Seminario di Astrofisica 1994 a cura di G. Bertin

SEMINARIO DI ASTROFISICA 1994

a cura di G. Bertin

APPUNTI DEI CORSI
TENUTI DA DOCENTI DELLA SCUOLA

SCUOLA NORMALE SUPERIORE
PISA

SEMINARIO DI ASTROFISICA 1994

a cura di G. Bertin

APPUNTI DEI CORSI
TENUTI DA DOCENTI DELLA SCUOLA

SCUOLA NORMALE SUPERIORE
1995

Preface

In spite of Pisa's long and strong tradition in Mathematics and Physics, only recently, mostly in the 70's, have several active groups grown in the area developing research projects and international collaborations related to Astrophysics. Probably the initial spark of interest in Astrophysics can be traced back to a relatively small workshop organized by L.A.Radicati at the Scuola Normale in the spring of 1969, following the recent discovery of *pulsars*, with the participation, among others, of G.Baym, S.Chandrasekhar, B.Coppi, T.Gold, F.Pacini, and D.Pines, and to the *Fermi Lectures* on *Neutron Stars and Pulsars* given by F.Dyson at the Scuola Normale in 1970.

The Astrophysics Group at the Scuola Normale Superiore is actively working on the dynamics of spiral and elliptical galaxies, especially in relation to the problem of the presence and distribution of dark matter, on the properties of the X-ray emission from cooling flows, on the dynamics of globular clusters, and on various problems involving astrophysical plasmas. The computing facilities include a Connection Machine CM2. An SPH code and several versions of N-body codes are available, so that interested students may have access to a laboratory for numerical simulations. Much of this work is carried out in close collaboration with other institutions, especially the Massachusetts Institute of Technology, the European Southern Observatory, the Kapteyn Institute in Groningen, and the Space Telescope Science Institute in Baltimore. A number of observational projects are in progress through ESO. Other collaborations are under way within Italy, especially with Arcetri, Bologna, Padova, and with the Teramo Observatory.

At the Department of Physics of the University of Pisa, theoretical research is carried out mostly on problems related to stellar evolution theory and to the chemical evolution of galaxies, while experimental groups within INFN are engaged in projects in high energy Astrophysics and in *VIRGO*, a large experiment aimed at the detection of gravitational waves. At the Department of Mathematics there is a very active group working in celestial mechanics.

Thus, the Pisa area offers many interesting research opportunities for the beginning astrophysicist, diversified into advanced projects well integrated at the international level. Still, since Astrophysics is a relatively new field in the Pisa area, students may find it hard to form an overall picture of the ongoing research and of the related opportunities for thesis projects. Therefore, graduate students planning to write their doctoral thesis on an astrophysically related subject are typically interested not only in finding an adequate choice of courses, in order to fulfill their requirements and to strengthen their background in Astrophysics, but also in getting

to appreciate what can be good thesis projects within the best possibilities available in the area.

In this respect, the regular courses of the graduate school provide excellent support in a few selected directions, but still they may not give a proper picture of the potential of the area. With the hope to overcome this difficulty, and thus to bridge the gap between the traditional courses and the specialized research seminars within the individual research groups, last year we have organized the *Seminario di Astrofisica* at the Scuola Normale as a series of lectures held on a weekly basis in the following way. A "core group" of people (about half of them were graduate students), all studying or working in the Pisa area, have met regularly and essentially each participant has contributed by giving one lecture on his own favorite subject. The working language was English, since a few participants were foreign postdocs. Each speaker was asked to address an audience of non-experts, not only because of the wide variety of backgrounds in Astronomy, Mathematics, and Physics required by the topics covered, but especially because the lectures were open to people outside the "core group" and announced as Astrophysics Colloquia. Thus the lectures were meant to be general introductions to a given research topic, rather than just informal seminars. The result has been a series of high quality and lively presentations. In addition, this has turned out to be a nice opportunity for many of us from different institutions to meet regularly and to exchange information on research work in progress.

The present little volume summarizes in detail the work of last year. The inhomogeneity of the topics covered, ranging from the origin of meteorites to gamma ray astronomy, reflects the direct interests of the participants in the *Seminario*. The set of lectures not only makes a good introduction to several hot topics in Astrophysics, but also gives a good overall picture, although obviously incomplete, of the current research projects in Astrophysics in the Pisa area. The bibliography for each contribution usually contains reference to some papers in preparation, but also a few general references that may be considered a good introduction to the subject described in the lecture.

The hope is that, besides being a concrete way to summarize our efforts of last year, this little volume may serve as a useful guide for present and prospective undergraduate and graduate students, and even post-doctoral fellows, to the current activity in Astrophysics in the Pisa area. On the other hand, for a general external reader it may offer an unusual survey of hot topics and a description of several original contributions in Astrophysics, in a form that is clearly not available in standard textbooks or in specialized workshop proceedings.

Giuseppe Bertin

Pisa, January 1, 1995

SEMINARIO DI ASTROFISICA 1994

Detection of gravitational wave by means of laser interferometers

Stefano Braccini[1,2]

January 28

1. Scuola Normale Superiore, Pisa, Italy
2. Istituto Nazionale di Fisica Nucleare, Sezione di Pisa, Italy

Gravitational waves and their effects

– *What is a gravitational wave?* – [1,2,3]

In the language of general relativity, a gravitational wave is a perturbation of the space-time metric, propagating at the speed of light. The evolution of this perturbation can be described by Einstein's field equations which determine the form of the metric tensor $g_{\mu\nu}$. In general, the nonlinearity of Einstein's equations makes the evaluation of this tensor impossible; nevertheless, the gravitational field is often weak enough to allow a linearization of the problem. In fact, in the weak field approximation, the small perturbation of the metric can be expressed as a small deviation ($h_{\mu\nu} \ll 1$) from a flat space-time metric, which is described by the pseudo–Euclideian tensor $\eta_{\mu\nu}$=Diag(1,1,1,-1). Furthermore, Einstein's equations in vacuum, with a suitable choice of gauge, take on the form of a wave equation which governs the propagation of the perturbation $h_{\mu\nu}$. Any solution can be written as a superposition of travelling plane waves which, with another suitable choice of gauge, can be written in the following form:

$$h_{\mu\nu} = (h_+ e^+_{\mu\nu} + h_X e^X_{\mu\nu}) \cdot e^{i(kz-\omega t)} \qquad (1)$$

where h_+ and h_X are the amplitudes of the two independent states of polarization of the wave described by the tensors $e^+_{\mu\nu}$ and $e^X_{\mu\nu}$. The meaning of these two polarization states can be easily understood by looking at the displacements that the two waves '+' and 'X', propagating in the positive z direction, induce on a ring of free masses lying in the x–y plane (see Fig.1). The relative displacement of each pair of masses turns out to be proportional to their initial separation. Therefore, experimental physicists usually evaluate the strength of

a gravitational signal by measuring the relative change h(t)=ΔL(t)/L of the distance (L) between two free masses. This quantity h is called the "dimensionless amplitude" of the gravitational wave.

– Sources of gravitational waves: expected signals – [2,3]

General relativity predicts that an accelerated system of masses emits gravitational waves. The main difference with respect to the case of electromagnetism is that the lowest order radiation is now quadrupolar and not dipolar. The strongest gravitational signals which can reach an observer are expected to be produced by astrophysical processes in which high velocity coherent motions of large masses take place. These astrophysical sources of gravitational waves can be distinguished as belonging to three main categories: burst sources, periodic sources, and stochastic background[1].

Burst sources produce gravitational signals for a short time only (lasting a few cycles). This category includes coalescing compact binaries (black holes and neutron stars), infalls of massive objects into black holes, supernovae, and so on. For instance, a typical supernova in the Virgo cluster is expected to radiate a gravitational impulse of a few milliseconds, having a dimensionless amplitude h, at the Earth, of 10^{-22} - 10^{-20}.

Periodic sources, such as pulsars and binary neutron stars, are expected to emit gravitational waves continuously[2]. Some thousands of pulsars in our galaxy are expected to produce signals in the frequency range from a fraction of a Hz to about one hundred Hz, with dimensionless amplitudes h, at the Earth, ranging from 10^{-28} to 10^{-24}.

Stochastic background is a superposition of random signals, coming from random directions at all frequencies[3]; since its amplitude is expected to be well below the detection threshold of the modern ground–based detectors this background is not interesting from an experimental point of view.

– Detection of gravitational waves via interferometric antennae – [4]

An ideal instrument for measuring the small displacements generated by a gravitational wave (see Fig.1) is a Michelson interferometer whose optical components are suspended. If the two arms of the interferometer are lying along the x and y axis, as can be easily understood by the schemes of Fig.1 and Fig.2, the incoming gravitational wave, with '+' polarization, will alternatively lengthen one arm of the interferometer and shorten the other. A detection of the gravitational wave can be provided by a measurement of the phase shift, $\Delta\phi$(t), arising from the variation ΔL(t) of length in the two arms:

[1]– The predictions which follow depend crucially on the adopted theoretical models and they are affected by rather large uncertainties (a few orders of magnitude).

[2]– The periodicity of these signals could favour their detection despite the very low amplitudes which are involved.

[3]– This background is produced by the radiation produced in the early evolution of the Universe, and by all the gravitational signals which come from unknown sources.

$$\Delta\phi(t) = \frac{4\pi\Delta L(t)}{\lambda} = \frac{4\pi h(t)L}{\lambda} \qquad (2)$$

where λ is the laser wavelength and L is the length of the two arms. One has to notice that the small phase shift to be detected is proportional to the length of the arms and to the dimensionless amplitude of the gravitational wave. Therefore, an interferometer with very long arms is desired for the purpose.

– The required sensitivity

In order to evaluate the sensitivity that the interferometric antenna has to have, let us consider the typical signal emitted by a supernova in the Virgo cluster. As described above, this signal is an impulse of a few milliseconds with dimensionless amplitude h ranging from 10^{-22} to 10^{-20}. If we refer to a 3 km arm interferometer and a laser wavelength of about 1 μm, the last equation shows that the passage of such a signal should induce an impulsive displacement of the mirrors of about 10^{-19} - 10^{-17} m and therefore an impulsive phase shift of 10^{-12} - 10^{-10} radians. It is clear that, in spite of its large size, the interferometer has to be characterized by an extremely high sensitivity.

How to develop a high sensitivity antenna

– Fabry-Perot interferometers – [4]

A way for increasing the phase shift generated by the gravitational signal is to increase the optical path of the light in the two arms of the interferometer. This result can be attained by allowing the light to bounce back and forth several times along the two arms before recombination on the beam splitter. In this way, the phase shift $\Delta\phi$ is magnified by the number of reflections made by the light in each arm. As shown in Fig.2, this increase of the optical path can be obtained by replacing each arm of the interferometer with a Fabry–Perot cavity resonating at the laser frequency. In this way each photon goes back and forth in each arm an average number of times related to the Finesse F of the cavity[4]; with this technique the relevant phase shift can be amplified by a factor F/π.

– Noise in the antenna – [4]

Several noise processes can generate spurious signals in the antenna, masking the effect induced by the gravitational wave. Some of these processes, such as seismic noise and thermal noise, generate displacements of the mirrors ("displacement noise"), while others, such as the noise induced by the frequency fluctuations of the laser, affect the phase of the optical rays even if a real

[4]– Typical values of the Finesse of the Fabry–Perot range, in interferometric antennae, from a few tens to about 100.

movement of the mirrors is not present ("phase noise"). In order to define the sensitivity of the antenna it is necessary to compare the real signal $h_{GW}(t)$ to each relevant fake signal $h_{Noise}(t)$. This comparison is usually expressed in terms of the so called "linear spectral density" $\tilde{h}(f)$, which is defined as the square root of the power spectrum of the signal[5].

– *High power beam* – [5]

Poisson statistical fluctuations of the number of photons in the light beam generate a shot–noise on the interference signal; the linear spectral density of this noise decreases like $1/\sqrt{P}$ where P is the beam power. The beam power can be increased by "recycling" the light which is back–reflected from the interferometer, i.e. by using a semi–transparent mirror located in the position indicated in Fig.2 in order to build a resonating cavity between this mirror and the whole interferometer. With this technique, and with a very high power laser source (for example a 20 W Nd:Yag laser), an impinging power of 1 kW on the beam splitter can be obtained. In this condition shot–noise is white noise, with linear spectral density $\tilde{h}_{Sh-noise}$ expected to be close to $10^{-23}\ 1/\sqrt{Hz}$. Any other source of noise which affects the antenna is normally compared to this level of noise, called "shot–noise limit".

– *Stabilization of the laser frequency* – [6]

The static path difference between the two arms of the interferometer generates noise through frequency fluctuations of the laser source. In fact, because of an unavoidable asymmetry between the two arms, the fields which interfere at the beam splitter are originated from the laser at different times, and a fluctuation of the laser frequency is converted into a fluctuation of the phase difference $\Delta\phi(t)$. In order to keep this noise below the shot–noise limit, it is adequate to develop a very efficient system for the stabilization of the laser frequency ν (the goal is $\delta\tilde{\nu}(f) \leq 10^{-6}\mathrm{Hz}/\sqrt{Hz}$ at all the interesting spectral frequencies) – see [6].

– *The vacuum in the interferometer* – [7]

Fluctuations of the gas density in the two arms generate a variation in the refractive index, and therefore in the speed of light in the two cavities. Since these fluctuations are not correlated, they will produce a fluctuating phase difference and thus a spurious signal. For this reason the 3 km arms of the interferometer must be kept under vacuum. As described in reference [7], a pressure of 10^{-8} mbar in the vacuum pipe which surrounds each arm of the interferometer is required.

[5]– In the case of dimensionless amplitudes, the linear spectral density can be therefore expressed in units of $1/\sqrt{Hz}$.

Low frequency detection

– *The main problem: the seismic noise* – [8]

By improving on the techniques described above, it is possible to develop an antenna of very high sensitivity able to detect signals at high frequencies (above a few tens of Hz). At lower frequencies the detection is hindered by seismic noise, which can cause a spurious vibration of the optical components of the interferometer. Measurements of the spectrum of the seismic noise have been performed by many groups and the linear spectral density of the seismic displacement of a point on the ground is found to be of the same order of magnitude in all directions and to vary, to a good approximation, as A/f^2 (where A ranges from 10^{-9} to 10^{-6} $\mathrm{mHz}^{3/2}$, depending on the location of the measurement). Since many gravitational sources, such as pulsars and coalescing binaries, are expected to emit low frequency signals (between a fraction of Hz and some tens of Hz) it is mandatory to develop an antenna able to cover also this frequency range. In order to keep the movement of the suspended optical components well below the displacements generated by the gravitational signal, it is necessary to design a suspension system able to attenuate the propagation of the seismic vibrations from the suspension points to the optical components.

– *The "superattenuator": an efficient suspension system* – [8,9]

The VIRGO project, approved by French and Italian authorities in September 1993, is the only interferometric antenna which incorporates in its design a suspension system providing a strong reduction of seismic noise both in the vertical and in the horizontal directions. The basic idea is to use as a suspension a multiple-stage pendulum ("superattenuator") with the masses supported by springs[6] (see Fig.2 of Losurdo, this volume). The superattenuator is able to isolate the optical component from seismic movement of the ground in a wide frequency band, provided that the resonances are below a few Hz.

– *Conclusions*

In Fig.3 the expected sensitivity of VIRGO is presented: a comparison with the expected sensitivity of the American interferometric antenna (LIGO) demonstrates the effectiveness of VIRGO at extending the detection band to low frequencies.

[6]– In a pendulum the motion of the suspension point, at frequency much higher than the resonance frequency, is trasmitted to the suspended mass reduced by a factor proportional to f^2. In an N-stages pendulum this frequency dependence becomes f^{2N}.

Fig.1 – Effects of the polarization states '+' and 'X' of a gravitational wave, propagating in the positive z direction, on a ring of free masses lying in the x–y plane.

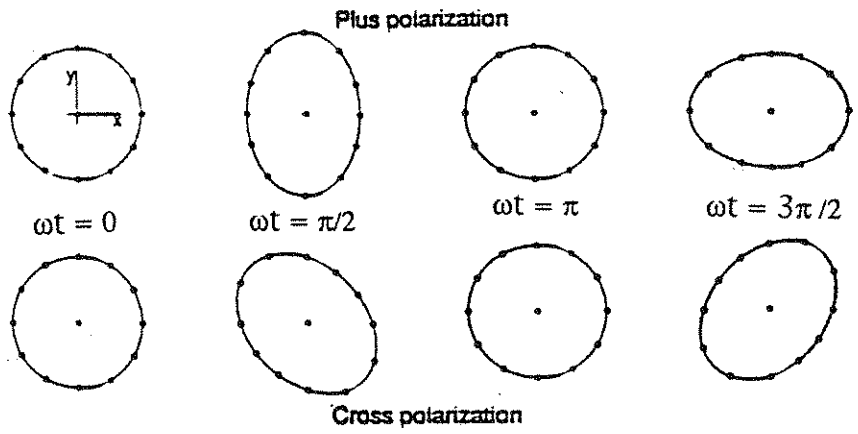
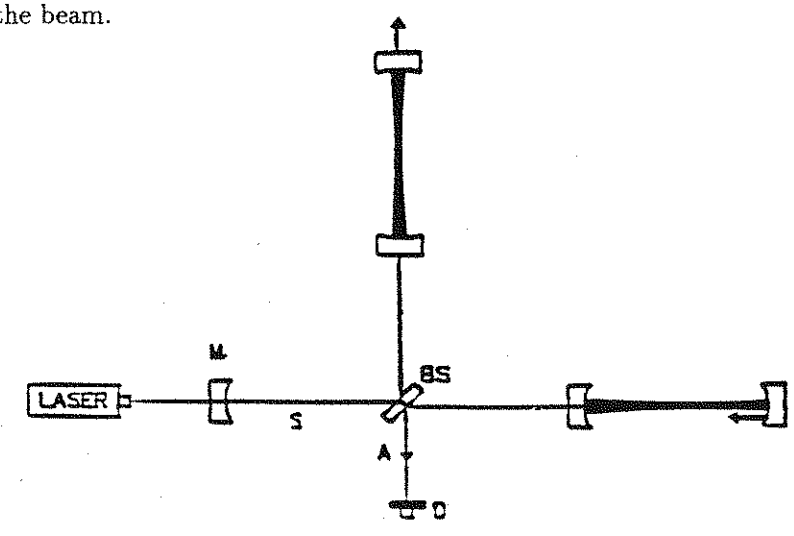


Fig.2 – Effect of the gravitational wave, propagating in the positive z direction and polarized '+', on an interferometer whose suspended arms are built along the x and y axis: alternately one arm is shortened while the other is lenghtened. The resulting phase shift oscillation is detected by the photodiode D which measures the interference signal. A Fabry–Perot cavity is inserted in each arm of the interferometer in order to increase the optical path lengths. The mirror M is the so called "recycling mirror" used for increasing the power of the beam.



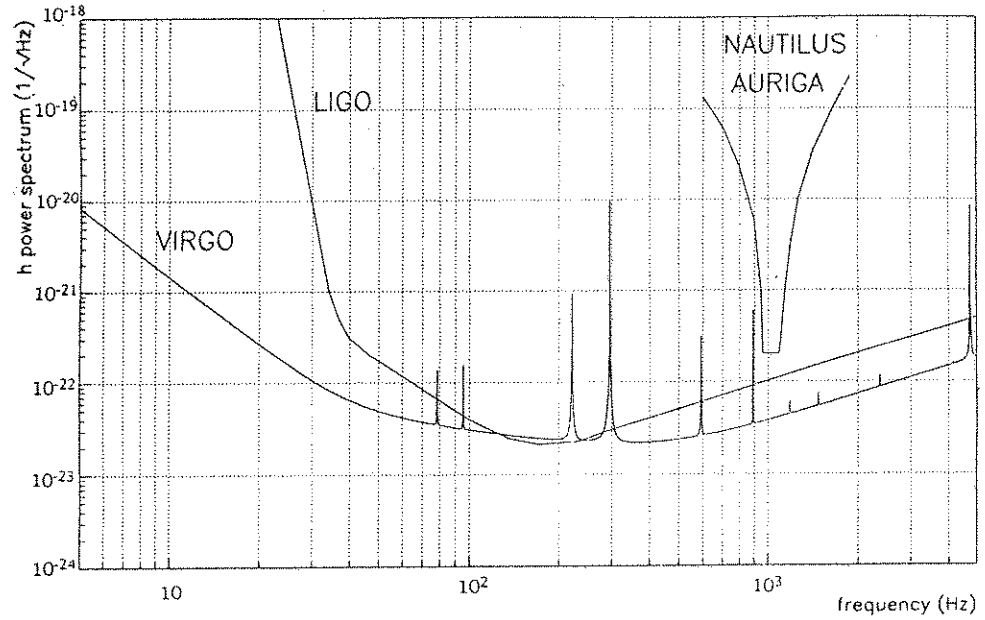


Fig.3 – Expected sensitivities of VIRGO and LIGO expressed as the linear spectral densities of the dimensionless amplitudes of the noise ($\tilde{h}(f)$).

References

1. Misner, C.W., Thorne, K.S. and Wheeler, J.: 1973, *Gravitation*, W.H.Freeman, S.Francisco.

2. Thorne, K.S.: 1987, "Gravitational radiation" in *300 years of gravitation* (S.W.Hawking and W.Israel, Eds.), **9**, 330–458, Cambridge University Press, Cambridge.

3. Shapiro, H.V. and Teukolsky, S.A.: 1984, *The physics of compact objects*, J.Wiley and Sons, New York.

4. Drever, R.W.P. et al.: 1991, "Fabry-Perot cavity gravity-wave detectors", in *The detection of gravitational waves* (D.G.Blair, Ed.), **12**, 306–328, Cambridge University Press, Cambridge.

5. Vinet, J.Y., Meers, B., Man, C.N. and Brillet, A.: 1988, *Phys. Rev. D*, **38–1**, 433–447.

6. Hough, J. et al.: 1991, "The stabilisation of lasers for interferometric gravitational wave detectors", in *The detection of gravitational waves* (D.G.Blair, Ed.),**13**, 329–352, Cambridge University Press, Cambridge.

7. Brillet, A.: 1985, *Ann. de Phys.*, **10**, 219–226.

8. Giazotto, A.: 1989, *Phys. Lett. A*, **182–6**, 367–424.

9. Del Fabbro, R. et al.: 1988, *Phys. Lett. A*, **132–5**, 237–240.

Massive black holes inside galaxy cores

Maurizio Cipollina*

February 4

* Scuola Normale Superiore, Pisa, Italy

Introduction

The complex phenomenology of Active Galactic Nuclei, or AGNs – i.e. quasars, radio galaxies, Seyfert galaxies etc. – is unified by a firmly established point: the power supply appears to be primarily of gravitational origin. AGNs are characterized by a powerful emission, up to 10^{46} erg/s on the whole spectrum (for quasars, if the emission is taken to be isotropic), huge amounts of energy (the lobes of radio loud galaxies store some 10^{61} erg), relativistic plasma motions, in the form of collimated jets emanating from the nucleus, and rapid variability – over periods of days, or even hours (cf. Stiavelli, this volume). As soon became clear, such a scenario is not compatible with a purely nuclear power source. The mass to energy conversion efficiency of nuclear reactions is at most 0.007, and 10^{61} erg produced by stellar processes, in a region as small as a few light hours, requires such a mass concentration that the corresponding gravitational energy turns out to be predominant [7]. AGNs can thus be modeled as the result of gas accretion onto a central growing black hole [8] (BH hereafter), with mass to energy conversion efficiencies of the order of 0.1. The relativistic motions inside jets, as revealed by the phenomenon of superluminal apparent motions of their features, also imply the existence of a relativistic potential well.

Even if accretion is only one of several phenomena that can be argued to characterize galactic nuclear activity (alternatives are, e.g., dense stellar clusters, massive stars, clusters of neutron stars etc.), the ultimate fate of a tightly bound gravitational system is gravitational collapse [8]. Thus, the endpoint of galactic nuclear activity has to be a central massive BH. The mass of the central BH, estimated either from the Eddington luminosity of quasars, or from the observed energy output of AGNs, is of the order of $10^6 - 10^8\ M_\odot$ ($M_\odot$ denotes the solar mass). On the other hand, the quasar number

density increases with redshift (i.e., as we look farther from us and back in time from the present epoch), showing that a significant fraction of the galaxies underwent, in some phase of their evolution, an active phase. This fact has a far reaching consequence. Most of the galaxies close to us should host, as a remnant of their past nuclear activity, a massive BH. A whole generation of astronomers has striven in the effort of detecting such BHs, identifying several good candidates (M31, M32, M87, NGC 3115, NGC 4594, NGC 1399 etc., see e.g. [9] for a review and specific references) but, so far, no definitely convincing evidence. [1]

Dynamical detection of black holes

The presence of a massive BH can be revealed through its dynamical influence on the stars around it. Stars orbiting close to a BH move fast, and their velocity (directly measured through the shift and the widening of spectral lines) acts as a probe of the mass distribution. In particular, for a spherically symmetric non rotating system with a central BH of mass $M_\bullet$, one can refer to the Jeans equation [1]

$$\frac{G\,M(r)}{r} + \frac{G\,M_\bullet}{r} \;=\; -\frac{d\ln(\rho c_r^2)}{d\ln r}\,c_r^2 + (c^2 - 3c_r^2),$$

where r is the galactocentric radius, $c(r)$ and $c_r(r)$ are the total and the radial velocity dispersion of stellar motions respectively, $M(r)$ and $\rho(r)$ are the cumulative mass and the mass density of stars. Note that ρ is actually given, up to a multiplicative factor, by the observed light density, if one assumes the mean mass to light ratio of the stellar population to be constant with radius. For small r, stars are bound to the BH and move on Keplerian orbits, so that

$$c^2 \;\sim\; G\,M_\bullet/r,$$

and we expect a steep rise in velocities in the inner region (Fig. 1). If the observational data of a spherical galaxy core are found to require $M_\bullet \neq 0$, the presence of this unresolved amount of central dark mass is ascribed to a central BH (the *actual* detection of a BH, which would be based on observations of relativistic star velocities, or extreme redshifts, is far more difficult).

In spite of the apparent simplicity of the equations governing the equilibrium of a stellar system, the dynamical detection of BHs is complicated by several factors:

- The observable quantities are only mean values along the line of sight, i.e. projected onto the sky. Since just very few stars close to the black hole move fast, a strong increase in the *projected* velocities is not to be expected.

[1] During the preparation of this volume an important discovery for the galaxy NGC 4254 has been announced by Miyoshi et al. (1995, *Nature*, 373, 127).

