J. Astrophys. Astr. (2007) **28,** 1–7

Oblique Propagation and Dissipation of Alfvén Waves in Coronal Holes

A. K. Srivastava & B. N. Dwivedi*
Department of Applied Physics, Institute of Technology, Banaras Hindu University, Varanasi 221 005, India.
**e-mail: bholadwivedi@yahoo.com*

Received 2006 March 31; accepted 2007 January 5

Abstract. We investigate the effect of viscosity and magnetic diffusivity on the oblique propagation and dissipation of Alfvén waves with respect to the normal outward direction, making use of MHD equations, density, temperature and magnetic field structure in coronal holes and underlying magnetic funnels. We find reduction in the damping length scale, group velocity and energy flux density as the propagation angle of Alfvén waves increases inside the coronal holes. For any propagation angle, the energy flux density and damping length scale also show a decrement in the source region of the solar wind ($< 1.05\,R_\odot$) where these may be one of the primary energy sources, which can convert the inflow of the solar wind into the outflow. In the outer region ($> 1.21\,R_\odot$), for any propagation angle, the energy flux density peaks match with the peaks of MgX 609.78 Å and 624.78 Å linewidths observed from the Coronal Diagnostic Spectrometer (CDS) on SOHO and the non-thermal velocity derived from these observations, justify the observed spectroscopic signature of the Alfvén wave dissipation.

Key words. Sun: Alfvén waves—coronal holes—solar wind.

1. Introduction

MHD waves are the prominent candidates for the heating of coronal holes and acceleration of the solar wind since the pioneering work of Parker (1965). Many spectroscopic studies show the importance of the damped and undamped Alfvén waves in the solar corona (Hassler *et al.* 1990; Harrison *et al.* 2002; O'Shea *et al.* 2005). The microscopic properties of the coronal hole plasma, e.g., viscosity and resistivity, etc., are important for conversion of the mechanical energy of Alfvén waves into thermal energy. Dwivedi & Srivastava (2006) have studied the importance of radially outward propagating Alfvén waves and their dissipation in coronal holes, which support the solar wind outflow starting at about 20 Mm as reported by Tu *et al.* (2005), and also the spectroscopic signature of Alfvén wave dissipation in the off-limb coronal hole plasma as reported by O'Shea *et al.* (2005).

In this paper, we consider the viscous and resistive dissipation of Alfvén waves propagating at various angles inside the inner part of polar coronal holes ($R < 1.35\,R_\odot$) and underlying magnetic funnels. At different propagation angles, we investigate the

combined effect of viscosity and magnetic diffusivity on the spatial variation of damping length scale, group velocity and mechanical energy flux density of Alfvén waves. In section 2, we describe the Alfvén wave propagation and dissipation. Results and discussion are given in the last section.

2. Oblique propagation and dissipation of Alfvén wave

We study the Alfvén wave propagation and dissipation in the viscous, magnetic diffusive, and incompressible coronal hole plasma. The MHD equations for viscous and resistive plasma are:

$$\rho \frac{\partial \mathbf{v}}{\partial t} + \rho \left(\mathbf{v} \cdot \nabla\right) \mathbf{v} = \frac{1}{\mu} \left(\nabla \times \mathbf{B}\right) \times \mathbf{B} + \rho \nu \nabla^2 \mathbf{v} \qquad \text{(momentum equation)},$$

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\mathbf{v} \times \mathbf{B}) + \eta \nabla^2 \mathbf{B} \qquad \text{(induction equation)},$$

and

$$\nabla \cdot \mathbf{B} = 0 \qquad \text{(magnetic flux conservation)},$$

where ρ, μ, η and ν respectively are mass density, magnetic permeability, magnetic diffusivity and coefficient of kinematic viscosity. When we linearize these equations, we get

$$\rho_{\mathrm{o}} \frac{\partial \mathbf{v}_1}{\partial t} = \frac{1}{\mu} \left(\mathbf{B}_{\mathrm{o}} \cdot \nabla\right) \mathbf{B}_1 + \rho_{\mathrm{o}} \nu \nabla^2 \mathbf{v}_1,$$

and

$$\frac{\partial \mathbf{B}_1}{\partial t} = \left(\mathbf{B}_{\mathrm{o}} \cdot \nabla\right) \mathbf{v}_1 + \eta \nabla^2 \mathbf{B}_1.$$

