an excess of points above the zero level of the empty Universe.

cosmological models. To calculate statistically the effect of interrelation we introduced an f stochastic background variable representing the relationship between $\Delta\mu$ distance moduli residuals and the A_V internal extinction:

$$A_V = A_0\, f + \varepsilon_A \tag{4}$$

$$s = s_0\, f + \varepsilon_s \tag{5}$$

where $s = \Delta\mu/\sigma_\mu$ is the standardized residuals of distance moduli, $A_0, s_0, \varepsilon_A, \varepsilon_s$ are constants and noise terms, respectively. Estimation of f and other quantities in this system of equations can be performed by using factor analysis which is a standard procedure in multivariate statistics (the details will be given elsewhere). Based on this solution we can remove the effect of the background variable responsible for the interrelation between the residual and internal extinction and obtain the correct luminosity distances appropriate for testing cosmological models. Fig. 5 shows the dependence of the uncorrected standardized residual on the corrected extinction represented by the f background variable.

With the procedure outlined above we ceased statistically the interrelation between the distance moduli residuals and internal extinction displayed in Fig 4. It is worth noting that these corrections have also an effect on the data-points of low extinction values which were the firm basis of arguing for the existence of a positive cosmological constant.

4 Cosmological constant revisited

As a result of the corrections mentioned above we obtain a new luminosity distance-redshift diagram. While former results suggested the existence of a hump in the data (Fig 6, left), and the best fit was an accelerating model with $\Omega_m = 0.3, \Omega_\Lambda = 0.7$, after our correction the situation became highly different. Fig. 6, right shows that the modified SNe Ia data prefer a standard solution which seems to be open and contains only matter. It is interesting to note that due to our treatment the scatter of SN Ia distance moduli residuals decreased, as one can infer from comparing the two panels in Fig. 6.

5 Conclusion

Recently, it is commonly accepted that the expansion of the Cosmos is accelerating. However, there are some papers, which argue against the introduction of a cosmological constant with different arguments. Mészáros (2002) used pure statistical arguments to show that the use of the cosmological constant is premature yet. We also have done

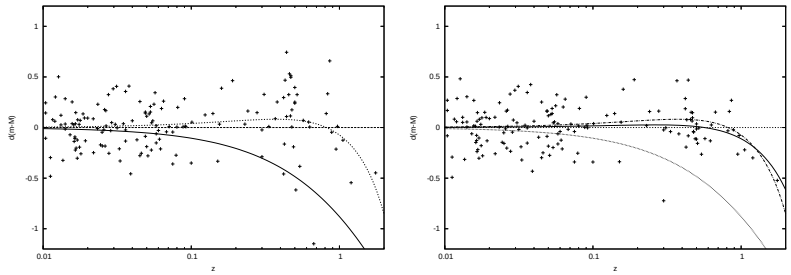


Figure 6: *Previous and new data plotted against z. As old sample favors a flat accelerating model (left) the new (right), statistically corrected one rather fits a matter dominated Universe. The labels are the same as Fig. 1, the new curve on the right panel is the fit.*

statistical study on the newest TONRY dataset. We investigated the absorption-luminosity distance relation statistically on those data which have $z > 0.25$. We found that there is an interrelation between the distance moduli residuals and the internal extinction of host galaxies in this part of the sample. We used statistical methods to calculate quantitatively this interrelation. After correcting for this effect the new residual Hubble-diagram does not fit a flat model with a positive cosmological constant instead, the best fit is an open matter-dominated model.

There is a need to explain the physical nature of the interrelation between distance muduli residuals and internal extinction of host galaxies. Does it have an astrophysical basis or it is only a byproduct of the data reduction?

Acknowledgement

The authors thank Zsolt Bagoly, Attila Mészáros, Béla Balázs, Zsolt Regály, and Szilárd Csizmadia for useful and friendly discussions and proposals. Technical help of András Pál and Emese Forgács-Dajka is also kindly appreciated.

References

Blanchard, A., Douspis, M., Rowan-Robinson, M., Sarkar, S. 2003, A&A 412, 35
Choudhury, T. R., Padmanabhan, T. 2004, A&A 429, 807
Mészáros, A. 2002, ApJ, 580, 12 (M02)
Perlmutter, S. et al. 1999, ApJ, 517, 565

Richardson, D., Branch, D., Casebeer, D., Millard, J., Thomas, R. C., Baron, E. 2002 AJ 123, 745

Riess, A.G. et al. 1998, AJ, 116, 1009

Riess, A.G. 2000, PASP, 112, 1284

Riess, A.G. et al. 2000, ApJ, 536, 62

Rowan-Robinson, M. 2002, MNRAS, 332, 352

Schmidt, B.P. et al. 1998, ApJ, 507, 46

Sullivan, M. et al. 2003, MNRAS, 340, 1057

Tonry, J. L., Schmidt, B. P., Barris, B., et al. 2003, ApJ 594, 1

PADEU | PADEU **15**, 169 (2005)
ISBN 963 463 557

Cosmic Rays Radiative Pre-acceleration in Wolf-Rayet Stellar Winds

A. S. Popescu

Astronomical Institute of Romanian Academy, Str. Cutitul de Argint 5, RO-040557 Bucharest, Romania

E-mail: sabinp@aira.astro.ro

Abstract
In the article "Cosmic Rays VIII" of Popescu et al. (2005) we gave an model for the observed cosmic rays between $5 \cdot 10^{15}$ and $3 \cdot 10^{18}$ eV. Their surse is presumed to be supernova of stars that explode in their winds. The observed cosmic abundance at the source are affected by spallation in the supernova shell, by the difference in ionization degree (being one two times ionized) at the injection in the supernova shock, the stars with initial masses $15M_\odot \leq M \leq 30M_\odot$ having a different contribution to them than the stars with $30M_\odot \leq M \leq 50M_\odot$ and this is 2:1 for the elements with $Z \geq 6$. Still, the abundances after these corrections are different by a factor Z_i/Z_{He} where Z_i is the atomic number for the element i and Z_{He} the one for He. This paper is dedicated to the explanation of this factor and its physical meanings.
Keywords: *cosmic rays, Wolf-Rayet stars, fractal density distribution*

1 Introduction

To understand why we need a preacceleration of cosmic rays (CRs) and a phase space dispersion before injection in the supernova shock of the particles found in the Wolf-Rayet (WR) wind we will do a breaf presentation of our previous results (Popescu et al. , 2005) on PeV-EeV energy range CRs.

These results are showing that the ions that can be found in the atmospheres of stars with the initial masses of $15M_\odot \leq M \leq 50M_\odot$ are possible to be the CRs observed particles with abundances affected by spallation and ionization losses. Also

Table 1: *k_i exponent values for even-Z elements in the case that $\alpha = \frac{2}{3}$. The preacceleration correction factor is $(Z_i/Z_{He})^k$ (Popescu et al. , 2005)*

elem. i	Z	k
H	1	-
C	6	0.876 ± 0.022
O	8	0.998 ± 0.0265
Ne	10	1.145 ± 0.022
Mg	12	1
Si	14	1
Fe	26	1

it was shown that the mass fraction for one even-Z element with an atomic number $Z \geq 2$ from CRs is $X_i = \alpha X_{i,RSG} + (1-\alpha)X_{i,WR}$.

We are having here a different contribution to CRs abundance of stars that explode as supernova in RSG stage or in WR stage, α factor being approximative equal with 2/3.

Also, ionisation loss is responsable for the underabundance in the observed CR elements with FIP$\geq$ 10 eV (the first ionization potential). We can consider Si as reference element for overabundance of elements with FIP $\leq$ 10 eV (Silberberg & Tsao , 1990). The FIP correction factor will be, in this case, $4.0088 \div 4.878$. In this way, we can see that the elements with FIP less than 10 eV are having an mass fraction larger than the elements with FIP greater than 10 eV (and relative to Si), by a factor of ~ 4. This happens to be exactly $Z^2_{injection}$, the initial degree of ionization squared. Therefore, cosmic ray particles of an element with an initial degree of ionization of $Z^2_{injection}$ are more likely to be injected by a factor of 4.

After introducing the spallation correction of the mass fractions in the massive stars atmospheres (Tsao et al. , 1998) we still remain with a difference between the all above effects corrected mass fractions in massive star winds and the observed CRs: $(X_i/X_{He})^{observed\ CRs} = f \cdot (X_i/X_{He})^{corrected\ wind}$, where:

$$f = (Z_i/Z_{He})^{k_i}. \tag{1}$$

The k values can be seen in Table 1.

Now we can wonder from where it's coming this factor. We will see that in a radiative instability (Owocki , 1994), thermal instability (Alfvén waves induced - Gonçalves et al. (1998)), forward (Lucy , 1982) or reverse (Moffat , 1994) shocks in the wind, that are driving turbulence, it is appearing a phase space separation. Also, in a radiative accelerated wind (CAK theory - Castor et al. (1975); MCAK theory -

Pauldrach et al. (1986); etc.), the ions are differently accelerated in function of the resonant line absorbtions of photons.

2 The Model

2.1 The Radiative Acceleration

A differentiated acceleration for the atomic species that can be found in the winds of WR stars (considered as having the same ionization degree) will give a phase space disperssion. This means that the acceleration as function of the local bulk velocity for oxygen must be bigger that the one for carbon, and the carbon acceleration bigger that the one for helium. The above affirmation is sustained by the radiation-driven wind model (Cassinelli , 1979), condition in which momentum is transferred from the radiation field to the gas by scattering of radiation in spectral lines. The radiative acceleration for a particular element found in the wind is given as the sum of all the radiative accelerations provided by single lines (Castor et al. , 1975). In CAK (after Castor, Abbott and Klein) theory it is shown that this sum can be parametrized by:

$$g_{rad} = \frac{1}{cN(r)}\sigma_{th}(r)\sigma_B T_{eff}^4 M_{CAK}(t), \tag{2}$$

where $M_{CAK}(t) = kt^{-\alpha}$ is the force multiplier and encapsulates the atomic physics of the line list for numerical computation, c - the speed of light, *N(r)* - the particle number density at the distance r from the base of the wind, σ_B - the Stefan-Boltzmann constant, T_{eff} - the effective temperature at the photospheric radius R_*. In our work the effective temperatures, and also the terminal velocities v_∞ and the stellar masses for Galactic WR stars were taken from García-Segura et al. (García-Segura et al. , 1996), from Crowther et al. (Crowther et al. , 1995), and from (van der Hucht , 2001) (from here we are making also a selection for WR stars without companion, after their spectral type). In the expression of the force multiplier the constants k and α represent the number of scattering lines and the ratio of weak to strong lines, respectively (Abbott , 1980) and, for computational purposes, were taken from (Pauldrach et al. , 1986). The depth parameter t is defined by (v_{th} is the thermal velocity of the carbon ion and σ_{Th} the Thompson total cross section for scattering of radiation):

$$t = \sigma_{Th} v_{th} (dv/dr)^{-1}. \tag{3}$$

The radiative acceleration in the equation (2) is written in the isothermal case. A more accurate radiative acceleration takes into account a temperature distribution for a spherically grey atmosphere in radiative equilibrium (Milne-Eddington temperature distribution) (Lucy & Abbott , 1993).

In our model we don't find useful to take a *T(r)* distribution, the radiative acceleration from (2) being taken in the isothermal approximation. This is because the particles are begining to be injected in the supernova shock very close to the stellar surface, and we are interested how those particles behave, before injection, till one to three stellar radius distance in the atmosphere, and if there appears a phase space separation as function of their type.

2.2 Velocity and Acceleration Laws

In the previous section we said that in order to have a phase space dispersion we must first have: $a_O(v_O) > a_C(v_C) > a_{He}(v_{He})$.

The most popular velocity law (so-called β-law - Castor & Lamers (1979)) which describes the bulk motion of the accelerated bulk material is:

$$v(r) = v_\infty(1 - R_*/r)^\beta. \tag{4}$$

In an aproximation of a point source stars, the CAK theory predicts a velocity law with $\beta = 0.5$. When are considered also the finite disk effects, the CAK theory yields $\beta = 0.8$ (Friend & Abbott , 1986). Puls et al. (Puls , 1996) succesfully used a value of $\beta = 1$ for the prediction of OB stars mass loss. Still, observations of LPV subpeaks in WR winds, suggested much larger values of β (Robert , 1994), the spectral analysis of WR spectra with a clumped wind model (Schmutz , 1997) being consistent with these values of $\beta \simeq 4 - 8$.

The acceleration law that follows from (4), as function of the local bulk wind velocity, is:

$$a(v) = \beta\frac{v^2}{R_*}\left[\left(\frac{v}{v_\infty}\right)^{-1/2\beta} - \left(\frac{v}{v_\infty}\right)^{1/2\beta}\right]^2. \tag{5}$$

In Lépine & Moffat (1999) is illustrated the $a(r)$ and $\beta a(v)$ behavior for different values of the parameter β. Appart from other important remarks that can be made related to these representations, it can be seen that, for the same particle velocities we have a spatial separation (into blobs - Moffat et al. (1988)) and a constant $\beta a(v)$. For our purposes it is also useful to see that, for a star with a specified photospheric radius R_*, at a constant r and different β values we have a phase space dispersion due to different particle velocities.

If the blobs follow a velocity law (4) the time spent between $x_0 = r_0/R_*$ and $x = r/R_*$ is: $T = \frac{R_*}{v_\infty}\int_{x_0}^{x}\frac{dx}{(1-1/x)^\beta}$, so that most time is spent near x_0 for likely β values. For $x_0 = 1$, t is infinite (unless $\beta < 1$) for all x. Statistically speaking, this means that for the (4) β-law blobs should never be observed other than at $x = 1$. In addition, any $v(r)$ law with $v = 0$ at $x = 1$ requires infinite wind density for mass flow continuity (Brown et al. , 1995). This difficulty can be avoided by allowance for an

initial velocity $v(x=1) = v_{th} = \epsilon v_\infty$, where v_{th} is the thermal velocity of the carbon ion near the "dynamical photosphere". Then, the modified β-law becomes:

$$v(r) = v_\infty[\epsilon + (1-\epsilon)(1 - R_*/r)^\beta]. \quad (6)$$

At the same ϵ values the acceleration in the first two stellar atmosphere radius distances is more efficient for small β values. The conclusions that follows from Lépine & Moffat (1999) and Brown et al. (1995) is that we can have a phase space separation for different atomic species that are present in the WR wind and if we want to have $a_O(r_O) > a_C(r_C) > a_{He}(r_{He})$ we must have first $\beta_O < \beta_C < \beta_{He}$, which means that (see Lépine & Moffat (1999)) we need that $a_O(v_O) > a_C(v_C) > a_{He}(v_{He})$ at the same velocity value.

2.3 Code Equations

For a phase space separation we take the factor f from the equation (1) equal with the ratio of accelerations (see eq. (5)):

$$\frac{a_C(v_C)}{a_{He}(v_{He})} = \frac{\beta_C}{\beta_{He}} \left(\frac{v_C}{v_{He}}\right)^2 \left[\frac{\left(\frac{v_C}{v_\infty}\right)^{-\frac{1}{2}\beta_C} - \left(\frac{v_C}{v_\infty}\right)^{\frac{1}{2}\beta_C}}{\left(\frac{v_{He}}{v_\infty}\right)^{-\frac{1}{2}\beta_{He}} - \left(\frac{v_{He}}{v_\infty}\right)^{\frac{1}{2}\beta_{He}}}\right] = \left(\frac{Z_C}{Z_{He}}\right)^{k_C} \simeq 3^{k_C}, \quad (7)$$

$$\frac{a_O(v_O)}{a_{He}(v_{He})} = \frac{\beta_O}{\beta_{He}} \left(\frac{v_O}{v_{He}}\right)^2 \left[\frac{\left(\frac{v_O}{v_\infty}\right)^{-\frac{1}{2}\beta_O} - \left(\frac{v_O}{v_\infty}\right)^{\frac{1}{2}\beta_O}}{\left(\frac{v_{He}}{v_\infty}\right)^{-\frac{1}{2}\beta_{He}} - \left(\frac{v_{He}}{v_\infty}\right)^{\frac{1}{2}\beta_{He}}}\right] = \left(\frac{Z_O}{Z_{He}}\right)^{k_O} \simeq 4^{k_O}, \quad (8)$$

where $k_C = 0.764 \pm 0.022$ and $k_O = 0.984 \pm 0.0265$ (see table 1) and the He velocity is a (6) β-law:

$$v_{He}(r_{He}) = v_\infty[\epsilon + (1-\epsilon)(1 - R_*/r_{He})^{\beta_{He}}]. \quad (9)$$

In (7), (8), and (9), β will take values between 4 and 8 (Schmutz , 1997) (with the specification that $\beta_O < \beta_C < \beta_{He}$). The photospheric radius will be determined from the relation:

$$R_* = \sqrt{L_{nuc}/(4\pi\sigma_B T_{eff})},$$

with T_{eff} the effective photospheric temperature, σ_B - the Stefan-Boltzmann constant, and L_{nuc} - the nuclear luminosity of the star, given by an empirical form of the mass-luminosity relation (Langer , 1989a), with the fitting coefficients found by Heger et al. (Heger & Langer , 1996).