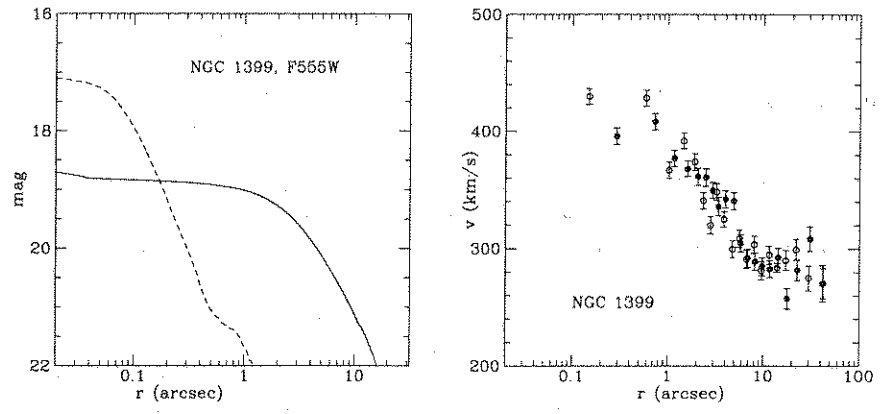


Figure 1: Photometry and kinematics of NGC 1399. The left frame shows the light profile (solid line) as measured with the HST, while the dashed line is the point spread function (i.e. the image of a point source at $r = 0$). Note the steep central rise in velocities, and the power-law increase in density inside the core. The data come from [10].

- Radial orbits can produce a rather steep sky-projected velocity dispersion profile mimicking the presence of dark mass. A similar effect is achieved by a suitably oriented non spherical star distribution. Note, however, that the orbital configuration cannot be taken arbitrarily; for example, if a maximum allowed fraction of radial orbits is exceeded, dynamical instability of the core takes place.

- *Seeing* (i.e., the effect of the atmosphere that spreads a point source over a region of 0.2 - 0.3 arcseconds in the best cases, and of a few arcseconds under unfavorable atmospheric conditions) strongly limits the resolution of ground-based observations. The Hubble Space Telescope has been limited, before being repaired with COSTAR in December 1993, by spherical aberration (Fig. 1).

- Often, the 'dark mass' of a BH is not dark at all. Most candidate cores show a residual nuclear activity, with central point sources that, as their spectra reveal, are found to be of non-stellar origin. In these cases, the point sources make it difficult to observe the stellar component in the innermost regions.

- Kinematical measurements require the resolution of spectral lines as a function of radius, with long integration times and difficult data reductions. On the other hand, a purely photometric detection of a BH is not possible, since, in the absence of significant kinematical constraints, the

observed light distributions can be modeled, in general, without a BH.

Black hole growth as a stellar dynamical problem

The quest for massive BHs can follow two complementary strategies. On the one hand, one can observe the closest and best resolved galaxy cores, hoping to find one or few among them with an unquestionable detection of a central dark mass. On the other hand, one can pursue a theoretical investigation of how the dynamics of a core with a BH should be, taking into account the evolution of both the AGN and of the host galaxy, thus obtaining constraints or predictions to be tested against a statistically significant sample of galaxies. Since the theory of AGNs is far from being complete, and we are just beginning to clarify some of the issues involved in galaxy evolution, no firmly established model of a core with central BH is available, so that we are forced to proceed by hypotheses. A simple and reasonable set of assumptions about BH growth at the center of a galaxy is given by the adiabatic hypothesis (see [11] for a more quantitative discussion):

- The BH grows on a timescale longer than the dynamical time of the core (the timescale of star orbits)

$$t_{dyn} \sim 1\ Myr$$

(this is the case unless the mass accretion exceeds by far the Eddington limit [8]), but shorter than the two-body relaxation time

$$t_{rel} \sim \frac{N}{logN}\, t_{dyn},$$

where N is the number of stars inside the core [1]. For time intervals shorter than t_{rel} the stellar system behaves as *collisionless*, and can be approximated as being continuous rather than a discrete distribution of point masses. The properties of the observed cores are found to be consistent with this hypothesis.

- While the BH grows, most of the core mass is in the form of stars, the gas component being negligible.

- The core is spherically symmetric, and the BH stays immobile at the center of the system. Since most galaxies are not spherical, this symmetry assumption should be regarded as a starting point rather than as a physical requirement. Note that motions of the BH are probably suppressed by dynamical friction.

- The mass accreted by the BH comes from the 'outside', in the form of infalling gas, or by capture of stars on radial orbits.

- The initial core is a non-singular isothermal sphere – an isotropic equilibrium model with a flat core, and a constant velocity dispersion (see Fig. 2, case $M_\bullet = 0$, and [1]). Such an initial distribution is suggested by the theory of violent relaxation, and confirmed by N-body simulations.

- The galaxy did not undergo any merging with other galaxies during its evolution. Since many cosmological theories consider merging as a frequent phenomenon, one should be cautious about this latter point.

Even if the adiabatic hypothesis might seem oversimplified, it provides at least a recipe to compute self-consistent models of a core with central BH, with the advantage of being supported by some physical framework.

Given these assumptions, the determination of the final star distribution after the BH has grown, becomes a non trivial and interesting dynamical problem, whose solution follows from the constraints of self-consistency and of adiabatic invariance. Let f be the phase space distribution of stars, so that $\int_D f(x,v)\, d^3x\, d^3v$ equals the mass of stars with positions and velocities (x,v) in the domain D. Each star moves in the potential $-G\,M_\bullet/r + \Phi(r)$, where Φ is the potential generated by the stars. In order for the system to be stationary, f has to be a function of the integrals of motion, i.e. the energy E and the angular momentum L (Jeans theorem [1]). In addition, f has to be self-consistent, so that the associated mass distribution actually generates Φ

$$\Delta\Phi = 4\pi G \int f\, d^3v. \qquad (1)$$

On the other hand, while the BH grows, each star orbits in a slowly varying potential, so that its adiabatic invariants L and I_r are conserved. The radial action I_r is given by

$$I_r(E,L) = \frac{1}{\pi}\int_{r_-}^{r_+} \sqrt{2[E-\Phi(r)+GM_\bullet/r] - L^2/r^2}\; dr$$

where r_+ and r_- are the radial turning points of the orbit. As a consequence, f turns out to be, as $M_\bullet$ increases, a *fixed* function of (I_r, L),

$$f(E,L,M_\bullet) \equiv g(I_r,L) \qquad (2)$$

where g is given by the initial conditions at $M_\bullet = 0$. Since I_r does depend on Φ, Eqs. (1) and (2) are a coupled system of non linear equations that can be solved for the unknowns f, Φ.

The adiabatic invariants are known only for integrable systems, and non spherical potential are generally *not* integrable: as a consequence, the adiabatic growth problem in flattened systems is far more difficult, and has not yet been solved (see [2] for a discussion; small ellipticity models, with potential approximated as spherical, are found in [6]).

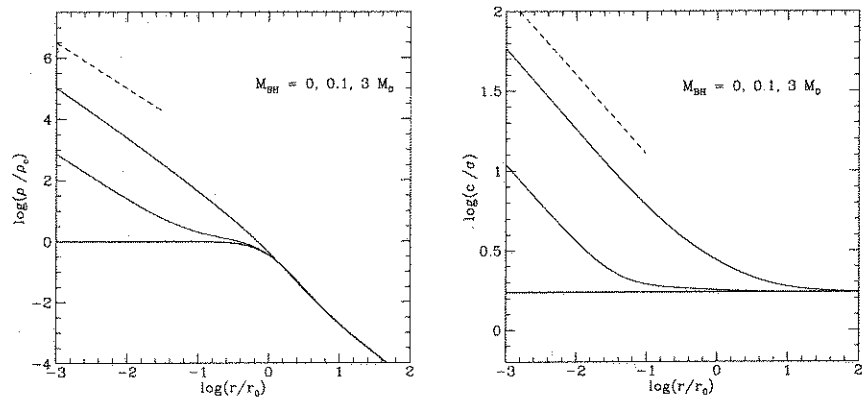


Figure 2: Intrinsic density and velocity dispersion profiles for the adiabatic models of a core with central BH, for $M_\bullet/M_0 = 0, 0.1, 3$ in order of increasing ρ and c; note that $M_\bullet/M_0 = r_\bullet/r_0$. The dashed lines are the power laws $r^{-3/2}$ (left) and $r^{-1/2}$ (right).

Adiabatic cusp models

The adiabatic growth equations, starting from an isothermal core, can be solved numerically or through analytical approximations (see [11], [3], [4]). The initial core is characterized by a central density ρ_c, a constant one–dimensional velocity dispersion σ and a radius

$$r_0 \;=\; \left(\frac{9\sigma^2}{4\pi G\rho_c}\right)^{1/2} \;=\; \frac{GM_0}{\sigma^2}$$

where the core mass is defined by $M_0 = (4\pi/9)\rho_c r_0^3$. The black hole models are characterized by a cusp, both in the density and in the velocity dispersion profile, which extends out to the black hole influence radius

$$r_\bullet \;=\; \frac{GM_\bullet}{\sigma^2}$$

as illustrated in Figure 2. For $r < r_\bullet$ the cusp is approximated by the power laws

$$\rho \;\sim\; \rho_c \left(\frac{r}{r_\bullet}\right)^{-3/2}, \qquad c \;\sim\; \sqrt{\frac{GM_\bullet}{r}}$$

where $c = [(1/\rho)\int f v^2\, d^3v]^{1/2}$ is the total velocity dispersion (so that, for $M_\bullet = 0$, $c^2 = 3\sigma^2$).

Several galaxy cores are found to be compatible with a black hole density cusp [5]. However, when compared to high resolution kinematical observations,

attention is restricted to those with a candidate BH – is too complex to fit the one parameter family of adiabatic growth models. By releasing some of the basic assumptions, e.g. starting form an initial non isothermal core, or growing the BH with part of its mass subtracted to the core stars rather than all coming from the outside, and eventually taking into account rotation, flattening or even triaxiality, one would obtain a wider family of models that might successfully fit the observations. This program is challenging from the point of view of theoretical and computational stellar dynamics, and will probably require a new generation of computers, algorithms, and methods. Further understanding of massive BHs inside galaxy cores should be accompanied by a much deeper knowledge about galactic formation, evolution, and activity, i.e., by a satisfactory physical framework.

References

1. Binney, J., Tremaine, S. 1987, *Galactic Dynamics*. Princeton University Press, Princeton.

2. Cipollina, M. 1992, Tesi di Laurea, Università di Pisa.

3. Cipollina, M., Bertin, G. 1994, *Astron. Astrophys.*, **288**, 43.

4. Cipollina, M., 1995, *Astron. Astrophys. Suppl.*, **110**, 1.

5. Crane, P., Stiavelli, M., King, I. R., Deharveng, J. M. et al. 1993, *Astron. J.*, **106**, 1371.

6. Lee, M. H., Goodman, J. 1989, *Astrophys. J.*, **343**, 594.

7. Lynden–Bell, D. 1967. *Month. Not. Royal Astr. Soc.*, **136**, 101.

8. Rees, M. J. 1984, "Black hole models for active galactic nuclei", *Ann. Rev. Astron. Astrophys.*, **22**, 471.

9. Richstone, D. 1993, *Ann. NY Acad. Sci.*, **688**, 716.

10. Stiavelli, M., Möller, P., Zeilinger, W. W. 1993. *Astron. Astrophys.*, **277**, 421.

11. Young, P. J. 1980, *Astrophys. J.*, **242**, 1232.

Cooling flow models for the X–ray emission from elliptical galaxies

Thomas Toniazzo*

February 11

*Scuola Normale Superiore, Pisa, Italy

X-ray observations with the *Einstein* observatory [9] first revealed the presence of extended X-ray halos around elliptical galaxies. It is generally agreed that they represent the emission of a hot, tenuous, diffuse ("coronal") gaseous component, of a total mass of the order of $10^{10}M_{\odot}$, which provides most of the X-ray emission in the X-ray-bright elliptical galaxies [8,21,23], with X-ray luminosities of $10^{41} \div 10^{42}$erg s^{-1} [11]. The emission profiles measured by the X-ray imaging detectors in these galaxies have no cut-off at galactocentric distances large with respect to the optical image, while in the inner parts they tend to be similar to the optical light profiles [23,24].

The relevant mechanisms responsible for the observed emission depend on electron-ion collisions (free-free and free-bound processes, one- and two-photon radiative decay of collisionally excited lines in Hydrogen- or Helium-like ions), so that, for a given chemical composition, the total radiated power per unit volume is proportional to the particle number densities squared, or, equivalently, to the mass density ρ^2. The proportionality factor is a function of the gas temperature, the so-called 'cooling function' $\Lambda(T)$ [14]. The coronal gas has typical parameters $kT \sim 1$ keV and $n \sim 10^{-2}$ cm^{-3} and characteristic cooling times at the centre of gas-rich galaxies are $3p/2\rho^2\Lambda(T) \sim 10^9$yr [21]. Possible sources of gaseous material are found within the galaxy, in the form of stellar winds and nova or supernova explosions [15], or outside it as an intergalactic medium (perhaps of cosmological origin [10]) which gets captured by the galactic gravitational potential well.

The basic, large-scale dynamics of the gas is determined by its energy budget, resulting from the radiative cooling and the stellar mass and energy injection rate. If the energy sources (primarily supernovae) are powerful enough, the gas within the galaxy gets heated and a low density, supersonic galactic wind forms. If, conversely, cooling is important, the gas slowly sinks inwards, towards the galactic centre, as the inner (higher-density) layers cool

down, approximatively maintaining hydrostatic equilibrium. In this latter case we speak af a "cooling–flow". Observationally, the two cases differ significantly in the resulting X-ray surface brightness, a cooling–flow being a strong, a wind a weak, X-ray source. Actually, a fairly large spread in the ratio of X-ray to optical luminosity L_X/L_B for elliptical galaxies is observed.

This dynamical description is complicated by the possible occurrence of instabilities of thermal type [13,12] and of transsonic flow velocities [13,18]. Moreover, the relevant parameters for the flow (especially the injection temperature of the gas) should evolve in time, which may give rise to significant time-dependent effects [7], doing the 'switching' from a cooling-flow to a wind regime. However, this idea is not confirmed [25] by any correlation of the ratio L_X/L_B with intrinsic optical properties of the galaxies. The possibility interaction of the the gas flow with a central black hole may be more promising [5].

The gas plays no significant role for the galactic dynamics, which is dominated by the stars and possibly by dark matter [16]. On the other hand, it can be a precious diagnostic tool for the galactic gravitational potential $\Phi(r)$ out to scales larger than those typical of the luminous matter. Furthermore, the X-ray emission spectra contain information on the heavy element abundances $[X_i]/[H]$ in the gas, which, if it is of stellar origin, can tell us something about the nucleosynthesis in elliptical galaxies.

These ideas raise the wish for spatially resolved X-ray spectra. In principle, from comparison with synthetic spectra with different input parameters (temperature, abundances and neutral hydrogen column density), and integrated over a gasdynamical model, one can find best-fit values for the $[X_i]/[H]$'s and the true amount of cooling in the gas. Therefore, a detailed study of the hot gas flows in elliptical galaxies can help providing useful constraints on the much-debated subjects of the amounts of Dark Matter and of the metal abundances associated with the stellar populations of early-type galaxies. Much work has been done in these directions in the last ten years [3,8,19,20,23,24]. However, a satisfactory and simple gasdynamical model, which should be self-consistent, in agreement with presently published data, and having some predictive power, is still missing [7,18].

Recently [4], the steady-state, spherically symmetric cooling–flow description has been tested against good quality X-ray data available today, by including the available stellar dynamical data [3,16], and by considering the possibility of a decoupling of cool gas from the bulk of the flow [17]. We have focussed on a set of 5 elliptical galaxies (NGC 1399 and NGC 1404 in the Fornax group, and NGC 4374, NGC 4472 and NGC 4636 in the Virgo cluster), for which detailed X-ray emission profiles and X-ray temperatures are published[1],

[1]After submission of our paper [4], *ASCA* spectral data for NGC1404, NGC4374 and NGC4636 have been published (Loewenstein et al. and Mushotzky et al. 1995, *Astrophys. J. Letters*, **436**, L75 and L79), generally consistent with our findings.

and for which we can refer to a specific stellar dynamical model [16].

A cooling–flow model with gas–mass loss

We describe the hot coronal gas by means of a set of equations under the assumption of a steady-state, spherically symmetric inflow in the presence of mass and energy source terms (related to mass and energy injection in the form of hot gas originating from stars and supernovae), of radiative losses (responsible for the observed X-ray emission), and of a mass sink term (describing the decoupling of cold gas from the hot phase [17]):

$$\frac{1}{r^2}\frac{\mathrm{d}}{\mathrm{d}r}(r^2\rho u) = \alpha\rho_\star - q\frac{\rho}{\tau_{dec}} \tag{1}$$

$$\rho u\frac{\mathrm{d}u}{\mathrm{d}r} + \frac{\mathrm{d}p}{\mathrm{d}r} = -\rho\frac{\mathrm{d}\Phi}{\mathrm{d}r} - \alpha\rho_\star u \tag{2}$$

$$\frac{1}{r^2}\frac{\mathrm{d}}{\mathrm{d}r}\left[r^2\rho u\left(\frac{1}{2}u^2 + \frac{5}{2}\frac{p}{\rho} + \Phi\right)\right] = $$

$$= -\rho^2\Lambda(T) + \alpha\rho_\star\left(\epsilon_{inj} + \Phi\right) - q\frac{\rho}{\tau_{dec}}\left(\frac{1}{2}u^2 + \frac{5}{2}\frac{p}{\rho} + \Phi\right). \tag{3}$$

Three unknown functions ρ, p, and u (mass density, pressure, and radial velocity) describing the hot, X-ray emitting gas are to be derived by solving the above equations under the appropriate boundary conditions (see below). The gas temperature T is obtained from the equation of state $p = nkT$, n being the number density and k Boltzmann's constant. We are ignoring the contribution of the possible presence of magnetic fields and of various transport processes within the gas. Let us call "q-models" the solutions of this set of equations.

The three functions $\rho_\star(r)$, $\sigma_\star(r)$, and $\Phi(r)$ are the *stellar* mass density, the one-dimensional stellar velocity dispersion (averaged over the three directions), and the gravitational potential of the galaxy (with the contribution of stars and dark matter, but, for simplicity, *without* the contribution of the gas). They were evaluated for each astronomical object using Saglia et al.'s [16] corresponding best-fit ("$f_\infty + f_\infty$") stellar dynamical model.

We take a "cosmic" chemical composition of the gas [1], with a mean mass per particle $\mu = 0.63$ amu, so that $n = \rho/\mu$. The stellar mass loss rate is [15] $\alpha = 4.75 \times 10^{-19}(M_\star/L_B)^{-1}$ sec^{-1} (with the mass-to-light $M_\star/L_B$ ratio expressed in solar units). The specific energy of the injected material ϵ_{inj} includes a contribution due to disordered stellar motions $3\sigma_\star^2/2$ and that of an effective temperature associated with the hot supernova ejecta mixing into the gas. For the supernova rate we test the three values of 0.245 [2], 0.055 [6] and 0.0137 SNIa per century per $10^{10}\mathrm{L}_{\odot B}$.

The cooling function $\Lambda(T)$ is obtained by numerical interpolation of the integral over a "Raymond" thermal spectrum [14], which is suitable for an optically thin thermal plasma at temperatures $10^5\ ^0\mathrm{K} < T < 10^8\ ^0\mathrm{K}$.

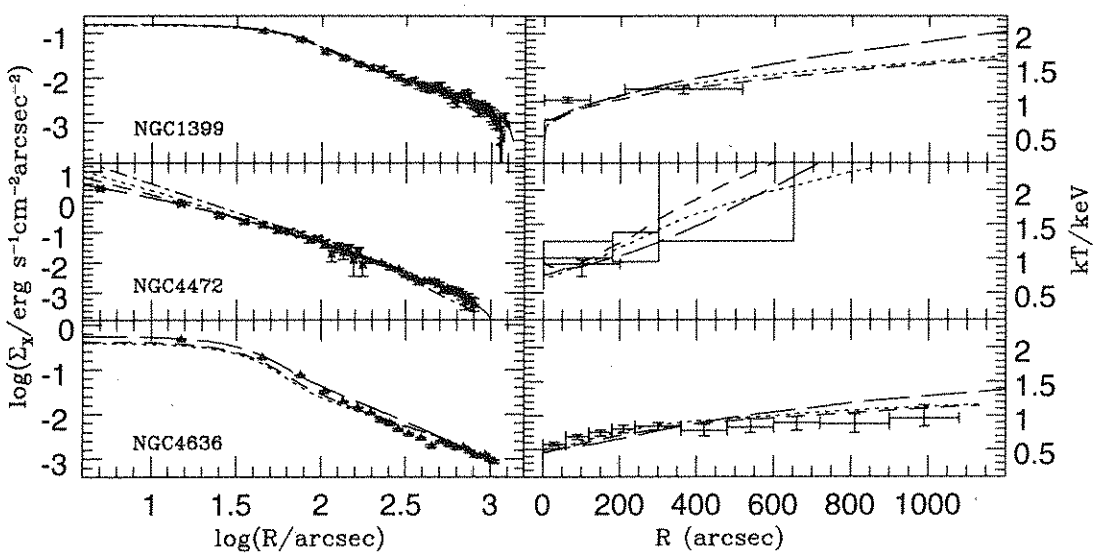


Figure 1: Emission and temperature profiles

If we want to derive $\Phi(r)$ using 'observed' $T(r)$ and $\rho(r)$, we must solve equations (2) and (3) for $\Phi(r)$ *and* $u(r)$, obtaining also the gas mass flux $\dot{m}(r) \equiv -4\pi r^2 \rho u$ (positive for inflow). This exercise leads to the result (well known in cluster cooling–flows) that $\dot{m}$ increases with r [20,21]. This is the reason for the introduction of the empirical mass–loss term in Eq. (1). Its form follows from the growth of linear thermal instabilities due to the fast cooling of the gas when its density increases [12,13,22], with timescale

$$\tau_{dec} = \frac{5}{2}\frac{p}{\rho^2 \Lambda(T)}\left(2 - \frac{\mathrm{d}\,\ln\Lambda}{\mathrm{d}\,\ln T}\right)^{-1} . \qquad (4)$$

The dimensionless parameter q is a free parameter of the model. In general, for any reasonable choice of $\Phi(r)$, it is well-known that models with $q = 0$ give too steep X-ray emission profiles [3,18]. Two additional free parameters are the boundary conditions $\dot{m}(r_{ext}) = \dot{m}_{ext}$ and $p(r_{ext}) = p_{ext}$ at the outer edge of the galaxy, determining the amount of non-stellar material accreted and the external confining pressure. (The third boundary condition is determined by the requirement of regularity of the solution at the sonic point [18]).

Figure 1 displays the emission and temperature profiles computed for the models selected for NGC 1399, NGC 4472, and NGC 4636 with the three different SNIa rate assumptions (dotted lines corresponding to the lowest rate, and long-dashed to the largest), compared to observational data (errorbars and

boxes [11,19,21,24]). The Table reports the model parameters (in ascending order of the SNIa rate) and the model total gas masses. In the plot of the emission profiles of NGC 4472 the dash-dotted line shows the model curve of Reference 3. Calculated emission profiles and temperatures in these models compare generally better with the observations than previously obtained. It turns out that a SNIa rate as suggested by the most recent analyses [6] gives generally slightly better fits.

NGC	q	$\dot{m}_{ext}$ ($M_\odot$/yr)	p_{ext} (10^3 ^{0}K/cm^3)	r_{ext} (kpc)	M_{gas} ($10^{10}M_\odot$)
1399	0.8/0.6/0.4	4.4/4.1/3.3	8.2/7.6/9.9	200	19.7/19.5/22.6
4472	0.8/0.8/0.3	1.0/0.6/0.3	10.1/12.6/12.6	120	4.4/4.5/6.0
4636	0.5/0.4/0.2	3.0/3.0/1.8	4.0/3.8/4.1	160	7.9/8.2/9.1

While the value of q influences the inner parts of the model profiles (it turns out [22] that for large q the X-ray surface brightness is proportional to the optical one [23]), their steepness in the outer region and the total model X-ray luminosity depend on the parameters $\dot{m}_{ext}$ and p_{ext}. As to the last point, a correlation between $\dot{m}_{ext}/\alpha M_*$ (accreted to injected gas mass) and L_X/L_B is expected, as shown in Fig. 2, right panel (the dotted line is a linear regression). If the intergalactic gas has lower metal abundances than that ejected by stars, an anticorrelation between $[X_i]/[H]$ and L_X/L_B should be observed. Further, we find also, rather surprisingly, a tight correlation between the extension of the galaxy's dark halo and the amount of accreted material (the case of NGC 1404 is a bit peculiar since the X-ray halo of which is embedded in that of NGC 1399). In conclusion, within the frame of simple spherical steady-state cooling–flow models we have found that sufficiently accurate results can be achieved provided that accretion of external material and decoupling of cold gas clumps from the flow are included in the description. These simple models may help constraining the values for the pressure of the intergalactic medium and lead to (at least qualitative) predictions for the depth of the galactic potential well and the metal abundances of elliptical galaxies.