Assuming the plane wave solutions with the variables having the phase-factor $\exp\left[i\left(k \cdot r\,|\cos\theta| - \omega t\right)\right]$, we get the following dispersion relation in linear MHD context,

$$P k^4 + Q k^2 - \omega^2 = 0, \tag{1}$$

where $P = \nu\eta \cos^4\theta$, $Q = \left\lfloor v_A^2 - i\omega\left(\nu + \eta\right)\right\rfloor \cos^2\theta$, and $v_A = \mathbf{B}/\sqrt{\mu\rho}$ is the Alfvén velocity. The angle of propagation with respect to normal outward direction is θ. We take $\rho\nu = 1.0045 \times 10^{-16} T_R^{2.5}$ g cm^{-1} s^{-1} (Spitzer 1962) and magnetic diffusivity $\eta = 1.144 \times 10^{13} T_R^{-1.5}$ cm^2 s^{-1} (Priest 1982). We take the density profile in a polar coronal hole as a function of radial height from the empirical relation of Doyle *et al.* (1999):

$$N_e(R) = \frac{1 \times 10^8}{R^8} + \frac{2.5 \times 10^3}{R^4} + \frac{2.9 \times 10^5}{R^2}\ (\mathrm{cm}^{-3}). \tag{2}$$

The mass density is $\rho_{\mathrm{o}}(R) = 0.6\, m_p\, N_e(R)$, where m_p is the proton mass. Using David *et al.* (1998) temperature measurements in coronal holes, the temperature profile fit is given by Pekünlü *et al.* (2002),

$$T_R = -2 \times 10^7 R^2 + 5 \times 10^7 R - 3 \times 10^7\ (\mathrm{K}). \tag{3}$$

The temperature below $1.05\,R_\odot$ and above $1.30\,R_\odot$ is obtained by the extrapolated points of the fit. The energy flux density of Alfvén wave is $W = \rho v_{\rm NT}^2(R)(\partial\omega/\partial k)$, where $(\partial\omega/\partial k)$ is group velocity and $v_{\rm NT}(R)$ is velocity equivalent of the non-thermal component of relevant spectral line at FWHM. Using Banerjee *et al.* (1998) measurements for Si VIII ion, the non-thermal velocity profile fit is given by Pekünlü *et al.* (2002),

$$v_{\rm NT}(R) = -1522.3R^4 + 8638.2R^3 - 18191R^2 + 16882R - 5786.5 \quad (\mathrm{km\ s^{-1}}). \tag{4}$$

The measurements start from 27 arcsec ($\sim 1.03\,R_\odot$). The values below $\sim 1.03\,R_\odot$ and above $\sim 1.26\,R_\odot$ are obtained from the extrapolated points of the fit.

Using the measurements of Tu *et al.* (2005); Banaszkiewicz *et al.* (1998) and Hackenberg *et al.* (2000), the empirical relation for the magnetic field in the inner coronal hole is given by Dwivedi & Srivastava (2006),

$$B = \frac{a}{1 + \exp\left[(R-b)/c\right]} + d, \quad \text{for } 1.00\,R_\odot < R < 1.05\,R_\odot, \ \text{(Gauss)}, \tag{5a}$$

where $a = 113.1348$, $b = 1.0035$, $c = 0.0024$, and $d = 8.7852$, and

$$B = 95.20344 - 129.39079\,R + 45.49857\,R^2, \quad \text{for } 1.05\,R_\odot < R < 1.35\,R_\odot. \tag{5b}$$

We have taken $1.05\,R_\odot$ as a reference height and made all the physical quantities (e.g., T, η, ν, v_A, ρ, B and $v_{\rm NT}$) dimensionless, by dividing the respective values at $R = 1.05\,R_\odot$, for example, $T_R \rightarrow (T_R/T_{R=1.05\,R_\odot})$, etc. We analyze the propagation and dissipation of high frequency Alfvén wave in the magnetic funnel for the time periods $\tau_A = 0.0001$ s in the short-wavelength assumption. We also consider the shortest period $\tau_{A_{\rm ref}} = 0.01$ s as a reference time to make the time periods dimensionless, i.e., $\tau_A = \tau_A/\tau_{A_{\rm ref}}$. We have used these dimensionless physical quantities in the dispersion relation (equation 1) and calculated the wave number (which would be actually $k_{\rm mod} \approx k_{\tau_A,R}/k_{\tau_{A\,\rm ref},R_{\rm ref}}$) in our model. Using values of $k_{\rm mod}$, we have calculated damping length scale (D), energy flux density (W), and group velocity (V_g). The dimensionless quantities can be expressed as follows:

$$D \approx \frac{2\pi}{k_{\rm mod}} \approx \frac{2\pi}{k_{\tau_A,R}/k_{\tau_{A\,\rm ref},R_{\rm ref}}}$$

$$\approx \frac{2\pi}{k_{\tau_A,R}} \times \frac{k_{\tau_{A\,\rm ref},R_{\rm ref}}}{2\pi} \times 2\pi$$

$$\approx 2\pi \times \frac{D_{\tau_A,R}}{D_{\tau_{A\,\rm ref},R_{\rm ref}}}.$$

Hence, we can write

$$D_{\tau_A,R} \approx \frac{1}{2\pi} \times D \times D_{\tau_{A\,\rm ref},R_{\rm ref}} \quad (\mathrm{cm}), \tag{6a}$$

$$W_{\tau_A,R} \approx \frac{1}{2\pi} \times W \times W_{\tau_{A\,\mathrm{ref}},R_{\mathrm{ref}}} \quad (\mathrm{ergs\ cm^{-2}\ s^{-1}}), \tag{6b}$$