(7), (8), and (9) will be the first three equations from a system of seven with seven variables. We will try to prove that at any given r, $30M_{\odot} \leq M \leq 50M_{\odot}$, ϵ (from the non-zero initial velocity condition in (6)) we are having a phase space dispersion, meaning a convergence for the system of equations. The system variables are:

- The distances at which the particles reach in the wind before their injection in the supernova shock. Assuming that C, O and He (the choice of these atomic species will be later explained) are becoming wind particles at the same moment of time and that are simultaneously injected in the supernova shock. The last approximation is possible because the bulk velocity in the wind is in the best case 0.1 from the shock velocity. So, the quantities that are to be determined are $(\Delta r)_1 \equiv a = r_C - r_{He}$ and $(\Delta r)_2 \equiv b = r_O - r_{He}$ (in (9) will be made the substitution $r_{He} = r_C - a$), where $r_C \equiv r$ will be the distance at which is reaching the carbon in the wind and it will be given (through an iteration loop) values from $1R_*$ to $3R_*$ with the step smaller than the Alfvénic damping lengths. The Alfvénic damping lengths can be understood as limits to the size of the formed blobs (self-similar domains) through thermal instability due to Alfvén waves (Gonçalves et al. , 1998). Gonçalves et al. found that the blob diameters must be $3 \times 10^5 \leq d_{blob} \leq 1.4 \times 10^6$ cm. Consequently, our correlation length (r_0) and our iteration step ($r_C \rightarrow r_C + \Delta r_C$) will be chosen in the interval defined by $r_0 \equiv d_{blob}/2$, and $1.5 \times 10^5 \leq \Delta r_C \leq 7 \times 10^5$ cm, respectively;
- particle densities, $N_O(r_O)$ and $N_C(r_C)$ in the volume of the sample of radius $R_{S,i}$. The sample volume will be considered spherical and its radius equal with the distance from the base of the wind at which the particle i (C, O or He) is injected in the supernova shock, $r_i \equiv R_{S,i}$. Also, the correlation length $r_0 < R_{S,i}$;
- The velocities $v_O(r_O)$, $v_C(r_C)$ and $v_{He}(r_{He})$.

Our "reduced" depth parameter will be $\tilde{t} = t(dv/dr)$, with t from (3). So, the wind equation in the assumptions made by Biermann et al. (Biermann & Cassinelli , 1993), after doing the integration, will look like:

$$y^2 - (1/y) = -\tilde{g}_{rad}(\ln x - x \ln x + x). \tag{10}$$

In the equation (10) the dimensionless value of $x \equiv r/[\mathrm{cm}]$. Re-writting (10) for i element (O, C, or He) at $x = r_{in}$ and $x = r_{fin}$ ($r_{fin} = r_{in} + \Delta r$; $\Delta r \leq r_0 = 7 \times 10^5$) distances, corresponding to $y_{in} \equiv v_{i,\ in}$ and $y_{fin} \equiv v_{i,\ fin}$, respectively, and making their substraction we get:

$$v_{i,\ fin}^2 - v_{i,\ in}^2 - \frac{1}{v_{i,\ fin}} + \frac{1}{v_{i,\ in}} = -\tilde{g}_{i,\ rad}(\ln r_{i,\ fin} - \ln r_{i,\ in} - r_{i,\ fin} \ln r_{i,\ fin} + \\ + r_{i,\ in} \ln r_{i,\ in} + r_{i,\ fin} - r_{i,\ in}). \tag{11}$$

Let now consider the case of two WR stars of equal masses, terminal velocities, and effective temperatures, one in WC pre-supernova sequence and the other in WO pre-supernova sequence. For this purpose we are using real stellar data from (van der Hucht , 2001; García-Segura et al. , 1996; Crowther et al. , 1995). We suppose that in the wind of the WR star of WC spectral type is no O, but just He and C, while in the wind of the WR star of WO spectral type is no C (just He and O). This is allowed knowing that a WR star passes through both, WC and WO sequences, excepting the case when the star explodes as supernova before that, but even so, due to the fact that we are observing in CRs the contribution of all (and is no reason to think that the number of WR stars which explode in WC sequence is larger than for the ones that explode in WN sequence) WR stars as a unique spectral slope, this is allowed.

At a r_C for the WC sequence star we compute the velocity $v_C(r_C)$ (eq. (11)). With this velocity we "go" to the WO sequence star ($v'_O(r'_O) = v_C(r_C)$) and we compute r'_O. Then, having r'_O, we are "coming back" at the WC sequence star ($r_O = r'_O$) and we compute the velocity $v_O(r_O)$. In this way, we will be able to write the equations for a WR star in the intermediary sequence WC/WO, in whose wind we find He, C and, also, O:

$$v_O^2 - v_{O,in}^2 - \frac{1}{v_O} + \frac{1}{v_{O,in}} = -\tilde{g}_{O,rad}[\ln(r_C + b) - \ln r_{O,in} - (r_C + b)\ln(r_C + b) +$$

$$+ r_{O,in} \ln r_{O,in} + (r_C + b) - r_{O,in}], \tag{12}$$

$$v_C^2 - v_{C,in}^2 - \frac{1}{v_C} + \frac{1}{v_{C,in}} = -\tilde{g}_{C,rad}[\ln r_C - \ln r_{C,in} - r_C \ln r_C + r_{C,in} \ln r_{C,in} + r_C - r_{C,in}], \tag{13}$$

$$v_{He}^2 - v_{He,in}^2 - \frac{1}{v_{He}} + \frac{1}{v_{He,in}} = -\tilde{g}_{He,rad}[\ln(r_C - a) - \ln r_{He,in} - (r_C - a)\ln(r_C - a) +$$

$$r_{He,in} \ln r_{He,in} + (r_C - a) - r_{He,in}], \tag{14}$$

where we consider that all the wind particles are originating at the stellar surface, $r_{O,in} = r_{C,in} = r_{He,in} = 1R_*$, and $r_{He} = r_C - a$ and $r_O = r_C + b$. These are the fourth, the fifth, and the sixth equations of our system of equations.

In the equation (13) for the WC sequence WR star (see also (2)and (3)):

$$\tilde{g}_{C,rad} = \frac{1}{c(N_C(r_C) + N_{He}(r_{He}))} \sigma_{Th} \sigma_B T_{eff}^4 [k(\sigma_{Th} v_{th})^\alpha].$$

Also, in (12) (for the WO sequence WR star):

$$\tilde{g}_{O,rad} = \frac{1}{c(N_O(r_O) + N_{He}(r_{He}))} \sigma_{Th} \sigma_B T_{eff}^4 [k(\sigma_{Th} v_{th})^\alpha],$$

and, in (14) (for the WC/WO sequence WR star):

$$\tilde{g}_{He,rad} = \frac{1}{c(N_O(r_O) + N_C(r_C) + N_{He}(r_{He}))}\sigma_{Th}\sigma_B T_{eff}^4[k(\sigma_{Th}v_{th})^\alpha].$$

In the approximation that all the WR wind particles are at least one time ionized, let us define a self-similar particle distribution (the ionic density in the in blobs). For this we must observe that the structure function (Farge , 1992; Lépine & Moffat , 1999; Muzy et al. , 1993) study of wavelet images of the WR surounding nebula is showing that there we have self-similar structures (Grosdidier et al. , 2001; Schertzer & Lovejoy , 1987) and that the power law of the energy dissipation is that of a compressible turbulence (Grosdidier et al. , 2001; Moffat , 1994; Sylvestre et al. , 1999). The distribution of the number of structures (blobs), N, with the number of free electrons in a blob, N_e, reads (Richardson et al. , 1996):

$$\frac{dN}{dN_e} \simeq \frac{\Delta N}{\Delta N_e} = \frac{(1-\gamma)N_0}{N_{e,max}^{1-\gamma} - N_{e,min}^{1-\gamma}} N_e^{1-\gamma},$$

with N_0 - the total number of blobs at any time, $N_{e,min}$ and $N_{e,max}$ - the minimum and, respectively, the maximum number of free electrons contained in any blob: $N_{e,max} = 10N_{e,min}$, $N_{e,max} = 2 \times 10^{46}$ (Richardson et al. , 1996). Also, $\gamma = n - D$ (n - the space dimensionality = 3; D - the fractal dimension = 2.32).

In our case:

$$N_C(r_C) - N_{He}(r_{He}) = \frac{(1-\gamma)(\bar{N}_{0,C}(r_C) + \bar{N}_{0,He}(r_{He}))}{N_{e,max}^{1-\gamma} - N_{e,min}^{1-\gamma}}[N_e(r_C) - N_e(r_{He})]\bar{N}_e^{-\gamma}, \quad (15)$$

$$N_O(r_O) - N_{He}(r_{He}) = \frac{(1-\gamma)(\bar{N}_{0,O}(r_O) + \bar{N}_{0,He}(r_{He}))}{N_{e,max}^{1-\gamma} - N_{e,min}^{1-\gamma}}[N_e(r_O) - N_e(r_{He})]\bar{N}_e^{-\gamma}, \quad (16)$$

where the equations (15) and (16) are for our WR star in WC sequence and, respectively, in WO sequence, when we consider that in the wind the spatial separation of C from He (WC case) and of O from He (WO case) is giving an self-similar particle distribution in blobs (each particle type is forming its own self-similar like blob structures). In the right side of (15) and (16), $\bar{N}_{0,C}(r_C)$, $\bar{N}_{0,}(r_C)$, and $\bar{N}_{0,C}(r_C)$ are the average C, O and He densities, and we can find them from the matter conservation law: $\dot{M}_i = 4\pi\bar{\rho}_i(r_i)r_i^2 v_{esc}$.

Then, the average density $\bar{\rho}_i(r_i)$ (in g/cm^3) for i wind element over the sample (blob) volume: $\bar{\rho}_i(r_i) = \dot{M}_i/(4\pi r_i^2 v_{esc})$, and if we consider that the total mass loss, $\dot{M}(X_j)$ ($j \in [1, i]$) of each stellar surface element isotope (Woosley et al. , 1995) (assuming that the surface mass fraction X_j is having a small time dependency - in the the supernova deflagration time scale, comparatively with the stellar age): $\dot{M}(X_j) = X_{j,surf}(t)\dot{M} = X_{j,surf}\dot{M}$, results that:

$$\bar{\rho}_i(r_i) = \dot{M}X_i/(4\pi r_i^2 v_{esc}), \quad (17)$$

In (17), $X_i = \sum_{j=1}^{i} X_j$ and the mass loss for Wolf-Rayet hydrogen depleted WC/WO pre-supernova stars with initial masses between 30-35 $M_\odot$ and 50 $M_\odot$ is described through the empirical formula given in (Langer , 1989b) and (Woosley et al. , 1993).

The X_j surface mass fractions are the Langer & Henkel (1995) stellar evolution computation results.

Since the average particle number density:

$$\bar{N}_{0,i}(r_i) = \bar{\rho}_i(r_i)/\bar{m} = \bar{\rho}_i(r_i)/(\mu m_H),$$

where the mean molecular weight $\mu \equiv \bar{m}/m_H$ (with $\bar{m}$, the average mass of the gas particle, m_H, the mass of a hydrogen atom) and $\mu_{ion} < \mu < \mu_{neu}$ (μ_{ion} and μ_{neu} are the mean molecular weights for the completely ionized and for completely neutral gas, respectively) (Carroll & Ostlie , 1996), we will be able to write, with the help of equation (17):

$$\bar{N}_{0,i}(r_i) = \dot{M} X_i/(4\pi r_i^2 v_{esc} \mu m_H). \tag{18}$$

Also, in (15) and (16):

$$\bar{N}_{0,i}(r_i) + \bar{N}_{0,He}(r_{He}) = [(\bar{N}_{0,i}(r_i)/\bar{N}_{0,He}(r_{He})) + 1]\bar{N}_{0,He}(r_{He}).$$

Observing from (18) that $\bar{N}_{0,i}(r_i)/\bar{N}_{0,He}(r_{He}) = (X_i/X_{He}) \cdot (m_{He}/m_i)$, we find:

$$\bar{N}_{0,i}(r_i) + \bar{N}_{0,He}(r_{He}) = [(X_i/X_{He}) \cdot (m_{He}/m_i) + 1]\bar{N}_{0,He}(r_{He}), \tag{19}$$

where the values of X_i are the sum of the isotopic mass fraction for C or O, and m_i and m_{He} are the atomic masses of i element and the atomic mass of He, respectively, expressed in grams.

Now, if we consider that the free electrons in the WR wind are catched in the same instability driven turbulent general motion as the ions and that they have a self-similar distribution in the blobs and a random distribution between blobs, we have:

$$N_e(r_i) = \bar{N}_e(f_{i,e}^2 A_{i,bb} r_i^{-\gamma} + 1 - f_{i,e}^2) \ , \ f_{i,e} < 1 \tag{20}$$

with: $f_{i,e} = \frac{K_i R_{i,S}^D}{K_i R_{i,S}^D + K_i' R_{i,S}^3}$, $A_{i,bb} = \frac{A_{i,ee}}{f_{i,e}^2}$ and:

$$A_{i,ee} = \left(1 - \frac{\gamma}{3}\right) \frac{R_{i,S}^{-\gamma}}{R_{i,S}^{-2\gamma} + \left(\frac{K'}{K}\right)_i^2 + 2R_{i,S}^{-\gamma}\left(\frac{K'}{K}\right)_i} \equiv r_{i,0}^{\gamma}, \tag{21}$$

and we get a full description of the densities that enter in the right side of (15) and (16) by making the (19) and (20), and (from (18)):

$$\bar{N}_{0,He}(r_{He}) = \frac{\dot{M} X_{He}}{4\pi r_{He}^2 v_{esc} \mu m_H}.$$

substitutions. In the above equation v_{esc} is given by (Castor et al. , 1975). For $\bar{N}_e$ we take the average value between $N_{e,min}$ and $N_{e,max}$ over the volume of radius $R_{i,S}$.

After substitution, one last problem arises in (15) and (16) by the apparition of the $(K'/K)_i$ ratios. This can be solved considering that the observed blobs are having an universal fractal distribution where the fractal behavior $\zeta(p)$ for the p order (which can be regarded as a spatial scaling of the analyzed image) of the structure function. Because also the correlation length $r_{i,0}$ for the i element in blobs is (like $\zeta(p)$ but with a much smaller generality) a measure of fractality in the system and because $R_{i,S}$ is for certain a spatial scaling measure, we conclude that in three dimensional space, in our particular case:

$$\zeta(p)/p \simeq r_{i,0}/R_{i,S}. \tag{22}$$

With $r_{i,0}$ from (21) and (22) we find:

$$(K'/K)_i = R_{i,S}^{-\gamma}[-1 + \sqrt{1-(\gamma/3)} \cdot (\zeta(p)/p)^{-\gamma/2}]. \tag{23}$$

Subtracting (16) from (15) we get the seventh equation of our system of equations for a WC/WO Wolf-Rayet star without companion.

3 Results and Conclusion

Our primal purpose was to show that the system of equations constructed on the above assumptions converges to a solution for any value of the $r \equiv r_C$ in the range 1-3R_*. Our numerical code, made in IDL 5.2, was tested for convergence when we independently varied quantities like the stellar initial mass ($30M_\odot \leq M \leq 50M_\odot$; with the corresponding effective temperatures, terminal velocities, mass fractions, etc.), molecular weight ($\mu_{ion} < \mu < \mu_{neu}$), ϵ ($v_{th} = \epsilon v_\infty$; $0.01 \leq \epsilon \leq 0.1$), or iteration step $\Delta r \equiv \Delta_C$ (when $r_C \rightarrow r_C + \Delta r_C$; $1.5 \times 10^5 \leq \Delta r_C \leq 7 \times 10^5$). Our unique model free parameter, p ($0.1 < p \leq 3$), had also the not to be ignored purpose to enable a rescaling of our computational lattice in such a way that the sample radius, $R_{i,S}$, never to touch the lattice border which would had meant a disaster ("explosion") for the system variables.

Indeed, the system converged in all the above situations, which meant that we had a pre-supernova phase space dispersion between C, O, and He in the wind seen in the observed CRs through the factor $(Z_i/Z_{He})^{k_i}$.