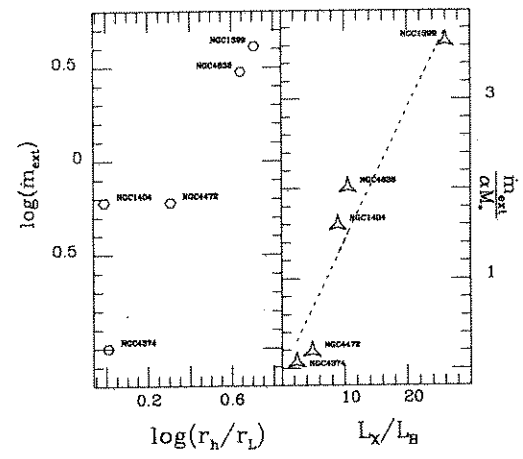


Figure 2: Model correlations

References

1. Allen, C.W. 1973, *Astrophysical Quantities*, Athlone, London
2. van den Bergh, S., Tammann, G. 1991, *Ann. Rev. Astron. Astrophys.*, **29**, 363
3. Bertin, G., Pignatelli, E., Saglia, R.P. 1993, *Astron. & Astrophys.*, **271**, 381
4. Bertin, G., Toniazzo, T. 1995, submitted to *Astrophys. J.*
5. Binney, J., Tabor, G. 1995, submitted to *Mon. Not. Roy. Astron. Soc.*
6. Cappellaro, E., Turatto, M., Benetti, S., Tsvetkov, D.Yu., Bartunov, O.S., Makarova, I.N. 1993, *Astron. Astrophys.*, **268**, 472
7. Ciotti, L., d'Ercole, A., Pellegrini, S., Renzini, A. 1991, *Astrophys. J.*, **376**, 380
8. Forman, W., Jones, C., Tucker, W. 1985, *Astrophys. J.*, **293**, 102
9. Giacconi, R., et al., 1979, *Astrophys. J.*, **230**, 540
10. Gunn, J.E., Gott, J.R. 1972, *Astrophys. J.*, **176**, 1
11. Kim, D.-W., Fabbiano, G., Trinchieri, G. 1992, *Astrophys. J. Supp.*, **80**, 645
12. Loewenstein, M. 1990, *Astrophys. J.*, **349**, 471
13. Mathews, W.G., Bregman, J.N. 1978, *Astrophys. J.*, **224**, 308
14. Raymond, J.C., Cox, D.P., Smith, B.W. 1976, *Astrophys. J.*, **204**, 290
15. Renzini, A. 1988, in: *Windows on galaxies* (G. Fabbiano, J.S. Gallagher, A. Renzini editors), Kluwer, Dordrecht, p. 255
16. Saglia, R.P., Bertin, G., Stiavelli, M. 1992, *Astrophys. J.*, **384**, 433
17. Sarazin, C.L., Ashe, G.A. 1989, *Astrophys. J.*, **345**, 22
18. Sarazin, C.L., White, R.E. III 1988, *Astrophys. J.*, **331**, 102
19. Serlemitsos, P.J., Loewenstein, M., Mushotzky, R.F., Marshall, F.E., Petre, R. 1993, *Astrophys. J.*, **413**, 518
20. Thomas, P.A. 1986, *Mon. Not. Roy. Astron. Soc.*, **220**, 949
21. Thomas, P.A., Fabian, A.C., Arnaud, K.A., Forman, W., Jones, C. 1986, *Mon. Not. Roy. Astron. Soc.*, **222**, 655
22. Toniazzo, T. 1993, Tesi di Laurea, Università di Pisa
23. Trinchieri, G., Fabbiano, G., Canizares, C.R. 1986, *Astrophys. J.*, **310**, 637
24. Trinchieri, G., Kim, D.-W., Fabbiano, G., Canizares, C.R. 1994, *Astrophys. J.*, **428**, 555
25. White, R.E., Sarazin, C.L. 1991, *Astrophys. J.*, **367**, 476

Nuclear activity in galaxies

Massimo Stiavelli[1,2]

February 18

1. Scuola Normale Superiore, Pisa, Italy
2. Space Telescope Science Institute, Baltimore, USA

The aim of this contribution is to review current observational results and ideas with the aim of illustrating a number of open problems in the study of Active Galactic Nuclei (hereafter AGNs). This field has reached a sort of maturity from the observational standpoint, and there are now a number of well defined issues that could be of interest for the plasma physicist as well as for the theoretical astrophysicist.

The plan of this contribution is the following. After introducing some basic classification criteria for AGNs in the general context of the so-called *Unified Schemes*, we will consider in more detail the synchrotron emission from radiogalaxies, illustrating the physical ideas which should underlie any model for these objects. Finally, we will describe some more recent theoretical and observational results.

Many good reviews have been written on AGNs and it would be hard to cite them all. A set of three good reviews covering all aspects of the problem can be found in a single volume of the Saas-Fee Lectures [2,12,19].

Many years of studies of AGNs have been necessary to arrange all the various classes and subclasses of AGNs into a coherent scheme. In the modern sense any galaxy characterized by nuclear activity significantly exceeding stellar-like sources at some wavelength of the electromagnetic spectrum may be included in the AGN category.

Optical observations of AGNs tend to look separately at the continuum and at the line emission. Broad band observations are adequate to establish the overall shape of the AGN continuum, while spectroscopic observations are required to determine the presence and the properties of emission lines. Most optical classification schemes rely on the properties of optical spectral lines. Indeed, most AGNs are characterized by a rich collection of emission

lines throughout their UV, optical, and near IR spectrum. Since optical spectroscopy is technically easier than spectroscopy in the ultraviolet or in the infrared, the rest-frame optical lines are the most studied in low redshift objects, while the rest-frame ultraviolet lines are the most commonly studied for high redshift objects.

It turns out that the profile of a generic AGN emission line is characterized by at least two components: a narrow component, with widths in the range of a few hundreds km/s, and a broad component, which may reach widths of several thousands km/s. Very often the broad lines have double-peaked profiles, indicating that the broadening is of non-thermal origin and is probably due to rotation. The strength ratio between these components may vary so much that often only one of these is observed.

Three broad classes of AGNs may be recognized, depending on their luminosity at optical wavelengths. AGNs with a power significantly lower than the host galaxy luminosity are classified as LINER (Low Ionization Narrow Emission Region). It is significant that no low power broad line analogues of LINERs are observed. Seyfert galaxies are instead objects characterized by an active nucleus with a brightness roughly comparable to that of the underlying galaxy. Seyfert I galaxies show broad emission lines in their spectra, while Seyfert II have spectra with narrow emission lines only. The typical luminosity of the AGNs in Seyfert galaxies are in the 10^{43} erg/s range. The last main category of AGNs is that of QSOs. In these objects the AGN is so bright that the galaxy is very hard to detect. The typical luminosity of QSOs is in the 10^{46} erg/s range, with the brightest being 10^4 times brighter. Such luminosities make QSOs the brightest known objects in the universe and, consequently, those that can be traced out to the farthest distances. All QSOs show the presence of both broad and narrow emission lines in their spectra.

Several other categories of AGNs are worth mentioning. One is that of radiogalaxies. These objects are often associated to giant elliptical galaxies (at least in the nearby universe where the underlying galaxy can be more easily detected) and are characterized by often inconspicuous nuclei in the optical (of the LINER type). However, it is at radio wavelengths that radiogalaxies reveal their nature, showing a range in luminosity going from that of Seyfert to that of QSOs. Radiogalaxies are classified as Fanaroff-Riley type I or II objects depending on the development and symmetry of their radio lobes. FR I radiogalaxies show irregular and non symmetric lobes, with often most of the radiopower concentrated in the central regions. In contrast, FR II objects (known also as *classical doubles*) show very extended and highly symmetric radiolobes radiating most of the object luminosity. Both kinds of objects may contain jets. The lobes of FR II galaxies are probably the most extended coherent structures in the universe, able to reach 4-5 Mpc in linear size, and contain regions, far from the nucleus and close to the radio lobe boundaries that are significantly brighter than their surroundings: the hot spots. The

fraction of QSOs which are radio-loud is comparable to the fraction of galaxies which are ellipticals.

All AGNs show some sign of variability. However, some do so on a grand scale. They are called BL Lacertae and OVV (Optical Violent Variable) QSOs. These objects have, respectively, the luminosities of an active galaxy or of a QSO, but are characterized by large variability and by a spectrum with a very large continuum showing almost no trace of emission lines.

Although additional subclasses have been proposed, the ones described above include most of the AGNs known.

Several *Unified Models* for AGNs have been proposed. Here we will focus on a Grand Unified (hereafter GU) version, aiming at including all the major known kinds of AGNs into the same basic picture. Although these models are still challenged by some, they appear to resist various statistical tests [14].

The main issue for bright AGNs is the nature of its energy source. Various arguments have been put forward showing that the energy source of QSOs has to be gravitational, since any other energy source would have to function in such an extreme environment that gravitational energy would be in any case very large and quickly able to dominate [6,13].

Alternative models have been proposed. Some interpret QSO redshifts as non cosmological, others explain activity in terms of ordinary stellar activity, starbursts, or supernova chains. The cosmological origin of QSO redshifts is a very solid result relying on detecting absorption lines in QSO spectra with redshift below the QSO one and on detecting faint galaxies around or in the neighborhood of QSOs. The stellar origin of activity is harder to rule out. Starbursts are indeed observed at the required luminosity level: these are the IRAS Far Infrared Ultra Luminous galaxies. However, a stellar mechanism would have great difficulties in explaining flat spectrum sources or radiogalaxies. On this basis, we will assume that the AGN energy source is gravitational without being able to prove it. For a review of the "heretic" stellar AGN models, see Refs [17, 18]

The standard mechanism to release gravitational energy is through an accretion disk. In a disk, energy can be dissipated and the disk collimation makes the formation of jets more natural. Spherical accretion has been proposed as a powering mechanism for AGNs, but its efficiency appears to be at least ten times lower than accretion in a disk.

Despite the fact that accretion is generally considered to be non spherical, one of the basic quantities considered for an AGN, its Eddington luminosity, is derived by assuming spherical symmetry. The Eddington luminosity is derived by imposing that radiation pressure on the electrons equals gravitational attraction on the protons. Should the radiation pressure be higher, electrons would be pushed away from the central engine, and due to the very large ratio of electromagnetic versus gravitational forces the protons would follow

them, thus stopping the accretion. Therefore, under the assumption of spherical symmetry, a given luminosity L implies that the accreting mass must be larger than:

$$M_{edd} = \frac{\sigma_T L}{4\pi G m_p c}, \quad (1)$$

where σ_T is the Thomson cross-section, and m_p the proton mass.

Such a simple picture has two drawbacks. One is that accretion is probably non spherical and the other is that, in the situation described above a Rayleigh-Taylor instability would probably arise, destroying the spherical symmetry. Despite this criticism, Eddington luminosities are always taken as an indication for the central engine mass. For most AGNs they end up being in the range $10^6 - 10^9$ $M_\odot$. Since for these masses no stable compact non-collapsed configuration can exist, it is assumed that the accreting mass is a supermassive black hole.

The accretion disk is able to transfer angular momentum to the black hole; despite this, most models are based on the Schwarzschild rather than the Kerr metric. The typical scales of the accretion disk are ~ 100 Schwarzschild radii or ~ 300 A.U.. On these and smaller scales the disk temperature can be high enough to allow for thermal X-ray emission.

The basic idea of the Grand Unified model is that all AGNs have essentially the same structure and that their properties are determined mainly by two parameters: the angle between the line of sight and the accretion disk axis and the absolute luminosity. While luminosity determines mainly the broad classification scheme of AGNs into LINERs, Seyferts, and QSOs, the angle to the line of sight plays a more subtle role. The basic assumption is that the central engine is partly screened by a dust torus, so that the ionizing radiation is emitted anisotropically in two ionization cones. When an AGN is seen from a direction close to the accretion disk axis, the innermost components of the central engine itself can be observed; these have the most rapid timescales and the smallest size, so that one is able to observe the most violent variability. In addition, the thermal and non-thermal emission from the accretion disk are directly observable. The possible presence of jets contributes to increase both variability and luminosity by boosting relativistically the emission from the approaching jet substructures and therefore by increasing the observed ratio of continuum to line emission and the optical luminosity compared to the radio emission (which is generated in the non-boosted radio lobes). These objects are classified as BL Lacertae or OVV depending on their luminosity.

When the viewing angle is such that the innermost portion of the central engine is screened, high ionization gas clouds can still be visible. These clouds will be moving at very high speed in the central black hole gravitational potential. All these clouds together make up the so called Broad Line Region (BLR). Some relativistic boosting may still contribute to increase the apparent optical luminosity as compared to the radio one. Objects seen from such an

angle are classified as Seyfert I galaxies or QSOs depending on luminosity.

Finally objects seen from directions lying close to the accretion disk plane will suffer from very strong extinction due to the dust torus surrounding the accretion disk. In these objects the BLR is almost completely screened and only the Narrow Line Region (NLR) is usually visible and the AGN is classified as Seyfert II. For these objects the ionization cones point away from us. If the radio lobes are fully developed the appearance is that of a FR II radiogalaxy.

It is estimated that in the radiogalaxies Centaurus A and Cygnus A the extinction to the central engine at optical wavelengths is of 30 and 50 magnitudes, respectively. Indeed, in the case of Centaurus A the true nucleus may have been detected at near IR wavelengths.

The Grand Unified model has been tested mainly on statistical grounds, by verifying that the various classes of AGNs indeed behave as if luminosity and viewing angle were the only parameters. In addition to statistical tests, one particularly impressive check of the GU model came from observations in polarized light from galaxies classified as Seyfert II. When the ionization cones point away from us and the broad emission lines from the BLR are screened by the dust torus, we can sometimes observe the emission coming from dust patches lying in the ionization cone. It is possible to obtain a spectrum of the incoming ionizing radiation, since the dust-reflected light is polarized. The polarized light turns out to have much stronger broad emission lines than the total unpolarized light, thus confirming the GU model and showing that galaxies can have different classifications depending on the viewing angle [1, 10].

An independent test of the hypothesis that all QSOs are beamed comes from studying the spatial distribution of neutral hydrogen clouds seen in absorption on QSO spectra. When a QSO lies very close to the line-of-sight toward a more distant QSO, if the emission was isotropic the nearer QSO would ionize the clouds in its vicinity. Since this is not observed, one concludes that QSOs are generally beamed [11].

The properties of radiogalaxies cannot be explained solely on the basis of viewing angle and luminosity; additional parameters, related to their interstellar medium (hereafter ISM) properties, appear to be required in order to describe their radio-properties.

FR I radiogalaxies are often seen nearly along the jet axis. This increases the core luminosity compared to the jet. In addition the fact that radiolobes are extended along the line of sight reduces their apparent spatial extension and their regularity. In contrast, FR II are often observed from angles close to the disk plane. However, these effects alone cannot fully explain the different properties of FR I and II. Indeed, the ISM also plays a crucial role. A given amount of energy injected into the ISM can originate quite different radio structures depending on the medium density. Elliptical galaxies with

high ISM densities and extended H_α emission, like NGC 4696, have much less extended radio structures than galaxies with lower ISM densities but similar luminosity, like NGC 3557. The ability of the ISM to confine the radiolobes is probably also the reason why spiral galaxies, that are gas rich systems, are very rarely radiogalaxies. The striking difference in appearance of the central regions of radiogalaxies is illustrated by the example of NGC 3862 and NGC 6251, observed with the Faint Object Camera on board of the Hubble Space Telescope [4]. NGC 3862 is an inconspicuous FR I galaxy, core dominated, and possessing an unresolved nuclear X ray source. NGC 6251 is, instead, a classical double with very extended radio jets and radio lobes. HST observations reveal the presence of an unresolved blue optical point source in the nucleus of both galaxies; however, NGC 3862 also possesses an optical synchrotron jet 0.6 arcsec long, previously unseen from the ground. It is tempting to attribute the visibility of this (one-sided) jet to relativistic boosting of an otherwise symmetric jet lying close to the line of sight. One of the best studied FR I galaxies is M87, which is also the best known example of bright optical synchrotron jet. This jet was discovered in 1918 [5] and since then was studied at the highest resolution, both from space and from the ground. These studies have allowed us to show that there are several reacceleration regions in the jet. Energy can be efficiently transported along the jet, which is bright along its limbs. It is only recently that it has been established that M87 possesses a hot spot in the east radiolobe, opposite to the jet [16]. The properties of this hotspot clearly indicate that it must be powered even now by an otherwise unseen eastern jet. Both the presence of symmetric jets and the presence of hot spots are characteristics traits of FR II radiogalaxies. Hotspots in radiogalaxies have often been interpreted as bow shocks at the end of jets. In the case of M87 the newly discovered hotspot does indeed have the shape of a bow shock. In the case of the southern hotspot of the FR II 3C33 radiogalaxy, it has been possible to establish through, Hubble Space Telescope Planetary Camera observations, that even hotspots may be characterized by multiple reacceleration regions. The 3C33S hotspot is also bow shock shaped and is unresolved in the direction perpendicular to the shock. Along the shock surface, the hotspot is resolved and its size is larger than the maximum path that can be traveled by the synchrotron emitting electrons during their lifetime [3]. Both for jets and hot spots it is usually assumed that the process locally accelerating the relativistic electrons is Fermi acceleration [9], although it is unclear how the observed spectrum is produced [8].

A number of well known basic physical processes occur in a radiogalaxy: synchrotron emission, gas cooling processes, relativistic kinematics, Fermi acceleration in shocks. However, it is unclear what is the precise interplay of all these processes in leading to the observed properties of radio galaxies. Moreover, the mechanisms by which the central engine is able to deliver huge amounts

of energy in a coherent way over megaparsec distances is totally unknown.

Numerical models of jet propagation are often merely hydrodynamical and axially symmetric. We know, especially for the M87 jet, that magnetic fields play a very important role, emission appears to be localized in filaments winding on the surface of a cone (see, *e.g.*, [15]). The intensity of emission appears to be correlated to the filaments pitch angle [9]. Despite the amount of data available, our understanding of MHD jet propagation in inhomogeneous media is still very limited. Radiogalaxies often host cooling flows. However, there has been so far no study of the entire radiogalaxy environment, even for relatively nearby objects like the Virgo cluster ellipticals. Most cooling flow models are stationary and spherically symmetric. In many radiogalaxies, however, we know that the radio lobes have carved cavities inside the ISM and it is to be expected that the cooling flow will not be axisymmetric. Cooling flow galaxies, like M87 and NGC 4696, often show ionized gas filaments and patches, and such gas appears to be able to confine the radio lobes. In a galaxy like M87 the radio power is comparable to the X ray luminosity and we know that radio emission is caused by ultrarelativistic electrons. Under these conditions, it would be odd if the relativistic electrons and the X-ray emitting gas did not interfere. An even more extreme scenario could be that cooling flows do not exist at all. The stars that supposedly should be produced in cooling flows have never been observed; on the other hand the energy required to reheat up the gas appears to be easily available. The lack of a strong correlation between radio and X-ray luminosity is not at all surprising. In fact, we have seen that high ISM or inter galactic densities appear to reduce the luminosity of radiolobes as shown by individual nearby galaxies and by the clustering properties of FR I and FR II galaxies. We know that radioactivity in ellipticals is episodic, since there are some galaxies where a new radiolobe is observed not aligned with an older one not completely dissipated. These properties could be explained easily by considering the fuelling mechanisms in radiogalaxies. It is not unlikely that AGNs in ellipticals are in a "starving" regime, where a supermassive black hole characterized by huge Eddington luminosity does not give rise to any nuclear activity due to a lack of fuel. Black holes more massive than 2×10^9 $M_\odot$ are unable to tidally disrupt main sequence stars and therefore would be fed only by infalling gas or occasional stripping of red giants. During these events the central engine would turn on and emit at very high luminosities, so high that the whole galaxy could be deprived of ISM. This would turn off the nuclear activity again until other gas, possibly due to galactic interactions or thermal instabilities in the cooling flow, feeds the black hole.

The quantitative investigation of scenarios of this type would certainly require a very significant observational, theoretical, and numerical effort over the years to come.

References

1. Antonucci, R., Barvainis, R., 1990, *Astrophys. J.*, **363**, L17.
2. Blandford, R.D., 1990, in *Active Galactic Nuclei*, Springer-Verlag, Berlin, p.161.
3. Crane, P., Stiavelli, M., 1992, *Mon. Not. R. astron. Soc.*, **257**, 517p.
4. Crane, P., Stiavelli, M., King, I.R., et al., 1993, *Astron. J.*, **106**, 137.
5. Curtis, H.D., 1918, *Lick Obs. Publ.*, **13**, 11.
6. Lynden-Bell, D., 1969, *Nature*, **223**, 690.
7. Lynden-Bell, D., 1978, *Physica Scripta*, **17**, 185.
8. Kardashev, N.S., 1962, *Sov. Astron. A.J.*, **6**, 317.
9. Meisenheimer, K., 1991, in *Physics of Active Galactic Nuclei*, Dusch, W.J., and Wagner, S.J., eds., Springer-Verlag, Berlin, p. 525.
10. Miller, J.S., Goodrich, B.F., 1990, *Astrophys. J.*, **355**, 456.
11. Møller, P., Kjærgaard, P., 1992, *Astron. Astrophys.*, **258**, 234.
12. Netzer, H., 1990, in *Active Galactic Nuclei*, Springer-Verlag, Berlin, p.57.
13. Rees, M.J., 1984, *Ann. Rev. Astron. Astrophys.*, **22**, 471.
14. Saikia, D.J., Kulkarni, V.K., 1994, *Mon. Not. R. astron. Soc.*, **270**, 897.
15. Stiavelli, M., Møller, P., Zeilinger, W.W., 1991, *Nature*, **354**, 132.
16. Stiavelli, M., Biretta, J., Møller, P., Zeilinger, W.W., 1992, *Nature*, **355**, 802.
17. Terlevich, R., 1989, in *Evolutionary Phases in Galaxies*, Beckman, J.E., and Pagel, B.E.J., eds., Cambridge Univ. Press., Cambridge, p. 149.
18. Terlevich, R., Melnik, J., 1988, *Nature*, **333**, 329.
19. Woltjer, L., 1990, in *Active Galactic Nuclei*, Springer-Verlag, Berlin, p.1.

Binary stars and dynamical evolution of globular clusters

Enrico Vesperini*

February 25

* Scuola Normale Superiore, Pisa, Italy

Dynamical evolution of globular clusters

Globular clusters are the most massive star clusters present in galaxies and consist of spherical distributions of $10^4 - 10^6$ stars bound by their self-gravitation. The dynamical evolution of globular clusters is largely driven by internal processes (gravitational two-body encounters between stars). Interaction with the disk and/or the bulge of the host galaxy may also play an important role (see [24,27] and references therein).

The time scale on which encounters become important and affect the evolution of a stellar system is called *relaxation time*. The relaxation timescale depends on the number density and on the velocity dispersion of the systems that usually are a strong function of the position inside the clusters.

For galactic globular clusters the relaxation times calculated at the half-mass radius, t_{rh}, span a range between 10^7 and 10^{10} years and the relaxation times in the core between 10^6 and 10^8 years (see, e.g., [4]). Thus collisional interactions act on a timescale shorter than the age of globular clusters, which is for most of them estimated to be generally close to 1.5×10^{10} years, and they are important in their evolution.

The evolution of a globular cluster as driven by collisional effects has been investigated in the last twenty years by different methods (see, e.g., [5] for a review of the numerical models used) all giving support to a scenario in which the evolution drives the system toward configurations with smaller and smaller cores, with increasing values of the central density, and with larger and larger diffuse 'halos'.[1] The physical process driving the evolution varies during the history of the cluster and the characteristic time scale of the collapse of the core varies accordingly. The first stage of the evolution is driven by the evaporation of stars from the core due to weak encounters. The stars

[1] This scenario refers to an isolated cluster without keeping track of the tidal field of the host galaxy on the loosely bound stars.

evaporated from the core populate the halo making it larger and larger and the core, by energy conservation, shrinks. If all the evolution were evaporation-driven, the full core collapse would take place in a time $\tau \sim 100 t_{rh}$. In reality, after $t \sim 5 t_{rh}$ the collapse of the core changes its nature and accelerates, driven by a thermodynamical instability (*gravothermal catastrophe*; [1,18]).

The process of core collapse does not proceed indefinitely. When the central density becomes sufficiently large, processes other than two-body relaxation may stop and reverse the core collapse. The phase of 'post-core collapse' requires the presence of an energy source compensating for the energy loss from the core to stop the core collapse and for the heat outflow across the half-mass radius once the collapse has been stopped, with a nearly isothermal region being established from the center out to the half-mass radius. The post-core collapse phase is likely to be supported by binaries (see, e.g., [16] for the observational evidence of binaries in globular clusters), even though other energy sources (black hole at the center of the cluster, star mergers) have been considered.

Theory of binary-single star interactions

The theory of binary-single star interaction is largely based on the analytical studies of Heggie [10] and the numerical investigations of Hut [14] and Hut & Bahcall [15] (see [12] for a recent review of all the analytical and numerical studies available so far). Analogous work for binary-binary interactions has been done by Mikkola [20,21,22], even though in this case the theory is far from being fully developed.

As a result of an interaction with a single star a binary can be disrupted (*ionization*), it can become more bound (*hardening*; in this case both the single star and the center of mass of the binary increase their kinetic energy and this corresponds to a heating of the cluster), or it can become less bound (*softening*; the single star and the center of mass of the binary decrease their kinetic energy and this corresponds to a cooling of the cluster).

Analytical studies confirmed by numerical simulations show that a binary with semi-major axis a and binding energy (in absolute value) $\epsilon = Gm^2/2a$ (where m is the mass of each component of the binary) interacting with a background of single stars with the same mass m and one-dimensional velocity dispersion v_s increases, on average, its binding energy if $x \equiv \epsilon/mv_s^2 > 1$ (*hard* binaries), while its binding energy decreases if $x < 1$ (*soft* binaries) [10]. Both hard and soft binaries can be ionized, that is disrupted by one single encounter, but the harder the binary is, the smaller is the probability that a single encounter can disrupt it.

Numerical simulations provide the following estimates for the ionization and the hardening/softening rate for a population of binaries with number density n_b interacting with a population of single stars with number density

n_s. The ionization rate is given by [15]

$$\frac{1}{n_b}\frac{\mathrm{d}n_b}{\mathrm{d}t} = -n_s \pi a^2 V_{th} R(x), \tag{1}$$

where $R(x)$ is a function of the hardness factor x (see Eq. (5.12) in [15]); if equipartition of energy between the binary and the single star holds, the relative thermal velocity V_{th} entering Eq. (1) is given by $V_{th} = \frac{3}{\sqrt{2}} v_s$.

With an analogous formalism, the hardening/softening rate can be written as

$$\frac{1}{\epsilon}\frac{\mathrm{d}\epsilon}{\mathrm{d}t} = n_s \pi a^2 V_{th} G(x), \tag{2}$$

(see Eq. (49) in [12] for the expression $G(x)$).

Binary stars and dynamical evolution of globular clusters

In this section we will describe the effects of binary stars in the dynamical evolution of globular clusters focussing our attention on their role in the post-core collapse phase. Figure 1 shows the histogram of the estimated ratio of the core to the half-mass radius, r_c/r_h, for galactic globular clusters classified as post core collapse (PCC) [26]. The core radius r_c is defined as the radius where the density projected along the line of sight of an infinite isothermal sphere falls to half of its central value.

Depending on the formation mechanism, binaries are classified as *three-body* binaries if born from the interaction of three single stars resulting in a binary and an escaping star taking away the energy necessary to leave the other two stars bound, *tidal* binaries if born from an interaction of two single stars passing so close to each other that orbital energy is dissipated in tidal deformation of the stars and a bound system is formed, and *primordial* binaries if they are produced at the time of star formation.