and

$$V_{g_{\tau_A},R} \approx \frac{1}{2\pi} \times V_g \times V_{g_{\tau_{A\,\mathrm{ref}},R_{\mathrm{ref}}}} \quad (\mathrm{cm\ s^{-1}}), \tag{6c}$$

where $D_{\tau_A,R}$, $W_{\tau_A,R}$, $V_{g_{\tau_A},R}$ are damping length scale, energy flux density and group velocity of the Alfvén wave of time period τ_A at any radial height R respectively in c.g.s. units and D, W, V_g are their corresponding dimensionless values. $D_{\tau_{A\,\mathrm{ref}},R_{\mathrm{ref}}}$, $W_{\tau_{A\,\mathrm{ref}},R_{\mathrm{ref}}}$, $V_{g_{\tau_{A\,\mathrm{ref}},R_{\mathrm{ref}}}}$ are damping length scale, energy flux density and group velocity for the Alfvén wave with reference time period 0.01 s at the reference height 1.05 $\mathrm{R}_\odot$ in c.g.s units.

3. Results and discussion

Equation (1) is used to study the combined effect of viscosity and magnetic diffusivity on Alfvén waves propagating at different angles with respect to the normal direction outwards to the Sun. We have also calculated the damping length scale, group velocity and mechanical energy flux density. The solution of the equation gives four roots, in which two are the complex conjugates of the other two. These calculations, using linear and incompressible MHD approximation, exhibit two wave modes. The one mode corresponds to slow-wave, which has lower wave velocity and damping length scale. These waves cannot reach into the corona. Only the fast-wave mode is considered here which propagates into the corona. The calculations and details for the radially propagating high-frequency Alfvén waves of different time periods are given in our companion paper Dwivedi & Srivastava (2006). In the present work, we analyze the propagation and dissipation of the high frequency Alfvén waves, which are propagating at a different angle.

Figure 1 shows the combined effect of viscosity and magnetic diffusivity on the spatial variation of damping length scale of high frequency Alfvén waves. The sharp decrement of damping length scale in this figure indicates an efficient viscous and resistive damping of Alfvén waves below 25 Mm in the magnetic funnel region. It is clear that damping length scale decreases as the propagation angle increases at a particular height. This implies that Alfvén waves, propagating at a larger angle with respect to outward normal direction, dissipate more efficiently.

Figure 2 shows the combined effect of viscosity and magnetic diffusivity on the spatial variation of energy flux density (W) of high frequency Alfvén waves. The sharp decrement of the mechanical energy flux density below 25 Mm indicates the efficient viscous and resistive dissipation of Alfvén waves. At any particular height, the energy flux density decreases with the increment of propagation angle of the Alfvén waves. At any particular propagation angle, the rate of increment of W with radial height approximately agrees with that of Pekünlü *et al.* (2002) in the region between 1.05 $\mathrm{R}_\odot$ and 1.15 $\mathrm{R}_\odot$ However, we find the decrement of Alfvén wave energy flux density in the region beyond 1.21 $\mathrm{R}_\odot$, which may likely indicate the Alfvén wave dissipation. O'Shea *et al.* (2005) measurements of MgX 609.78 Å and 624.78 Å lines from the CDS on SOHO, provide the variation of linewidths and line ratio in the regions

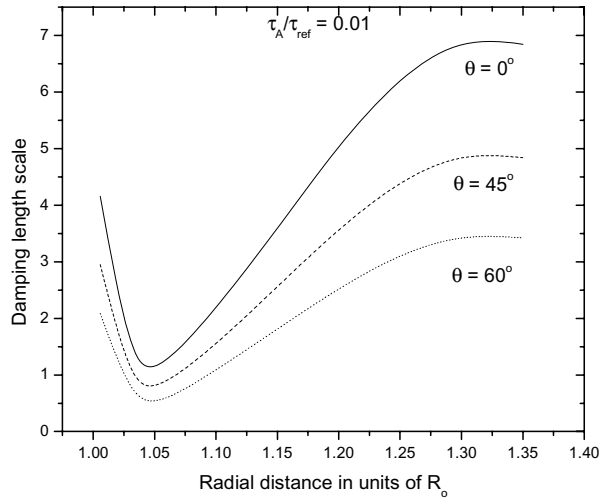

Figure 1. Spatial variation of damping length scales of Alfvén wave under combined effect of viscosity and magnetic diffusivity.