References

Abbott, D.C. 1980, ApJ, 242, 1183

Biermann, P.L., Cassinelli, J.P. 1993, A&A, 277, 691

Brown, J.C., Richardson, L.L., Antokhin, I., et al. 1995, A&A, 295, 725

Carroll, B.W., Ostlie, D.A. 1996, An Introduction to Modern Astrophysics, Addison-Wesley Publishing Company, New-York, 323-326

Cassinelli, J.P. 1979, Ann. Rev. Astr. Ap., 17, 275

Castor, J.I., Abbott, D.C, Klein, R.I. 1975, ApJ, 195, 157

Castor, J.I., Lamers, H.G.J.L.M. 1979, ApJS, 39, 481

Crowther, P.A., Smith, L.J., Hillier, D.J., et al. 1995, A&A, 293, 427

Farge, M. 1992, Ann. Rev. Fluid Mech., 24, 395

Friend, D.B., Abbott, D.C. 1986, ApJ, 311, 701

García-Segura, G., Langer, N., Low, M.-M.M. 1996, A&A, 316, 134

Gonçalves, D.R., Jatenco-Pereira, V., Opher, R. 1998, ApJ, 501, 797

Grosdidier, Y., Moffat, A.F.J., Blais-Ouellette, S., et al. 2001, ApJ, 562, 753

Heger, A., Langer, N. 1996, A&A, 315, 428

Langer, N. 1989a, A&A, 210, 93

Langer, N. 1989b, A&A, 220, 135

Langer, N., Henkel, C. 1995, Space Sci. Rev., 74, 343

Lépine, S., Moffat, A.F.J. 1999, ApJ, 514, 909

Lucy, L.B. 1982, ApJ, 255, 278

Lucy, L.B., Abbott, D.C. 1993, ApJ, 405, 740

Moffat, A.F.J. 1994, RvMA, 7, 51M

Moffat, A.F.J., Drissen, L., Lamontagne, R., et al. 1988, ApJ, 334, 1038

Muzy, J.F., Bacry, E., Arneodo, A. 1993, Phys. Rev. E, 47, 875

Owocki, S. 1994, in Proc. Workshop on Instability and Variability of Hot-star Winds, eds. A.F.J. Moffat et al., Dordrecht: Kluwer, 1995

Pauldrach, A., Puls, J., Kudritzki, R.P. 1986, A&A, 164, 88

Popescu, A.S., Biermann, P.L., Langer,N. 2005, A&A (submitted)

Puls, J. 1996, A&A, 305, 171

Richardson, L.L., Brown, J.C., Simmons, J.F.L. 1996, A&A, 306, 519

Robert, C. 1994, Ap&SS, 221, 137

Schertzer, D., Lovejoy, S. 1987, J. Geophys. Res., 92, 9693

Schmutz, W. 1997, A&A, 321, 268

Silberberg, R., Tsao, C.H. 1990, ApJ, 352, L49

Sylvestre, J., Khoury, J., Lovejoy, S., et al. 1999, in Turbulent Bipolar Outflows in Young Stellar Objects: Multifractal Universality Classes and Generalized Scaling, (astro-ph/9907442.pdf)

Tsao, C.H., Silberberg, R., Barghouty, A.F. 1998, ApJ, 501, 920

van der Hucht 2001, 7th Catalog of Galactic Wolf-Rayet stars, VizieR On-line Data

Catalog: III/215, Originally published in: New Astronomy Reviews, 45, 135
Woosley, S.E., Langer, N., Weaver, T.A. 1993, ApJ, 411, 823
Woosley, S.E., Langer, N., Weaver, T.A. 1995, ApJ, 448, 319

PADEU | PADEU **15**, 181 (2005)
ISBN 963 463 557

New dynamical description of the rotation curves of galaxies

Sinziana Paduroiu[1] and Mircea Rusu[2]

University of Bucharest, Atomistilor 405 Ro-76900, Bucharest-Magurele, Romania

E-mail: [1]sinziana.paduroiu@gmail.com, [2]mrusu@dnt.ro

Abstract
Observations of dynamics of galaxies and plasma in the universe led us to the conclusion that there exists a huge amount of matter that is invisible, the so called "dark matter". Rotation curves of galaxies represent important evidences for introducing this concept. In this article we present the data interpretation for the rotation curve of NGC6503, introducing a new density distribution function in order to describe the behaviour at large distances from the center of the galaxy. One of the result's possible implications is a change in the classical approach regarding the matter and its movement in galaxies, as an attempt to explain the rotation curves of galaxies.
Keywords: *dark matter, Newtonian Dynamics, rotation curves, density distribution*

1 Introduction

Together with the cosmological issues regarding the actual state of our universe, observations of galaxies and clusters lead to the necessity of introducing new physics.

For example, in the observations of polar ring galaxies, measurement of the rotation velocities within the disc and within the ring results in flat curves in both cases. This implies that the velocity for large radii remains constant, if we assume a spherical mass distribution. Moreover, nearly all-luminous elliptical galaxies contain about 10^{10} $M_\odot$ of gas in the form of gas halos, with a size of at least 50 kpc. Due to the X-ray emission of this hot gas its temperature could be estimated to be about 10^7 to 10^8 K, which implies a velocity of the gas particles that lies far above the escape velocity derived from the visible mass. If this gas should be gravitationally bound, a lot more mass is needed.

One of the problems comes from the explanation of the rotation curves of the galaxies. The Newtonian gravitational theory has been expected to describe in excellent terms the dynamics in the extragalactic regime, but the estimated acceleration of stars and gas are much larger than those generated by the visible matter. The amount of light starts falling off near the edge of the galaxy, but the rotation speed doesn't, as one would expect, instead it remains almost constant and highly above the limit given by the gravitational field produced by the visible matter. This phenomena was explained by the existence of dark matter.

An attempt to explain the rotation curves of spiral galaxies could be to reconsider the invisible mass distribution around the galactic center, without introducing a new kind of matter, but an ordinary one, having a low density, but spread over a large distance from the galactic center. In the following section we will present an alternative approach computed for the particular case of NGC 6503, with the possible generalized form.

2 Galaxies dynamics interpretation

2.1 Observations of rotation curves of galaxies

Spiral galaxies usually consist of two components, a flat, large disk which often contains a lot of interstellar matter (visible sometimes as reddish diffuse emission nebulae, or as dark dust clouds) and young (open) star clusters and associations, which have emerged from them, often arranged in conspicuous and striking spiral patterns and/or bar structures, and an ellipsoidally formed bulge component, consisting of an old stellar population without interstellar matter, and often associated with globular clusters. The pattern structures in the disk are most probably transient phenomena only, caused by gravitational interaction with neighboring galaxies. A typical spiral galaxy contains 100 billion stars and measures 100 000 light-years in diameter.

Rotation curves are usually obtained by combining observations of the 21cm line with optical surface photometry. Using 21-cm emission, the circular velocities of clouds of neutral hydrogen can be measured as a function of r, the distance from the center of the galaxy. The optical rotation curves provide high spatial resolution in the visible disk, and in particular in the center, to trace central mass concentrations, while only the HI gas extend far enough in radius to trace the outer parts, where dark matter is supposed to be dominating. Using 21-cm emission, the circular velocities of clouds of neutral hydrogen can be measured as a function of r, the distance from the center of the galaxy. Observed rotation curves usually exhibit a characteristic flat behavior at large distances, out towards, and even far beyond, the edge of the visible disks. The fact that for most spiral galaxies, the dark matter is not dominant within the optical disk, comes as a result of recent observations. A non-baryonic matter would not follow the spiral instabilities in the disk. The dark matter component seems to become important only for large radii.

2.2 The data interpretation for the rotation curve of NGC6503

In almost all cases, after a rise near r=0, the velocities remain constant out as far as can be measured. In the following section we will present the data set for NGC6503. Fig. 1 shows the rotation curve for the spiral galaxy NGC6503.

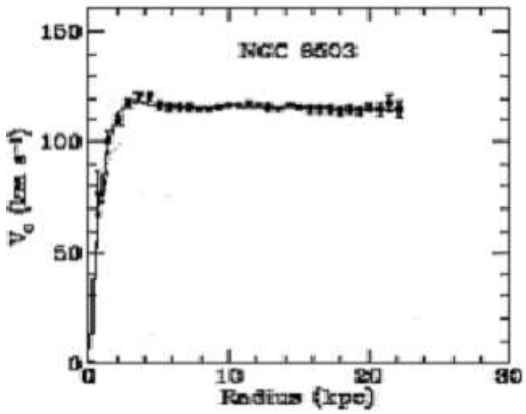


Figure 1: *Rotation curve for NGC6503. The points are circular rotation velocities as a function of distance from the center of the galaxy (Ref.2).*

In Newtonian dynamics the circular velocity is expected to be

$$v(r) = \sqrt{GM(r)/r}. \tag{1}$$

where, $M(r) \equiv 4\pi \int \rho(r) r^2 dr$, and $\rho(r)$ is the density profile, and is expected to fall $\sim 1/\sqrt{r}$ beyond the optical disc. The fact that $v(R)$ is approximately constant implies $M(r) \sim r$ and $\rho \sim 1/r^2$.

The first segment of the slope, until approximately 2.5 kpc, can be fitted as a linear dependence of the velocity with the distance and can be treated as a rigid rotator (see Fig.2). This means that until that distance, the matter within the galaxy moves as a whole.

If we assume that the stars have a circular orbit around the galactic center, the rotation velocity of a star can be computed considering the centripetal acceleration for a circular movement equal to the central force acceleration. We will consider an expression for the central force as power law dependent on the distance, with an unknown exponent, β:

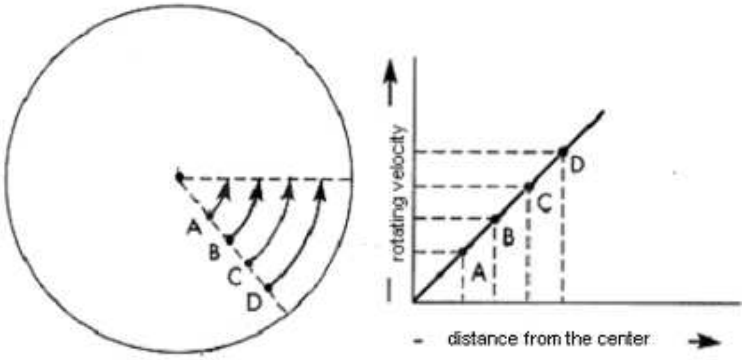


Figure 2: *Left:Rigid rotator Right:Rotation curve for a rigid rotator*

$$F_G = \frac{GmM_r}{r^\beta} = \frac{mv^2}{r} = F_z. \tag{2}$$

From Eq.(2) it follows that:

$$v(r) = \sqrt{\frac{GM_r}{r^{\beta-1}}}. \tag{3}$$

where M_r is the mass within the orbit of radius r, (in our case, within approximately 2.5 kpc). The forces lying outside the orbit compensate exactly for cylindrically and spherically distribution.

Assuming a spherically symmetric bulge with constant density ρ , then

$$M_r = \rho \cdot V_r = \rho \frac{4}{3}\pi r^3. \tag{4}$$

Because up to r_1, the innermost part of the galaxy has a rotation curve of $v(r) \sim r$, we have to choose $\beta = 2$:

$$v(r) = \sqrt{\frac{G\rho\frac{4}{3}\pi r^3}{r^{\beta-1}}} \sim \sqrt{\frac{4}{3}G\rho} \cdot r, for \beta = 2. \tag{5}$$

Considering that at distances larger then r_1, the radius of the bulge, the mass is negligible we will obtain a different dependence. If the total mass of the galaxy, $M_{gal} = M_r$, therefore:

$$v(r) = \sqrt{\frac{GM_{gal}}{r^{\beta-1}}} = \sqrt{GM_{gal}} \cdot r^{-1/2}, for \beta = 2. \tag{6}$$

For that part (between 2.5 - 3.5 kpc), we obtain the usual Keplerian motion (Fig. 3), $v(r) \sim r^{-1/2}$, for the value of 2 of the exponent.

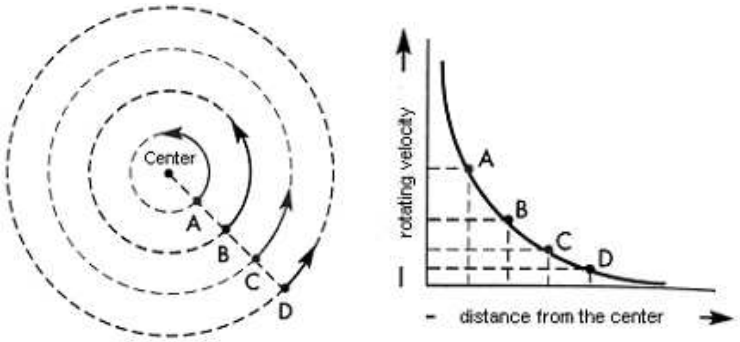


Figure 3: *Left:Planetary revolution rotation Right:Rotation curve-Keplerian dependence*

But, for large r, as shown for many of the galaxies, the value of 2 for the exponent does not give a correct velocity dependence. A choice could be to choose a new value for β, but that will be a highly unphysically solution. However, we have another possibility, i.e. to try to find if any physically reasonable variable mass density distribution would be able to give the desired velocity distribution, far form the center.

2.3 The effect of the density distribution function at large distance from the galactic center, on the velocity function

After approximately 3.5 kpc, the rotation curve tends to remain almost constant. Considering different density distributions, we can compute the rotation velocity of a body at a given distance from the center.

In Table 1 we summarize the results obtained. If we consider an exponentially decreasing density that is expressed by an exponent, α, then a general relation can be computed for the velocity function:

$$v(r) = \sqrt{\frac{1 - e^{-\alpha r} - \alpha r(1 - \frac{\alpha r}{2})e^{-\alpha r}}{r^2}}. \tag{7}$$

If we consider that function and we calculate the corrections for luminosity and mass/luminosity ratio parameters, we obtain

$$\overline{L} = 0.46 \cdot 10^{10} L_{\odot}. \tag{8}$$

Table 1: *Rotation velocities for different density distributions*

Geometry/density	Rotation velocity	
$\rho = \frac{m}{4\pi R^3/3}$	$v(r) \sim r, r < R$	Spherical geometry
	$v(r) \sim r^{-1/2}, r > R$	
$\rho = a/\sqrt{r}$	$v(r) \sim r^{3/4}, r < R$	Spherical geometry
	$v(r) \sim ct., r > R$	
$\rho = a/\sqrt{r}$	$v(r) \sim r^{3/4}, r < R$	Cylindrical geometry
	$v(r) \sim r^{-1/2}, r > R$	
$\rho = \rho_0 e^{-\alpha r}$	$v(r) = \sqrt{\frac{1-e^{-\alpha r}-\alpha r(1-\frac{\alpha r}{2})e^{-\alpha r}}{r^2}}$	Spherical geometry

and

$$\overline{M/L} = 1.66 M_\odot/L_\odot. \tag{9}$$

The results are in concordance with the predictions.

3 Conclusions

Until approximately 2.5 kpc distance from the nucleus, the galaxy moves like a rigid rotator, with a rotation curve $v(r) \sim r$. Between approximately 2.5-3.5 kpc , the galaxy has a rotation curve $v(r) \sim r^{-1/2}$. After approximately 3.5 kpc, the rotation curve may be described by the dependence from Eq.7, for a density distribution $\rho = \rho_0 e^{-\alpha r}$. The corrections of luminosity parameter and mass-to-light ratio parameter for this density function are in accordance with the observations and predictions, but at large radii, this function shows a slowly decay of the rotation velocity with the distance from the galactic center, while the observations show a rotation curve almost constant. Taking into account the difficulties in observations and radius estimations, this discrepancy is not representing an impediment.

The explanation for the rotation curve of NGC6503 that has been done considering that the large distance behaviour is due to the existence of a massive dark halo is shown in Fig.4, where the dashed and dotted curves are the contribution to the rotational velocity due to the observed disk and gas, respectively, and the dot-dash curve is the contribution from the presumed dark halo.

If we consider α as unknown parameter, we can adjust its value to fit the measured velocity function for different galaxies. If it is possible to have just one value for α ,

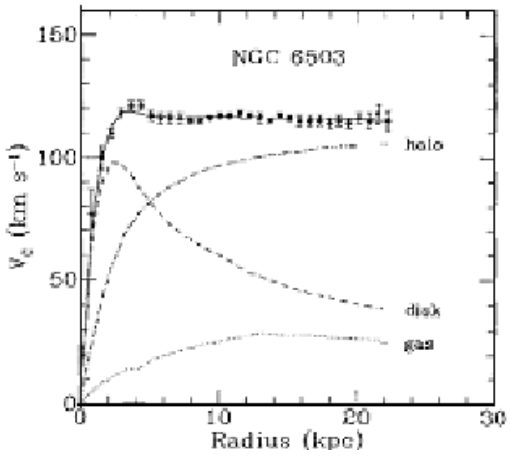


Figure 4: *Rotation curve including the contribution of a dark halo (Ref.1)*

and the whole velocity function could be fitted with enough accuracy, we can suppose that it is not necessary to consider a new kind of matter (dark matter) to describe the velocity function. It is enough to consider that extremely low density gases or dust (that are so faint and therefore below the present observational limits), exponentially radial distributed around the bulge, could give the same velocity function like the observed one. This could be a reasonable hypothesis just considering the simplest idea of gravitational accretion of dust and gases around a massive bulge. Our computation shows that it is possible to find such value for α for at least several cases we examined. A general dependence of the radial velocity function on distance for different values of the α parameter is presented in Fig.5.

We will try to analyze in details also for other galaxies this dependence, in order to find values for densities around the bulge at large distances. It will be a supplementary proof if the values are indeed so low that it is beyond today's observable limits. Also we will compute, for that values the drag effect of the dust. If all these results could be in reasonable limits, this hypothesis could be considered with much confidence.

References

Begeman, k.g., Broeils, A.H., Sanders, R.H., 1991, MNRAS 249, 523

Bekenstein, J.D.,2004, in Relativistic gravitation theory for the MOND paradigm, astro-ph/0403694 v1

Figure 5: *General dependence between radial velocity and distance, for different values of* α

Bertone, G., Hooper, D., Silk, J., 2004 in Particle Dark Matter: Evidence, Candidates and Constraints, hep-ph/0404175 v1 21

Bissantz, N., Debattista, V.P., and Ortwin G.,2004 ApJ, 601, L155-L158

Bottino A., Fornengo N., in Dark Matter and its Particle Candidates

Caroll, B.W., Ostlie, D.A.,1996 , in An introduction to modern astrophysics, Addison-Wesley Publishing Company, NY

Cline, D.B., Searching for dark matter, Scientific American, 5/2003, Univers Publishing

Combes, F., 2004, in Properties of dark matter haloes, New Astronomy Reviews 46, 755-766

Debattista, V.P., Carollo, C.M., Mayer, L., Moore, B.,2004, 604:L93-L96

Dolgov, A.D.,2004, in Problems of Vacuum Energy and Dark Energy, hep-ph/0405089 v1 11

Kaufmann, W., 1990, Discovering the Universe, Ed. W.H. Freeman and Company, NY

Kluperd, K., Zber, H., Particle Physics, 2000, I.O.P.