Immediately after the end of core collapse, thermal equilibrium inside the half-mass radius (conventionally taken as the radius of the bulk of the cluster) is quickly established. Thereafter, in agreement with the scenario first proposed by Hénon [13], the evolution of the system is determined by the balance between energy production in the core and energy conduction at the half-mass radius.

The flow of energy into the outer parts of the cluster can be written as

$$\dot{E} = \frac{1}{\gamma}\frac{|E_{tot}|}{t_{rh}}, \tag{3}$$

where E_{tot} is the total energy of the cluster, t_{rh} is the half-mass relaxation time and γ is a dimensionless parameter derived from numerical simulations and shown to be approximately constant in the post collapse phase ($\gamma \sim 10$) (see, e.g., [6]).

The size of the core during the post-core collapse phase can be estimated by equating the energy flux at the half-mass radius to the energy production in the core. As for three-body binaries, the energy production in the core, given by the product of three-body binary formation rate and the total energy input per binary (see, e.g., [3]), yields $r_c \propto r_h M^{-2/3}$. For typical values of the mass of globular clusters, this energy source thus gives extremely concentrated clusters ($r_c/r_h \approx 10^{-3}$).

Numerical simulations [25] show that tidal binaries are able to halt core collapse before three-body processes become important, but still at very high central density ($n_c \approx 10^9 \mathrm{pc}^{-3}$). During the post-collapse expansion driven by tidal binaries the cluster is still very concentrated ($r_c/r_h \leq 10^{-2}$).

Since tidal binaries are characterized by semi-major axes of a few stellar radii, they have extremely high values of x. Thus when a tidal binary and a single star interact close enough to modify the binding energy of the binary, the recoil energy of both the binary and of the single star is much larger than the escape energy from the cluster and they are both ejected from the cluster after the encounter. Tidal binaries can thus heat the cluster only indirectly, by supplying potential energy by mass ejection.

Recent studies have cast doubts on the possibility for tidal binaries to survive long enough and to be an efficient energy source. In fact it has been shown (see, e.g., [17,23]) that in most cases the less compact star involved in the process of a tidal capture is disrupted as a result of the dissipation of the tidal energy deposited during the close encounter.

As for primordial binaries, many investigations on their role in the dynamical evolution of globular clusters have been carried out (see, e.g., [11,19] and references therein) and the emerging scenario is that primordial binaries are able to delay core collapse and to stop it at a much lower central density than 'dynamically made' binaries which, in contrast with the case of primordial binaries, need very high central densities to form.

We now describe some the results of an analytical calculation (see [28] for further details) showing how the ratio r_c/r_h in the post-core collapse phase depends on the abundance of binaries and on the parameters of the binary binding energy distribution.

We express the heating rate per unit volume as $e = \frac{G^2 m^3}{v_s} n^2 (\mu_{bs} + \mu_{bb})$ where n is the total number density (single stars plus binaries). The dimensionless quantities μ_{bs} and μ_{bb} are essentially the average rate at which energy is released into the core by binary-single and binary-binary star interactions. We express μ_{bs} as

$$\mu_{bs} = \frac{n_b n_s}{n^2} \int f(x) g(x) h(x) \mathrm{d}x. \tag{4}$$

where $f(x)$ is the distribution of binding energies of the binaries, $g(x) = 1.66 G(x)/x$ is the average rate of change of binding energy of a binary in a sea of single stars. An additional contribution arises from the possibility of

ionization of the binary and we assume that all the stars involved are retained by the core. However the contribution is not just an integral of $f(x)g(x)$ over x because some stars interact so strongly that they are kicked out of the cluster. When ejection occurs the binding energy of the cluster is decreased as a result of mass loss. The function $h(x)$ takes this possibility into account.

The binary encounter rate can be described by an analogous expression

$$\mu_{bb} = \frac{n_b^2}{n^2} \int f(x_1) f(x_2) g(x_1, x_2) H(x_1, x_2) \mathrm{d}x_1 \mathrm{d}x_2, \tag{5}$$

where $H(x_1, x_2)$ and $g(x_1, x_2)$ are the analogues of $h(x)$ and $g(x)$. The detailed derivation of the functions $h(x)$, $H(x_1, x_2)$ and $g(x_1, x_2)$ is described in [28].

Balancing the energy production in the core with the power needed for the expansion of the cluster (see Eq. (3)) after some simple algebra we finally get

$$\begin{aligned} \frac{r_c}{r_h} &= \frac{0.1872}{\log_{10}(0.4N)} \frac{1}{(1+\phi_b)^4} \left[\phi_b (1-\phi_b) \left(\sum_i f_i g_i h_i - \sum_i 1.66 \frac{f_i R_i}{x_i} \right) \right. \\ &+ \left. 12.5 \phi_b^2 \left(\frac{1}{2} \sum_i f_i^2 H_{ii} + \sum_i \sum_{j<i} f_i f_j \left[\frac{0.5(x_i + x_j)}{x_j} \right] H_{ij} \right) \right] \\ &\times \left(\frac{v_c}{v_h} \right)^3 \left(\frac{\gamma}{10} \right) \end{aligned} \tag{6}$$

where v_h is the one-dimensional velocity dispersion at the half-mass radius, v_c is the central one-dimensional velocity dispersion, $\phi_b = n_b/n$, $f_i \equiv f(x_i)\Delta x_i$, $h_i = h(x_i)$, $g_i \equiv g(x_i)$, $R_i \equiv R(x_i)$, $H_{ij} \equiv H(x_i, x_j)$ and the sum for $j < i$ means $x_j < x_i$.

Equation (6) gives r_c/r_h as a function of (1) the fraction of binaries, (2) the slope of the binary binding energy distribution, and (3) the upper and lower cutoffs on the binary binding energy. The specific numerical values for some of the parameters are uncertain (γ) and for others vary from cluster to cluster ($N, v_c/v_h$). We have assumed $\gamma = 11.5$ [11], $v_c/v_h = \sqrt{2}$ [8] and $N = 10^6$.

We assume a power-law distribution for the binary semi-major axes of the form $f(a) \propto a^{-n}$ or, in terms of the hardness x, $f(x) \propto x^{n-2}$ over a range $[x_l, x_u]$.

Figure 2 displays r_c/r_h as a function of ϕ_b for three power-law distributions ($n = 0$, 1 and 2) and a fixed range in x ($10 < x < 100$). For large values of ϕ_b, the core is large in size because the energy generation mechanism which balances a fixed flux at the half-mass radius only requires a small central density. Of course, at any given instant the abundance of binaries will be history dependent.

The binary abundance in the core is the single most important parameter governing the core size (see [28] for plots showing the dependence on the choice of the upper and the lower limit in the distribution $f(x)$).

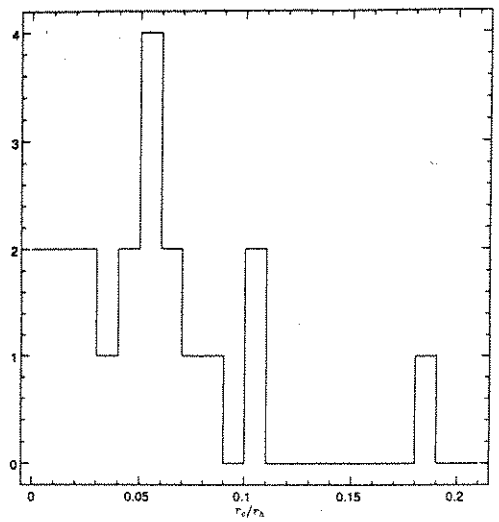


Figure 1: Distribution of the values of r_c/r_h for galactic PCC clusters (data from Ref. [26])

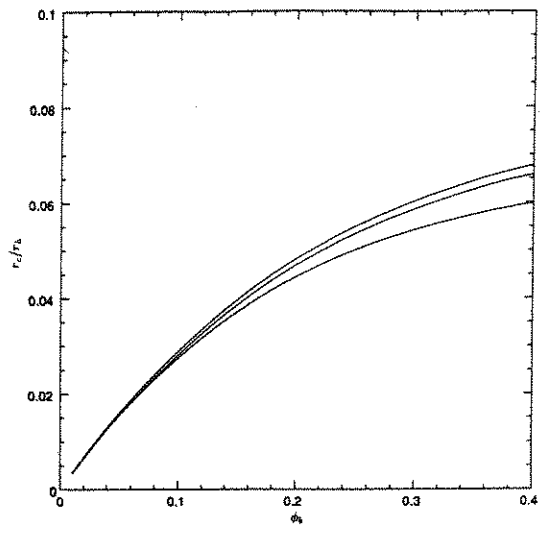


Figure 2: r_c/r_h for $x_l = 10$, $x_u = 100$ vs. the abundance of the binary population ϕ_b for three different indexes of the power-law distribution of binding energy $f(x) \propto x^{n-2}$. From the upper to the lower curve $n = 1, 0, 2$

Figure 2 shows that the slope of the binary distribution function plays only a minor role for the adopted range of hardness $10 < x < 100$ in the example. The situation illustrated is particularly simple because little indirect heating occurs for $x < 100$. The rate of core heating is the rate of binary hardening and this is independent of x for very hard binaries. Hence, r_c/r_h is not very sensitive to the distribution of x when direct heating by hard binaries dominates. The small differences between power-laws seen in Fig. 2 are the result of binary-binary contributions to the heating rate. These interactions give a net heating rate that depends on $(x_1 + x_2)/\min(x_1, x_2)$, no longer independent of the distribution of x.

The results are indicative of the range of r_c/r_h that may be plausibly obtained during the evolution of a single cluster in the post-collapse phase, as ϕ_b varies, and of the range that a group of clusters with different initial binary abundances and distributions might display. The calculation done indicates that the spread in r_c/r_h shown in Fig. 1 for the Galactic PCC globular clusters could be explained by variation in the properties of the binary distribution.

References

1. Antonov V.A., 1962, *Vest. Leningrad Univ.*, **7**, 135 translated in *Dynamics of Star Clusters* IAU Symp. 113, p. 525, eds. J. Goodman, P. Hut

2. Binney J., Tremaine S., 1987, *Galactic Dynamics*, Princeton University Press

3. Cohn H., Hut P., Wise M., 1989, *Astrophys. J.*, **342**, 814

4. Djorgovski S.G., 1993, in *Structure and Dynamics of Globular Clusters*, eds. S.G. Djorgovski & G. Meylan, Astron. Soc. of Pacific, p. 373

5. Elson R., Hut P., Inagaki S., 1987, *Ann. Rev. Astron. Astrophys.*, **25**, 565

6. Gao B., Goodman J., Cohn H., Murphy B., 1991, *Astrophys. J.*, **370**, 567

7. Goodman J., 1989, in *Dynamics of Dense Stellar Systems*, ed. D. Merritt, Cambridge University Press, p. 183

8. Goodman J., Hut P., 1989, *Nature*, **339**, 40

9. Goodman J., Hut P., 1993, *Astrophys. J.*, **403**, 271

10. Heggie D.C., 1975, *Mon. Not. Roy. Astron. Soc.*, **173**, 729

11. Heggie D.C., Aarseth S.J., 1992, *Mon. Not. Roy. Astron. Soc.*, **257**, 513

12. Heggie D.C., Hut P., 1993, *Astrophys. J. Suppl.*, **85**, 347

13. Hénon M., 1961, *Annals d'Astrophys.*, **24**, 369

14. Hut P., 1983, *Astrophys. J.*, **268**, 342

15. Hut P., Bahcall J.N., 1983, *Astrophys. J.*, **268**, 319

16. Hut et al., 1992, *Publ. Astron. Soc. of Pacific*, **104**, 981

17. Kochanek C.S., 1992, *Astrophys. J.*, **385**, 604

18. Lynden-Bell D., Wood R., 1968, *Mon. Not. Roy. Astron. Soc.*, **138**, 495

19. McMillan S.L.W., Hut P., 1994, *Astrophys. J.*, **427**, 793

20. Mikkola S., 1983a, *Mon. Not. Roy. Astron. Soc.*, **203**, 1107

21. Mikkola S., 1983b, *Mon. Not. Roy. Astron. Soc.*, **205**, 733

22. Mikkola S., 1984, *Mon. Not. Roy. Astron. Soc.*, **207**, 115

23. Rasio F.A., Shapiro, S.L., 1991, *Astrophys. J.*, **377**, 559

24. Spitzer L., 1987, *Dynamical Evolution of Globular Clusters*, Princeton University Press

25. Statler T.S., Ostriker J.P., Cohn H.N., 1987, *Astrophys. J.*, **316**, 626

26. Trager S.C., Djorgovski S.G., King I., 1993 in *Structure and Dynamics of Globular Clusters*, eds. S.G. Djorgovski & G. Meylan, Astron. Soc. of Pacific, p. 343

27. Vesperini E., 1994, *Ph. D. Thesis*, Scuola Normale Superiore, Pisa

28. Vesperini E., Chernoff D.F., 1994, *Astrophys. J.*, **431**, 231

The Perturbation Particles method applied to stellar dynamics

Francine Leeuwin*

March 4

* Scuola Normale Superiore, Pisa, Italy

Perturbations in galaxies

A motivation for studying perturbations to dynamical models of galaxies, although these models are not well defined [4], is given by the apparent ubiquity of "fine structure" in galaxies. This is a well-known observational fact for spiral galaxies, which exhibit, in addition to their spiral arms, other features such as stellar bars, warps of their disk, ripples in the distribution of stars ... For elliptical galaxies, improvements in CCD techniques in the last decade have allowed to detect perturbations in their images [13] that were not seen earlier due to their low luminosity level (a few percent of the total light). They consist of ripples, sometimes organized in remarkably ordered shell systems (the best example is NGC 3923), or of isophotal deviations (apparently the rule rather than the exception) from pure ellipses towards either disky or boxy shapes.

The hope is then to extract the information contained in these sub-structures and to learn something about the geometry of the host galaxy, from which they originate, since models are so difficult to determine directly. Further motivation is given by the desire to eventually relate the history and formation of these fine structures to those of the galaxy; indeed, they are sometimes considered as possible clues for the formation of elliptical galaxies as the result of the merging of two disk galaxies [1].

Perturbations in N-body systems

The dynamics of galaxies, as collisionless stellar systems, can be described by a distribution function of stars in phase-space $f(\mathbf{r}, \mathbf{v})$ governed by the Collisionless Boltzmann equation (C.B.E.):

$$f(t)_{|H} = f(0)$$

where $H = \mathbf{v}^2/2 + \Phi(\mathbf{r})$ is the single-star Hamiltonian. The equation states that the density of stars is conserved along the orbits dictated by the potential

Φ, which is in turn related, in a fully self-gravitating system, to the density of stars through the Poisson equation:

$$\nabla^2\Phi = 4\pi G \int d^3\mathbf{v} f(\mathbf{r},\mathbf{v}).$$

In a small perturbation calculation, one usually divides the total distribution function f into two terms:

$$f(t) = f_0 + f_1(t),$$

where $f_0 >> f_1$, the former being a model representing the initial state of the galaxy, the latter being the evolving perturbation. In practice, even the linear calculations are very difficult to carry out analytically, and demand an impressive artillery of different sets of base functions. Only in a homogeneous medium is a complete solution easily derived, and this explains why many problems have been treated with a local approximation, i.e. under the assumption of a constant density. Some problems can nonetheless be solved semi-analytically using expansions into appropriate bases, and this approach has been used at the Scuola Normale [2] to study the modes associated with the radial orbit instability for a family of models of elliptical galaxies reproducing realistic density profiles [3]. I'll come back to this point in the last section.

In an alternative approach, numerical N-body simulations can in principle lead to full non-linear solutions, but are subject to another *caveat*: the present computers are unable to handle realistic numbers of stars in a galaxy ($\sim 10^{11}$), being limited to $\sim 10^5 - 10^6$ particles. This discretisation introduces a "particle noise" which makes them unsuitable to resolve weak perturbations. Hence the analytical and the numerical solutions look at two distinct sides of the question, i.e. the birth of the disturbance (where a linear approximation is valid), and the large amplitude perturbations. Each of these methods cannot provide a full description. More specifically, realistic time-scales for the global evolution of a mode are missing.

The Perturbation Particles method

Because of the limitations of each of these techniques, analytical or numerical, we turned to a mid-way scheme, which blends :

- an analytical model f_0 to describe the initial unperturbed state of the system
- a numerical solution for the perturbation f_1.

The point of this separation into analysis and simulation is to allow for a non-linear solution while concentrating the numerical cost onto the small part of the system which is evolving. This 'Perturbation Particles' scheme (PP to abbreviate) was first proposed by Rybicki in 1972 (*unpublished*), but did not seem to attract much attention, except for an application to a homogeneous

medium [12]. On the other hand, similar techniques have been used in the meantime in plasma physics [7].

In practice, one chooses a model f_0 for the initial state (see sketch below); then spreads the N particles of the simulation over the model phase-space, following an arbitrary distribution function $f_S(\mathbf{r},\mathbf{v})$. The S subscript stands here for 'Sampling', since the duty of these PP's is to sample the perturbation which might develop in the system. For this purpose, the particles carry a mass which is not proportional as usual to the global density, but to the perturbation f_1. As a consequence, the masses can be either positive, if the PP stands on a local excess of density, or negative, if locally $f_1 < 0$. The simulation runs very much as a standard one, by integrating the orbits for the particles, except that at each time-step one integrates the CBE for f_1, to adjust the masses:

$$f_1(t)_{|H} = f(0) - f_0(\mathbf{r}(t),\mathbf{v}(t)).$$

By replacing in this equation H by H_0, one gets the linear solution, which can be compared directly to analytical estimates when available.

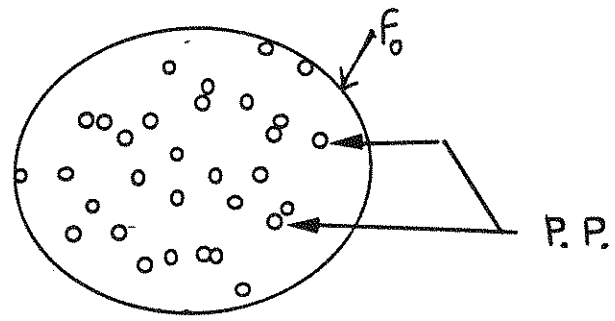


Figure 1: Setting up a PP simulation

Sampling distribution function

The main difficulty in using the PP scheme is the choice of the initial distribution $f_S(\mathbf{r},\mathbf{v})$ for the particles. In standard N-body codes, all the particles have the same mass, and reproduce the model density profile and velocity field by following $f_0(\mathbf{r},\mathbf{v})$. Here they can be initially picked up arbitrarily, but:

- (i) one would like them to roughly follow f_1, thus will try to sample more densely the regions where the perturbation is expected to grow, if this is at all known;
- (ii) to avoid fast relaxation of this distribution of particles inside the potential Φ_0, one should choose f_S to be a function of the integrals of the motion for the unperturbed model (energy E, angular momentum square $\mathbf{L}^2$, etc); then from Jeans' theorem, this distribution will remain to lowest order constant in time.

The masses will depend on how the particles have been spread over the system phase-space, more isolated particles having in charge larger phase-space regions:

$$m_i(t) = \frac{f_1(\mathbf{r}(t), \mathbf{v}(t))}{f_S(t)}, \quad i = 1, \ldots N$$

Fortunately the particle distribution $f_S(t)$ does not have to be re-evaluated in the course of the simulation: the Liouville theorem which, written for the mass distribution, gives the CBE, is here applied to the number distribution. It stipulates that particle density is conserved along the orbits:

$$f_S(t)_{|H} = cst = f_S(0).$$

When applying the method, one soon realizes that troubles come from point (i) above. If one has sampled only the initial phase-space system, evolution of the system might distort its boundaries, and all the phase-space may not be covered anymore with PP's (see Fig. 2). Hence one might wish to sample a phase-space larger than the initial one, but if this can indeed solve the problem for one or two dimensional systems [8], it obviously increases the numerical cost by requiring more particles. Therefore, the method is neither safe nor advantageous except for weak perturbations. A bad sampling of the perturbation, where some of it is missing, will translate into non-conservation of the total mass $M_1 \equiv \sum_i m_i \neq 0$; this quantity is to be checked during the simulation.

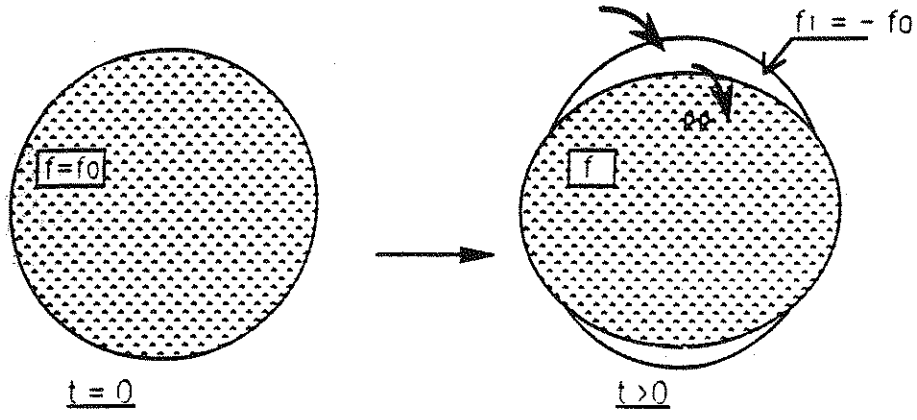


Figure 2: Phase-space distortion leading to a failure in the sampling procedure

Applications to perturbation modes

When we started this work, the only known applications of the PP method to stellar dynamics were an attempt to calculate the pulsating mode of a homogeneous infinite slab of matter, by Zimmerman (1973, *unpublished*), and a simulation of a plane wave travelling in a homogeneous infinite medium, in

very good agreement with the theory [12]. Recently, Wachlin *et al.* [16] have performed another simulation of a travelling plane wave.

Our first aim was to understand why the first attempt had failed; it used a very small number of PP's, but even with larger numbers ($\sim 10,000$) the total mass was not conserved untill we used phase-space samplings covering domains larger than that of f_0, for the reasons just explained. Larger perturbations require more and more particles, but once this requirement is satisfied, very good agreement with the analytical predictions is obtained.

We then turned to a few other mode calculations in simple systems for which analytical solutions are available, in order to test the method [8]. An application of greater astrophysical interest, concerning dynamical friction between galaxies, is described below.

Dynamical friction and merging of galaxies

Numerical simulations, stimulated by the pioneering work of Toomre & Toomre [15], have shown that the galaxy produced by the merging of two disk galaxies resembles an elliptical galaxy. The question of how many elliptical galaxies may have formed in this way has therefore gained a central position in galactic astronomy. The process that drives the merging can be described within the frame of dynamical friction [6]; it would be very valuable to have time scales for it, in order to estimate the fraction of ellipticals thus formed. Unfortunately, the detailed mechanisms involved in these interactions are poorly known, either analytically, because the interactions between two galaxies do not obey the hypotheses of Chandrasekhar's formulation for dynamical friction[1], meant for a different problem, or numerically, because of the artificial particle noise mentioned above.

As in many articles available in the literature [10, 5, 11, 13...], we restricted our simulations to the simplified problem of a rigid satellite falling onto a larger galaxy, both modeled as spherical systems - the primary galaxy being here simulated with the PP scheme. The reason why all previous investigations have only computed the end of the merging, starting from the moment in which the satellite is already touching the primary, is that in typical simulations the undesired particle noise is stronger than the perturbation when the galaxies are far apart from each other (this is visible in Fig. 3(a), taken from [13]). This unrealistic noise acts onto the satellite orbit as an artificial destabilizing perturbation, causing quick infalls.

For the weak perturbation regimes, the PP method can be applied with the effect of a much reduced, more realistic noise level (Fig. 3(b)). We have finally computed complete mergings, starting from separations between the two galaxies that were comparable to a separation inside a group of galaxies (60 kpc). The computed merging times are model dependent, and many sim-

[1] namely first-order interactions between an intruder and the stars of an infinite homogeneous medium

plifying assumptions are to be re-examined (see next section). Still, when the friction is very weak, and when the orbital period of the satellite is longer than the dynamical time of any star of the primary, allowing efficient phase-mixing of the perturbation, the infall is very slow (Fig. 4). This result could not be found in classical simulations, which obviously under-estimated the merging times. We conclude that the so-called compact group paradox, i.e. the strange longevity observed for small dense groups of galaxies while simulations predicted a violent instability towards merging, must have been exaggerated.

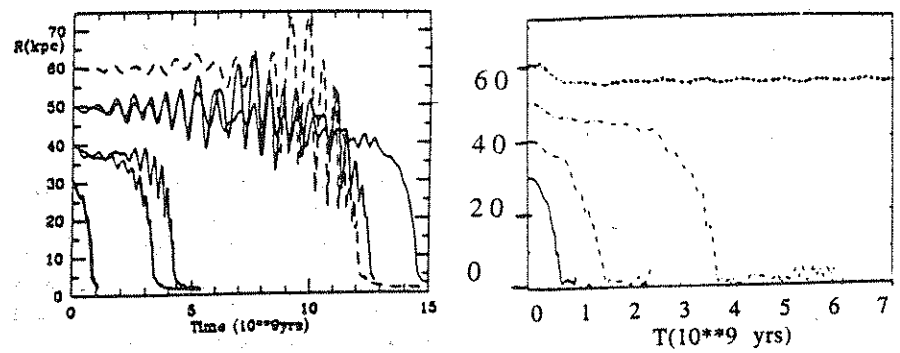


Figure 3: Time-evolution of the satellite-primary distance for various initial separations: (a) reproduced from [13], (b) simulations with PP's

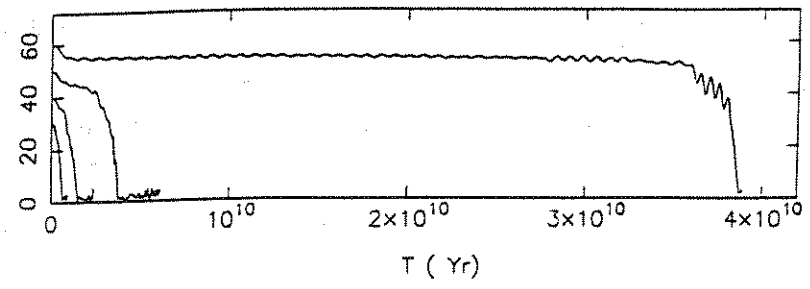


Figure 4: Enlarged view of fig. 3(b)

Prospects and work in progress

Much work is still needed to derive realistic merging rates inside the various environments: field, clusters, or groups of galaxies. Because it reduces the numerical noise, which leads to over-estimating these rates, the method described above could be valuably applied to simulations of merging between two or more "live" galaxies, without modeling any as a rigid body. One should account as well for the various possible distributions of dark-matter, especially for groups or clusters.