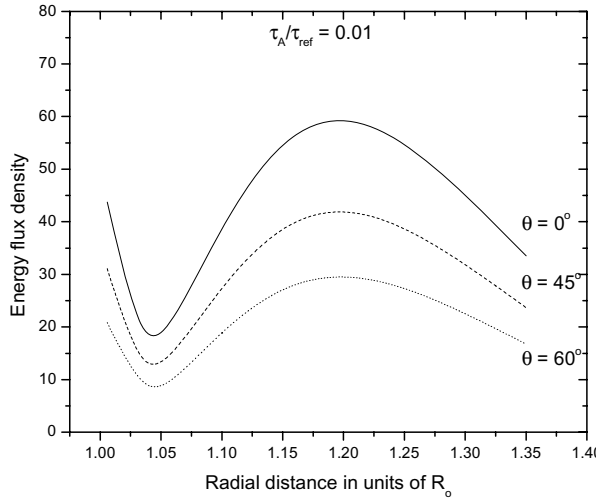

Figure 2. Spatial variation of energy flux density of Alfvén wave under combined effect of viscosity and magnetic diffusivity.

far off-limb at northern polar coronal holes. They find that linewidths show a decrement above $\sim 1.21\,R_\odot$. Dwivedi & Srivastava (2006) have calculated the non-thermal velocity from the linewidth measurements of O'Shea *et al.* (2005) at the effective ion temperature of 10^6 K (see Fig. 3). For any propagation angle, our theoretically calculated Alfvén wave energy flux density shows a decrement beyond $1.21\,R_\odot$ where non-thermal velocity, calculated by Dwivedi & Srivastava (2006) from the linewidth measurements of O'Shea *et al.* (2005), also starts reducing. These results indicate the dissipation of outwardly propagating Alfvén waves, which causes the reduction in the

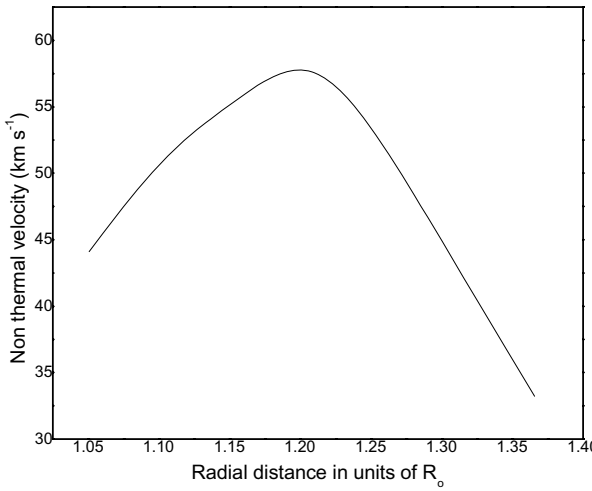

Figure 3. Spatial variation of non-thermal velocity, calculated by using the linewidth observations of O'Shea *et al.* (2005).

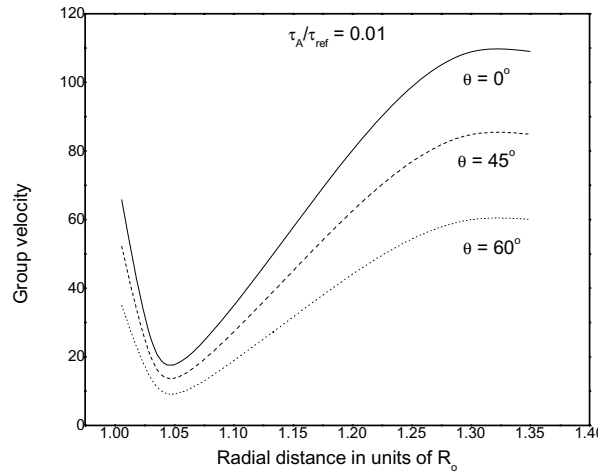

Figure 4. Spatial variation of group velocity of Alfvén wave under combined effect of viscosity and magnetic diffusivity.

non-thermal component of the observed linewidths. This result is also supported by the reduction of our theoretically calculated Alfvén energy flux density for all propagation angles and their peaks match with the peaks in the radial profiles of observed linewidths and non-thermal velocity.

Figure 4 shows the combined effect of viscosity and magnetic diffusivity on the spatial variation of group velocity of high frequency Alfvén waves. We see that the group velocity of Alfvén waves decreases in the region below 25 Mm. As the propagation angle increases, the group velocity of Alfvén wave frequency decreases at any particular height.

Tu *et al.* (2005) made the correlation of the Doppler-velocity and radiance maps of spectral lines emitted by various ions (e.g., Si II, C IV and Ne VIII) with the force-free magnetic field as extrapolated from photospheric magnetograms to different altitudes. They have reported that the solar wind outflow starts between 5 Mm and 20 Mm above the photosphere. Our results show that the region where the damping length scale and energy flux density of Alfvén waves propagating at different angles reduce sharply, coincide with the region where the solar wind outflow starts in the magnetic funnel. Therefore, we suggest that viscous and resistive dissipation of Alfvén waves below 25 Mm may be one of the primary energy sources for the solar wind outflow. As the propagation angle increases, energy flux density reduces at any particular height. At the higher propagation angle, damping length scale also reduces. This shows that although the energy flux density carried over by Alfvén waves which is propagating more steeply, is less at any particular height yet it dissipates quickly via viscous and resistive coronal hole plasma and has a significant role for the conversion of inflow of the solar wind into outflow, below 25 Mm in the underlying magnetic funnels inside the coronal holes.