Padmanabhan, T.,2003, in Cosmological constant-the weight of the vacuum, Physics Reports 380 235-320

Perlmutter, S.,2004, in New Data on Dark Energy,DM 2004, Marina del Rey

Primack, J.,2004,in Precision Cosmology, DM 2004, Marina del Rey

Ramachers, Y.,2002,in WIMP direct detection overview,astro-ph/0211500 v1

Sahni, V.,2004,in Dark Matter and Dark Energy,astro-ph/0403324 v1

Sanders, R.H., McGaugh, S.S.,2002, Modified Newtonian Dynamics as an Alternative to Dark Matter, Annu. Rev. Astron. Astrophys, 40:263-317

Spiro, M., Desperately Seeking Dark Matter, CERN, 1995-1996

Treu, T.,2004,in The Dark Matter Halo of Spheroids and the Hubble constant, Marina del Rey

PADEU | PADEU **15**, 191 (2005)
ISBN 963 463 557

Embedding Diagrams for Black Holes

Paul A. Blaga

"Babeş-Bolyai" University, Faculty of Mathematics and Computer Science, 1, Kogălniceanu Street, 3400 Cluj-Napoca, Romania

E-mail: pablaga@cs.ubbcluj.ro

Abstract
We discuss two types of embedding for some submanifolds of a spacetime containing a black hole into an Euclidean space and a Minkowski spacetime, respectively. We comment on their meaning and usefulness and on the perspectives that the second one opens for further investigations.
Keywords: *black holes; embedding diagrams, geodesics*

1 Introduction

The embedding diagrams for black holes are almost as old as the notion of black hole. Although Oppenheimer and Snyder shown already in the late 30th that if a spherically symmetric star is massive enough, its gravitational collapse cannot be stopped by anything and a singularity of spacetime will be created, the rigorous notion of a black hole was introduced only in the sixtieth by people like Penrose and Hawking who were able to show that such singularities might exist without making assumption of symmetry and showing that all of them have in common the fact that they are surrounded by a special surface, called *event horizon*. The name of *black hole* was coined, as far as we know, by John Wheeler, in the late sixtieth and he was, also, one of the first to consider embedding diagrams, made popular by the classical book Misner et al. (1971). In the classical embedding diagrams, usually, the equatorial plane of the spacetime containing a black hole is embedded into the three dimensional Euclidean space. Quite recently, Donald Marolf considered another kind of diagram, in which another piece of the black hole spacetime is embedded into the three dimensional Minkowski spacetime. In many respects, this embedding is more useful, because it offers more information on the physical peculiarities of the spacetime. The aim of this

paper is to review the general notion of embedding and then to shortly discuss the two particular classes we mentioned already.

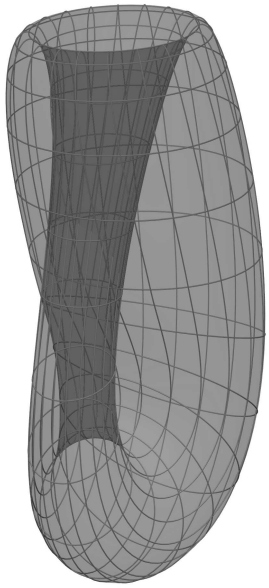
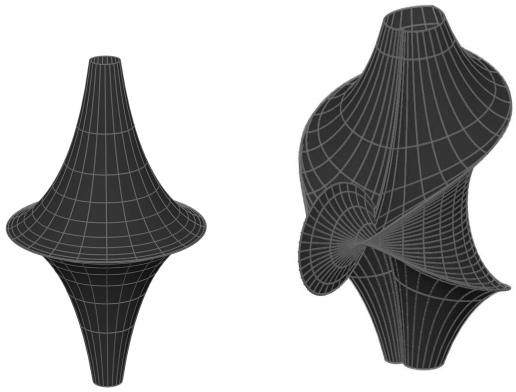


Figure 1: *The Klein's bottle*

2 Embeddings: what and why?

The spacetime containing a black hole, as most of the spacetimes considered in classical general relativity, is four-dimensional. On the other hand, we are living into three dimensional space. Therefore, we don't have an intuitive picture of the entire spacetime. What we can do, nevertheless, is to take lower dimensional pieces of spacetime (in our case they will always be two-dimensional) and represent them as subsets (surfaces in our case) of the three-dimensional Euclidean space we are living in. As a result, we should get subsets of a very particular form. In the case of surfaces, we should have no self-intersections or singular points (for instance corners or edges). A counterexample is the cone, which is not a smooth surface (unless we remove its vertex). In most papers dealing with the problem of embedding diagrams, the existence of embedding is taken as granted. However, this is far from reality. An arbitrary two-dimensional surface cannot be embedded into the Euclidean three-space. For instance, the widely

known *bottle of Klein* (see figure 1) cannot be embedded into $\mathbb{R}^3$. As we can see, it has self-intersections. In fact, the Klein's bottle is defined, initially, by factorization, by identifying some subset of $\mathbb{R}^3$ and it can be embedded only in Euclidean spaces of dimension at least four. All we can claim, in the general situation, is that it can



(a) The pseudosphere (b) The surface of Kuen

Figure 2: *Surfaces of constant negative curvature*

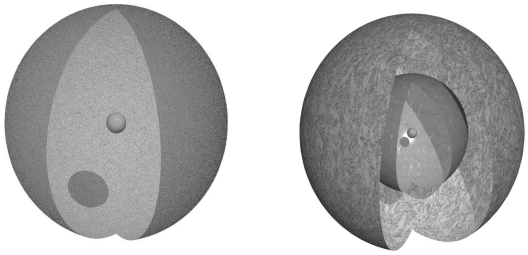
be embedded into $\mathbb{R}^4$. But this is only half of the story. The pieces of spacetime we consider come equipped with a metric, which is induced by the metric of the ambient spacetime. What we would like is to get an embedding that preserves the metrics. In other words, if we put on the embedded surface (provided there is one) the metric induced from the ambient Euclidean or Minkowski space, the embedding should be an isometry. But, as a modern version of a celebrated theorem of John Nash claims, a two-dimensional surface can only be embedded isometrically into $\mathbb{R}^6$ or, if it is endowed with an undefined metric, into the Minkowski spacetime of dimension $6 + 1$. Thus, the existence of embedding into the three-dimensional Euclidean space or into the (2+1)-dimensional Minkowski spacetime seems to be rather the exception than the rule. It should be emphasized, also, that the isometries we are speaking about are, usually, local. Two surfaces can have the same coefficients of the first fundamental form, but their shape can be different. For instance, the surfaces having the same constant curvature are locally isometric, but as the two surfaces from the figure 2 suggest, their shape might be very different.

3 Spherically symmetric black holes

Generally speaking, as we said before, a black hole spacetime is characterized by the existence of a singularity (a point at which at least some of the components of the curvature tensor become infinite), surrounded by an *event horizon*, i.e. a closed surface which has the property that no information (not even light) can escape from its interior towards the infinity. Some black holes (for instance the Reissner-Nordström black hole, or the axially symmetric black hole), may have more than one horizon, although they have, usually, different characteristics (see figure 3). Black holes spacetimes do exist in theory of gravitation different from Einstein's, but we shall confine to this in this paper. As such, the metric of a black hole spacetime should be a solution of *Einstein's field equations*:

$$R_{ij} - \frac{1}{2} g_{ij} R = k T_{ij},$$

where R_{ij} is the Ricci tensor, g_{ij} is the metric, R is the scalar curvature, k is a constant (*Einstein's gravitation constant)* and T_{ij} is the energy-momentum tensor, describing the matter content of the spacetime. We will be interested, in particular, only in



(a) The Schwarzschild BH (b) The Reissner-Nordström BH

Figure 3: *Spherically symmetric Black holes*

two spherically symmetric black hole solution of the Einstein's equations. The first one is the simplest one, the so-called *Schwarzschild solution*, depending on a single parameter, the mass of the body producing the black hole:

$$ds^2 = -\left(1 - \frac{2M}{r}\right) dt^2 + \frac{dr^2}{1 - \frac{2M}{r}} + r^2 d\theta^2 + r^2 \sin^2 \theta d\varphi^2, \tag{1}$$

and the Reissner-Nordström solution, depending on two parameters, the mass and the charge of the body:

$$ds^2 = -\left(1 - \frac{2M}{r} + \frac{Q^2}{r^2}\right) dt^2 + \frac{dr^2}{1 - \frac{2M}{r} + \frac{Q^2}{r^2}} + r^2 d\theta^2 + r^2 \sin^2 \theta . d\varphi^2, \qquad (2)$$

We have to mention, however, that the equations we just mentioned actually describe the metric in the *exterior* of the black hole or, to be more precise, in the exterior of the event horizon of the black hole. Nevertheless, these metrics have analytical extensions which are valid also on the other side of the horizon. The spacetime obtained through the extension for the case of the Schwarzschild metric is called the *Kruskal spacetime*, after the name of the mathematician who obtained this extension, in the fifties. What it is, usually, embedded, however, is exactly a slice of the exterior part of the black hole, therefore we shall not discuss these extensions.

4 Classical embeddings

Before the Marolf work, what was embedded was a spacelike slice of a black hole spacetime. We shall exemplify on the particular case of the Schwarzschild spacetime. The spacelike part of the spacetime is obtained, in this particular case, by just letting $t = const.$ We get, thus, a three-dimensional Riemannian space with the metric given by:

$$ds^2 = \frac{dr^2}{1 - \frac{2M}{r}} + r^2 d\theta^2 + r^2 \sin^2 \theta d\varphi^2. \qquad (3)$$

This is, however, impossible to visualize, therefore we shall content to embed into the Euclidean space $\mathbb{R}^3$ the "equatorial plane", i.e. we let, also, $\theta = const(= \pi/2)$. We are, left, thus, with a two-dimensional submanifold of the original Schwarzschild spacetime, with the positively-defined metric

$$ds^2 = \frac{dr^2}{1 - \frac{2M}{r}} + r^2 d\varphi^2. \qquad (4)$$

To embed this submanifold into the Euclidean space means, in fact, to find a two-dimensional submanifold of $\mathbb{R}^3$ (in other word, a surface in the intuitive space) such that the metric induced on this surface by the metric of the Euclidean space, i.e. its first fundamental form to be exactly the metric (4). As shown in Misner et al. (1971), such a surface can be described, in cylindrical coordinates, by

$$z = \sqrt{8M(r - 2M)} \qquad (5)$$

The plot can be seen in the figure 4. The event horizon corresponds to the bottom of the diagram (as it corresponds to $r = 2M$). We will call this kind of embedding, when a spacelike slice of the spacetime is embedded, a *classical embedding diagram*. This

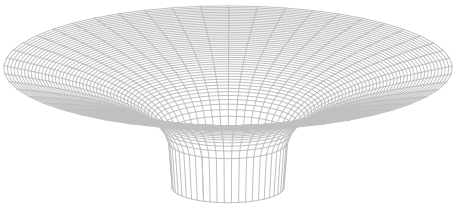
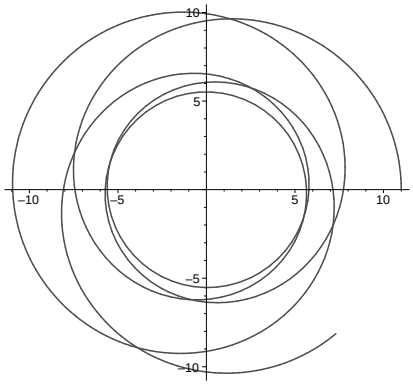
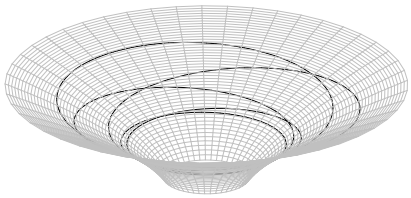


Figure 4: *An embedding diagram for the Schwarzschild black hole for* $M = 1$

kind of diagram is quite useful for the visualization of some phenomena. For instance, in most books of relativity, a picture of an equatorial geodesic in the Schwarzschild spacetime typically looks like that in the figure 5(a) while, instead, it should be



(a) The plane version



(b) On the embedded surface

Figure 5: *A geodesic around a black hole*

viewed as a geodesic on the embedded equatorial plane, i.e. it should look like in the figure 5(b). Thus, the classical embedding diagram does help the intuition and makes things clearer. However, it emphasizes only the curvature of the *space* instead of emphasizing the curvature of the *spacetime*. In other words, generally speaking, what has physical significance is the curvature of the spacetime rather than the curvature of the spacelike section. For instance, the four-dimensional Minkowski spacetime has zero curvature. Nevertheless, we can choose coordinates in such a way that the section $t = const$ are *curved* three-dimensional Riemannian spaces. Still, this has nothing to do with physics, it just reflects a particular choice of coordinates. The four-dimensional spacetime curvature, instead, is either different from zero in any coordinate system, either zero in any coordinate system. It would be useful to have, therefore, also a

way of visualizing two-dimensional slices for which the induced metric is non-defined. Clearly, such slices can only be embedded into a Minkowski spacetime. It is exactly what the Marolf's embedding diagrams are doing and we shall dedicate the next section to them.

5 Marolf's embedding diagrams

The new kind of embedding were introduced in Marolf (1999) and discussed in more details in Giblin et al. (2004). As we said previously, the idea is two embed a 2-dimensional submanifold with a Lorentzian metric into the (2+1)-dimensional Minkowski spacetime rather than into the Euclidean space. We mention that, as is the case with the classical embeddings, the construction of a Marolf embedding is quite delicate and it is only possible to be done in special situation, for instance when the spacetime has spherical symmetry and, moreover, it is static. The black hole having this properties are endowed with a metric of the form

$$ds^2 = -\phi dt^2 + \phi^{-1} dr^2 + r^2 d\theta^2 + r^2 \sin^2\theta d\varphi^2,$$

where ϕ is a function depending only on r, due to the spherical symmetry and to the static character of the spacetime. What is intended is to embed the (t, r)-part of the spacetime, with the metric

$$ds^2 = -\phi dt^2 + \phi^{-1} dr^2.$$

This should be embedded into the (2+1) Minkowski spacetime, with the metric given by

$$ds^2 = -dT^2 + dX^2 + dY^2.$$

In fact, there are some technicalities that we are not going to discuss here (see Giblin et al. (2004)), related to the fact that, usually, we cannot use the same formulae to embed the entire submanifold, therefore we divide the Minkowski spacetime into several regions and then embed different pieces of the submanifold in different region and then we "past" them together to get the overall picture. The trick is, again, to use cylindrical coordinates, but this time they are *hyperbolic* (as the metric of the surface is Lorentzian, rather than Riemannian). In the figure 6 we represent the diagram obtained for the Schwarzschild spacetime (Marolf (1999)). As one can see, it is very different from the classical embedding diagram. In particular, one might have difficulties to locate the event horizon on this diagram which is a smooth surface. It turns out that the horizon correspond to $Y = 0$, while the singularity $r = 0$ is inside the horizon, corresponding to infinite values of T (see Marolf (1999) for the argumentation). The Marolf's diagram is useful because it emphasizes the curvature of spacetime and, also, because on this diagram one can represent the wordlines of particles, instead of geodesic corresponding to constant values of the time coordinate, as is the case for the classical embedding diagrams.

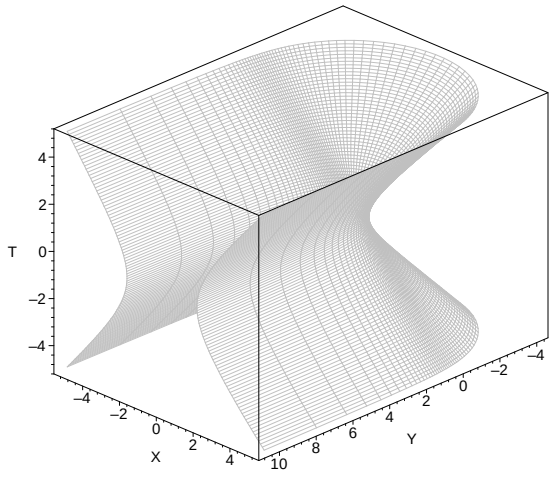


Figure 6: *The Marolf's diagram for the Schwarzschild black hole*

For the Reissner-Nordström black hole (figure 7, Giblin et al. (2004)) the situation is more complicated, because of the presence of two event horizons, corresponding to the two solutions of the equation $\phi(r) = 0$. It turns out that only for the part of the 2-submanifold lying outside the exterior event horizon the embedding is possible and this is the one appearing in the figure.

6 Final notes and perspectives

The embedding diagrams are very useful tools both for teaching general relativity and for a better understanding of different aspects of the geometry and physics of black holes. In particular, the Marolf's diagram should provide a lot of insight. Much remain to be done in this respect. In particular, it would be nice to have, also, such diagrams for relativistic stars and to attempt to study their evolution during the gravitational collapse. A detailed study of the geodesics on these surface is also something that has to be done. Some of these problems will be touched in Blaga (2005).