A more fundamental application under study within the astrophysics group at the Scuola Normale consists in the attempt to link the existing analytical methods and numerical simulations for modal analyses; while each of these techniques aims at one of the two ends of a global mode life, our scope is to follow a mode inside a stellar system from birth to maturity. A direct comparison with semi-analytical results is planned for the modal analyses for models of elliptical galaxies described in Ref [2].

References

1. Bender, R., 1989. In *Dynamics of Interacting Galaxies*, ed. R. Wielen, Springer, Heidelberg, p. 232.

2. Bertin, G., Pegoraro, F., Rubini, F., Vesperini, E., 1994, *Astrophys. J.*, **434**, 94.

3. Bertin, G., Stiavelli, M., 1984, *Astron. Astrophys.*, **137**, 26.

4. Binney, J.J., Tremaine, S., 1987 *Galactic Dynamics*, Princeton University Press, Princeton.

5. Bontekoe, Tj.R., van Albada, T.S., 1987, *Mon. Not. Roy. Astron. Soc.*, **224**, 349.

6. Chandrasekhar, S., 1943, *Astrophys. J.*, **97**, 255.

7. Denton, R.E., Kotschenreuther, M., 1995, *J. Comp. Phys.*, in press

8. Leeuwin, F., Combes, F., Binney, J., 1993, *Mon. Not. Roy. Astron. Soc.*, **262**, 1013.

9. Leeuwin, F., Combes, F., 1994, *Mon. Not. Roy. Astron. Soc.*, submitted

10. Lin, D.N.C., Tremaine, S., 1983, *Astrophys. J.*, **264**, 364.

11. Hernquist, L., Weinberg, M.D., 1989, *Mon. Not. Roy. Astron. Soc.*, **238**, 407.

12. Merritt, D., 1986 In *Structure and Dynamics of Elliptical Galaxies*, IAU symp. No **127**, ed. T. de Zeeuw, Reidel, Dordrecht, p. 315 .

13. Prugniel, P., Combes, F., 1992, *Astron. Astrophys.*, **259**, 25.

14. Schweizer, F., 1989. In *Dynamics of Interacting Galaxies*, ed. R. Wielen, Springer, Heidelberg, p. 60.

15. Toomre, A., Toomre, J., 1972, *Astrophys. J.*, **178**, 623.

16. Wachlin, F., Rybicki, G.B., Muzzio, J.C., 1993, *Mon. Not. Roy. Astron. Soc.*, **262**, 1007.

VIRGO sensitivity to low frequency gravitational waves

Giovanni Losurdo[1,2]

March 11

1. Scuola Normale Superiore, Pisa, Italy
2. Istituto Nazionale di Fisica Nucleare, Sezione di Pisa, Italy

Introduction

In a few years, interferometric antennas for the detection of gravitational waves (GW) will be operating. With respect to the bar antennas they will have greater sensitivity and, what is more striking, they will be *wideband* detectors. Expected sources of gravitational radiation can be divided in two main groups: those which emit *periodic* (or quasi-periodic) radiation, such as binary systems and rotating neutron stars, and those which emit bursts of GW. Bursts are emitted typically in catastrophic collapses, such as supernova explosions, coalescence of compact binaries, fall of stars into black holes (for a complete review about sources of gravitational waves and detection techniques see Ref.[11]).

The radiation emitted by these events will cover a very wide range of frequencies. The main purpose of the interferometric antenna VIRGO is to investigate the window between a few Hz and a few kHz (see Ref.[5] and references therein).

Rotating neutron stars as sources of gravitational waves

Among the possible sources of periodic waves, those which can be detected by ground-based antennas are the rotating neutron stars [2,10]. Non-axisymmetric neutron stars, that is stars whose mass distribution is not symmetric with respect to the axis of rotation, are expected to emit GW at frequency $f_{GW} = 2f_{spin}$. The predicted luminosity is [12]:

$$L_{GW} = \frac{32}{5}\frac{G}{c^5}(\epsilon_0\theta_W)^2 J^2 (2\pi f_{spin})^6,$$

where J is the moment of inertia with respect to the axis of rotation, the parameter ϵ_0 measures the oblateness of the star and the wobble angle θ_W

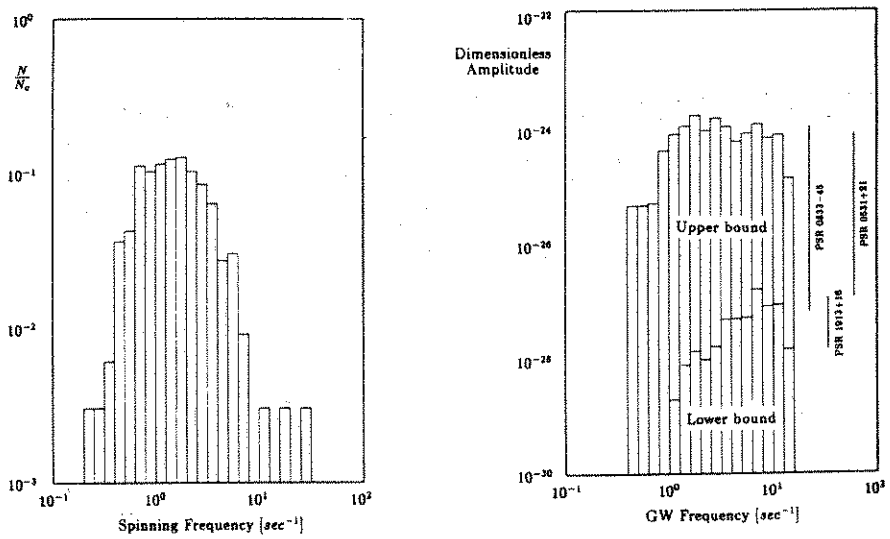


Figure 1: Sample of 334 galactic pulsars: a) frequency distribution, b) expected h on Earth (from Ref.[1]).

gives the departure from axisymmetry. At distance R from Earth, gravitational waves emitted with luminosity L_{GW} should have a dimensionless amplitude:

$$h \approx \sqrt{\frac{G}{\pi^2 c^3}} \frac{\sqrt{L_{GW}}}{f_{GW} R}.$$

Estimates of h for known galactic pulsars are affected by uncertainties of several orders of magnitudes, mostly related to the poor knowledge of the parameter $\epsilon_0 \theta_W$. Figure 1 shows the characteristics of a sample of 334 galactic pulsars: frequency distribution and expected h on Earth. The upper bound of expected GW amplitude is calculated assuming that the observed spindown is entirely due to emission of GW. The lower bound is calculated on the basis of conservative estimates of ϵ_0 [12,1].

In Table 1 the properties of a subsample of pulsars are listed.

The super-attenuator: "VIRGO" sensitivity at low frequency waves

The spectrum of the seismic noise, usually measured in units of $m/\sqrt{Hz}$, behaves approximately as:

$$\tilde{x}(f) \sim \frac{A}{f^2}, \qquad (1)$$

where $A \sim 10^{-6} \div 10^{-7}\, m/Hz^{3/2}$. The VIRGO goal in the low frequency region is to reach a spectral strain sensitivity of $\tilde{h} = 10^{-21}\, 1/\sqrt{Hz}$ at 10 Hz. In order to achieve this goal the seismic noise must be reduced at least by a factor 10^9.

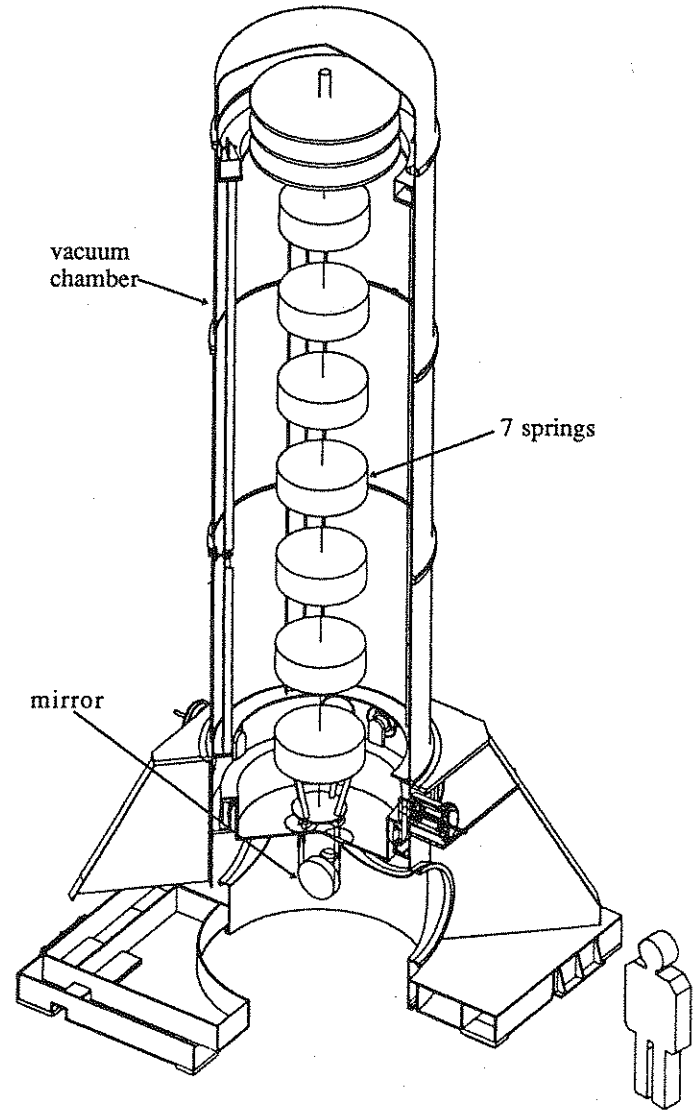


Figure 2: Front view of the super-attenuator.

pulsar	$P(sec)$	f_{GW}	$R(pc)$	h_{max}	h_{min}
PSR 0833-45	0.08925	22	500	$1.2 \cdot 10^{-24}$	$6.7 \cdot 10^{-28}$
PSR 0531+21	0.03313	60	2000	$0.9 \cdot 10^{-24}$	$1.2 \cdot 10^{-27}$
PSR 1929+10	0.22652	9	800	$0.4 \cdot 10^{-24}$	$0.6 \cdot 10^{-27}$
PSR 1916+14	1.18088	1.5	760	$0.3 \cdot 10^{-24}$	$0.2 \cdot 10^{-29}$
PSR 1822-09	0.76896	2.5	560	$0.2 \cdot 10^{-24}$	$0.8 \cdot 10^{-29}$
PSR 1133+16	1.18791	1.5	150	$0.2 \cdot 10^{-24}$	$1.3 \cdot 10^{-27}$
PSR 0950+08	0.25306	8	90	$0.2 \cdot 10^{-24}$	$0.4 \cdot 10^{-27}$
PSR 0740-28	0.16675	12	1500	$0.1 \cdot 10^{-24}$	$0.6 \cdot 10^{-28}$
PSR 1055-52	0.19711	10	920	$0.9 \cdot 10^{-25}$	$0.7 \cdot 10^{-28}$
PSR 0450+55	0.34073	6	450	$0.9 \cdot 10^{-25}$	$0.5 \cdot 10^{-28}$
PSR 1937+21	0.00156	1300	2000	$0.5 \cdot 10^{-24}$	$0.3 \cdot 10^{-26}$
PSR 1855+09	0.00536	370	350	$0.3 \cdot 10^{-24}$	$0.3 \cdot 10^{-26}$
PSR 1953+29	0.00613	326	3500	$0.2 \cdot 10^{-25}$	$0.4 \cdot 10^{-27}$
PSR 1913+16	0.05903	34	5200	$0.1 \cdot 10^{-26}$	$0.1 \cdot 10^{-27}$
PSR 1830-08	0.08528	23	14000	$0.1 \cdot 10^{-25}$	$0.3 \cdot 10^{-28}$

Table 1

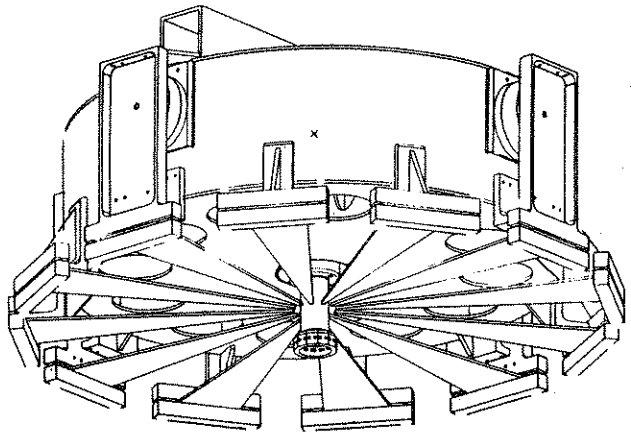


Figure 3: View of a single filter of the super-attenuator. The triangular blade springs which provide the vertical attenuation are visible.

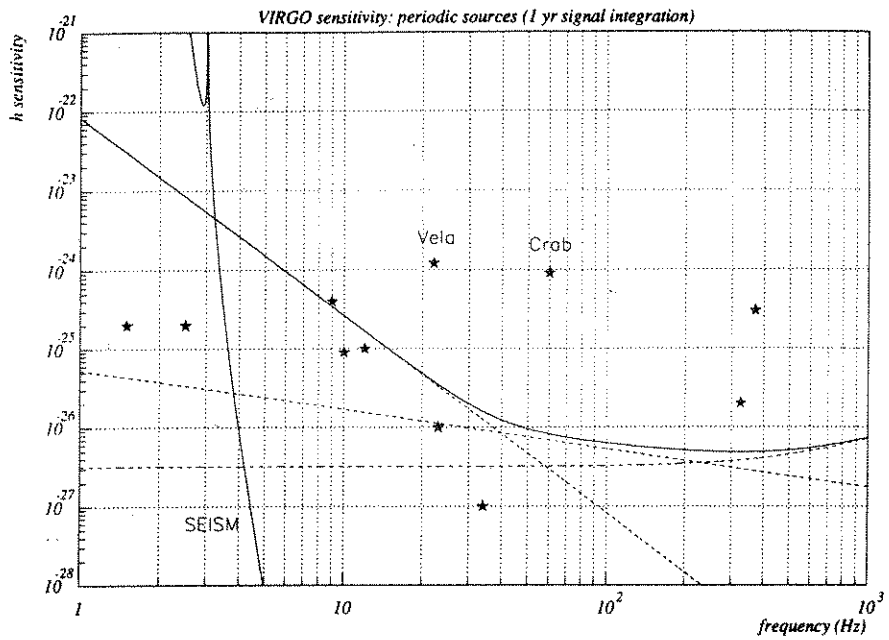


Figure 4: VIRGO h sensitivity to periodic sources (1 yr signal integration). Notice that the seismic noise dominates at very low frequencies, but it is "cut" for $f < 4$ Hz: this clearly illustrates the action of the super-attenuator. The dashed lines represent the suspension thermal noise (which dominates at $4 < f < 40$ Hz), the mirror thermal noise ($40 < f < 200$ Hz), and the laser shot noise ($f > 200$ Hz). The "stars" mark the expected h_{max} of the pulsars listed in the Table.

Each mirror of the VIRGO interferometer will be suspended by a special suspension system, called *super-attenuator* [7,8,9,13]. As shown in Fig.2, the super-attenuator is a seven-fold pendulum about 10 meter high, which acts as a very steep low pass filter. Due to the mechanical coupling among the different degrees of freedom of the suspension, vibrations in the vertical direction can produce mirror displacements in the laser beam direction, thus lowering the antenna sensitivity. In order to avoid this, each stage of the multipendulum is suspended from the upper one through a system of blade springs (see Fig.3). This means that the super-attenuator acts as a cascade of harmonic oscillators both in the transverse and in the longitudinal direction with respect to the laser beam, thus acting as a low pass filter for both the horizontal and the vertical seismic noise (for details, see Ref.[13,4]).

The VIRGO strain sensitivity to periodic GW (based on a 1 year integration of the signal), in the range 1 Hz–1 kHz, is shown in Fig.4, compared with the expected h_{max} of the pulsars listed in the table.

Further developments

The super-attenuator is successful in making the mirrors "free" from seismic noise at frequencies $f > 5$ Hz. Nevertheless, the mirrors would oscillate with amplitude $\delta x \sim 30\,\mu m$ at the frequencies of normal modes of the multipendulum. Such oscillation amplitude cannot be tolerated by the interferometer global control system, that is, the feedback system which keeps the optical cavities in resonance and the interferometer on the dark fringe. Therefore, it is necessary to damp the super-attenuator normal modes [6].

A six dimensional active damping system, which makes use of accelerometers as sensors of the resonant vibrations of the suspensions, is being tested now. A very low noise wideband accelerometer, suitable for the purpose, has been designed and built by the VIRGO group in Pisa [3]. Six accelerometers (one for each degree of freedom) are rigidly fixed on one of the super-attenuator stages. Their signals are used to derive six feedback forces on the stages, whose effect is to damp the super-attenuator resonant motions.

In conclusion, the VIRGO antenna should be able to detect GW from periodic sources such as rotating neutron stars. A significant effort is currently being made in order to reduce the suspension thermal noise and to increase the sensitivity of the antenna around 10 Hz (where most pulsars are expected to emit).

The construction of the VIRGO antenna will start in 1995.

References

1. Barone, F., et al., 1988, *Astron. Astrophys.*, **203**, 322.

2. Blair, D.G., 1991, in *The Detection of Gravitational Waves*, D.G. Blair, ed., Cambridge Univ. Press, Cambridge, p. 17.

3. Braccini, S., et al., 1994, *Rev. Scient. Instr.*, to be published.

4. Braccini, S., et al., 1994, *Phys. Lett. A*, submitted.

5. Braccini, S., 1995, this volume.

6. Bradaschia, C., et al., 1989, *Phys. Lett. A*, **137**, 329 .

7. Del Fabbro, R., et al., 1987, *Phys. Lett. A*, **124**, 253.

8. Del Fabbro, R., et al., 1988, *Phys. Lett. A*, **132**, 237.

9. Del Fabbro, R., et al., 1988, *Phys. Lett. A*, **133**, 471.

10. Schutz, B.F., 1988, *Ann. New York Acad. Sci.*, **571**, 27.

11. Thorne, K.S., 1987, in *300 Years of Gravitation*, S.W. Hawking and W. Israel, eds., Cambridge Univ. Press, Cambridge, p. 330.

12. Zimmermann, M., 1978, *Nature*, **271**, 524.

13. VIRGO Collaboration, 1992, *VIRGO Final Conceptual Design*, INFN, Pisa.

Chaotic dynamics in high-dimensional Hamiltonian systems

Lapo Casetti[1,2]

March 18

1. Scuola Normale Superiore, Pisa, Italy

2. Istituto Nazionale di Fisica Nucleare, Sezione di Firenze, Italy

Classical Hamiltonian dynamical systems, with $N \gg 1$ degrees of freedom, described by Hamiltonian functions of the form

$$\mathcal{H}(\mathbf{p},\mathbf{q}) = \frac{1}{2}\sum_{i=1}^{N} p_i^2 + V(q_1,\ldots,q_N) \qquad (1)$$

where V is a nonlinear potential, arise in many physical situations, ranging from astrophysics to plasma and condensed matter physics.

The statistical approach to systems like (1) by-passes the dynamics by means of a probabilistic hypotesis, the so-called *ergodic hypotesis*, which dates back to Boltzmann and states that asymptotic time averages of any observable $O(\mathbf{p},\mathbf{q})$ taken along a dynamical trajectory are equal to phase averages, i.e. to static averages computed with a distribution function $\varrho(\mathbf{p},\mathbf{q})$ which is uniform on the constant energy hypersurface in phase space (microcanonical measure). This approach is effective and powerful in most physical situations. Nevertheless the deep relation between the statistical approach and "true" dynamics, i.e. the dynamics given by (1) through Hamilton's equations is still far from being understood [14,17]. A sufficient condition for the validity of the ergodic assumption is *mixing*: roughly speaking, a dynamical system is mixing if any subset of the accessible phase space is transformed by the dynamics in such a way that after a sufficiently long time it fills the whole constant energy hypersurface. On the other hand, phase space volume is conserved by the dynamics (Liouville theorem), hence mixing can appear as the result of repeated *stretchings*, due to the exponential instability of the dynamical trajectories, and *foldings*, due to the fact that such an instability takes place in compact spaces. These features are the signature of deterministic chaos, and represent

the link between chaotic dynamics and the dynamical foundations of statistical mechanics, as it was first realized by Krylov [13]. Moreover, the average rate of exponential instability of the trajectories, measured by the Lyapunov exponent λ, quantifies the degree of chaoticity of a dynamical system and can be computed in a numerical simulation [1].

Chaotic dynamics is also relevant to understand how equilibrium is reached and whether there are any observable phenomena related to the existence of different mixing time scales in physical systems. The key example in this perspective is the famous Fermi-Pasta-Ulam (FPU) problem [8]: in a chain of linear oscillators coupled by a weak anharmonicity, as the energy density $\varepsilon = E/N$ is below a threshold value, mixing times become so long that the dynamics – which is indeed chaotic at any ε – can be easily mistaken for a regular one. There is also numerical evidence that such phenomena could be relevant for realistic condensed matter systems [5]. A complete explanation of these puzzling features of Hamiltonian dynamics is still lacking. Nevertheless, it is clear that such an explanation is deeply connected with the problem of the origin of chaos in Hamiltonian systems [3,4,15].

The dynamics of integrable systems, where N independent integrals of motion exist, is regular, but chaos suddenly appears once a small perturbation is added to an integrable Hamiltonian. This is a very weak kind of chaos, and the explanation of its origin dates back to Poincaré [16]: such an explanation is suitable for low-dimensional systems,[1] and becomes of rather difficult interpretation as the number N of degrees of freedom grows, or as the perturbation is no longer small. Another explanation of the origin of chaos, which is completely different from Poincaré's approach, exists and applies to the case of abstract systems like the so-called Anosov flows, where the degree of chaoticity is extremely high. Such systems are studied by the ergodic theory and belong to pure mathematics rather than to theoretical physics: here the origin of chaos – grasped by Krylov – lies in the intrinsic instability of the geodesics on a manifold whose curvature is everywhere negative. Unfortunately, it is extremely difficult to extend these explanations towards Hamiltonian systems of physical interest, which lie somewhere in the middle between these two extremes, whence the need for theoretical tools bridging the two limits. Some results in this direction have been achieved with a geometric approach along the lines of Krylov's original intuitions. These advances will be briefly reviewed in the following, together with a discussion of the relevance of such concepts for the gravitational N-body problem.

Newtonian dynamics, described by Hamiltonians of the form (1), or equivalently by the Lagrangian

[1] Poincaré discovered chaos in Hamiltonian dynamics studying the three-body problem in celestial mechanics.

$$\mathcal{L}(\dot{\mathbf{q}}, \mathbf{q}) = \frac{1}{2}\sum_{i=1}^{N} \dot{q}_i^2 - V(q_1, \ldots, q_N) \,, \tag{2}$$

can be geometrized because the dynamical trajectories are obtained by means of variational principles. In fact, the natural motions are the extrema, among synchronous paths with fixed endpoints in configuration space, of the Hamiltonian action functional

$$\mathcal{S}_H = \int_{t_0}^{t_1} \mathcal{L}(\dot{\mathbf{q}}, \mathbf{q})\, dt \,, \tag{3}$$

or of the Maupertuis' action

$$\mathcal{S}_M = 2\int_{\gamma(t)} T(\dot{\mathbf{q}}, \mathbf{q})\, dt \tag{4}$$

among the isoenergetic paths $\gamma(t)$; here T is the kinetic energy. The geodesics of a Riemannian manifold are given by the stationary length principle $\delta\ell = 0$, where

$$\ell = \int_{P_0}^{P_1} ds \tag{5}$$

and $ds^2 = g_{\mu\nu} dq^\mu dq^\nu$, hence a suitable choice of the metric tensor allows the identification of the arc length (5) with either $\mathcal{S}_H$ or $\mathcal{S}_M$. In the latter case the metric is given by a conformal transformation of the kinetic energy metric a_{ij} associated with the free ($V = 0$) Lagrangian[2],

$$(g_J)_{ij} = [E - V(\mathbf{q})]\, a_{ij} \tag{6}$$

and is referred to as the Jacobi metric. If equipped with g_J, the accessible configuration manifold $\{q^1, \ldots, q^N : V < E\}$ is given a proper Riemannian structure and the geodesics are just the extrema of $\mathcal{S}_M$. A description of the extrema of Hamilton's action (3) as geodesics of a manifold can be obtained using Eisenhart's metric [7] on an enlarged configuration spacetime ($\{t \equiv q^0, q^1, \ldots, q^N\}$ plus one real coordinate q^{N+1}), whose arc-length is

$$ds^2 = -2V(\mathbf{q})(dq^0)^2 + a_{ij} dq^i dq^j + 2dq^0 dq^{N+1} \,. \tag{7}$$

The manifold acquires a Lorentzian structure and the dynamical trajectories are recovered as the geodesics satisfying the condition $ds^2 = C dt^2$ with $C > 0$ [3]. In the geometrical framework, the stability of the trajectories is mapped on the stability of the geodesics, hence it can be studied by the Jacobi equation for geodesic deviation

$$\frac{D^2 J^\mu}{ds^2} + R^\mu_{\nu\rho\sigma} \frac{dq^\nu}{ds} J^\rho \frac{dq^\sigma}{ds} = 0 \,, \tag{8}$$

[2] In a Cartesian reference frame $a_{ij} = \delta_{ij}$.

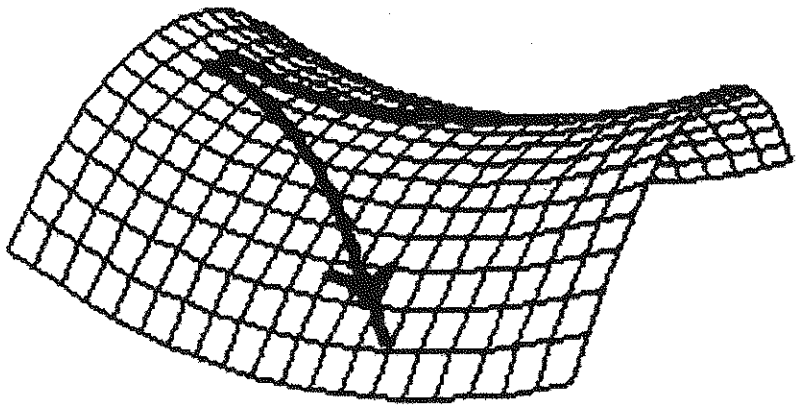


Figure 1: Instability of the geodesics on a saddle-shaped surface, the simplest example of negative-curvature manifold.

where R is the Riemann curvature tensor and J, which measures the deviation between nearby geodesics, is referred to as the Jacobi field.