In conclusion, we have emphasized the importance of the Alfvén waves propagating at different angles with respect to the normal outward direction, and their dissipation in coronal holes, supporting that solar wind outflow starts between 4 Mm and 20 Mm as reported by Tu *et al.* (2005), and also the spectroscopic signature of Alfvén wave dissipation in the off-limb coronal hole plasma as reported by O'Shea *et al.* (2005).

Acknowledgements

We express our gratitude to an anonymous referee for his valuable comments. This work was supported by the Indian Space Research Organisation under its RESPOND programme.

References

Banaszkiewicz, M., Axford, W. I., McKenzie, J. F. 1998, *Astron. Astrophys.*, **337**, 940.
Banerjee, D., Teriaca, L., Doyle, J. G., Wilhelm, K. 1998, *Astron. Astrophys.*, **339**, 208.
Doyle, J. G., Teriaca, L., Banarjee, D. 1999, *Astron. Astrophys.*, **349**, 956.
David, C., Gabriel, A. H., Bely-Dubau, F., Fludra, A., Lemaire, P., Wilhelm, K. 1998, *Astron. Astrophys.*, **336**, 90.
Dwivedi, B. N., Srivastava, A. K. 2006, *Solar Phys.*, **237**, 143–152.
Hackenberg, P., Marsch, E., Mann, G. 2000, *Astron. Astrophys.*, **360**, 1139.
Harrison, R. A., Hood, A. W., Pike, C. D. 2002, *Astron. Astrophys.*, **392**, 319.
Hassler, D. M., Rottman., G. J., Shoub, E. C., Holzer, T. E. 1990, *Astrophys. J.*, **348**, L77.
O'Shea, E., Banerjee, D., Doyle, J. G. 2005, *Astron. Astrophys.*, **436**, L35.
Parker, E. N. 1965, *Space Sci. Rev.*, **4**, 666.
Priest, E. R. 1982, *Solar Magnetohydrodynamics*, Dordrecht, Holland: D. Reidel Publ. Co.; Hingham.
Pekünlü, E. R., Bozkurt, Z., Afsar, M., Soydugan, E., Soydugan, F. 2002, *MNRAS*, **336**, 1195.
Spitzer, L., Jr. 1962, *Physics of Fully Ionized Gases,* Interscience Publ., NY.
Tu, C.-Y., Zhou, C., Marsch, E., Xia, L.-D., Zhao, L., Wang, J.-X., Wilhelm, K. 2005, *Science*, **308**, 519.

J. Astrophys. Astr. (2007) **28,** 9–16

Homotopy Continuation Method of Arbitrary Order of Convergence for Solving the Hyperbolic Form of Kepler's Equation

M. A. Sharaf[1,*], M. A. Banajh[2] & A. A. Alshaary[2]
[1]*Department of Astronomy, Faculty of Science, King Abdul Aziz University, Jeddah, Saudi Arabia.*
[2]*Department of Mathematics, Girls College of Education, Jeddah, Saudi Arabia.*
[*]*e-mail: sharaf_adel@hotmail.com*

Received 2005 July 18; accepted 2007 January 10

Abstract. In this paper, an efficient iterative method of arbitrary integer order of convergence ≥ 2 has been established for solving the hyperbolic form of Kepler's equation. The method is of a dynamic nature in the sense that, moving from one iterative scheme to the subsequent one, only additional instruction is needed. Most importantly, the method does not need any prior knowledge of the initial guess. A property which avoids the critical situations between divergent and very slow convergent solutions that may exist in other numerical methods which depend on initial guess. Computational Package for digital implementation of the method is given and is applied to many case studies.

Key words. Homotopy method—hyperbolic Kepler's equation—initial value problem—orbit determination.

1. Introduction

Many instances of hyperbolic orbits occur in the solar system and recently, among the artificial satellites, lunar and solar probes. Moreover, in some cases of orbit determination for an elliptic orbit, it may very well happen (Escobal 1975) that during the solution process (usually iteration), the eccentricity e becomes greater than unity and the orbit becomes hyperbolic. Also, in the interplanetary transfer, the escape from the departure planet and the capture by the target planet involve hyperbolic orbits (Gurzadyan 1996). On the other hand, in orbit determination of visual binaries provisional hyperbolic orbits are used to represent the periastron section of high-eccentricity orbits of long and indeterminate period (Knudsen 1953). In fact, we should handle hyperbolic orbits frequently when integrating a perturbed motion with the initial condition of nearly parabolic orbits (Fukushima 1997).

From the above, it is then clear that the hyperbolic orbits not only exist naturally, but can also be used to solve some critical orbital situations.