References

Blaga, P.A., in preparation

Giblin Jr., J.T., Marolf, D., Garvey, R., 2004 *Gen. Rel. Grav*, 36, 83

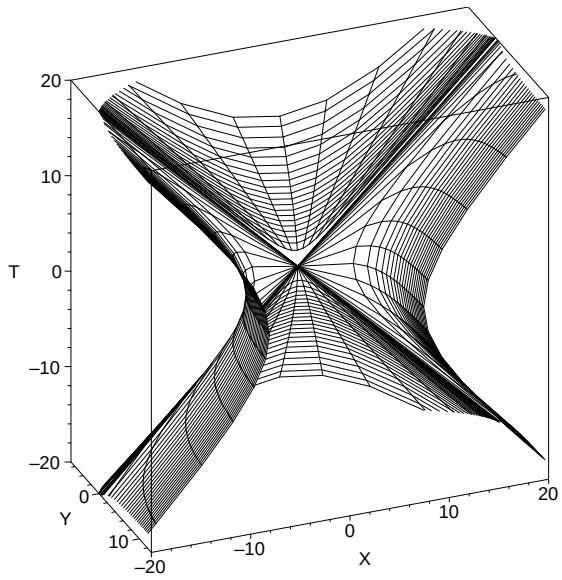


Figure 7: *The Marolf's diagram for the Reissner-Nordström black hole*

Marolf, D., 1999, *Gen. Rel. Grav*, 31, 919

Misner, C., Thorne, K., Wheeler, J., *Gravitation* (W.H. Freeman and Co., New York, 1971)

PADEU | PADEU **15**, 201 (2005)
ISBN 963 463 557

Small radial perturbations of magnetic polytropes

Cristina Blaga

Babeş-Bolyai University, Faculty of Mathematics and Computer Sciences, Department of Astronomy and Mechanics, Cluj-Napoca, Kogălniceanu 1, Romania

E-mail: cpblaga@math.ubbcluj.ro

Abstract
The paper is devoted to a qualitative analysis of the nonlinear, radial oscillations of magnetic polytropes. The magnetic field is assumed to be purely toroidal. The small adiabatic perturbations are investigated using the normal forms method. The nonadiabatic effects, described with the aid of two additional terms, related to the sources of the energy and energy damping, respectively, are analyzed with the aid of the dynamical systems theory. The numerical examples confirm and complete the qualitative investigation.
Keywords: *Stars: polytropes, magnetic field, stability*

1 Introduction

In this paper we study the stability of the polytropic stars in a *weak toroidal magnetic field*. To investigate the radial oscillations of a star we use the radial approximation of the Lorentz force proposed by Monaghan (1968). The study of the nonlinear radial *adiabatic* oscillations of magnetic polytropes is made through the *normal forms method*. The same problem was investigated using the multiple scales method by Das et al (1994). The nonadiabatic effects are described using additional terms connected to the energy production and loss. Their influence on the nonlinear radial pulsation of magnetic polytropes is investigated with the aid of *dynamical systems theory*.

2 Small perturbations of magnetic polytropes

The influence of a weak magnetic field on stellar oscillations can be obtained using a perturbative method, so we consider the Lagrangean perturbation of the hydromagnetic equilibrium equation:

$$\nabla P - \rho\nabla\Phi + \frac{1}{8\pi}\nabla B^2 - \frac{1}{4\pi}\left(\vec{B}\cdot\nabla\right)\vec{B} = 0. \tag{1}$$

where P is the pressure, Φ – the gravitational potential and B – the induction of the magnetic field. After long, but straightforward computations we obtain the equation of small oscillations of a magnetic star:

$$\rho\partial_t^2\xi_i = L_{ij}\xi_j, \tag{2}$$

where the right hand side of the equation is

$$\begin{aligned} L_{ij}\xi_j &= \nabla_i(\Gamma_1 P\nabla_j\xi_j) - (\nabla_j\xi_j)\nabla_i P + (\nabla_i\xi_j)\nabla_j P + \rho\xi_j\nabla_j\nabla_i\Phi + \rho\nabla_i\delta\Phi \\ &- \frac{1}{8\pi}(\nabla_j\xi_j)\nabla_i B^2 - \frac{1}{4\pi}\delta B_k\left(\frac{\partial B_k}{\partial x_i} - \frac{\partial B_i}{\partial x_k}\right) - \frac{1}{4\pi}B_k\left(\frac{\partial\delta B_k}{\partial x_i} - \frac{\partial\delta B_i}{\partial x_k}\right) \\ &- \frac{1}{8\pi}\nabla_i(\xi_k\nabla_k B^2) + \frac{1}{4\pi}\left(\xi_j\nabla_j B_k\nabla_k B_i + \xi_j B_k\nabla_j\nabla_k B_i\right). \end{aligned} \tag{3}$$

To obtain the normal modes equation we consider the peculiar solution of the former equation $\xi(r,t) = \xi(r)e^{i\sigma t}$ (Anand and Kushwaha, 1962).

3 Radial oscillations of a polytropic star in a toroidal magnetic field

The magnetic field destroys the spherical symmetry of the star, but in the first approximation to investigate the problem of stellar oscillation in a weak magnetic field, we can use the radial approximation of the Lorentz force (Monaghan, 1968) which is

$$\int_{-1}^{1}\int_{0}^{2\pi} F(r)d\nu d\varphi = \int_{-1}^{1}\int_{0}^{2\pi}\left(-\frac{1}{8\pi}\nabla B^2 + \frac{(\vec{B}\cdot\nabla)\vec{B}}{4\pi}\right)d\nu d\varphi \tag{4}$$

where $\nu = \cos\theta$, $F(r)$ is the radial approximation of the Lorentz force and B – the induction of the magnetic field.

Assuming that the magnetic field of the star is purely toroidal and that the nonzero component of the induction of the magnetic field is $B_\varphi = \Gamma\rho r\sin\theta$ (Roxburgh, 1966), the equation of radial adiabatic pulsations becomes:

$$\begin{aligned} \rho_0 a\ddot{r}_1 &= -(1+r_1)^2\left[P_0(1+r_1)^{-3\gamma}(1+ar_1'+r_1)^{-\gamma}\right]' - \rho_0 g_0(1+r_1)^{-2} \\ &- \frac{\Gamma^2\rho_0}{6\pi}(1+ar_1'+r_1)^{-1}\left[\rho_0 a^2(1+ar_1'+r_1)^{-1}\right]', \end{aligned} \tag{5}$$

where the dot stands for the time derivative, the prime – for the derivative with respect to the radial variable (here the radius of the unperturbed configuration a), zero index emphasizes that it is considered the value of the function in the unperturbed configuration. The distance to the stars center in the perturbed case, denoted by r, is

$$r = a(1 + r_1) \tag{6}$$

where a is distance to stars center in equilibrium and r_1, the adimensional radius, is the difference between the distances to star's center in unperturbed stellar radius.

Comparing this equation to that derived by Rosseland (1949) for radial adiabatic oscillation of a star without rotation or magnetic field we conclude that the last term from the left hand side appears because of the existence of the magnetic field.

4 Nonlinear radial oscillations of magnetic polytropes

Expanding in Taylor series the right hand side of the equation (5) and keeping the terms up to the first order Roxburgh and Durney (1967) obtained the equation of linear pulsations of magnetic polytropes.

Keeping in the series expansion the terms up to third order in r_1 we obtain the following equation

$$\rho_0 a\ddot{r}_1 = L(r_1) + Q(r_1) + S(r_1), \tag{7}$$

where the functions contain the terms of order one, two, three in r_1. Their expressions are two long to be reproduced here. If we consider that r_1 is

$$r_1 = \xi_1(a)q_1(t), \tag{8}$$

with ξ_1 is the eigenfunction corresponding to the fundamental mode of the equation (5) we obtain

$$\ddot{q_1} + q_1 = Aq_1^2 + Bq_1^3, \tag{9}$$

where the variable is proportional to time and the coefficients A and B depend on the polytropic index and on the ratio between the magnetic and gravitational energy. This equation (9) is the equation of the radial nonlinear oscillation of magnetic polytropes obtained by Das et al (1994). They investigated it by using the multiple scales method. We will find its approximate solution using the normal forms method.

5 Adiabatic oscillations of magnetic polytropes through normal forms method

5.1 Short description of the method

We are looking for an approximate solution of the pulsations equation. The small parameter we choose is the amplitude of the initial oscillation, because from observational data it is between 0.05 and 0.15 for classical Cepheids, $0.01-0.08$ for RR Lyrae stars and $0.10-0.30$ for W Vir (Buchler, 1990; Stothers, 1981).

Let $q_1 = \lambda Q_1$, where λ is the initial amplitude and Q_1 the unknown function, then the equation (9) becomes

$$\ddot{Q_1} + Q_1 = A\lambda Q_1^2 + B\lambda^2 Q_1^3, \tag{10}$$

where $Q_1 = 1.0$ and $\dot{Q_1} = 0.0$ for the initial moment $t = 0$.

We replace the real unknown function Q_1 with a complex unknown function of complex variable to reduce the order of the differential equation we have to solve. The complex variables $\xi, \bar{\xi}$ are introduced by

$$\xi = \frac{1}{2}(Q_1 - i\dot{Q_1}), \quad \bar{\xi} = \frac{1}{2}(Q_1 + i\dot{Q_1}). \tag{11}$$

In these new variables the equation 10) becomes

$$\dot{\xi} = i\xi - \frac{i\lambda A}{2}(\xi + \bar{\xi})^2 - \frac{i\lambda^2 B}{2}(\xi + \bar{\xi})^3, \tag{12}$$

equation which is written in a simpler form using a power series expansion with respect to a new variable η. Let

$$\xi = \eta + \lambda h_1(\eta, \bar{\eta}) + \lambda^2 h_2(\eta, \bar{\eta}) + \lambda^3 h_3(\eta, \bar{\eta}) + ... \tag{13}$$

and

$$\dot{\eta} = i\eta + \lambda g_1(\eta, \bar{\eta}) + \lambda^2 g_2(\eta, \bar{\eta}) + \lambda^3 g_3(\eta, \bar{\eta}) + ..., \tag{14}$$

where $h_i (i = 1, 2, 3)$ are smooth function in η and $\bar{\eta}$ and $g_i (i = 1, 2, 3)$ contain the resonant terms (*i.e.* $\sim e^{it}$). We mention that after we have determined and replaced the functions which appear in (14), it became the normal form of equation (12).

5.2 Approximate solution of the radial nonlinear pulsations equation

Identifying the coefficients of the terms that contain λ at equal power, we are able to specify the form of the functions g_i and to write down the normal form of equation (12) as

$$\dot{\eta} = i\eta - \lambda^2 i \left(\frac{5A^2}{6} + \frac{3B}{2}\right)\eta^2\bar{\eta}. \tag{15}$$

with the initial condition $\eta = 0$ for $t = 0$. The solution of this equation is

$$\eta = \frac{1}{2} e^{i\left[t-\lambda^2\left(\frac{5A^2}{24}+\frac{3B}{8}\right)t\right]}, \tag{16}$$

and the oscillations period is $2\pi + O(\lambda^2)$. To simplify the form of the solution we introduce the notation

$$\tau^+ = t - \lambda^2 \left(\frac{5A^2}{24} + \frac{3B}{8}\right) t. \tag{17}$$

The approximate solution of equation (10) is

$$\begin{aligned} Q_1 &= \cos\tau^+ + \lambda\left(-\frac{A}{6}\cos 2\tau^+ + \frac{A}{2}\right) + \frac{\lambda^2}{4}\left(\frac{A^2}{24} - \frac{B}{8}\right)\cos 3\tau^+ \\ &+ \frac{\lambda^3}{8}\left[\left(-\frac{A^3}{540} + \frac{AB}{12}\right)\cos 4\tau^+ + \left(-\frac{7A^3}{216} - \frac{31AB}{12}\right)\cos 2\tau^+\right. \\ &- \left.\frac{19A^3}{9} - 5AB\right]. \end{aligned} \tag{18}$$

5.3 Concluding remarks

The numerical evaluations of (18) for different values of the polytropic index (n) and ratio between magnetic and gravitational energy (h) revealed us, as expected, that the precision of the computation is highly dependent on the initial amplitude, and less sensitive at the values of n or h.

6 Radial nonlinear nonadiabatic oscillations of magnetic polytropes

The dissipative phenomena are described with the aid of two terms introduced in the equation (10). This idea and the form of these terms belong to Krogdahl (1955), who use them to explain the shape of the light curves observed at Cepheids. The terms added are $\mu\frac{dq_1}{dt}$, with $\mu > 0$, μ being a constant related to the energy sources of the star and $-\frac{\mu}{\lambda}q_1^2\frac{dq_1}{dt}$, where λ is proportional to the energy loss.

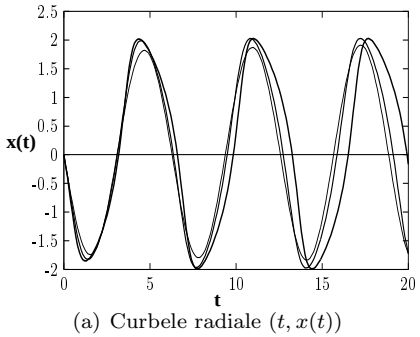
If we substitute $q_1 = \lambda Q_1$, the equation (10), in which we have added the two terms describing the energy production and loss, becomes

$$\ddot{Q}_1 + Q_1 = A\lambda {Q_1}^2 + B\lambda^2 {Q_1}^3 + \mu(1 - Q_1^2)\dot{Q}_1, \tag{19}$$

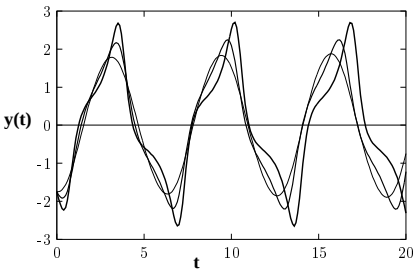
which for $\lambda = 0$ is a van der Pol equation[1]. A qualitative study of the equation (19) was done by Blaga (1998). It reveals that the number of the equilibrium points of

[1]In this case λ is a parameter connected to energy production and dissipation, not to the initial amplitude as in the case of normal form method discussed above.

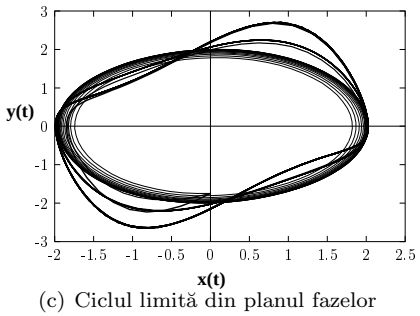
the equation (19) depends on the polytropic index and the ratio between magnetic and gravitational energy and their nature depends on the energy dissipation through μ. For $\mu > 0$ exists at least one periodic solution. For small positive values of μ the origin is unstable, but there exists a stable limit cycle. From the physical point of view, this means that, no matter which the initial conditions are the solution tends asymptotically to that periodic solution. In the amplitude-frequency relation μ plays no role as long as it is small.



(a) Curbele radiale $(t, x(t))$



(b) Curbele vitezei radiale $(t, y(t))$



(c) Ciclul limită din planul fazelor

Figure 1: *Reprezentarea grafică a soluţiei ecuaţiei pulsaţiilor neliniare neadiabatice pentru* $n = 3$ *şi* $h = 0.004$

We conclude that the presence of the energy sources ($\mu \neq 0$, $\mu > 0$) in a polytropic star does not contradict the existence of periodic orbits for the equation of radial nonlinear pulsations. For $\mu = 0$ (*i.e.* adiabatic pulsations) and $\mu \neq 0$, $\mu > 0$ (*i.e.* nonadiabatic oscillations) the motion in the phase plane looks different. In the first case the solution depends on the initial conditions and in the second case the amplitude of the pulsation is independent of the initial conditions.

We emphasize that for small values of parameters μ şi λ, the magnetic field, through

A and B, the nonadiabatic processes, through the damping coefficient, tend to diminish the frequency of the oscillations. This thing could be observed from the figures 1(a) and 1(b) in which we have represented the radial curves and the radial velocity curves $(t, x(t))$, respectively $(t, y(t))$ for the first order differential system corresponding to the second order differential equation (19). These were obtained solving it numerically for $n = 3$ and $h = 0.004$ (where h is the ratio between the magnetic and gravitational energy) and $\mu \in \{0.05, 0.5, 1.0\}$. In figure 1(c) we have represented the limit cycle. For small values for μ it is symmetric, this quality is lost for bigger values for μ, as could be observed from the radial velocity curves (figure 1(a)).