The relation between chaos and geometry is summarized in the Jacobi equation (8) which links the stability – or instability – of the dynamics with the curvature properties of the underlying manifold. In the particular case of *constant curvature* manifolds, Eq. (8) becomes very simple [2]

$$\frac{D^2 J^\mu}{ds^2} + K\, J^\mu = 0 \ , \tag{9}$$

and has bounded oscillating solutions $J \approx \cos(\sqrt{K}\, s)$ or exponentially unstable solutions $J \approx \exp(\sqrt{-K}\, s)$ according to the sign of the constant sectional curvature K, which is given by

$$K = \frac{K_R}{N-1} = \frac{\mathcal{R}}{N(N-1)} \ . \tag{10}$$

where $K_R = R_{\mu\nu}\dot{q}^\mu\dot{q}^\nu$ is the Ricci curvature and $\mathcal{R} = R^\mu_\mu$ is the scalar curvature; $R_{\mu\nu}$ is the Ricci tensor. Manifolds with $K < 0$ are considered in abstract ergodic theory, and, according to Eq. (9), the geodesics on such manifolds are extremely chaotic (see Fig. 1), with a Lyapunov exponent $\lambda = \sqrt{-K}$. Krylov [13] originally proposed that the presence of some negative curvature could be the mechanism actually at work to make chaos in physical systems,

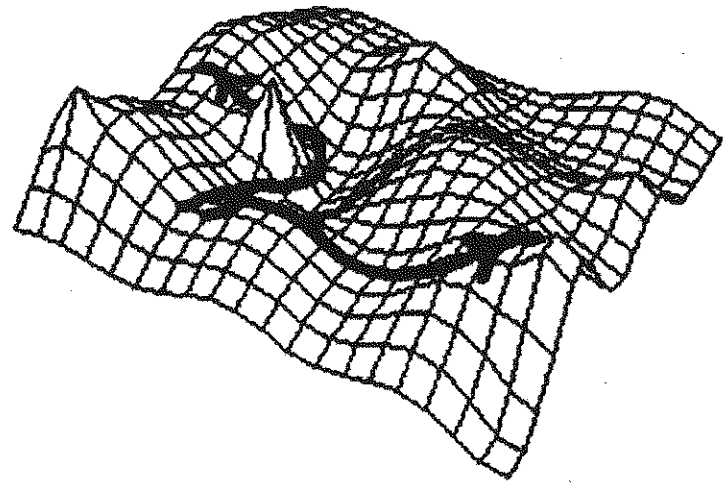


Figure 2: Destabilization of geodesics on a bumpy surface via parametric resonance.

but in realistic cases the curvatures are neither found constant nor everywhere negative, and the straightforward approach based on Eq. (9) does not apply. This is the main reason why Krylov's ideas remained confined to abstract ergodic theory with few exceptions.

In spite of these major problems, some approximations on Eq. (8) are possible even in the general case, and lead to new insight in the problem of the nature of chaos in Hamiltonian dynamics, with the aid of numerical simulations [3,4,15]. The key point is that negative curvatures are not strictly necessary to make chaos, and that a subtler mechanism related to the *bumpiness* of the ambient manifold is actually at work. From Eq. (8) an evolution equation for the norm of J is derived. Then such an equation is cast in the Hill form

$$\frac{d^2\psi}{dt^2} + Q(t)\,\psi = 0 \ , \tag{11}$$

where $\psi \propto \|J\|$ and $Q(t)$ is tightly related to the sectional curvature K which is no longer a constant. In the case of Eisenhart's metric (7) $Q(t) = K(t)$ is a wildly fluctuating function taking mostly *positive* values (in some cases like the FPU model $K > 0$), whence the solutions of Eq. (11) can be subject to *parametric instability*. Curvature fluctuations can produce chaos even if no negative curvature is experienced by the geodesics (see Fig. 2). As the sectional curvature is no longer constant, K_R and $\mathcal{R}$ are respectively averages of K over the direction of J and over both the direction of J and the direction of

the reference geodesic. Equation (10) no longer holds, nevertheless it is a first order approximation to which an estimate of the curvature fluctuations can be added to obtain a stochastic model of $K(t)$ independent of the dynamics of the system. This model leads to an analytical estimate of the Lyapunov exponent λ, which is correct at least for the FPU model in the limit $N \to \infty$ [4].

The gravitational N-body system is one of the few cases, among physically relevant systems, where a reduction of the problem of relaxation and mixing to the geometric framework of ergodic theory has been attempted [9-12]. The system is defined by the following Hamiltonian

$$\mathcal{H}(\mathbf{p},\mathbf{q}) = \sum_{i=1}^{N} \frac{1}{2m_i}(p_{i_x}^2 + p_{i_y}^2 + p_{i_z}^2) - G\sum_{i<j} \frac{m_i m_j}{|\mathbf{r}_i - \mathbf{r}_j|} \; ; \tag{12}$$

in the above cited works it is claimed that such a system is practically equivalent to a geodesic flow on a manifold whose curvature is everywhere negative. Such a claim stems from an estimate of the sectional curvature, in the framework of the Jacobi metric (6), in terms of the scalar curvature of the configuration manifold. The scalar curvature $\mathcal{R}$ is indeed almost everywhere negative for the Hamiltonian (12), provided N is large enough. A collective relaxation time $\tau_{\rm coll}$ related to $\mathcal{R}$ can be introduced, and for a dense stellar system (e.g. a globular cluster) of linear extension D it is approximately given by

$$\tau_{\rm coll} \simeq \frac{D}{d}\tau_{\rm dyn} \; , \tag{13}$$

where d is the typical distance between stars and $\tau_{\rm dyn} = D^{3/2}/(GmN)^{1/2}$ is the typical dynamical time scale of the system. The collective relaxation time turns out to be much smaller than the relaxation time due to binary collisions

$$\tau_{\rm bin} \simeq \frac{D}{r_* \log N}\tau_{\rm dyn} \; , \tag{14}$$

where $r_* = (Gm)/v^2 \ll d$ is the effective gravitational radius of a star; evidence for a possible relevance of collective relaxation has been found in globular clusters [18].

Nevertheless, a more detailed analysis of the gravitational N-body problem along the lines discussed previously [6] has recently pointed out that such a heuristic approach can be misleading. In fact $\mathcal{R} < 0$ does not ensure $Q(t) < 0$ in Eq. (11), first because the scalar curvature $\mathcal{R}$ is not a good estimate of the actual curvature as long as the latter is not a constant, and then because other important terms enter $Q(t)$ and can affect its sign. The results of accurate numerical simulations clearly show that, even if the scalar curvature is mostly negative, the mechanism where chaos comes from is once more a sort of parametric resonance due to wild oscillations of $Q(t)$.

The gravitational N-body system is thus an illuminating example of how the relation between chaotic dynamics and Riemannian geometry is deep but

nontrivial. Naive attempts to borrow analytical tools from abstract ergodic theory can lead to misleading conclusions if not supported by a deep analysis where the role played by numerical simulations is still essential. This is a key point, because rigorous mathematical results can be worked out only in the case of constant curvature manifolds. Nevertheless, numerical simulations can suggest new insight in the problem of chaotic dynamics, as it actually happened in the case of the discovery of the relevance of parametric instability due to curvature fluctuations, and can provide new hints towards a Riemannian theory of Hamiltonian chaos.

References

1. Benettin, G., Galgani, L., and Strelcyn, J.M. 1976, *Phys. Rev. A*, **14**, 2338
2. do Carmo, M.P. 1992, *Riemannian geometry*, Birkhäuser, Boston
3. Casetti, L., and Pettini, M. 1993, *Phys. Rev. E*, **48**, 4320
4. Casetti, L., Livi, R., and Pettini, M. 1995, *Phys. Rev. Lett.*, **74**, 375
5. Casetti, L., Livi, R., Macchi, A., and Pettini, M. 1995, *Phys. Rev. Lett.*, submitted
6. Cerruti-Sola, M., and Pettiní, M. 1994, *Phys. Rev. E*, **51**, 53
7. Eisenhart, L.P. 1929, *Ann. of Math. (Princeton)*, **30**, 591
8. Fermi, E., Pasta, J., and Ulam, S. 1955, Los Alamos Report LA-1940, in Segré, E. (ed.) 1965, *Collected works of Enrico Fermi*, Chicago University Press, Chicago
9. Gurzadyan, V.G., and Savvidy, G.K. 1986, *Astron. Astrophys.*, **160**, 203
10. Gurzadyan, V.G., and Kocharian, A.A. 1987, *Astrophys. Space Sci.*, **135**, 307
11. Kandrup, H.E. 1989, *Phys. Lett. A*, **140**, 97
12. Kandrup, H.E. 1990, *Astrophys. J.*, **364**, 420
13. Krylov, N.S. 1942, PhD thesis, in Krylov, N.S. 1979, *Works on the foundations of statistical physics*, Princeton University Press, Princeton
14. Ma, S.K. 1985, *Statistical Mechanics*, World Scientific, Singapore
15. Pettini, M. 1993, *Phys. Rev. E*, **48**, 828
16. Poincaré, H. 1892, *Les méthodes nouvelles de la mécanique céleste*, Gauthier-Villars, Paris
17. Toda, M., Kubo, R., and Saitô, N. 1991, *Statistical physics I: equilibrium statistical mechanics*, 2nd ed., Ch. 5, Springer-Verlag, Berlin
18. Vesperini, E. 1992, *Europhys. Lett.*, **17**, 661

Induced star formation and non-linear multipopulation models for galactic evolution

Federico Ferrini*

April 22

* Sezione di Astronomia, Dipartimento di Fisica, Università di Pisa, Italy

Nonlinear multiple population models for galactic evolution are required to understand the complex dynamics of the individual phases associated with real galaxies; in turn, these models suggest the existence of complex star forming histories for the galaxies. This review outlines the physical idea at the basis of the nonlinear treatment of global star formation processes and how the approach changes the way that one views the ecology of a galactic system. These models serve as a framework for including detailed dynamical, chemical, and radiative microphysics in an integrated picture of galactic star formation. A thorough discussion of all the relevant aspects of the problem can be found in [14].

The first attempts to determine the dependence of the rate of galactic star formation on the density of interstellar gas immediately suggested the existence of nonlinear relations. Schmidt [12,13], in a series of papers that are the "mother of all nonlinear models", established this observationally by the simplest of procedures. He looked at the space density of stars in different regions of the Galaxy and compared it with the number density of neutral hydrogen obtained from the 21 cm line emission. At that time, there was no hint that a molecular component was actually the hidden player in galactic evolution, the actual phase out of which stars are formed. Schmidt found that the number of OB stars was a function of the gas density ρ_g following the relation $\psi \sim \rho_g^n$, where n turned out to be around 2. Even the earliest attempt to understand this empirical law required some form of interaction between the different components of the Galaxy. For instance, Field and Saslaw [7] pointed out that a cloud-cloud collision mechanism would imply the required density dependence, although the details of how such events could produce stars were unclear.

By the early 1970s, the "closed box" model for galactic evolution, which dominated all discussions of the consequences of the Schmidt star formation law, had already met some difficulties in relation to some observational facts. As an example, the star formation rate inferred at current epochs would lead to unacceptably high rates of star formation in the past.

The idea that stars can stimulate new star formation was best promoted by Elmegreen and Lada [3]. The most important aspect of this influential paper was its focus on the *normal* formation and evolution of an H II region at the periphery of a massive, self-gravitating (or nearly so) molecular cloud. They pointed out that, much like a flame igniting new fires in the presence of adequate fuel supplies, star formation can propagate into the molecular cloud. This process can take place as a natural by-product of ordinary star forming activity. In their picture, it is inevitable that star formation should be able to propagate and that the cloud phase carries within it the seed for its own destruction.

Stimulated star formation refers to the direct action of stars on the interstellar medium (ISM). The stellar population drives most of the energetics of the ISM, resupplies consumed mass, pumps radiation out through nuclear reactions, accelerates cosmic rays, produces supernovae and stellar winds that are important sources of kinetic energy and turbulence, and generally determines the structure of the global physical environment of a galaxy. It is therefore intuitively clear that any model for galactic evolution must include some interaction between the stars thus formed and the surrounding gaseous components of the system. Nonlinearity means that a galaxy evolves as a result of the mutual interconversion and interaction of different phases. It also means that a limit cycle behavior, bursts, and stochasticity may all appear.

Nonlinear population models. Predation, illness, and conversion: building up the formal framework.

In order to introduce the idea of stimulated processes in multipopulation models, I begin with an example which everyone is familiar with. We all have had the experience of coming down with a cold. One day you are perfectly healthy. You go into the office only to find that someone there has all the symptoms of a cold, but you still have to get your work done so you stay around. A few days later, you have all the raging symptoms yourself. You *were* well and now you *are* sick and the cause was some external trigger that depended on contact of some sort (*e.g.* [1]). Further, in this new state you are capable of spreading the infection. This is what is meant by a "stimulated process".

Consider a box (the office, for instance) composed of several populations (*i.e.*, sick and well), $\{X_i\}$, all of which can inter-convert (*i.e.* become ill or recover). The rate at which one population transforms into another is given by some constant a_{ij} (the infection and recovery rates). The assumption that the rates are constant in time is central to the formalism. Such a model is called

autonomous. It is fictitious, of course, unless the rates are examined over only a short time interval compared with the rate at which the environment changes. Pairwise interactions between populations provide the non–linear term for the processes. Then:

$$\frac{dX_i}{dt} = \sum_j b_{ij} X_j + \sum_{j \neq i} a_{ij} X_i X_j \tag{1}$$

is a model for the evolution of the system. The equilibrium solution, given by $\dot{X}_i = 0$, will be designated by $X_{0,i}$. This state is not one that the system need ever achieve (that is, not everyone in the office is necessarily going to be simultaneously ill or recovered and immune). If the system at the time of first observation is in equilibrium – and this state may take a long time to reach – then its subsequent evolution following some perturbation is given by the solution of Eq. (1). To continue for a moment with the disease example, imagine that everyone in the office is initially well and a visitor arrives who is ill. This is the perturbation. The rates are still assumed to be the same in the perturbed state as in equilibrium. Then writing $X_i(t) = X_{0,i} + x_i(t)$, where x_i is a small deviation taken to be a simple function of time, for instance $x_i(t) \sim e^{\lambda t}$, we find that:

$$\frac{dx_i}{dt} = \sum_j b_{ij} x_j + \sum_{j \neq i} a_{ij} (X_{0,j} x_i + X_{0,i} x_j). \tag{2}$$

The evolution is thus given by the solution to an eigenvalue problem for λ. A limit cycle results from the existence of periodic solutions. In this case, the fixed point is a focal point around which the state of the system oscillates. An attractor is one for which the system at equilibrium is overstable, such that the periodic mode grows with time and any small departure immediately moves from equilibrium toward the periodic state. Whatever the initial state of the system, bifurcated solutions can also exist such that within some domain of the relevant parameter space, the system will decay while, in other regions, there may be either an attractor or a repeller. In the two-state disease example, there is no limit cycle if the individual develops immunity once he recovers. The illness runs its course through the population at some rate that depends on the infection rate compared with the recovery rate. If the infection rate is low enough and/or the recovery/immunity rate is high enough, the illness will stop propagating. If recovery confers no immunity, then the disease may continue for quite a long time because each recovered individual is a potential infectee after some time.

An autonomous system has the advantage that the coefficients are independent of time. This is equivalent to an instantaneous recycling scheme, since only the properties of the populations at any given moment determine the fate of the system. If the system instead has a memory, the interaction may be

cast in the form of a convolution integral:

$$\frac{dX_i}{dt} = \sum_j \int_0^t b_{ij}(t')X_j(t-t')dt' + \sum_{j\neq i} \int_0^t a_{ij}(t')X_i(t')X_j(t-t')dt' \qquad (3)$$

In the disease model, this is equivalent to an incubation period during which an infected individual is a carrier but otherwise shows no apparent symptoms. An ecosystem may depend on the gestation period for the prey or the predator population. For a galaxy, this may be the time it takes to build molecular clouds.

Building a galaxy: What are the phases?

Since the earliest days of extragalactic astronomy galaxies, the "island universes", have been recognized as complex communities [8]. The individual phases are often viewed as separate, and, because of a wide–spread reductionism, the relationship of one phase to another are often systematically ignored. Here I will briefly discuss each phase, *i.e.* population.

The diffuse gas.

By diffuse gas, I mean the diffuse material that is not incorporated into self-gravitating structures regardless of temperature. It is characterized by a wide range of densities and temperatures, largely because its density is so low that it cools only inefficiently. It has several diverse sources. The first is the material shed by stars. This source is actually an average over the stellar initial mass function. The second source is the evaporation of molecular clouds; this source can be seen as part of the "book–keeping". Finally, diffuse gas, as well as clouds, can rain into the disk from the halo.

The molecular clouds.

The molecular component is the best studied from the observational point of view, partly because it is the most easily identified. This is the galactic phase that is most important for star formation, at least in our Galaxy and in other spirals. Stars are formed in the densest parts of the clouds. They subsequently disperse the medium through the combined effects of stellar winds and supernovae, and they are also able to locally heat the clouds and form expanding internal H II regions. Unlike the gas, which is directly affected by the stars in a linear way, the molecular component is cycled through a complex interaction between the phases. A comprehensive picture of the field is provided by the review articles in [2].

The stars.

The central theme of the stimulated star formation picture, which is the foundation of our nonlinear approach, is that stars drive the evolution of galaxies. This is seen in many diverse actions of the stars on clouds and gas. The active stellar phase dies by producing inert remnants. No other component of

the galaxy does so. Gas can always be recycled among the different phases, and the molecular component probably vanishes at some stages, but the stars are always there. Even if the stars are not the sole cause of new star-forming activity, they are certainly not mere spectators in the process. This is why the analyses of simple star formation laws may be intrinsically unrealistic. If the stellar population is neglected in the proposed power laws for ψ, then the correlations of gas density with the number of massive stars is biased *a priori*.

Principles for representing the population interactions.

All the nonlinear galactic population models that have been explored so far in the literature make a few basically similar assumptions. First, galaxies are recognized to be multi-phase systems consisting of stars (separate, pointlike self-gravitating units), diffuse gas (atomic and warm or hot gaseous material unable to self-shield from the interstellar radiation field), and molecular clouds. As far as we can tell, stars form in clouds. Thus, there is a difference between the diffuse gas that fills the interstellar medium and the matter that coalesces into large scale (greater than a few parsec) cold structures. These large scale structures have finite lifetimes. Finally, stars die, violently or quietly, and return some fraction of their consumed gas to the medium. All of these steps are easily observed. What is not known directly from the observations is the interaction process. The models *assume* that stars are responsible for triggering new star formation, for destroying the clouds, and for driving much of the dynamics of the diffuse medium.

Another assumption is that clouds, and cloud complexes, are comparatively fragile objects. Clouds, once self-gravitating, will form stars. But as soon as the stars start to form, they produce significant changes in their environments. This marks the first appearance of the phase interactions. The molecular material, once dissociated and ionized, ceases to be an effective star formation source. The internal star formation is partly responsible for this. In addition, the increase in any internal turbulence, generated by the interaction of the winds from newly formed stars and the cloud, increases the cloud internal pressure and reduces the efficiency of subsequent star formation. Therefore, the star formation efficiency, and hence the rate, depends on the nonlinear coupling between the stellar and the molecular phase. Finally, the ionization that results from both bulk and surface star formation, and the formation of H II regions, will contribute to the destruction of the cloud even while the lower mass stars are being formed.

Clouds form from the gas. The gas can become unbound from the galaxy (in the case of ram pressure induced stripping, which is an extrinsic process), expelled by strong stellar winds and supernovae (galactic winds, intrinsic), or it can suffer direct removal as a result of galaxy collisions (extrinsic). The diffuse gas is distributed ubiquitously over the disk and the halo of a galaxy. The clouds are only present where they can come into pressure equilibrium with

the diffuse gas and when they are self-shielding. They require the formation of molecules. The clouds are more like ballistic objects than the gas and therefore can be expelled like stars during galactic collisions. They are not easily removed by fountains or ram stripping. In contrast, the only way to get rid of the stars is to disrupt the whole galaxy.

The goals of the models.

The nonlinear models are designed to meet several empirical features. These are the same as the ones that constrain all models for galactic star formation. The tracers of current star formation are far infrared emission, ultraviolet emission, and thermal (H II regions) and nonthermal (supernova remnants and diffuse gas) centimeter radio emission. The history of global star formation is recorded in the metallicity distribution within the stellar and gas phases of a galaxy. Unfortunately, none of the signatures of star formation, or of its history, give direct, unique answers. One of the problems with the standard treatments is that they are constrained to produce metals and remnants based on some *assumed* (somewhat arbitrarily) law for the star formation in the course of galactic history. While the choice of such laws may be common sense, once they have been specified the entire calculation becomes a matter of book–keeping. The models are self-consistent only to the extent that they produce metals and population distributions in accord with the chosen law for star formation. The inherent advantage of the nonlinear models is that they are able to produce, from general interactions, a history of star formation without any ad hoc assumed formal time dependence.

The solar neighborhood and nonlinear treatments.

Models for the evolution of spiral galaxies have been developed by several groups within the general theoretical framework of the multiphase nonlinear approach. The most detailed treatments have concentrated on the analysis of the solar neighborhood. This region is defined as a cylindrical zone with the symmetry axis centered on the Sun and extending in the galactic plane to the typical scale length of about one kiloparsec. The vertical extent of this region incorporates both the disk and the halo.

A complete model for the solar neighborhood requires that the nonlinear evolution equations contain the memory function for the stellar population through the Initial Mass Function (IMF). This is necessary because of the enormous range of ages that must be considered once the halo has been properly included in the model:

$$\frac{ds_{1,i}}{dt} = H_{1,i}c_i^2 + a_{1,i}c_i s_{2,i} - D_{1,i} \tag{4}$$

$$\frac{ds_{2,i}}{dt} = H_{2,i}c_i^2 + a_{2,i}c_i s_{2,i} - D_{2,i} \tag{5}$$

$$\frac{dg_i}{dt} = -\mu_i g_i^n + a_i' c_i s_{2,i} + H_i' c_i^2 + W_i + \mathcal{F}_i \tag{6}$$

$$\frac{dc_i}{dt} = \mu_i g_i^n - (a_{1,i} + a_{2,i} + a_i')c_i s_{2,i} - (H_{1,i} + H_{2,i} + H_i')c_i^2 \tag{7}$$

$$\frac{dr_i}{dt} = D_{1,i} + D_{2,i} - W_i \tag{8}$$

The subscript i refers to halo (H) or disk (D); a detailed and complete list of definitions is given in Ref. [4]. The fraction of mass in diffuse gas is g, and c refers to clouds. Equation (8) follows the remnants of stellar evolution, r_i, because their production depends on the stellar population at any time. This equation for the remnants could be replaced, however, by one for the IMF as a function of population and metallicity, if such a representation were known. The function $\mathcal{F}_i$ describes the coupling between halo and disk and hence is a loss term $-fg_H$ in the set of equations for the halo; it appears in the evolution equation for the disk gas as a gain term with the opposite sign. For convenience, the stars have been divided in two separate populations. The fraction of stars massive enough to induce further star formation are called $s_{2,i}$ (*i.e.*, those more massive than 4 $M_\odot$), while those of lower mass incapable of triggering the interstellar-medium to further condense into protostars are called $s_{1,i}$. The death rates $D_{j,i}$ represent the fraction of mass that leaves the star phase at time t for low–mass and massive stars; the restitution rates W_i give the fraction of mass expelled into the ISM at time t summing the contribution for each element (processed or unprocessed).

The equations for the element abundances, X_j, are written according to the standard approach (see, *e.g.*, [15]) as:

$$\frac{d(X_{j,i}(g_i + c_i))}{dt} = -\psi_i X_{j,i} + \mathcal{F}_i X_{j,i} + W_i \tag{9}$$

Application of specific nonlinear models.

The multiphase approach has been applied to several problems of galactic evolution. The earlier applications to the evolution of the solar neighborhood [4, 10] have been extended to the whole disk, to understand the nature of abundance gradients [5]. The presence of the thick disk component [11] as well as of the bulge [9] have been analyzed and their effects have been studied in detail. Another interesting extension of this type of models is to the case of elliptical galaxies [6].

References

1. Bailey, N. T. J. 1975, *The Mathematical Theory of Infectious Disease*, Haffner, New York

2. Combes, F. & Casoli, F. (eds.) 1991, *Dynamics of Galaxies & their Molecular Cloud Distributions – IAU Symp. 146*, Kluwer, Dordrecht

3. Elmegreen, B.G. & Lada, C. J. 1977, *Astrophys. J.*, **214**, 725

4. Ferrini, F., Matteucci, F., Pardi, M. C., & Penco, U. 1992, *Astrophys. J.*, **387**, 138

5. Ferrini, F., Mollá, M., Pardi, M.C., & Diaz, A.I. 1994, *Astrophys. J.*, **427**, 745

6. Ferrini, F. & Poggianti, B. M. 1993, *Astrophys. J.*, **410**, 44

7. Field, G. B. & Saslaw, W. C. 1965, *Astrophys. J.*, **142**, 568

8. Hubble, E.P. 1930, *Astrophys. J.*, **71**, 231

9. Mollà, M. & Ferrini, F. 1995, *Astrophys. J.*, submitted

10. Pardi, M.C., & Ferrini, F. 1994, *Astrophys. J.*, **421**, 491

11. Pardi, M.C., Ferrini, F. & Matteucci, F. 1995, *Astrophys. J.*, **444**, May 1st issue

12. Schmidt, M. 1959, *Astrophys. J.*, **129**, 243

13. Schmidt, M. 1963, *Astrophys. J.*, **137**, 758

14. Shore, S. N. & Ferrini, F. 1995, *Fund. Cosmic Physics*, in press

15. Tinsley, B. M. 1980, *Fund. Cosmic Physics*, **5**, 287

Recent progress in gamma ray astrophysics (100 MeV to 10 TeV)

David A. Smith*

May 6

* Istituto Nazionale di Fisica Nucleare, Sezione di Pisa, Italy

The launch of the Compton Gamma Ray Observatory (CGRO) on the Space Shuttle in 1991 coincided with breakthroughs at ground-based gamma ray telescopes, sparking renewed interest in astrophysical gamma rays from 100 MeV to 10 TeV. This talk begins with a few of the striking results from the Egret instrument on the CGRO and from the ground based gamma ray telescopes, which include observations of pulsars and Active Galactic Nuclei (AGN's). The AGN results will be interpreted in terms of gamma ray absorption by intergalactic photons through e^+e^- pair production.