The position–time relation in hyperbolic orbits is known as Kepler's equation for the hyperbolic case and is given as

$$M = e \sinh G - G; \quad 1 \leq e < \infty; \quad 0 \leq M < \infty, \tag{1.1}$$

where G is the eccentric anomaly for a hyperbolic orbit, and M is the mean anomaly (Danby 1988).

Equation (1.1) is transcendental and is usually solved by iterative methods, which in turn need: (a) initial guess and (b) an iterative scheme. In fact, these two points are not separated from each other but there is a full agreement that even accurate iterative schemes are extremely sensitive to initial guess. Moreover, in many cases the initial guess may lead to a drastic situation between divergent and very slow convergent solutions.

In the field of numerical analysis, very powerful techniques have been devoted (Allgower & George 1993) to solve transcendental equations without any prior knowledge of the initial guess. These techniques are known as homotopy continuation methods.

In the present paper, an efficient iterative method of arbitrary positive integer order of convergence ≥ 2 has been established for solving Kepler's equation (1.1) for hyperbolic orbits using homotopy continuation technique. The method does not need any prior knowledge of the initial guess, a property which avoids the critical situations between divergent and very slow convergent solutions, that may exist in other numerical methods depending on initial guess.

Computational package for digital implementation of the method is given and applied for many cases.

2. Basic formulations

2.1 *Homotopy continuation method for solving* $Y(x) = 0$

Suppose one wishes to obtain a solution of a single non-linear equation in one variable x (say)

$$Y(x) = 0, \tag{2.1}$$

where $Y : \mathbf{R} \to \mathbf{R}$ is a mapping which, for our application is assumed to be smooth, i.e., a map has as many continuous derivatives as it requires. Let us consider the situation in which no prior knowledge concerning the zero point of Y is available. Since we assume that such a prior knowledge is not available, then any of the iterative methods will often fail to calculate the zero $\bar{x}$, because poor starting value is likely to be chosen. As a possible remedy, one defines a homotopy or deformation $H : \mathbf{R} \times \mathbf{R} \to \mathbf{R}$ such that

$$H(x, 1) = Q(x); \qquad H(x, 0) = Y(x),$$

where $Q : \mathbf{R} \to \mathbf{R}$ is a (trivial) smooth map having known zero point and H is also smooth. Typically, one may choose a convex

$$H(x, \lambda) = \lambda Q(x) + (1 - \lambda)Y(x) \tag{2.2}$$

and attempt to trace an implicitly defined curve $\Phi(z) \in H^{-1}(0)$ from a starting point $(x_1, 1)$ to a solution point $(\bar{x}, 0)$. If this succeeds, then a zero point $\bar{x}$ of Y is obtained.

The curve $\Phi(z) \in H^{-1}(0)$ can be traced numerically if it is parameterized with respect to the parameter λ, then the classical embedding methods can be applied (Allgower & George 1993).

2.2 *One-point iteration formulae*

Let $Y(x) = 0$ such that $Y : \mathbf{R} \to \mathbf{R}$ smooth map and has a solution $x = \xi$ (say). To construct iterative schemes for solving this equation, some basic definitions are to be recalled as follows:

1. The error in the kth iterate is defined as
$$\varepsilon_k = \xi - x_k.$$
2. If the sequence $\{x_k\}$ converges to $x = \xi$, then
$$\lim_{k\to\infty} x_k = \xi. \tag{2.3}$$
3. If there exists a real number $p \geq 1$ such that
$$\lim_{i\to\infty} \frac{|x_{i+1} - \xi|}{|x_i - \xi|^p} = \lim_{i\to\infty} \frac{|\varepsilon_{i+1}|}{|\varepsilon_i|^p} = K \neq 0, \tag{2.4}$$
we say that the iterative scheme is of order p at ξ. The constant K is called the asymptotic error constant. For $p = 1$, the convergence is linear; for $p = 2$, the convergence is quadratic; for $p = 3, 4, 5$, the convergence is cubic, quartic and quintic, respectively.
4. One-point iteration formulae are those which use information at only one point. Here, we shall consider only stationary one-point iteration formulae which have the form
$$x_{i+1} = R(x_i); \quad i = 0, 1, \ldots \tag{2.5}$$
5. The order of one-point iteration formulae could be determined either from:
   (a) The Taylor series of the iteration function $R(x_n)$ about ξ (e.g., Ralston & Rabinowitz 1978) or from
   (b) The Taylor series of the function $Y(x_{k+1})$ about x_k (Danby & Burkard 1983).