References

Anand S. P. S., Kushwaha R. S., 1962, Ann. Astrophys., 25, 310

Blaga, C., 1998, C.R. Acad. Sci. Paris, t.326, série IIb, 219

Buchler, R.J., 1990, Ann. NY Acad. Sci., Nonlinear Astrophysical Fluid Dynamics, 617, 17

Das, M. K.,Mollikutty, O. J., Tavakol, R. K.,1994 ApJ, 433, 786

Krogdahl, W.S., 1955, ApJ, 122, 43

Monaghan, J.J., 1968, Z.Astrophys., 69, 146

Rosseland, S., The Pulsation Theory of Variable Stars, Oxford, Clarendon Press, 1949

Roxburgh, I. W., 1966, M.N.R.A.S., 132, 347

Roxburgh, I. W., Durney, B. R., 1967, M.N.R.A.S., 135, 329

Stothers, R., 1981, ApJ, 247, 941

PADEU | PADEU **15**, 211 (2005)
ISBN 963 463 557

Stability of (exo)planetary systems

Zs. Sándor, Á. Süli and B. Érdi

Eötvös University, Department of Astronomy, H-1518 Budapest, Pf. 32, Hungary

E-mail: Zs.Sandor@astro.elte.hu

Abstract
Dynamical stability of the habitable zone of exoplanetary systems are investigated by a recently introduced chaos indicator (RLI). Study of individual systems, and the idea of a stability catalogue of habitable zones for exoplanetary systems are presented.
Keywords: *exoplanets, habitability, stability, chaos detection*

1 Introduction

The existence of extrasolar planets was a long-standing open issue of astronomy till 1995, when the first exoplanet has been discovered around the star 51 Pegasi by Mayor & Queloz (1995). Since this discovery we know more than 133 extrasolar planets, which form around their hosting star 117 exoplanetary systems. The majority of these systems are single systems in the sense that they are consisting of only one planet, and 13 are multiple systems consisting of two or even more planets.

Exoplanetary research focusses mainly on the observation of exoplanets. However, there are other very important questions related to this field: one of these is the investigation of formation scenarios of planetary systems, and another one, which is addressed in this paper, is the dynamics and stability of such systems. Also a very important question arising in the context of exoplanetary research is the problem of habitability. Beside the very important atmospheric and geophysical characteristics of an exoplanet, its orbital (dynamical) properties are also essential in studying its habitability. From a human point of view, a planet is habitable, if the temperature on its surface is high enough to keep water in liquid phase, see Kasting et al. (1993) for a more detailed definition. We note however, that the question of habitability defined above, has sense only in the case of terrestrial planets.

Unfortunately, the exoplanets observed by now are gas-giants. This is the consequence of the fact that by using radial-velocity measurements, which is the most effective ground-based observing technique, there is no chance to detect Earth-sized planets. Thus there are space missions aiming at the detection of Earth-like exoplanets in planning phase, which hopefully will be launched in the near future.

The paper is organized as following: first we review shortly the basic characteristics of exoplanets and exoplanetary systems discovered by now, then we describe briefly the ground-based observing techniques, and the above mentioned space missions. After discussing our method of stability investigation, we present our results obtained by the stability investigation of hypothetical terrestrial extrasolar planets.

2 Characterization of exoplanets and their methods of detection

2.1 Properties of exoplanetary systems

Studying the orbital characteristics of exoplanets (see: http://www.exoplanets.org) one can conclude that the exoplanetary systems discovered until now differ very much from our planetary system, where the Jupiter-like gas giants are at larger distances (from the Sun), and between the gas giants and the Sun there are the Earth-like planets. In the case of the Solar System the orbital eccentricities of giant planets are rather small, or moderate.

It is expected however, that by using other observing techniques Earth-like exoplanets will be detected, and such planetary systems that are more similar to the Solar System. In what follows we present the most important observing methods and the space missions planned for observation of Earth-like exoplanets.

2.2 Methods of investigation

There are a couple of methods to observe extrasolar planets, the most efficient is based on the Doppler-shift of the hosting star's spectral lines. Another very promising method is the transit photometry. Although there are some results provided by the OGLE program, a breakthrough is expected after launching space instruments using transit photometry. Finally, we mention the interferometric methods. There are grandious space missions as Terrestrial Planet Finder (NASA) or Darwin (ESA) planned to launch in the future, which may use interferometric measurements. These instruments will be able to detect even some signs (if there are any) of possible extraterrestrial life. In what follows, we summarize briefly the Doppler-method and the transit photometry.

2.2.1 Radial velocity measurements

By now, this is the most efficient method to observe Jupiter-like exoplanets around a star. The basic principle of this method is that the star and an unseen planet move around their common barycenter, which results in the periodic displacement of the star's spectral lines. From this effect the radial velocity curve of the star can be calculated, and various physical properties of the unseen planet can be deduced: mass ($m \sin i$), semi-major axis, eccentricity, etc. This method can be applied to detect 3-10 m/s change in radial velocity. We note that Jupiter causes a 12 m/s radial velocity change in Sun's motion.

2.2.2 Transit photometry and space instruments

Transit photometry can be applied if the unseen planet, its hosting star, and the observer are approximatelly in the same plane. Planetary transit then results in the periodic dimming of the star's light intensity, however, a terrestrial planet causes only a $\Delta m = 10^{-5}$ change in the light intensity! Clearly, this method can be used to detect Earth-like planets only from space, thus there are space missions for observing Earth-like planets in planning phase. Such a mission is COROT (COnvection, ROtation and planetary Transit) sponsored mainly by CNES, partly by ESA and other countries. Its planned launching date is 2006. COROT is a space telescope with a 30 cm diameter mirror and an array of CCD's as detectors. It has two scientific aims: stellar seismology and the detection of few times larger terrestrial planets than Earth. The minimal expectations of the program after observing 3×10^4 stars, and supposing that 5% of them have Earth-like planets, are the detection of 10 exoplanets having radius $R = 2R_\oplus$ and 6 exoplanets with radius $R = 1.58R_\oplus$.

Another space instrument, which is devoted entirely to observe Earth-like exoplanets by using the transit photometry is KEPLER (NASA). Practically this is a large Schmidt-telescope with 95 cm aperture, its mirror diameter is 1.4 m, and it has 42 CCD chips. The planned life-time of this instrument is 4 years, thus during this (in ideal case) three transits with period 1.33 year can be detected. KEPLER is calibrated to detect a transit of an Earth-sized planet with semi-major axis $a = 1AU$ around a star with $m_v = 12^m$ apparent brightness.

Analysing the light curve and the period of the transit the orbital parameters of an exoplanet can not be derived uniquely. More accurate orbital parameters can be obtained if dynamical constraints are present. Such a dynamical constraint could be a Jupiter-like gas giant, which results in appearing dynamically unstable regions in the system, which are similar to the Kirkwood-gaps in the Solar System. If a hypothetical object is placed in an unstable region its orbit would be chaotic. Thus by using a chaos-detection method, orbital solutions which result in chaotic orbits can be avoided. This can help in chosing the right orbital parameters of the observed exoplanet by using the transit photometry.

3 Method of stability investigation

A traditional way to detect chaotic behaviour of orbits in dynamical systems is the calculation of the maximum Lyapunov characteristic exponent (LCE). The LCE of a trajectory emanating from an initial point $\mathbf{x}^*$ of the phase space is defined as the limit

$$L^1(\mathbf{x}) = \lim_{t\to\infty} \frac{1}{t} \log \frac{||\xi_t||}{||\xi_0||}, \tag{1}$$

where ξ_t is the image of an initial infinitesimal small deviation vector ξ_0 after time t. The evolution of ξ_t can be calculated by numerical integration of the equations of motion together with their linearized equations:

$$\begin{aligned} \dot{\mathbf{x}} &= \mathbf{f}(\mathbf{x}), \\ \dot{\xi} &= D\mathbf{f}(\mathbf{x})\xi, \end{aligned} \tag{2}$$

where $D\mathbf{f}(\mathbf{x})$ is the Jacobian matrix evaluated at $\mathbf{x}$. In the case of Hamiltonian systems (dynamical systems describing the behaviour of planetary systems are Hamiltonian systems) if $L^1(\mathbf{x}) = 0$, the orbit evolving from $\mathbf{x}$ is regular, if $L^1(\mathbf{x}) > 0$ it is chaotic. A serious disadvantage of the calculation of the LCE is that it can not be obtained after finite integration time, thus its value can only be extrapolated, which makes the identification of weakly chaotic orbits very uncertain. Furthermore, weak chaos plays an essential rôle in understanding the long-term stability of exoplanetary systems (including the Solar System as well). Thus in recent years there has been a growing interest in the development and application of fast chaos detection methods. Being aware of the incompleteness of the references below, we refer to the FLI method of Froeschlé et al. (1997), to the method of spectral distance by Voglis et al. (1998, 1999), and to the SALI method of Skokos (2001). In Sándor et al. (2000) we introduced and in Sándor et al. (2004) we refined the concept of the Relative Lyapunov Indicator (RLI), which has also been proved in our investigations an efficient tool of chaos detection. Before the definition of the RLI, we recall the definition of the finite-t Lyapunov indicator originating from an initial point $\mathbf{x}_0$:

$$L(\mathbf{x}_0, t) = \frac{1}{t} \log \frac{||\xi_t||}{||\xi_0||}. \tag{3}$$

The RLI method is based on the fact, that the finite-time approximation of two neighbouring orbits, $L(\mathbf{x}, t)$ and $L(\mathbf{x}+\delta\mathbf{x}, t)$ evolves similarly (as a function of t) for regular orbits, and differently for chaotic orbits. In order to quantify the time-evolution of $L(\mathbf{x}, t)$ for neighbouring regular and chaotic orbits, we introduced the idea of the Relative Lyapunov Indicator (Sándor et al., 2000)

$$\Delta L(\mathbf{x}, t) = |L(\mathbf{x} + \delta\mathbf{x}, t) - L(\mathbf{x}, t)|, \tag{4}$$

where $\delta\mathbf{x} << 1$. Our investigations on different dynamical systems such as the restricted three-body problem (Sándor et al., 2000); the 2D, 4D symplectic mappings,

and the stability of certain exoplanetary systems Sándor et al. (2004); Érdi et al. (2004); and Érdi & Sándor (2005) show that the curve $\Delta L(\mathbf{x}, t)$ exhibits typical behaviour for regular and chaotic orbits, which differ essentially from each other.

Since the chaotic behaviour of orbits can be detected after a relatively short time numerical integration, the method of the Relative Lyapunov Indicators enables us to study a large set of initial conditions, thus to discover the regular and chaotic regions of the phase space. In order to separate regular and chaotic regions of the phase space we calculate the average value of $\Delta L(\mathbf{x}, t)$ for a given integration time t^*:

$$\langle \Delta L(\mathbf{x}^*) \rangle_{t^*} = \frac{1}{t^*} \sum_{j=1}^{t^*/\Delta t} \Delta L(\mathbf{x}, j\Delta t). \tag{5}$$

If $\mathbf{x}$ is in a regular region $\langle \Delta L(\mathbf{x}) \rangle_{t^*}$ is small, otherwise, if $\mathbf{x}$ is in a chaotic region $\langle \Delta L(\mathbf{x}) \rangle_{t^*}$ will be larger.

4 Dynamical stability of the habitable zones of exoplanetary systems

In this section we present our results obtained in studying the dynamical habitability of extrasolar planetary systems. After discussing the stability of the habitable zones of individual systems, we present our concept on a stability catalogue for hypothetical Earth-like planets. Finally, we discuss the case of Earth-like exotrojans. (The habitable zone is that region around a star where water can exist in fluid phase on a surface of a planet.)

4.1 Stability of individual systems

The investigation of the stability of the habitable zones of exoplanetary systems began with the work of Jones et al. (2001). They investigated four systems by long-time numerical integration, and found that two of these systems could have stable Earth-like orbits in their habitable zone. More recently, Menou & Tabachnik (2003) have studied the stability of the habitable zones for 85 exoplanetary systems. They found that one fourth of these systems might have stable habitable zones. In the above cases the method of investigation was the numerical integration of the equations of motion of fictitious Earth-like planets.

Instead of numerical integration of individual orbits, we have explored the semi-major axis – eccentricity plane of some exoplanetary systems having two giant objects. In our research (Sándor et al., 2004; Érdi et al., 2004) we studied the systems HD 38529, HD 169830, and HD 168443. In these cases the habitable zones are between the two giant planets.

4.1.1 HD38529

The habitable zone of this system is between 1.4 - 3.0 AU. In Figure 1, left panel we displayed the dynamical structure of the $a - e$ plane between $a = 0.6 - 2.0$ AU. Figure 1 shows how a small Earth-like planet would behave if it was started with initial conditions corresponding to the points of the $a - e$ plane of the plot. The dark region on the right hand side indicates strongly chaotic motion, the light region on the left ordered, thus stable motion. The grey strips correspond to the different mean motion resonances. Studying Figure 1 one can see that near the inner edge of the habitable zone a third planet with negligible mass can exist.

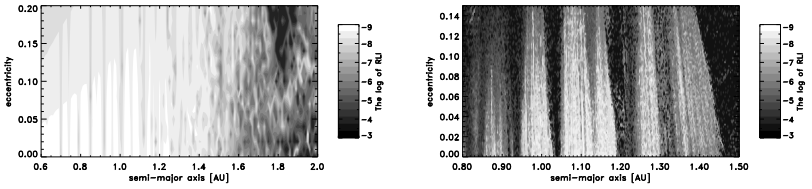


Figure 1: *The $a - e$ planes of HD38529 (left panel) and HD168443 (right panel)*

4.1.2 HD 168443 and HD 169830

The $a-e$ plane of HD 168443 systems are shown in Figure 1, right panel. The habitable zone for HD 168443 is between $0.7 - 1.3$ AU, while for HD 169830 is between $1.4 - 3.0$ AU. In Figure 1 (right panel) one can see that the $a - e$ plane is very chaotic. There are some lighter regions, which seem to allow stable motion for the Earth-like planet, but our careful analysis showed that all orbits are unstable.

The $a - e$ plane of HD 169830 is very similar to HD 168443 therefore we do not show it. The habitable zone of this system is very chaotic, thus it is very unlikely to host Earth-like planets with stable orbits.

4.2 A stability catalogue of Earth-like exoplanets in exoplanetary systems

As we mentioned before, by using data from transit photometry, the orbital parameters of an Earth-like planet can be calculated only with some error limits. The presence of another giant planet results in appearing chaotic (and therefore unstable) regions in the system, thus it means a dynamical constraint, which can help in deriving more accurate orbital elements. Our idea is to compile a catalogue of dynamical stability for exoplanetary systems consisting of a giant Jupiter-like planet and a small Earth-like

planet. This stability catalogue can be used to establish immediatelly the stability properties of the habitable zones of the known exoplanetary systems too.

The planned catalogue uses as dynamical modell the elliptic restricted three-body problem (ERTBP), assuming that the giant planet moves around the star in an elliptic orbit with semi-major axis normalized to unity: $a_p = 1$, eccentricity e_p, periastron argument ω_p, and mean anomaly at the epoch M_p. ERTBP depends on two parameters: the eccentricity of the massive planet, and the mass parameter $\mu = m_p/(m_s + m_p)$, where m_p is the mass of the giant planet and m_s is the mass of the star. By using the RLI we shall calculate the structure of the $a - e_p$ plane for several values of the mass parameter μ, where a is the semi-major axis of the hypothetical Earth-like planet and e_p is the eccentricity of the giant planet. First we shall fix $e = 0$, $\omega = 0$ $M = 0$, $\omega_p = 0$, and $M_p = 0$. Later on we shall change M_p between $0^\circ - 360^\circ$.

The compilation of the above described catalogue has begun, in what follows we present one "page" of it. Figure 2 (left panel) shows the structure of the $a - e_p$ plane for $\mu = 0.001$ (the case of the Sun–Jupiter system), when the (normalized) semi-major axis of the hypothetical Earth-like planet is larger than a_p. The light regions correspond to ordered, therefore stable orbits. There are grey "V"–shaped strips, which correspond to different mean-motion resonances. Increasing e_p these resonances overlap each other and a strongly chaotic, thus unstable behaviour appears.

By using this catalogue one shall easily decide whether the habitable zone (HZ) of a given system is dynamically stable, or not. If the physical properties of the central star, the eccentricity and the mass of the giant planet are known, it is easy to place the system's habitable zone on the $a - e_p$ plane. In Figure 2 (left panel) the HZ of four systems are shown. Two of them is entirely in the chaotic region, but the HZ of HD121504 is almost ordered, thus stable. The HZ of HD52265 is mainly in chaotic region, but it contains stable regions too.

Finally we note that the dynamical properties of the $a - e_p$ planes are governed by the different mean motion resonances between the giant planet and the Earth-like planet. The most important resonances are marked in Figure 2 (left panel). These resonances can represent both ordered (stable) or weakly chaotic (becoming unstable after very long time) orbits, depending on the initial angular positions of the planets.

4.3 Stability of co-orbital Earth-like exoplanets

We speak of co-orbital motion when two planets move in nearly the same orbits. In this case they are in a 1/1 mean-motion resonance. In the Solar System there are many examples for this type of motion: the best known representatives are the Jupiter's Trojan asteroids. Thus it can be expected that co-orbital objects exist also in exoplanetary systems. In our investigation (Érdi & Sándor, 2005) we studied that case when the giant planet moves in the habitable zone of the system. Then habitable Earth-like planet could exist only in the vicinity of the stable Lagrangian points of the giant planet – star system since around a stable Lagrangian point there is a region

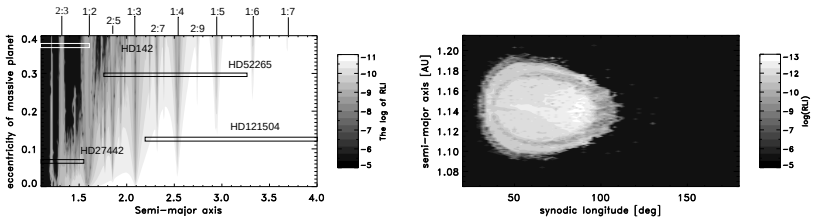


Figure 2: *One page of the stability catalogue (left picture). HD177830: the stability around L_4 (right picture).*

of non-linear stability. Test particles starting from this region librate around the corresponding Lagrangian point in stable orbits. We investigated 5 systems, in which the giant planet is always in the habitable zone, and 4 systems, in which the giant planet, due to its larger eccentricity, can leave temporarly the habitable zone. For our computations we used the modell of the elliptic restricted three-body problem. We found that in all systems co-orbital Earth-like planets can exist.