The conclusion is that measuring the shape of the gamma ray energy spectra in the window between the upper reach of the satellites (about 10 GeV) and the lower threshold of the ground-based detectors (currently about 400 GeV) probes many topics of fundamental astrophysical interest: the shape of the spectrum in that window tests AGN models, which are related to questions about the orgin of the cosmic radiation around the 'ankle' at 10^{18} eV. The spectra for many sources of different redshifts z gives a measurement, for fixed value of the critical cosmological density Ω_0, of $n(\epsilon)/H_0$, where $n(\epsilon)$ is the number density for diffuse extragalactic photons of energy ϵ (near infrared to the near ultraviolet), and H_0 is the Hubble constant. The soft photon density $n(\epsilon)$, in turn, probes details of galactic evolution in the early universe and of mixed cold and hot dark matter models.

The rest of the talk discusses how to open the window $20 < E_\gamma < 200$ GeV. First I review current gamma ray telescopes, and then outline how to use existing solar power plants as gamma ray detectors.

Results from gamma ray telescopes

The Compton Gamma Ray Observatory was launched in June, 1991 [3,17]. Of the four detectors on board, Egret, the Energetic Gamma Ray Telescope, measures 10 MeV to 30 GeV gamma rays at 5 arc-minute angular resolution with spark chambers and a sodium iodide crystal. A scintillator housing rejects

charged particle backgrounds. Perhaps the most striking result from the Egret instrument is that most of the discrete sources for $E_\gamma > 100$ Mev are extremely distant. Over 40 AGN's, all "blazars", have been detected [13] [1]. There is a comparable list of marginal detections and of unidentified sources at high galactic latitude.

Figure 1 shows the gamma ray spectrum from the Crab nebula over the energy range from Comptel to Themistocle (more on Thémis later). A two-component model for the Crab nebula, dominated by synchrotron emission up to 100 MeV and then by inverse Compton scattering through 10 TeV, fits the data. The long lever arm that comes from combining the ground-based and satellite measurements is a powerful tool for discriminating between different models [6]. As shown in Fig. 1, the Crab nebula has been observed by a number of Čerenkov telescopes. The first was the Whipple telescope at Mt. Hopkins, Arizona, using the imaging technique [16]. The Themistocle and Asgat [5] experiments pioneered the time sampling method at Thémis in the French Pyrenees. Since then, other groups have also seen the Crab, mainly using the imaging method. Only one other galactic source has been detected by a ground-based gamma-ray telescope: the pulsar PSR 1706–44, visible in the Southern hemisphere.

Gamma ray absorption by starlight

The list of reproducible ground-based observations is completed with only one more entry: the blazar Markarian 421 [11]. Mrk 421 is both the dimmest and the closest extragalactic source in the Egret catalog (redshift $z = 0.031$). The casual observer might think that the field is quite poor, with only three detectable sources. But the detection of Mrk 421, when so many brighter sources are invisible, is quite remarkable. For example, the blazar 3C279 has been seen by Egret to flare to 30 times the luminosity of Mrk 421 [8]. Why isn't 3C279 seen, while Mrk 421 is? The most reasonable explanation is that TeV gamma-rays above some energy are absorbed while crossing space through the mechanism $\gamma\gamma \rightarrow e^+e^-$. This process will be discussed below. An alternate explanation, that the acceleration mechanism in AGN's (except for Mrk421!) 'runs out of gas' between satellite and ground energies (like the roll-over in the Crab spectrum in Fig. 1), would be tested by measuring the energy roll-over for many sources at different redshifts. AGN models are described by [13].

A gamma-ray of energy $E_0(1+z)$ from an AGN at a redshift z travelling to Earth traverses a column of 'soft' photons of energy ϵ and density $n(\epsilon)$. The probability for the gamma to survive the trip is $e^{-\tau}$, where the 'optical

[1] A blazar is a core-dominant radio source with a highly variable, usually rather red, polarized optical component. BL Lac objects and Optically Violent Variable quasars (OVV's) are sub-classes of blazars. OVV's have an observable featureless continuum and emission lines, while for the BL Lacs the red polarized component is bright enough to mask the continuum and lines.

depth' τ is, schematically, $\tau \propto \sigma r n(\epsilon)$, for a distance $r = cz/H_0$. The relevant cross-section for pair production, $\gamma\gamma \rightarrow e^+e^-$, can be written as

$$\sigma(\beta) = \sigma_0(1-\beta^2)\Big[2\beta(\beta^2-2) + (3-\beta^4)\ln(\frac{1+\beta}{1-\beta})\Big]$$

with $(1-\beta^2) = \epsilon_1/\epsilon_0$ and $\sigma_0 = 0.125$ barn^2. Here ϵ_1 is the threshold energy,

$$\epsilon_1 = \frac{2m_e^2}{E_0 x(1+z)^2},$$

$x = (1-\cos\theta)$ gives the angle between the two photons, E_0 and ϵ_0 are the energies in the rest frame of the observers, and $E(z), \epsilon(z) = \epsilon_0(1+z)$ are the redshifted photon energies. The general expression for the optical depth is then

$$\tau(E_0, z) = \frac{c}{H_0}\int_0^z dz' \frac{(1+z')}{(1+\Omega_0 z')^{1/2}} \int_0^2 dx \frac{x}{2} \int_{\epsilon_1}^\infty d\epsilon_0 n\Big(\epsilon(z')\Big)\sigma(\beta),$$

where β depends on x, ϵ_0 and z'. Absorbing photons in the near infrared (2 μm $\simeq$ 0.6 eV) dominate. For lower energies (50 GeV) and very large redshifts ($z > 3$) there is a contribution from visible and near UV. The crux of the matter is the target photon density, $n(\epsilon)$. No reliable measurements have been made around 2 μm, only upper limits. Subtraction of local IR sources, *e.g.* planets, dust, and the Milky Way, complicates the measurement.

Nevertheless, the extragalactic IR spectrum is rich with information. The theory of early galactic evolution is closely tied in with the adopted cosmological model: some mixed cold and dark matter models make specific predictions for the IR density [9]. So it is worth the exercise to choose a spectrum shape, such as a power law, $n(\epsilon) = k\epsilon^{-\gamma}$ and to calculate the amount of absorption. In that case, and choosing $\Omega_0 = 1$ for the present day mass density of the universe, the expression for τ simplifies to

$$\tau(E_0, z, \gamma) = 2\frac{c}{H_0}\frac{k}{\gamma+1}\Big(\frac{E_0}{m_e^2}\Big)^{\gamma-1}\frac{(1+z)^{2\gamma-\frac{1}{2}}-1}{2\gamma-\frac{1}{2}}\int_0^1 d\beta \quad \beta(1-\beta^2)^{\gamma-2}\sigma(\beta).$$

Calculating the optical depth as a function of redshift for some different values of E_0, using $n(\epsilon) = 0.001\ \epsilon^{-2.55}$ cm^{-3} (see Ref.[12] and references therein), shows that at 1 TeV and $z = 0.031$, Mrk 421 suffers almost no attenuation. 3C279 on the other hand should be invisible ($z = 0.54$). Extending the energy sensitivity into the gap between ground devices and satellites will permit measurement of the roll-off zone, and hence will constrain $n(\epsilon)/H_0$ and, to a lesser extent, Ω_0.

[2] 1 barn $= 10^{-24}$ cm^2.

Gamma ray telescopes: satellites and ground-based detectors

The high energy reach of a satellite detector is limited by the relatively small mass and volume that can be sent aloft. For an observation period of 10 days on a typical source, Egret acquires only a few photons in the highest energy bins. On the other hand, satellite-based instruments have very low background, and a broad field-of-view so that large (30° squared) pieces of the sky can be viewed at once. This permits searches for unexpected sources, and also means that diffuse or extended sources can be studied. Ground-based techniques are sensitive only to point sources.

Two successors to Egret are being studied. Agate is an extension of the Egret design: a much larger spark chamber, useful up to 50 GeV or so [7]. Glast is a completely new device, based on silicon strips [4]. It would reach up to 300 GeV. However, the only satellite project that has been funded is Integral, which runs up to about 3 MeV. It seems unlikely that either Agate or Glast will fly before 2010.

Beyond the reach of satellites, the only way to have a large collection area is to allow the primary particle to shower in the atmosphere, and to reconstruct the energy, direction, and species of the primary particle by sampling some part of the cascade of secondary particles. Above 10^{16} eV, nitrogen fluorescence allows the "Fly's Eye" technique, with enormous sensitive areas and full longitudinal reconstruction of the shower that allows composition studies. Above about 20 TeV, enough charged particles reach the ground to enable use of scintillator arrays. At 1 TeV and lower, the only part of the shower that reaches the ground is the Čerenkov light generated by ultra-relativistic charged secondaries near shower maximum. A Čerenkov telescope is therefore the only ground-based technique that can reach down to satellite energies.

The observations of the Crab, PSR 1706–44, and Mrk 421 have proved the power of the Atmospheric Čerenkov Technique (ACT). It took over twenty years to achieve these results, basically due to two difficult-to-manage backgrounds. The first is the night sky light. A Čerenkov telescope minimizes this background with fast photomultipliers and a narrow angular aperture. The telescopes can only run on dark, clear nights, which limits observation time. The other background is from cosmic rays: the charged particle flux is typically 1000 times more than the gamma source (depending on the telescope design and the source). Hadron/gamma separation is possible because in a pure electromagnetic shower, both the spatial and angular distributions of Čerenkov light on the ground are uniform and narrow. The light from hadron showers is more uneven. Good angular resolution also enhances signal-to-noise: even a weak point source stands out against an isotropic background if the angular binning is fine enough.

Imaging versus timing

The Whipple Observatory, run by the Smithsonian Institute at Mt. Hopkins

in Arizona, includes a 10 m Čerenkov telescope. Light from showers impinges on the phototube array in the focal plane of the mirror. The particle shower begins to develop at about 20 km above sea level, and is biggest around 10 km. Since $\theta_C \simeq 10$ to 15 mr, the light pool has a radius of about 120 m. The 70 m^2 Whipple telescope samples less than 1% of the light pool.

The core of the shower generating light is about 20 m wide, so the angular width of the image is only a couple of milliradians. On the other hand, the shower develops over several kilometers, so the longitudinal image depends on where the impact point of the shower is relative to the mirror. If the shower points down the optic axis, the image is round. If the shower is offset, then the projection of the cigar-like shower in the focal plane is up to 20 mr long. This determines the scale of both the pixel size and the mirror aperture.

The focal plane of the Whipple mirror contains an array of 108 phototubes. The shower image is broader for hadron showers than for gamma showers, allowing about a factor of 100 background rejection. The orientation of the image is a good estimator of the shower direction.

The other approach is to measure the photon arrival times at several places in the light pool. The times are fit to a conical surface, giving the shower direction. Because of the irregular structure of the hadronic showers, a quality cut on the cone fit rejects background.

Other groups are developing different techniques to use atmospheric Čerenkov light for gamma-ray astronomy. At the INFN-Pisa, along with collaborators from Padova, Trieste, and Naples, we have built the CLUE experiment and installed it at the mountaintop observatory on the island of La Palma in the Canaries [1]. CLUE uses the imaging technique, with some innovations: wire chambers sensitive to ultraviolet light (220 nm) are used in place of photomultipliers, and multiple 1.8 meter mirrors are used instead of a single larger mirror. The advantage of this approach is that night sky background is completely absorbed by the ozone layer above the cascade. The disadvantage is that much of the Čerenkov signal from the shower core is also absorbed, by oxygen, so that the tail rather than the core of the cascade is sampled. Sampling the shower tail makes angle and energy reconstruction as well as hadron/gamma separation more difficult.

How to go to 10 GeV: solar farms?

Scaling the mirror area of a Whipple-style detector to 10 GeV suggests that 20,000 m^2 would be needed. However, new telescopes now under construction (CAT at Thémis [3], and the Whipple upgrade) use faster phototubes and better optics and, if they perform as expected, scaling to 10 GeV will require "only" 5,000 m^2 of mirrors.

Simulations show that below about 100 GeV, hadron/gamma separation is simplified by two effects. First, the secondary electrons in the hadron showers

[3]The author now works on CAT, the Čerenkov Array at Thémis.

tend to be below Čerenkov threshold, so the hadron shower produces too little light to trigger the apparatus. Furthermore, the "lumpiness" of the light distribution is even more pronounced at lower energies, so that rejection is achieved at the trigger level by using a narrow angular aperture and requiring a high coincidence multiplicity.

However, at these low energies, the cosmic ray electron background is higher than it is in the TeV range. Since the only way to separate the gamma signal from the electron background is by fine angular binning, the angular resolution of the apparatus becomes its limiting characteristic.

An imaging telescope for such a large mirror area requires very many phototubes, raising the cost. The timing approach is more complicated because at low energy the conical wavefront has become spherical. Fitting the arrival times to a sphere determines the location of the shower maximum but *not* the direction. The direction is determined if the impact point of the shower on the ground is known, *i.e.* the center of the light pool. This requires pulse height information (ADC) in addition to the timing (TDC) information, raising the cost. The angular precision is limited by the measurement of the impact point. Faced with designing a large, expensive detector based on difficult extrapolations of present performance, the possibility of running a proof-of-principle experiment with existing equipment is quite attractive. Mirror farms at solar power plants may provide the basic instrument.

Solar power at the industrial scale was explored in the 1980's. In the United States, France, and the Soviet Union, large mirror arrays tracking the sun were focused on boilers at the top of central receiver towers. The projects were abandoned when the price of oil dropped late in the decade.

The angular size of the Sun is close to that of an air shower. Hence, while the solar arrays are not perfect for gamma ray astronomy, they aren't completely wrong either, and they represent an investment of over a hundred million dollars. Two of these plants are under serious study for use as Čerenkov telescopes. Solar-1, near Barstow, California, is the focus of the LACE collaboration, and tests of the tracking, background light, optical quality, etc are now underway [14]. Similar work is proceeding at Thémis. The two sites have different strengths. Secondary optics will have to be mounted near the boiler, to collect the photons from each heliostat independently to keep timing tight (see [15]). Table 1 summarizes results of a simulation of a 20 GeV array using the Thémis solar plant. A 7σ signal from the Crab could be acquired in about 5 minutes, and a 5σ signal for a 50 mCrab source would take 10 hours. Any Egret source in the field-of-view is detectable. The Thémis angular resolution is better than for Egret, and so, for example, the northern of the eight unidentified high galactic latitude sources could be scrutinized. The spectra roll-off for several sources in the Egret AGN catalog would allow the absorption-based $n(\epsilon)/H_0$ measurement, which in turn constrains the cosmological models.

We are on the verge of understanding some basic mysteries of astrophysics.

	Raw rate ($10^{-8} cm^{-2} s^{-1}$)	After cuts (Hz)
$\gamma_\circ$ (Crab)	1	0.3
$e_\circ$ (Electrons)	2.5 deg^{-2}	0.08
$h_\circ$ (Hadrons)	600 deg^{-2}	0.03
$\gamma_\circ/(e_\circ + h_\circ)$	1/600	2.7/1

Table 1: Simulated event rates at the Thémis array [10].

Ground-based gamma ray astronomy provides a big piece to the puzzle. It is limited in some ways: intergalactic absorption restricts our view of the most distant objects; Čerenkov telescopes cannot measure diffuse sources, or map large pieces of the sky; gamma rays probe only the outer layers of the source. But every technique has its limits. When information from satellite-based and ground-based gamma ray telescopes will be combined with results from the planned neutrino telescopes and the very large scintillator arrays, several questions will certainly be answered.

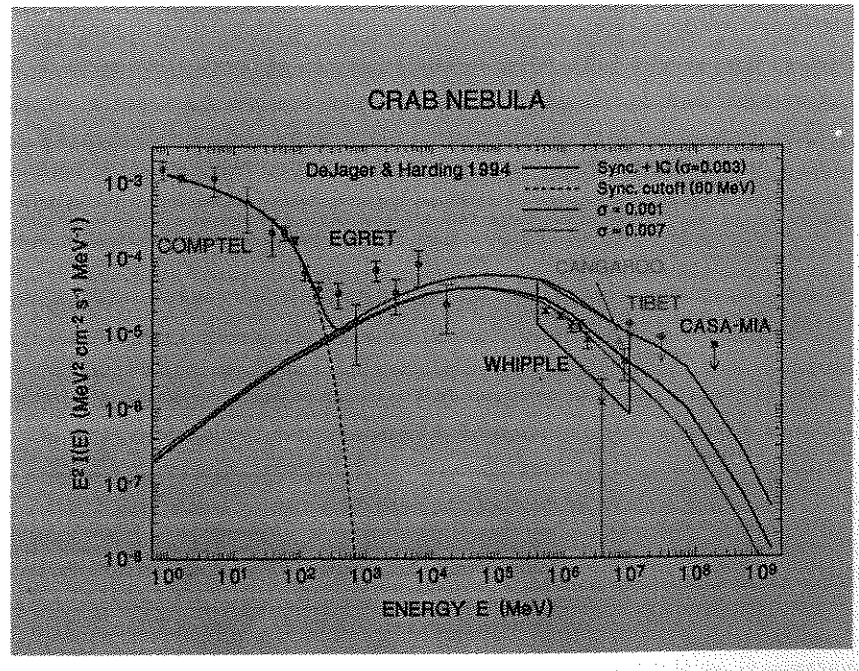


Figure 1: Gamma ray spectrum from the Crab nebula, measured by CGRO and by several ground-based experiments [6]

References

1. Anassontzis, E. et al. (The CLUE Collaboration): 1992, "A background-free detector for cosmic ray showers in the atmosphere", *Nucl. Instr. Meth.* **A315**, 267.

2. Fegan, D.J.: 1992, "The atmospheric Čerenkov technique: overview of 2nd generation detection systems". In *Towards a major atmospheric Čerenkov detector* (P. Fleury and G. Vacanti, Eds.), 3–41, Editions Frontieres, Palaiseau.

3. Gehrels, N., et al.: 1993 (December issue), "The Compton Gamma Ray Observatory", *Scientific American*, **269**, 68-77.

4. Godfrey, G.L.: 1993, "Gamma Large Area Silicon Telescope (GLAST)". In *Towards a major atmospheric Čerenkov detector - II* (R.C. Lamb, Ed.), 145–153, Iowa State University.

5. Goret, P. et al.: 1993, "Observations of TeV gamma rays from the Crab Nebula", *Astron. Astrophys.* **270**, 401-406.

6. Harding, A.K. and DeJager, O.C.: 1992, "The expected high-energy to ultrahigh-energy spectrum of the Crab Nebula", *Astrophys. J.* **396**, 161-172.

7. Hunter, S.D.: 1993, "The Advanced Gamma-ray Astronomy Telescope Experiment (AGATE): New satellite instrumentation beyond EGRET". In *Towards a Major Atmospheric Čerenkov Detector - II* (R.C. Lamb, Ed.), 135–144, Iowa State University.

8. Kniffen, D.A. et al.: 1993, "The variability in the gamma-ray emission of 3C 279", *Astrophys. J.* **411**, 133-136.

9. MacMinn, D. and Primack, J.: 1994, "The infrared extragalactic background and a new cosmological test", *Nature*, submitted.

10. Paré, E. et al.: 1994, "Gamma-ray astronomy: extension beyond the GeV domain by ground-based observations", *Phys. Lett.*, submitted.

11. Punch, M. et al.: 1992, "Detection of TeV photons from the active galaxy Markarian 421", *Nature* **358**, 477-478.

12. Salamon, M.H., Stecker, F.W., and DeJager, O.C.: 1994, "A new method for determining the Hubble constant from sub-TeV gamma-ray observations", *Astrophys. J. Lett.* **423**, L1-L5.

13. Stiavelli, M.: 1994, this volume. Antonucci, R.: 1993, "Unified models of AGN's", *Ann. Rev. Astron. Astrophys.* **31**, 473-521.

14. Tumar, T. et al.: 1991, "Proposal for a gamma ray telescope at Solar-1" (University of California at Riverside).

15. Travis, J.: 1994, "Solar farms may reap gamma rays", *Science* **266**, 30-31.

16. Vacanti, G. et al.: 1991, "Gamma-ray observations of the Crab Nebula at TeV energies", *Astrophys. J.* **377**, 467-479.

17. Weekes, T.C.: 1988, "Very high energy gamma-ray astronomy", *Phys. Reports* **160**, 1-121.

Supersonic outflow and accretion in stellar winds

Marco Velli*

May 13

*Dipartimento di Astronomia e Scienze dello Spazio, Università di Firenze, Italy

Many stars and other astrophysical objects continuously eject hot plasma into space. In a fundamental series of papers, Parker [1,2] demonstrated that, if the pressure of the interstellar medium is negligible, a steady state flow emanating from the object in question must become supersonic beyond some critical radius, hence predicting the existence of the solar wind. The reason for such a supersonic expansion is that a spherically symmetric, static, extended atmosphere, with a temperature profile decreasing with distance less rapidly than $1/r$, requires a finite pressure to be confined at great distances; the same holds true if the atmosphere 'evaporates' a subsonic flow, or breeze.

Mestel (quoted in [3]) first remarked that it would not take a large fall in coronal temperature for the pressure of the local interstellar medium (ISM) to be sufficient to suppress the solar wind entirely. Current estimates for the total pressure of the ISM are of the order of $p_{ISM} \simeq 1.24\ 10^{-12}$ dyne/ cm^2 [4]: for an isothermal heliosphere, this would be the pressure at infinity required to confine a 4 10^5 K static corona with base density 10^9 cm^{-3}.

In this contribution we show that the reaction of spherically symmetric, non-rotating, extended atmospheres with flows to changes of the external conditions is not so straightforward, but rather follows a cyclical behaviour involving abrupt transitions from winds to accretion and back. In the following, we classify the steady state solutions in terms of a given pressure difference (or better, enthalpy jump) between the atmospheric (or coronal) base and the distant medium. For simplicity we consider isothermal flows, since polytropic flows behave essentially in the same way. We will then show that the stratification due to breeze-type outflows is generally unstable, unless a physical boundary is present close enough to the coronal base. The stability of subsonic flows has been extensively studied before [5,6,7,8]. The different results obtained appear to be mainly due to the varying position of the applied boundary

conditions, as will be clear from our discussion of the physical mechanism of the instability.

It will immediately follow that the possible *stable* equilibrium flows are either

a) a supersonic shocked wind, if the relative pressure of the distant medium is low enough, or

b) a supersonic shocked accretion, if it is too high, or

c) a supersonic shocked wind *or* a subsonic accretion breeze, for an intermediate range of pressure differences.

The equations of motion for one-dimensional, spherically symmetric, stationary isothermal flows may be written in the form

$$\frac{\partial}{\partial r}(\rho u r^2) = 0 \tag{1}$$

$$u\frac{\partial u}{\partial r} = -\frac{1}{\rho}\frac{\partial p}{\partial r} - \frac{GM_s}{r^2} \tag{2}$$

where u, ρ, p, are velocity, density, and pressure. M_s is the mass of the star or other central object and GM_s/r^2 is the gravitational acceleration. We assume an "isothermal" equation of state $p = c^2\rho$ where c is the constant sound speed. The atmospheric self-gravity has been neglected. Hereafter we scale distances to the base radius R_0, and introduce the dimensionless gravity $g = GM_s/R_0c^2$.

For a static atmosphere, the density and pressure profiles are given by $(p, \rho) = (p_0, \rho_0)\exp(-g + g/r)$ implying an asymptotic value for the pressure at large distances $p^s_\infty = p_0 e^{-g}$. In terms of the Mach number $M = u/c$, the flow equations may be written in the standard way (a prime denoting dimensionless radial derivatives from now on)

$$\left(M - \frac{1}{M}\right)M' = \frac{2}{r} - \frac{g}{r^2}\,, \tag{3}$$

which may be integrated and expressed in two equivalent ways

$$\frac{1}{2}(M^2 - M_0^2) - ln(\frac{M}{M_0}) = 2lnr + \frac{g}{r} - g, \tag{4}$$

or

$$M^2/2 + lnp - g/r = M_0^2/2 + lnp_0 - g, \tag{5}$$

where M_0 is the base Mach number.

The well known structure of the (M, r) phase plane is illustrated in Fig. 1, for both positive and negative (accretion) values of M. In the positive M halfspace, 4 regions are delimitated by the intersection at the critical, or sonic, point of the upward accelerating transonic and downward decelerating transonic curves, defined by the solutions of Eq. (5), which pass through $M = 1$ for $r = r_c = g/2$. Single valued continuous flow profiles $M(r)$ which are

subsonic for all r, the breezes, lie below both transonic curves (region I). Above both transonic curves are flow profiles which are supersonic for all r (region III) and double valued solutions (regions II, IV) are confined to the right or to the left of the critical point between the accelerating and the decelerating transonic "separatrix".

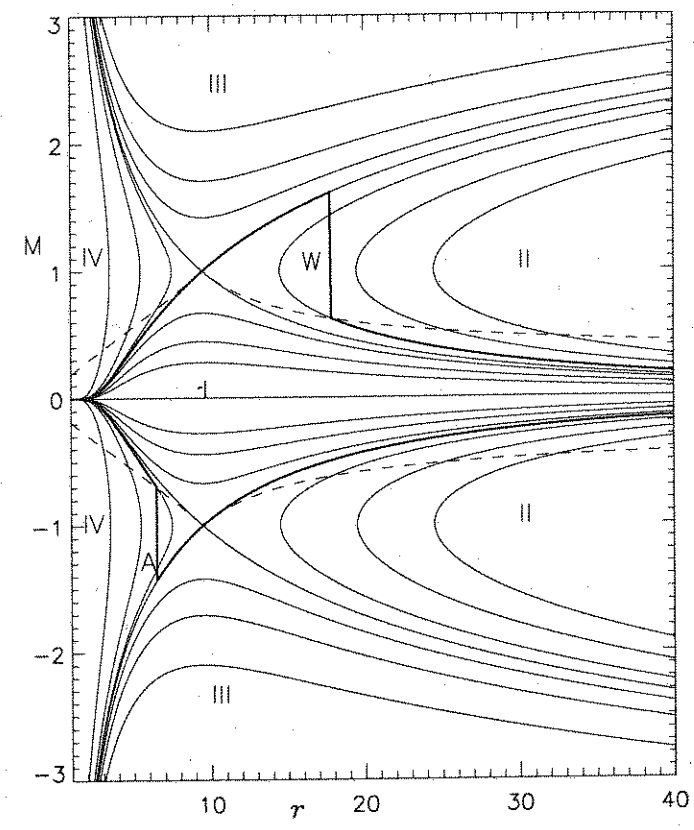


Figure 1: The (M, r) phase plane. The dashed line intersection with double valued curves defines the shock position for winds (region II, curve W) or accretion flows (region IV, curve A).