   On the basis of the second approach mentioned above [point (b)] it is easy to form a class of iterative formulae containing members of all integral orders (Sharaf & Sharaf 1998) to solve equation (2.1) as
$$x_{i+1} = x_i + \delta_{i,m+2}; \quad i = 0, 1, 2, \ldots; \quad m = 0, 1, 2, \ldots \tag{2.6}$$
   where
$$\delta_{i,m+2} = \frac{-Y_i}{\sum_{j=1}^{m+1} (\delta_{i,m+1})^{j-1}\, Y_i^{(j)}/j!}; \quad \delta_{i,1} = 1; \quad \forall i \geq 0, \tag{2.7}$$
$$Y_i^{(j)} \equiv \left.\frac{d^j Y(x)}{dx^j}\right|_{x=x_i}; \quad Y_i \equiv Y_i^{(0)}. \tag{2.8}$$

The convergence order is $m + 2$ as shown from the error formula

$$\varepsilon_{i+1} = -\frac{1}{(m+2)!} \frac{Y^{(m+2)}(\zeta)}{Y_i^{(1)}(\zeta_1)} \varepsilon_i^{m+2} \tag{2.9}$$

where ζ is between x_{i+1} and x_i and ζ_1 is between x_{i+1} and ξ.

3. Computational developments

3.1 *Embedding methods*

The basic idea of the embedding methods referred to at the end of subsection 2.1 is explained in the following algorithm for tracing the curve $\Phi(z) \in H^{-1}(0)$ from, say $\lambda = 1$ to $\lambda = 0$.

3.1.1 *Computational algorithm 1*

Purpose: To solve $Y(x) = 0$ by embedding method.
Input:

(1) The function $Q(x)$ with defined root x_1 such that $H(x_1, 1) = 0$,
(2) positive integer m.

Output: Solution x of $Y(x) = 0$.

Computational sequence:

(1) Set $x = x_1$, $\lambda = (m-1)/m$, $\Delta\lambda = 1/m$.
(2) For $i := 1$ to m, do
begin
Solve $H(y, \lambda) = \lambda Q(y) + (1-\lambda)Y(y) = 0$ iteratively for y
using x as starting value.
$x = y$
$\lambda = \lambda - \Delta\lambda$
end.

3.2 *Solution of the hyperbolic form of Kepler's equation*

Recalling equation (1.1) as

$$Y(G) = e \sinh G - G - M = 0. \tag{3.1}$$

Two notes are to be recorded as follows:

(1) From equation (2.9) it is clear that an iterative scheme for solving equation (3.1) includes derivatives of Y as much as the order of the scheme. On the other hand, the higher the order of an iterative scheme, the higher its accuracy and rate of convergence will be. Regarding this last fact, the remarkable simplicity of the derivative formulae of Y which are

$$Y^{(1)} = e \cosh G - 1, \tag{3.2.1}$$

$$Y^{(2k)} = e \sinh G; \quad k \geq 1, \tag{3.2.2}$$

$$Y^{(2k+1)} = e \cosh G; \quad k \geq 1, \tag{3.2.3}$$

enables us to find derivatives of $Y(G)$ as many as we need.
(2) Homotopy continuation method (see subsection 2.1) is very powerful technique for solving $Y(G) = 0$ without prior knowledge of the initial guess.
From these two notes, we can now establish for the solution of equation (3.1), an iterative algorithm of any positive integer order $l \geq 2$. Moreover, the algorithm

does not need prior knowledge of the initial guess. According to equation (2.9), the algorithm is of a dynamic nature in the sense that it includes iterative schemes up to the lth order such that, going from one scheme to the subsequent one, only additional instruction is needed.

This algorithm is illustrated in what follows with algorithm 1 augmented to it, together with the Q function of the homotopy H (equation 2.2) as $Q(x) = x - 1$, so that $H(x_1, 1) = 0$, where $x_1 = 1$.

3.2.1 *Computational algorithm 2*

Purpose: To solve the hyperbolic form of Kepler's equation by iterative schemes of quadratic up to lth convergence orders without prior knowledge of the initial guess using homotopy continuation method with $Q(G) = G - 1$.
Input: M, e, l, m (positive integer), *Tol* (specified tolerance)
Computational Sequence:

1. Set $G = 1$, $\Delta\lambda = 1/m$, $\lambda = 1 - \Delta\lambda$
2. For $i := 1$ to m do
   begin $\{i\}$
   $\Phi = 1 - \lambda$
   $Y = \lambda(G - 1) + \Phi(e \sinh G - G - M)$
   $Y^{(1)} = \lambda + \Phi(e \cosh G - 1)$
   $\Delta G = -Y/Y^{(1)}$
   If $l = 2$ go to step 3
   $Y^{(2)} = \Phi e \sinh G$
   $H = Y^{(1)} + \Delta G * Y^{(2)}/2$
   $\Delta G = -Y/H$
   If $l = 3$ go to step 3
   $Y^{(3)} = \Phi e \cosh G$
   $H = Y^{(1)} + \Delta G * Y^{(2)}/2 + (\Delta G)^2 * Y^{(3)}/6$
   $\Delta G = -Y/H$
   If $l = 4$ go to step 3
   $L = l - 1$
   For $k : 4$ to L do
   begin $\{k\}$
   Set $Y^{(k)} = Y^{(k-2)}$; $n = k - 1$; $H = Y^{(1)}$; $B = 1$
   For $j := 1$ to n do
   begin $\{j\}$
   $B = \Delta G * B/(j + 1)$
   $H = H + B * Y^{(j+1)}$
   end $\{i\}$
   $\Delta G = -Y/H$
   end $\{k\}$
3. If $|\Delta G| \leq Tol$ go to step 4
   $G = G + \Delta G$
   $\lambda = \lambda - \Delta\lambda$
   end $\{j\}$
4. end.