Figure 2 (right panel) shows such a stability region around the Lagrangian point L_4 of the system HD177830. On the $x-$axis there is the synodic longitude τ, which is the difference between the mean longitudes of the hypothetical Trojan and the giant planet; on the $y-$axis there is the semi-major axis of the hypothetical Trojan planet. We take initial τ and a values from a regular grid on this plane, and assuming that the remaining orbital elements (eccentricity, argument of periastron) are zero, we calculate the RLI for each orbit originating from these initial conditions. It can be seen that around L_4 there is a stability region, which may host stable Earth-like exotrojan planets for very long time in a very stable orbit.

5 Summary

The most efficient ground-based observation method of exoplanets is based on the radial velocity measurements of a star. Unfortunately, this method is unable to detect Earth-like planets. The existence of Earth-like planets, however, is a very important question regarding the formation theories of the Solar System, where both rocky planets and gaseous giant planets have been formed. Thus there are space missions aiming at the detection of Earth-like exoplanets in planning phase, which will be launched in the near future.

In this paper we have investigated the dynamical stability of Earth-like planets in exoplanetary systems. We have found that there are exoplanetary systems, which can host Earth-like exoplanets even in such an exotic case that the Earth-like planets are co-orbital companions of the gas giants. We have begun the compilation of a stability

catalogue of Earth-like planets in exoplanetary systems, which can be used to establish the stability properties of the habitable zones of the known exoplanetary sytems. This catalogue will help in determining more accurate orbital elements of Earth-like planets detected by transit photometry too.

Acknowledgement

This research was supported by the Hungarian Scientific Research Fund under the grants OTKA D048424 and OTKA T043739.

References

Érdi, B., Dvorak, R., Sándor, Zs. et al. 2004, Mon. Not. R. Astron. Soc., 351, 1043

Érdi, B. & Sándor, Zs., 2005, Celest. Mech. & Dyn. Astron (accepted)

Froeschlé, C., Lega, E., & Gonczi, R. 1997, Celest. Mech. & Dyn. Astron., 67, 41

Jones, B. W., Sleep, P. N. & Chambers, J. E., 2001, A&A, 366, 254

Kasting, J. F., Whitmire, D. P., & Reynolds, R. T. 1993, Icarus, 101, 108

Mayor, M., & Queloz, D. 1995, Nature, 378, 355

Menou, K. & Tabachnik, S., 2003, ApJ, 583, 473

Sándor, Zs., Érdi, B., Efthymiopoulos, C., 2000, Celest. Mech. & Dyn. Astron., 78, 113

Sándor, Zs., Érdi, B., Széll, A., et al. 2004, Celest. Mech. & Dyn. Astron., 90, 127

Skokos, Ch., 2001, J. Phys. A (Math. Gen.), 34, 10029

Voglis, N., Contopoulos, G., Efthymiopoulos, C., 1998, Phys. Rev. E, 57, 372

Voglis, N., Contopoulos, G., Efthymiopoulos, C., 1999, Celest. Mech. & Dyn. Astron. 73, 211

PADEU | PADEU **15**, 221 (2005)
ISBN 963 463 557

Pulsating zero velocity surfaces and capture in the elliptic restricted three-body problem

F. Szenkovits[1,2] and Z. Makó[1,3]

[1] Babeş-Bolyai University, Department of Applied Mathematics, RO-400048 Cluj-Napoca, Kogalniceanu 1, Romania
[2] Eötvös University, Department of Astronomy, H-1518 Budapest, Pf. 32, Hungary
[3] Sapientia University, Department of Mathematics and Informatics, RO-530104, Miercurea-Ciuc, Romania

E-mail: [1]fszenko@math.ubbcluj.ro, [3]makozoltan@sapientia.siculorum.ro

Abstract
Zero velocity surfaces are deduced in the gravitational restricted three-body problem by using the Jacobi-integral. These surfaces are the boundaries of the Hill's regions: regions where the motion of the third, massless particle around the two primaries is possible. V. Szebehely generalized this result for the planar elliptic restricted three-body problem. In a recent paper – Makó and Szenkovits (2004) – the authors presented a generalization of this result for the spatial elliptic restricted three-body problem, where the existence of an invariant relation was proved. From this invariant relation the equation of the zero velocity surfaces can be deduced. In this paper we discuss the pulsation and the change of the type of these zero velocity surfaces and we present applications to the phenomenon of the gravitational capture. In the model of the spatial elliptic restricted three-body problem criteria of the capture are deduced by using the pulsating Hill's regions.
Keywords: *Elliptic restricted three-body problem, zero velocity surfaces, gravitational capture.*

1 Introduction

It is well known that the equations of motion of a body of infinitesimal mass moving under the gravitational attractions of two massive bodies, which move in circles about

their common centre of gravity (the circular restricted three-body problem, CRTBP) admit of an integral, called after its discoverer, the Jacobi integral (Jacobi, 1836). The Jacobi integral plays an important role in the study of the CRTBP, since it makes possible certain general, qualitative statements regarding the motion without actually solving the equations of motion. It has been used for example to construct surfaces of zero velocity (ZVS) which limit the so called Hill's regions of space in which the small body, under given initial conditions, can move, and to derive a criterion (the Tisserand criterion) for the re-identification of a comet whose orbit has suffered perturbations by a planet.

Several authors have also used the circular Jacobi as an approximate integral of the elliptical restricted three-body problem (ERTBP), when the two primaries revolve on elliptical orbits. Thus Hill (1878) treating the Moon as an infinitesimal body in the Sun–Earth–Moon system, and neglecting the eccentricity of the Earth's orbit about the Sun, concluded that the Moon's distance from the Earth has a superior limit.

Ovenden and Roy (1961) starting from the integrals of the general three-body problem (expressed in non-uniformly rotating rectangular coordinates), obtained formal expressions for the Jacobi integral and the angular momentum integrals of the ERTBP in terms of certain auxiliary functions. Using these expressions they concluded that "the Jacobi integral of the CRTBP for long-term predictions in any real case where the two massive bodies' relative orbit has a finite eccentricity (however small) is without justification even if the mass of the third body is infinitesimal".

Szebehely and Giacaglia (1964) obtained in the planar ERTBP a simple form of the equations of motion – similar to that in case of the CRTBP –, by using the true anomaly of the primaries as the independent variable and by introducing a special set of dimensionless variables describing the position of the third body. They also deduced an invariant relation, the generalization of the Jacobi integral and proved the pulsation of the zero velocity curves in the planar case.

Contopoulos (1967) deduced two integrals of motion in the plane ERTBP for orbits with small eccentricity near the primaries. These integrals, given in the form of formal series, depend periodically on the time, with frequency equal to that of the second body.

Vrcelj and Kiewiet de Jonge (1978) derived an invariant relation – containing also a nonintegrable term – generalizing the Jacobi integral to the ERTBP on the basis of the classical perturbation theory and by making use of energy and angular momentum integrals. This invariant relation was reduced and applied to calculate the Jacobi constant for asteroids (Vrcelj and Kiewiet de Jonge, 1978a; Vrcelj, 1979).

In a recent paper (Makó and Szenkovits, 2004) the authors generalized Szebehely's (1964) result concerning the pulsating Hill's regions to the spatial ERTBP. By using the obtained invariant relation, necessary conditions of the gravitational capture of small bodies was also deduced, in case of small eccentricities of the primaries and bounded motion of the test particle.

The gravitational capture of small bodies by major planets is an interesting phe-

nomenon in planetary systems, having applications to the study of comets, asteroids and moons, and it can be studied by using different models of the celestial mechanics. Several authors studied this problem, introducing different concepts of capture, like weak capture (Belbruno, 1999; Belbruno and Marsden, 1997), temporary capture (Brunini, 1996), longest capture (Vieira and Winter, 2001), resonant capture (Yu and Tremaine, 2001), etc. Brunini, Giordano and Orellana (1996) studied the conditions of the capture in the restricted three-body problem. Murison (1989) pointed out connections between the gravitational capture and chaotic motions. An exciting study has been dedicated to the capture of irregular moons – with non-circular orbits - by giant planets (Astakhov et al., 2003). The authors confirmed with three-dimensional Monte Carlo simulations that irregular satellites are captured in a thin spatial region, where orbits are chaotic and that the resulting orbit is either prograde or retrograde depending on the initial energy. In (Astakhov and Farrelly, 2004) the authors, using the model of the ERTBP, show that the mechanism of the capture, presented in the beforehand cited paper, survives perturbations due to the ellipticity of the planet's orbit, however, the planet's ability to capture moons decreases with increasing orbital eccentricity.

In this paper we put in evidence some properties of the Hill's regions in the case of the spatial ERTBP. The pulsation and the change of the type of these zero velocity surfaces are evidenced. We present also applications to the phenomenon of the gravitational capture. In the model of the spatial elliptic restricted three-body problem criteria of the capture are deduced by using the pulsating Hill's regions.

2 Invariant relation in the spatial ERTBP

In the elliptic restricted three-body problem (ERTBP) two massive primaries (planetary objects) P_1 and P_2, with masses m_1 and m_2 revolve on elliptical orbits under their mutual gravitational attraction and the motion of a third, massless body $P_3, (m_3 = 0)$ is studied. The orbit of P_2 around P_1, in an inertial system is

$$r = \frac{a\left(1 - e^2\right)}{1 + e\cos f}, \tag{1}$$

where r is the mutual distance, a and e are the semimajor axis and the eccentricity of the elliptical orbit ($e < 1$), and f is the true anomaly.

There are several systems of reference that can be used to describe the elliptic restricted three-body problem. In our study a nonuniformly rotating and pulsating coordinate system is used. In this system of reference (Figure 1) the origin O is in the center of mass of the two massive primaries (Sun and Earth for example), and the $\tilde{\xi}$ axis is directed towards m_2. The $\tilde{\xi}\tilde{\eta}$ coordinate-plane rotates with variable angular velocity, in such a way, that the two massive primaries are always on the $\tilde{\xi}$ axis, and

the period of the rotation is 2π. Besides the rotation, the system also pulsates, to keep the primaries in fixed positions $\left(\tilde{\xi}_1 = -\mu, \tilde{\eta}_1 = \tilde{\zeta}_1 = 0, \tilde{\xi}_2 = 1-\mu, \tilde{\eta}_2 = \tilde{\zeta}_2 = 0\right)$.

In this system the equations of motion of the third massless particle are:

$$\begin{cases} \tilde{\xi}'' - 2\tilde{\eta}' = \frac{\partial \omega}{\partial \tilde{\xi}}, \\ \tilde{\eta}'' + 2\tilde{\xi}' = \frac{\partial \omega}{\partial \tilde{\eta}}, \\ \tilde{\zeta}'' = \frac{\partial \omega}{\partial \tilde{\zeta}}, \end{cases} \tag{2}$$

where the derivatives are taken with respect to the true anomaly f, and

$$\omega = (1 + e\cos f)^{-1}\,\Omega,$$

with

$$\begin{aligned} \Omega \quad &= \tfrac{1}{2}\left(\tilde{\xi}^2 + \tilde{\eta}^2 - e\tilde{\zeta}^2\cos f\right) + \frac{1-\mu}{\sqrt{\left(\tilde{\xi}+\mu\right)^2 + \tilde{\eta}^2 + \tilde{\zeta}^2}} + \\ &+ \frac{\mu}{\sqrt{\left(\tilde{\xi}-1+\mu\right)^2 + \tilde{\eta}^2 + \tilde{\zeta}^2}} + \tfrac{1}{2}\mu\,(1-\mu)\,. \end{aligned} \tag{3}$$

Performing the same operations, which in the RTBP leads to the Jacobi-integral, in the case of the spatial ERTBP we obtain an invariant relation of the form:

$$\begin{aligned} \left(\frac{d\tilde{\xi}}{df}\right)^2 + \left(\frac{d\tilde{\eta}}{df}\right)^2 + \left(\frac{d\tilde{\zeta}}{df}\right)^2 \quad &= 2\omega - e\int_{f_0}^{f} \frac{\tilde{\zeta}^2 \sin h}{1+e\cos h}dh - \\ &-2e\int_{f_0}^{f} \frac{\Omega \sin h}{(1+e\cos h)^2}dh - C_0. \end{aligned} \tag{4}$$

This is the generalization of Szebehely's invariant relation (Szebehely (1967), pp. 595) for the spatial ERTBP. Unforunately (4) is not an integral of motion, because it contains not only functions of the coordinates and velocities (and the true anomaly f) at any point of the orbit, but it depends also on the values of these quantities along the whole arc of the orbit from the initial position up to a given position. The value of the Jacobi constant C_0 can be calculated at the initial moment $t = t_0$, when $f = f_0$ and when the two integral terms in (4) are vanishing:

$$C_0 = 2\omega\left(\tilde{\xi}_0, \tilde{\eta}_0, \tilde{\zeta}_0, f_0\right) - \left[\left(\frac{d\tilde{\xi}}{df}\right)^2 + \left(\frac{d\tilde{\eta}}{df}\right)^2 + \left(\frac{d\tilde{\zeta}}{df}\right)^2\right]_{f=f_0} \tag{5}$$

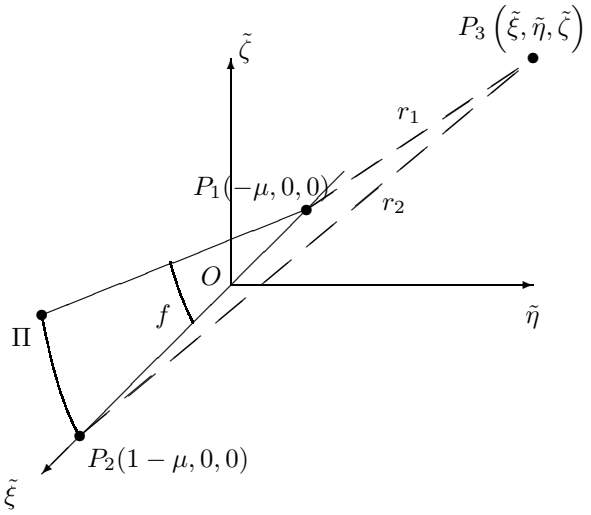


Figure 1: *The spatial ERTBP.*

3 Zero velocity surfaces in the ERTBP

The zero velocity surfaces (ZVS) in the ERTBP are given by:

$$\frac{2\Omega}{1+e\cos f}-e\int\limits_{f_0}^{f}\frac{\tilde{\zeta}^2\sin h}{1+e\cos h}dh-2e\int\limits_{f_0}^{f}\frac{\Omega\sin h}{(1+e\cos h)^2}dh=C_0. \tag{6}$$

These surfaces delimite the Hill's regions, in which the motion of the third particle is possible. Equation (4), depending also on the true anomaly f, is equivalent with

$$2\Omega\left(\tilde{\xi},\tilde{\eta},\tilde{\zeta},f\right)=C\left(f\right), \tag{7}$$

where

$$C\left(f\right)=\left(1+e\cos f\right)\left(C_0+e\int\limits_{f_0}^{f}\frac{\tilde{\zeta}^2\sin h}{1+e\cos h}dh+2e\int\limits_{f_0}^{f}\frac{\Omega\sin h}{(1+e\cos h)^2}dh\right). \tag{8}$$

Equation (7) shows us that the ZVS are changing their dimensions, depending on the variation of $C\left(f\right)$, i.e. they are pulsating. Analyzing the expression (3) of Ω we can

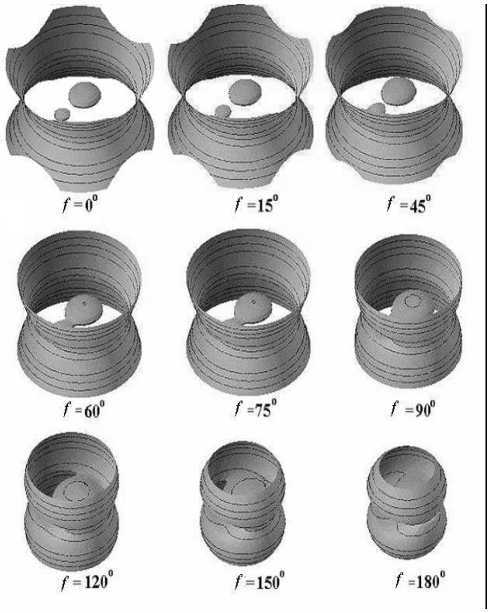


Figure 2: *Zero velocity surfaces in the ERTBP.*

see that the type of the ZVS is also changes due to the variation of the true anommaly. If the third body is far away from the primaries, a good approximation of Ω is

$$\Omega \quad = \tfrac{1}{2}\left(\tilde{\xi}^2 + \tilde{\eta}^2 - e\tilde{\zeta}^2 \cos f\right) + \tfrac{1}{2}\mu\left(1-\mu\right), \tag{9}$$

showing us that the ZVS in this region is of type of an ellipsoid, cylinder or hyperboloid, depending on the sign of $\cos f$. In conclusion, geometrically it means that at every time – or at every value of the true anomaly f – a different set of surfaces of zero velocity are to be constructed. The shape and dimension of these zero velocity surfaces vary in time (see Fig. 2). It is easy to see that for any value of the true anomaly these surfaces admits two planes of symmetry: $\tilde{\eta} = 0$ and $\tilde{\zeta} = 0$.