Among the flows which are subsonic at the atmospheric base (a condition we will assume throughout), the accelerating transonic has the special property [1] that density and pressure tend to zero at large distances. Because of the small, but finite, values of the pressure of the ambient 'external' medium, a terminal shock transition, connecting to the lower branch of the double

valued solutions filling region II, will in general be present (see, e.g., [9]). One such solution is given by the thick curve W. For given base values of the pressure, the position of the shock r_s is uniquely determined by the pressure p_∞ of the interstellar medium, and the distance from the critical point to the shock decreases as the pressure increases [10]. When p_∞ reaches a value $p_\infty^c = p_0 \exp({M_*}^2/2 - g)$, (here M_* is the base Mach number of the upward transonic), the shock distance $r_s = r_c = g/2$ and the discontinuity in the flow velocity reduces to a discontinuity in the derivative of the profile. This is the fastest possible, or critical, breeze, made up of the section of upward transonic below r_c and the section of downward transonic beyond r_c.

Now for the breeze solutions, with a base Mach number M_0 such that $M_* > M_0 \geq 0$, the asymptotic pressure is easily calculated to be

$$p_\infty = p_0 \exp(M_0^2/2 - g) \geq p_\infty^s. \tag{6}$$

It follows that the pressure required to confine a breeze *increases* with increasing base Mach number and is greater, if only slightly, than that of a static atmosphere. The limiting value of p_∞ is again p_∞^c. For a given base pressure and asymptotic pressures $p_\infty^c > p_\infty \geq p_\infty^s$, two possible steady state outflow solutions exist, a supersonic shocked wind and a subsonic breeze.

However, breeze solutions are unstable to low frequency standing sound waves which leave the pressure (and density) unperturbed at the atmospheric base and infinity. An analytic proof of this may be found in [10]. Here we will only show the results from numerical calculations of the growth rate of the instability (the dimensionless growth rate shown is the physical one multiplied by the sound crossing time $\tau_c = R_0/c$) as a function of base Mach number (Fig. 2, left panel). The growth rate is largest for high values of the base Mach numbers, while both the static atmosphere ($M_0 = 0$) and the critical breeze ($M_0 = M_*$) are marginally stable. In the latter case the perturbation equations become singular at the sonic point, because the phase speed of the inward propagating wave vanishes there. An additional regularity condition must be imposed in the steady state equations, effectively isolating the region inside the sonic point from the region beyond it. This is the mathematical reason behind the stability of flows with a continuous subsonic/supersonic transition. The breeze instability is driven by the unfavourable stratification induced by the flow. Imagine a static atmosphere, and let the pressure at infinity increase: clearly an *inflow*, not an outflow, is expected to result. That accretion breezes are stable follows immediately from the fact that the steady state flow equations are symmetrical in M (Fig. 1), while the perturbation equations are invariant under a change in sign of both M *and* γ.

Given that breezes are unstable, we see that even in the pressure range $p_\infty^c > p_\infty \geq p_\infty^s$ the only stable outflow is a supersonic wind with a terminal shock. What happens if the pressure difference between the coronal base and the distant medium varies? As p_∞ increases, the shock moves inward, decreasing in amplitude as the critical point is approached. When the critical breeze is reached, there is no neighbouring outflow solution capable of sustaining a higher pressure at infinity. The only possibility for the flow is to collapse into its symmetrical ($M \to -M$) critical breeze accretion profile, which is also marginally stable. As the pressure is increased further, an accretion shock is formed *inside* the sonic point, connecting the symmetrical of the downward transonic to one of the double valued curves in region IV, (as shown by the curve labelled A in Fig. 1). For $p_\infty > p_\infty^c$ there is a unique shocked accretion flow (see [11], for a general proof in the case of polytropic flows), and the shock position moves inward from the critical point as the pressure is increased beyond p_∞^c (if the pressure is too high, the shock may occur below the normalisation radius chosen). Consider now what happens if, starting from a shocked accretion flow, the pressure at the stellar surface increases, or alternatively the pressure of the ISM decreases. The shock moves outwards, but this time, as p_∞ decreases below p_∞^c, the flow can evolve with continuity into subsonic accretion. As p_∞ decreases further, the accretion-breeze velocities decrease, but when p_∞ decreases below p_∞^s, the flow must accelerate again into a supersonic shocked wind.

The stratification produced by breezes, though globally unstable, is not locally unstable everywhere. For example, below the critical point the pressure in breezes decreases with height more rapidly than in the static case. Inspection of Eq. (5) actually shows that this is true out to the radius r_f where the Mach number of the flow has decreased to the same level as the base Mach number.

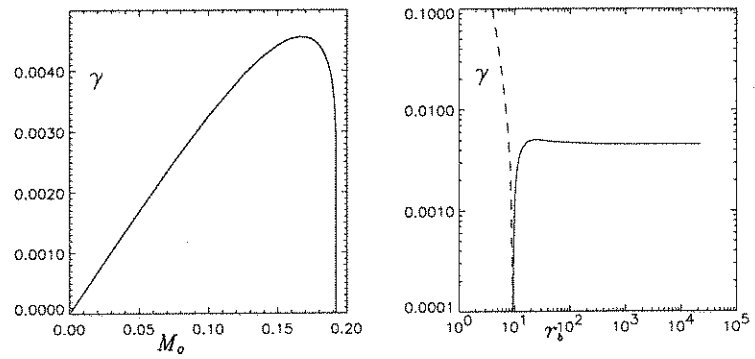


Figure 2: Left: Growth rate γ of breeze instability as a function of base Mach number ($g = 5.0$). Marginal stability is obtained for $M_0 = 0, M_0 = M_*$. Right: Maximal growth rate for breezes (continuous line) and subsonic accretion (dashed line) as a function of the position of the outer boundary r_b ($g = 5.0$) The two lines join in $r_b = r_f$, where $\gamma = 0$.

which may be calculated by imposing $M = M_0$ in Eq. (4), i.e. $2lnr_f + g/r_f - g = 0$. This equation is independent of base Mach number, which also means that at this height the pressure is the same for all breezes, while inside this radius, the pressure at a given height is a monotonically *decreasing* function of base Mach number. As the the boundary conditions are imposed at closer and closer distances r_b, the growth rate of the instability is reduced, and marginal stability is obtained when $r_b = r_f$. Imposing boundary conditions below this

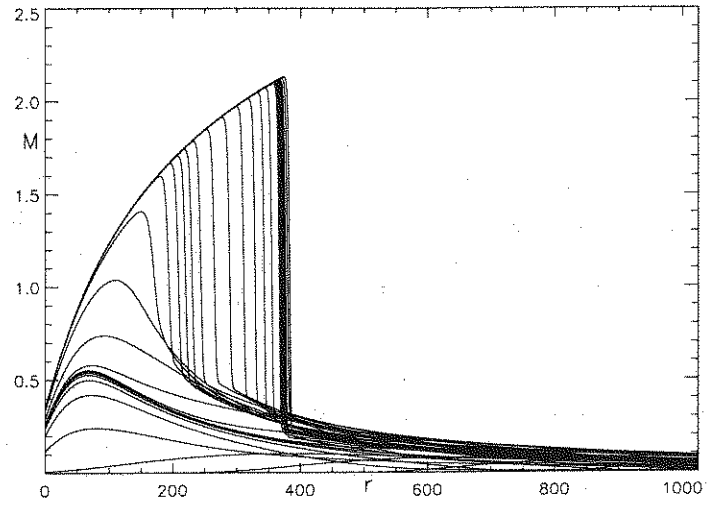


Figure 3: Snapshots of the flow profile set up by a pressure perturbation at the outer boundary; the x-coordinate is the number of grid points. The pressure is first decreased according to the formula $p_b = p_b^s(1 - 0.1t^2/(1+t^2))$, generating an expansion wave which propagates to the coronal base, is reflected and grows in amplitude developing into a shocked wind. Once this flow becomes steady state, the pressure at the outer boundary is increased back to its initial value in an analogous way. The shock is pushed slightly inward, but remains the equilibrium solution.

radius stabilizes the breezes, but consequently destabilizes subsonic accretion, as is shown in Fig. 2 (right panel), where the maximal growth rate for breezes (continuous line) and accretion (dashed line) as a function of r_b is plotted. For large values of g, this marginal stability radius depends exponentially on g as $r_f \simeq \exp(g/2) - g/2$.

When the boundary is at $r_b < r_f$ the equilibrium flows still present an hysteresis-type cycle in terms of the enthalpy jump between r_b and the coronal base, but in a reversed order with respect to that described previously.

Supersonic accretion is blown into supersonic winds as the base pressure is increased beyond a critical value, while an outflow breeze phase exists before the collapse to accretion as the pressure at the outer boundary is increased beyond the value appropriate to a static atmosphere.

As discussed in [10], the above result is robust and does not depend on the isothermal approximation, but rather on the positivity of the total energy of the flow. Some numerical simulations of the formation of winds including a full energy equation with distributed mechanical heating, thermal conduction and radiation losses, were performed by Korevaar ([12] and references therein). When increasing the pressure at the outer edge of the numerical domain, Korevaar found a continuous transition from supersonic shocked winds to breezes and then a 'rapid onset' of supersonic shocked accretion. His result may be understood in terms of the boundary conditions and dimensions of the numerical box, as the outer pressure boundary condition was imposed well below the marginal stability radius defining the second hysteresis cycle (unstable subsonic accretion) discussed at the end of the previous paragraph. The time-scale necessary for the inversion of the flow as the critical pressure levels are exceeded is an important point that requires a time-dependent analysis for investigation, but an order of magnitude estimate may be obtained from the sound travel time between the atmospheric base and the sonic point, in the isothermal case $\tau_l = GM_s/2c^3$, while an upper limit is probably given by the time it takes for a sound-wave to communicate the existence of an unfavourable pressure gradient to the inner parts of the atmosphere, or $\tau_u = R_0 r_f/c$. We are presently carrying out numerical simulations [13] aimed at a detailed understanding of the time evolution of the flows as the pressure jump is varied. In Fig. 3 we show a series of snapshots of the flow profile as the pressure at the outer boundary p_b is first decreased from the value appropriate to a static atmosphere, and then increased again back to the static value. Here the resolution was of 1024 grid points, and the viscosity (necessary to resolve the shocks) was $\nu = 0.0025$. We see that, as predicted by our analysis, the result of such an external perturbation is to set up a supersonic shocked wind, even though the final external pressure would be sufficient to confine a static atmosphere.

Let us now consider some applications of our result. Transitions from accretion to wind should take place during the early phase of stellar evolution. The kinematics described by our model may not be directly relevant in most cases given the important role played by angular momentum and accretion disks. One interesting point, however, is that the spherically symmetric calculations allow the presence of both accretion and supersonic outflows for a well defined range of pressure differences (of the order of that required to maintain static equilibrium) between the atmospheric base and the distant medium, in this case a dense molecular cloud. If subsonic accreting flows coexisted with supersonic outflows in physically distinct regions of the star (polar wind, equatorial

accretion for example), then, because the accretion breeze has a larger pressure at a given radius (below the terminal shock height), the outgoing wind could be "inertially" collimated by the accretion flow, and it would not be unreasonable to expect an equilibrium in which accretion rate and mass loss would be well correlated (Kelvin-Helmholtz instabilities and magnetic fields should also play an important role in the transverse dynamics).

Recent work on the dynamics of the galactic halo [8,14] shows that the instability we have discussed exists even in a subsonic flow with a faster than spherical expansion, and dominates the thermal instability, with a continuous transition to the thermal instability branch when the base Mach number tends to zero. Our analysis seems to imply that the corresponding stable flow should be a supersonic galactic wind connecting, via a shock transition, to the preexisting halo material, if steady state conditions can prevail at all. Such shock transitions should appear also in galactic chimney models [15].

To conclude, we have discussed the behaviour of steady, spherically symmetric flows with gravity, as the pressure jump responsible for the flow is varied. We have seen that the transition from accretion to wind and back is necessarily of an abrupt (catastrophic) nature, and that very special (and hence rare) conditions are necessary to establish steady state quasi-static or subsonic flows: in other words, when it rains, it pours.

References

1. Parker, E. N. 1958, *Astrophys. J.*, **128**, 664

2. Parker, E. N. 1963, *Interplanetary Dynamical Processes*, Interscience; New York

3. Roberts, P. H. and Soward, A. M. 1972, *Proc. Royal Soc. London*, **328A**, 185

4. Frisch, P. C., 1990, *Physics of the Outer Heliosphere*, S. Grzedzielski and D. E. Page eds., COSPAR Colloquium, p. 19

5. Petterson, J. A., Silk, J., & Ostriker, J. P. 1980, *Mon. Not. Roy. Astron. Soc.*, **191**, 571

6. Stellingwerf, R. F. & Buff, J. 1978, *Astrophys. J.*, **221**, 661

7. Theuns, T. and David, M. 1992, *Astrophys. J.*, **384**, 587

8. Ferrara, A. and Einaudi, G. 1992, *Astrophys. J.*, **395**, 475

9. Holzer, T. E. and Axford, W. I. 1970, *Ann. Rev. Astron. Astrophys.*, **8**, 30

10. Velli, M. 1994, *Astrophys. J.*, **432**, L55

11. McCrea, W. H. 1956, *Astrophys. J.*, **124**, 461

12. Korevaar, P. 1989, *Astron. Astrophys.*, **226**, 209

13. Velli, M., Grappin, R. and Cavillier, E. 1995, in preparation

14. Einaudi, G. and Ferrara, A. 1991, *Astrophys. J.*, **371**, 571

15. Norman, C. A. and Ikeuchi, S. 1989, *Astrophys. J.*, **345**, 372

Clusters of galaxies and the anti-proton lifetime

Daniel J. O'Connor*

May 20

* Istituto Nazionale di Fisica Nucleare, Sezione di Pisa, Italy

The manuscript was not provided by the Author

References

1. Becker-Szendy, R., et al. 1990, *Phys. Rev. D*, **42**, 2974

2. Gabrielse, G., et al. 1990, *Phys. Rev. Lett.*, **65**, 1317

3. Ormes, J. F., Streitmatter, R.E., 1991, "Searching for primordial anti-matter", NASA preprint 91-047

4. Steigman, G. 1976, *Ann. Rev. Astron. Astrophys.*, **14**, 339

5. Turner, M., Kolb, E. 1990, *The early universe*, Addison-Wesley, New York

Where do meteorites come from?

Paolo Farinella[1,2]

May 27

1. Gruppo di Meccanica Spaziale, Dipartimento di Matematica, Università di Pisa, Italy

2. Observatoire de la Côte d'Azur, Dept. Cassini, Nice, France

Every day, tons of interplanetary material fall into the Earth's atmosphere. Occasionally fragments heavier than a few kilograms partially survive atmospheric crossing, reach the Earth's surface, and become meteorites. The fall of larger bodies also occurs, with a frequency rapidly decreasing with size. As a consequence of the size distribution of the interplanetary bodies, the typical interval between two impacts of "projectiles" larger than D is roughly proportional to D^2 — that is, a few centuries for 50–meter sized bodies, causing a 10 Mton of TNT explosion like the one which devastated the Siberian *taiga* at Tunguska in 1908; some 10^5 yr for km–sized objects; and $\approx 10^8$ years for an asteroid or comet 10 km across, causing a global climatic and ecological catastrophe, such as that recorded at the Kretaceous–Tertiary boundary. Averaged over geologic time, a total influx of about 1.7×10^8 kg/yr hits the Earth over the mass range from 10^{-21} to 10^{15} kg [1].

Where does this material come from? Until the early nineteenth century, the fall of stones from outer space was widely believed by scientists to be a matter of superstition. On the other hand, until about 60 years ago, meteorites were thought to be interstellar objects travelling on hyperbolic orbits. Today, we know that cometary nuclei and debris contribute to the interplanetary complex of material (though apparently not to the small solid bodies reaching the Earth's surface), and there is evidence that a tiny fraction of the meteorites come from the Moon and Mars. However, the birthplace of most meteorites and near–Earth asteroids (NEAs) has been recognized to lie in the main asteroid belt, where the growth of a full–sized planet was aborted during the early evolution of the solar system (see [2,3,4] for comprehensive reviews). Therefore, for scientists interested in the earliest history of the Earth and of the planetary system, meteorites are veritable Rosetta stones that include a wide variety of materials which have survived almost unaltered since 4.5×10^9 years ago.

However, until recently serious problems remained in understanding the asteroid–meteorite connection. As a consequence of interasteroidal impact events, fragments are usually ejected at velocities relative to their parent bodies of the order of 100 m/s , far too small for directly causing the drastic orbital changes needed to achieve planet–crossing orbits (typical asteroid orbital velocities are 15 to 20 km/s). On the other hand, impulsive velocity increments of several km/s would induce strong shock effects and thermal modifications in the metoritic minerals, which in general are not observed. Thus the interplay between collisions and subtle dynamical effects is required to transport material from the asteroid belt to the Earth. A long–standing issue in planetary science is that of identifying and estimating the efficiency of these dynamical *routes*. Also, the meteorite delivery models should be consistent with the observed properties of meteorites, such as the pre–fall orbits (in the four cases in which they have been determined), the distribution of fall hours during the day, the time span over which the surfaces have been exposed to cosmic rays, and the inferred (dis)similarities in composition and thermal history among the diverse meteorite and asteroid classes.

In the last few years, in collaboration with several colleagues at the Nice and Turin Observatories and the University of Pisa, I have carried out a simulation of all the stages of the transport process from the main asteroid belt to Earth–crossing orbits, showing that plausible collisional physics and dynamical mechanisms can easily generate significant numbers of Earth–approaching asteroid fragments. The ejection of fragments from cratering and break–up events, involving thousands of asteroids orbiting in the inner part of the asteroid belt, has been simulated through a numerical model, consistent with the current estimates on asteroid collision rates [5] and with the experimental results on the outcome of high–velocity ($\approx$ 5 km/s) impact events [6,7]. For each asteroid, estimates have been obtained for the average flux of fragments dropped into the chaotic regions of the phase space associated with the 3/1 mean motion Jovian resonance (i.e., orbits whose period is close to 3 times that of Jupiter, leading to enhanced and accumulating Jovian perturbations) and with the so–called ν_6 secular resonance (in which the asteroid's and Saturn's average perihelion longitude rates are the same); these are indeed the most important dynamical *routes* from the inner part of the asteroid belt to the near–Earth region [8]. These calculations have provided quantitative indications about the most effective potential meteorite–delivering asteroids and the overall efficiency of the transport of fragments from the main asteroid belt to the Earth [9,10]. Most of the meteoritic material appears to come from a small number of relatively large, resonance–bordering asteroids, which are possibly non–representative of the general main belt population from the chemical and mineralogical points of view. Numerical integrations over time spans of a few Myr of the orbits of fictitious fragments from these main–belt asteroids show that many of them evolve to orbital elements closely resembling those of real

NEAs and of the largest observed fireballs, such as that associated with the Příbram meteorite fall [11].

Furthermore, we have estimated the NEA/meteorite production potential of the very energetic asteroidal collisions that generated the most populous asteroid families [12,13], identified from the similarity of their orbital parameters (once freed of planetary perturbations, namely transformed into *proper* elements [14,15,16]). In particular, very interesting results have been obtained for the case of the Vesta family, a populous grouping of bodies where spectrophotometric data indicate a clear genetic relationship between a large main–belt asteroid (Vesta itself), a number of minor family members lying close to it in the proper element space [17], a few km–sized NEAs, and a specific meteorite type (the basaltic achondrites). The existence of all these related objects can be explained if one or a few large impacts formed giant craters on the surface of Vesta, ejecting sprays of fragments of its basaltic crust [18]. Other families have been probably formed by collisional break–up of their parent bodies, occurred in some cases as recently as less than 50 Myr ago [19].

Numerical studies of the orbital evolution of real NEAs over millions of years have revealed unexpected connections and phenomena. For instance, secular and mean motion resonances can frequently increase the orbital eccentricity to almost unity, thus resulting into Sun–grazing episodes and even solar collisions [20,21]. On the other hand, dynamical *routes* exist between orbits typical of NEAs and of comets and meteoroid streams (such P/Encke and the Taurids), so that dynamics alone cannot distinguish in the near–Earth population between rocky, asteroidal objects and extinct or "dormant" cometary nuclei [22]. In addition, while the near–Earth objects form a quasi–stationary population with dynamical and collisional lifetimes ranging from $\approx 10^5$ to 10^8 yr, large stochastic fluctuations appear possible, as a result of the disruption of sizeable asteroids close to resonant zones in the orbital element space [23]. Thus the cratering rate on the Earth may vary in a considerable way over geologic time.

So far, the evolutionary mechanisms transferring asteroidal material to Earth have been investigated and modelled separately from each other, with a step–by–step approach. In the next stages of our research project, we plan to develop a more comprehensive *scenario* for the dynamical evolution of planet–crossing objects generated in the asteroid belt, taking into account the interactions and synergies among many different processes and mechanisms: resonances with the outer planets, close encounters with the inner ones, disruptive collisions with other asteroids/meteoroids, the planets and the Sun, hyperbolic ejections onto comet–like orbits. The final aim is to predict the orbital and size distribution, the time variability, the lifetimes and the compositions of the NEA/meteoroid population.

References

1. Ceplecha, Z.: 1992, "Influx of interplanetary bodies onto Earth", *Astron. Astrophys.* **263**, 361–366.

2. Wetherill, G.W. and Chapman, C.R.: 1988, "Asteroids and meteorites". In *Meteorites and the Early Solar System* (J.F. Kerridge and M.S. Matthews, Eds.), 35–67, Univ. of Arizona Press, Tucson.

3. Lipschutz, M.E., Gaffey, M.J. and Pellas, P.: 1989, "Meteoritic parent bodies: Nature, number, size and relation to present–day asteroids". In *Asteroids II* (R.P. Binzel, T. Gehrels, and M.S. Matthews, Eds.), 740–777, Univ. of Arizona Press, Tucson.

4. Greenberg, R. and Nolan, M.C.: 1989, "Delivery of asteroids and meteorites to the inner solar system". In *Asteroids II* (R.P. Binzel, T. Gehrels, and M.S. Matthews, Eds.), 778–804, Univ. of Arizona Press, Tucson.

5. Farinella, P. and Davis, D.R.: 1992, "Collision rates and impact velocities in the main asteroid belt", *Icarus* **97**, 111–123.

6. Davis, D.R., Farinella, P., Paolicchi, P., Weidenschilling, S.J. and Binzel, R.P.: 1989, "Asteroid collisional history: Effects on sizes and spins". In *Asteroids II* (T. Gehrels, Ed.), 805–826, Univ. of Arizona Press, Tucson.

7. Petit, J.–M. and Farinella, P.: 1993, "Modelling the outcomes of high–velocity impacts between small solar system bodies", *Celest. Mech.* **57**, 1–28.

8. Bertotti, B. and Farinella, P.: 1990, *Physics of the Earth and the Solar System*, Chs. 11 and 15, Kluwer, Dordrecht.

9. Farinella, P., Gonczi, R., Froeschlé Ch. and Froeschlé, C.: 1993, "The injection of asteroid fragments into resonances", *Icarus* **101**, 174–187.

10. Morbidelli, A., Gonczi, R., Froeschlé, Ch. and Farinella, P.: 1994, "Delivery of meteorites through the ν_6 secular resonance", *Astron. Astrophys.* **282**, 955–979.

11. Farinella, P., Gonczi, R. and Froeschlé, Ch.: 1993, "Meteorites from the asteroid 6 Hebe", *Celest. Mech.* **56**, 287–305.

12. Zappalà, V., Cellino, A., Farinella, P. and Knežević, Z.: 1990, "Asteroid families, I. Identification by hierarchical clustering and reliability assessment", *Astron. J.* **100**, 2030–2046.

13. Zappalà, V., Cellino, A., Farinella, P. and Milani, A.: 1994 "Asteroid families, II. Extension to unnumbered multi–opposition asteroids", *Astron. J.* **107**, 772–801.

14. Milani, A. and Knežević, Z.: 1990, "Secular perturbation theory and computation of asteroid proper elements", *Celest. Mech.* **49**, 347–411.

15. Milani, A. and Knežević, Z.: 1992, "Asteroid proper elements and secular resonances", *Icarus* **98**, 211–232.

16. Milani, A. and Knežević, Z.: 1994, "Asteroid proper elements and the dynamical structure of the asteroid belt", *Icarus* **107**, 219–254.

17. Binzel, R.P. and Xu, S.: 1993, "Chips off Vesta and a near–resonance source for basaltic achondrite meteorites", *Science* **260**, 186–190.

18. Davis, D.R., Ryan, E.V. and Farinella, P.: 1994, "Asteroid collisional evolution: Results from current scaling algorithms", *Planet. Space Sci.* **42**, 599–610.

19. Milani, A. and Farinella, P.: 1994, "The age of the Veritas asteroid family deduced from chaotic chronology", *Nature* **370**, 40–42.

20. Farinella, P., Froeschlé, Ch., Froeschlé, C., Gonczi, R., Hahn, G., Morbidelli, A. and Valsecchi, G.B.: 1994, "Asteroids falling into the Sun", *Nature* **371**, 314–317.

21. Froeschlé, Ch., Hahn, G., Gonczi, R., Morbidelli, A. and Farinella, P.: 1994, "Secular resonances and the dynamics of Mars–crossing and near–Earth asteroids", *Icarus*, submitted.

22. Valsecchi, G.B., Morbidelli, A., Gonczi, R., Farinella, P., Froeschlé, Ch. and Froeschlé, C.: 1994, "The dynamics of objects in orbits resembling that of P/Encke", *Icarus*, in press.

23. Menichella, P., Paolicchi, P. and Farinella P.: 1994, "The main belt as a source of near–Earth asteroids", *Earth, Moon and Planets*, in press.

Elenco dei volumi della collana
"Appunti dei corsi tenuti da docenti della Scuola"
pubblicati dall'Anno Accademico 1994/95

MARIO TOSI, *Introduction to Statistical Mechanics and Thermodynamics*, 1994

MARIO TOSI, *Introduction to the Theory of Many–Body Systems*, 1995

Seminario di Astrofisica a cura di GIUSEPPE BERTIN, 1995