Assuming that the eccentricity $0 < e < 1$ is small (as it is in the case of the Sun–Earth system, $e = 0.017$), and the motion of the third particle is bounded, the sum of the two integral terms in the equation (4) is smaller than the term 2ω (In the case of the Sun–Earth–Moon system see Figure).

Neglecting these small terms, we have the approximate equation of the surfaces of zero velocity (4):

$$2\Omega\left(\tilde{\xi}, \tilde{\eta}, \tilde{\zeta}, f\right) - C_0\left(1 + e\cos f\right) = 0. \tag{10}$$

Using this approximate equation of the ZVS we deduced some necessary conditions of the capture in the model of the ERTBP (Makó and Szenkovits (2004)).

4 Necessary conditions of the capture

To give necessary conditions of the capture, we approximate the zero velocity surfaces (**??**) with the equations

$$2\Omega^{\circ}\left(\tilde{\xi}, \tilde{\eta}, \tilde{\zeta}\right) = C^*. \tag{11}$$

This approximation is possible when e is small, and the third body moves near to the plane of the primaries. In this case the term $e\tilde{\zeta}^2 \cos f$ in (3) may be neglected, and we can write $\Omega = \Omega^{\circ}$. Equation (11) and (**??**) show that the zero velocity surfaces pulsate, and so an do the Hill-zones delimited by them.

Suppose that, for $f = f_0$ the position and velocity of the third body is given, and

$$C^* = C_0^* = C\left(1 + e\cos f_0\right). \tag{12}$$

Then, from (**??**) and (12) we have

$$C^*\left(f\right) = C_0^* \frac{1 + e\cos f}{1 + e\cos f_0}. \tag{13}$$

As we have seen in Section 2, as time as

$$C_2 < C^*,$$

the third body can not near to m_2 if initially is not inside of the Hill-region surrounding m_2. By using this property we are able to give a necessary condition to the close approach of one of the primaries by the massless body.

If the massless body in the moment corresponding to f_0 was not in the Hill-zone surrounding m_2, and satisfied the condition

$$C_2 < C_0^* \frac{1-e}{1+e\cos f_0} \tag{14}$$

then it never enter in this zone, and it can not be captured by m_2.

An other condition can be formulated in the next form:

If the massless body in the moment corresponding to f0 is in the exterior of the cylinder, and satisfies the condition

$$C_1 < C_0^* \frac{1-e}{1+e\cos f_0} \tag{15}$$

then it never enter in the Hill-zone around m_2.

The advantage of this conditios consist in fact that is not necessary the integration of the equations of motion of the third-body, a simple evaluation of the expression (14) or (15) is only necessary.

Acknowledgement

This work was supported by grants from the Hungarian Academy of Sciences through 'János Bolyai' grant and the Research Programs Institute of Foundation Sapientia, Romania.

References

Astakhov, S. A., Burbanks, A. D., Stephen, W. and Farrelly. D. 2003, *Nature*, 423, 464

Astakhov, S. A. and Farrelly. D. 2004, *Mon. Not. R. Astron. Soc.*, 353, 971

Belbruno, E. 1999, in *The Dynamics of Small Bodies in the Solar System, A Major Key to Solar System Studies*, ed. by Bonnie A. Steves and Archie E. Roy. Kluwer Academic Publishers, p. 37

Belbruno, E. and Marsden, B. G. 1997, *The Astronomical Journal*, 113, 1433

Brunini, A., Giordano, C. M. and Orellana, R. B. 1996, *Revista Mexicana de Astronomia y Astrofisica Serie de Conferencais*, 4, 111

Brunini, A. 1996, *Celest. Mech. Dynam. Astronom.*, 64, 79

Contopoulos, George 1967, *The Astronomical Journal*, 72, 669

Hill, G. W. 1878, *Am. J. Math.* 1, 129
Jacobi, C. G. J. 1836, *Compte Rendus de l'Acad. des Sciences*, 59, 1836
Makó, Z. & Szenkovits, F. 2004, *Celest. Mech. Dynam. Astronom.* 90, 51
Ovenden, M. W. & Roy, A. E. 1961, *Mon. Not. R. Astron. Soc.*, 123, 1
Murison, M. A. 1989, *The Astronomical Journal*, 98, 2346
Szebehely, V. & Giacaglia, G. E. O. 1964, *Astronomical Journal* 69, 230
Szebehely, V. 1967, *Theory of orbits.* Academic Press, New-York
Vieira, N. E. and Winter, O. C. 2001, *The Astronomical Journal*, 122, 440
Vrcelj, Zoran & Kiewiet de Jonge, Joost H. 1978, *The Astronomical Journal* 83, 514
Vrcelj, Zoran & Kiewiet de Jonge, Joost H. 1978a, *The Astronomical Journal* 83, 664
Vrcelj, Zoran 1979, *The Astronomical Journal* 84, 1072
Yu, Q. and Tremaine, S. 2001,*The Astronomical Journal*, 121, 1736

PADEU | PADEU **15**, 231 (2005)
ISBN 963 463 557

Hill's stability of the Moon in the spatial elliptic restricted three-body problem

Zoltán Makó

Babeş-Bolyai University, Department of Applied Mathematics, M. Kogalniceanu nr. 1 RO-400084 Cluj-Napoca, Romania

E-mail: zmako@math.ubbcluj.ro

Abstract
Hill's (G. Hill, 1878) global and nonlinear stability theory has the advantage of being applicable to a great variety of dynamical systems, including those occurring in the solar system. He used his method originally to study the stability of the Moon as influenced by the Earth and the Sun. V. Szebehely (V. Szebehely , 1978) showed that in the model of circular restricted three-body problem the measure of stability for the Earth's Moon is very low. Using the invariant relation of the spatial elliptic restricted three-body problem we show that the measure of stability for the Earth's Moon oscillate above stability critical value.
Keywords: *Hill's stability, Restricted three-body problem*

1 Introduction

Consider a dynamical system with an integral of motion (such as the Jacobi integral in the circular restricted three-body problem) given by

$$v^2 = C - V(x, y, z),$$

where v is the velocity, V is a generalized potential, and C is the constant of integration. For a given set of initial conditions (x_0, y_0, z_0, v_0) we find

$$C_0 = v_0^2 - V(x_0, y_0, z_0).$$

The relation

$$v^2 = C_0 - V(x, y, z)$$

must hold during the motion.

Definition 1 *The equation of* zero velocity surface *(*ZVS*) according to initial conditions* (x_0, y_0, z_0, v_0) *is*

$$C_0 = V(x, y, z).$$

In general the ZVS separate those points in the space for which

$$C_0 < V(x, y, z)$$

from those for which

$$C_0 \geq V(x, y, z).$$

In the moment $t = 0$ consider a test particle inside a closed ZVS, $C = V(x, y, z)$. If its integration constant C changes slightly, by some outside disturbing effect, then its ZVS will change also. If this new region is still inside a simple closed surface, the stability of the system will not change qualitatively. But if new region does not represent the inside a simple closed surface, then the test particle may depart from the system and its behavior may change suddenly.

Definition 2 *The* C_1 *value of integration constant is a* critical or bifurcation *of the ZVS if at this value the topology of the ZVS changes.*

If the actual value of the constant C is far removed from the bifurcation value C_1, we conclude that the system is more stable than if it is very close since, when $C - C_1 \approx 0$, small perturbations may change the stability characteristics.

Definition 3 *(V. Szebehely , 1978) The difference between the actual value of the integration constant and bifurcation value is a* measure of the stability *of the system:*

$$M_{st} = C - C_1.$$

2 Zero velocity surfaces in the ERTBP

In the elliptic restricted three-body problem (ERTBP) the two massive primaries P_1 and P_2, with masses $m_1$and m_2 revolve on elliptical orbits under their mutual gravitational attraction and the motion of a third, massless body is studied. The orbit of P_2 around P_1, in an inertial system is

$$\|P_1P_2\| = \frac{a\left(1 - e^2\right)}{1 + e\cos f}, \tag{1}$$

where $\|P_1P_2\|$ is the mutual distance, a and e are the semimajor axis and the eccentricity of the elliptical orbit, and f is the true anomaly.

There are several systems of reference that can be used to describe the elliptic restricted three-body problem. In our study a nonuniformly rotating and pulsating coordinate system is used. In this system of reference the origin is in the center of mass of the two massive primaries (Sun and Earth for example), and the $\tilde{\xi}$ axis is directed towards P_2. The $\tilde{\xi}\tilde{\eta}$ coordinate-plane rotates with variable angular velocity, in such a way, that the two massive primaries are always on the $\tilde{\xi}$ axis, and the period of the rotation is 2π. Besides the rotation, the system also pulsates, to keep the primaries in fixed positions $\left(\tilde{\xi}_1 = -\mu, \tilde{\eta}_1 = \tilde{\zeta}_1 = 0, \tilde{\xi}_2 = 1-\mu, \tilde{\eta}_2 = \tilde{\zeta}_2 = 0\right)$. In this system the equations of motion of the third massless particle are:

$$\begin{cases} \tilde{\xi}'' - 2\tilde{\eta}' = \frac{\partial \omega}{\partial \tilde{\xi}}, \\ \tilde{\eta}'' + 2\tilde{\xi}' = \frac{\partial \omega}{\partial \tilde{\eta}}, \\ \tilde{\zeta}'' = \frac{\partial \omega}{\partial \tilde{\zeta}}, \end{cases} \tag{2}$$

where the derivatives are taken with respect to the true anomaly f, and

$$\omega = (1 + e\cos f)^{-1}\,\Omega,$$

with

$$\begin{aligned} \Omega\left(\tilde{\xi}, \tilde{\eta}, \tilde{\zeta}, f\right) &= \frac{1}{2}\left(\tilde{\xi}^2 + \tilde{\eta}^2 - e\tilde{\zeta}^2\cos f\right) + \frac{1-\mu}{\sqrt{\left(\tilde{\xi}+\mu\right)^2 + \tilde{\eta}^2 + \tilde{\zeta}^2}} + \\ &+ \frac{\mu}{\sqrt{\left(\tilde{\xi}-1+\mu\right)^2 + \tilde{\eta}^2 + \tilde{\zeta}^2}} + \frac{1}{2}\mu\,(1-\mu)\,. \end{aligned} \tag{3}$$

Performing the same operations, which in the restricted three-body problem leads to the Jacobi-integral, in the case of the spatial ERTBP we obtain an invariant relation of the form (Z. Mako and F. Szenkovits , 2004)

$$v^2 = 2\omega - e\int_{f_0}^{f} \frac{\tilde{\zeta}^2 \sin h}{1 + e\cos h}dh - 2e\int_{f_0}^{f} \frac{\Omega \sin h}{(1+e\cos h)^2}dh - C,$$

where v is the velocity of the third massless particle. For a given set of initial conditions $(\tilde{\xi}_0, \tilde{\eta}_0, \tilde{\zeta}_0, v_0, f_0)$ we find

$$C_0 = \frac{2\Omega\left(\tilde{\xi}_0, \tilde{\eta}_0, \tilde{\zeta}_0, f_0\right)}{1 + e\cos f_0} - v_0^2.$$

The zero velocity surfaces in the ERTBP according to initial condition $(\tilde{\xi}_0, \tilde{\eta}_0, \tilde{\zeta}_0, v_0, f_0)$ are

$$\frac{2\Omega}{1+e\cos f} - e\int_{f_0}^{f} \frac{\tilde{\zeta}^2 \sin h}{1+e\cos h}dh - 2e\int_{f_0}^{f} \frac{\Omega \sin h}{(1+e\cos h)^2}dh = C_0. \tag{4}$$

These surfaces delimite the Hill-regions, in which the motion of the third particle is not possible. In three dimension space it means that at every time – or at every value of the true anomaly f – a different set of surfaces of zero velocity are to be constructed. The shape of these ZVSs vary in time. Therefore we might speak about pulsating surfaces of zero velocity.

3 Measure of stability of the Moon in the Sun-Earth-Moon system

In the case of the Sun–Earth system the eccentricity $e = 0.0167$ is small. Due to the variation of f, these regions can pulsate, and near to the critical values they can change they type. The critical points of the pulsating surfaces of zero velocity (4) correspond approximately to the equilibrium solutions of circular restricted three-body problem given by

$$C_i = 2\Omega^{\circ}(L_i), \qquad i = 1, \ldots, 5, \tag{5}$$

where L_i are the Lagrange-points. For these constants we have

$$3 = C_4 = C_5 \leq C_3 \leq C_1 \leq C_2 \leq 4.25$$

in generally, and in the case of the Sun–Earth system the critical value for L_2 between the two primaries is

$$C_2 = 3.000893278,$$

and in L_1, the Lagrange-point outside of the Earth the critical value is

$$C_1 = 3.000889276.$$

For $C > C_2$ the ZVSs delimit three regions where the motion of the small body is possible (Figure 1). Two of these regions are closed around the primaries, the third one is the exterior of the exterior surface. Between these regions the communication is impossible.

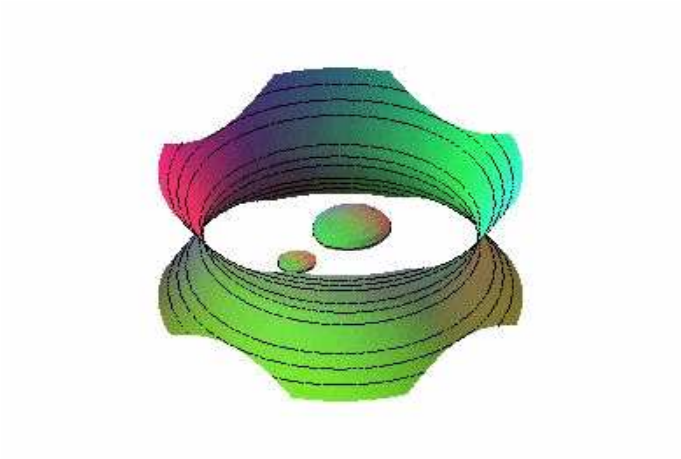


Figure 1: *The ZVSs in ERTBP if $C > C_2$.*

The measure of stability of the Moon in the Sun-Earth-Moon system considered in ERTBP model is

$$M_{st}\left(\tilde{\xi}(f),\tilde{\eta}(f),\tilde{\zeta}(f),f\right) = \frac{2\Omega\left(\tilde{\xi},\tilde{\eta},\tilde{\zeta},f\right)}{1+e\cos f} - e\int_{f_0}^{f}\frac{\tilde{\zeta}^2\sin h}{1+e\cos h}dh$$
$$-2e\int_{f_0}^{f}\frac{\Omega\left(\tilde{\xi},\tilde{\eta},\tilde{\zeta},f\right)\sin h}{(1+e\cos h)^2}dh - C_2,$$

where $(\tilde{\xi},\tilde{\eta},\tilde{\zeta})$ is solution of differential equation (2) at initial conditions of the Moon.

In figure 2 we show the variation of the M_{st} and variation of distance between Earth and Moon, notated by r_2. We observe that the measure of stability for the Moon vary from 0.0004 to 0.00055 and the maximal value is in the pericenter and minimal value is in the apocenter.

4 Conclusion

For the initial conditions of the Moon in the nonuniformly rotating and pulsating coordinate system the Hill region around to Earth is bounded by closed ZVS. The elliptical orbit of the Earth is not change the stability characteristics of the Moon.

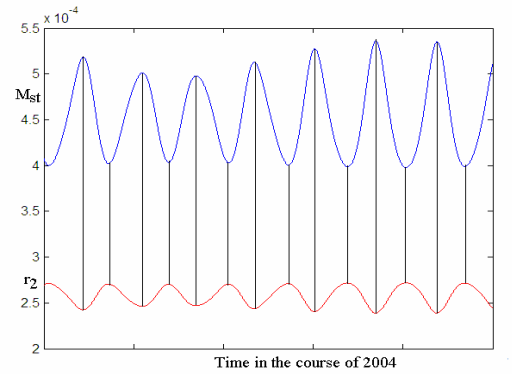


Figure 2: *The variation of measure of stability.*

We have been found that the measure of stability for the Moon vary from 0.0004 to 0.00055. The maximal value is in the pericenter and minimal value is in the apocenter. Comparatively to other satellites around of other planets, the measure of stability for the Moon is very low. For example the measure of stability for Mars's Phobos and Demos are approximately 0.0025 and 0.012.

Acknowledgement

This work was supported by the Research Programs Institute of Foundation Sapientia under grand 1416/25 October 2004.

References

Hill, G. 1878, Am. J. Math. 1, 5-26.

Hill, G. 1878, Am. J. Math. 1, 129-147.

Hill, G. 1878, Am. J. Math. 1, 245-260.

Mako, Z. and Szenkovits, F. 2004, Celestial Mech. Dyn. Astr. 90, 51-58.

Szebehely, V. 1967, Theory of orbits, Academic Press, New-York.

Szebehely, V. 1978, Proc. Natl. Acad. Sci. USA, Vol. 75, No. 12, 5743-5745.