3.3 *Numerical applications*

3.3.1 *Hyperbolic limitations on mean anomaly*

Because a trajectory in hyperbolic orbit approaches the asymptotes at infinite distance, the true anomaly ν is limited within a certain range. These values of ν are

$$-180^\circ + \cos^{-1}\left(\frac{1}{e}\right) < \nu < 180^\circ - \cos^{-1}\left(\frac{1}{e}\right),$$

from this inequality and the known equation

$$\tanh\left(\frac{G}{2}\right) = \sqrt{\frac{e-1}{e+1}}\tan\left(\frac{\nu}{2}\right),$$

we get using equation (1.1) Table 1, which shows some representative values of the limitations on the mean anomaly of a hyperbolic orbit whose eccentricity varies.

3.3.2 *Numerical results*

For a given values of M and e (as obtained in Table 1) there exists an optimum pair (m^*, l^*) that can determine the solution of equation (1.1) more accurately than that determined from other pairs. Consequently we may define an acceptable solution set to hyperbolic form of Kepler's equation for a given M and e as

$$G = \{G : Tol \leq \varepsilon, m = m^*; l = l^*\},$$

where *Tol* is a given tolerance.

The above computational algorithm 2 is applied for

$$e = 1.5, 2.0, 3.0, 4.0, 5.0, 9.0, 10.5, 13.5, 16.0, 19.0, 21.0, 25.5$$

and some values of M with $Tol = 10^{-8}$. The elements of the acceptable solution set of the present method are listed in Table 2 for the given M and e.

Table 1. Hyperbolic limitations on mean anomaly.

Eccentricity	Minimum M	Maximum M
1.5	− 11171.2	11171.2
2.0	− 17311.2	17311.2
3.0	− 28274.9	28274.9
4.0	− 38720.5	38720.5
5.0	− 48980.4	48980.4
9.0	− 89433.3	89433.3
10.5	−104513	104513
13.5	−134620	134620
16	−159678	159678
19	−189727	189727
21	−209752	209752
25.5	−254794	254794

Table 2. Elements of acceptable solution set of the hyperbolic form of Kepler's equation with $Tol = 10^{-8}$.

M	m^*	l^*	G
Case 1: $e = 1.5$			
−11151.0	20	3	−9.60783
11171.0	5	3	9.60962
Case 2: $e = 2.0$			
6311.0	6	3	8.75144
−17000.0	12	3	−9.74154
Case 3: $e = 3.0$			
2827.0	8	3	7.54417
−3500.0	18	3	−7.75727
Case 4: $e = 4.0$			
3700.2	7	3	7.52503
−370.2	6	4	−5.23497
Case 5: $e = 5.0$			
48970.4	7	3	9.88288
−3200.0	13	3	−7.15685
Case 6: $e = 9.0$			
89333.3	7	3	9.89616
−103.8	8	4	−3.17024
Case 7: $e = 10.5$			
145.31	19	3	3.34464
−104511.0	14	3	−9.89891
Case 8: $e = 13.5$			
1345.21	9	3	5.29872
−124520.0	17	3	−9.82276
Case 9: $e = 16.0$			
11154.2	6	3	7.24078
−154.2	9	4	−2.98053
Case 10: $e = 19.0$			
1997.5	7	3	5.35106
−180.0	9	4	−2.96066
Case 11: $e = 21.0$			
17500.5	6	3	7.41903
−4582.51	16	3	−6.07996
Case 12: $e = 25.5$			
12.85	14	3	0.502235
−1000.98	12	3	−4.36772

4. Conclusion

In concluding the paper, an efficient and simple algorithm for the solution of the hyperbolic form of Kepler's equation has been established. Its efficiency is due to factors such as:

- It is of a dynamical nature, in the sense that it includes iterative schemes of any positive integer order ≥ 2 such that, moving from one scheme to the subsequent one, only one additional instruction is needed.

- Does not depend on initial guess – a property that avoids it from falling in critical situations between divergent and very slow convergent solutions that may exist in other numerical methods which depend on initial guess.
- It includes two controllable parameters only, m and l by means of which one can certainly determine an optimum solution to any desired accuracy by what we call the acceptable solution set.

References

Allgower, E. L., George, K. 1993, *Numerical Continuation Methods*, Springer-Verlag, Berlin.
Danby, J. M. A. 1988, *Fundamental of Celestial Mechanics*, 2nd edn, William-Bell, Inc., Richmond, Virginia, USA.
Danby, J. M. A., Burkard, T. M. 1983, *Celestial Mechanics*, **31**, 95.
Escobal, P. R. 1975, *Methods of Orbit Determination*, Wiley, New York.
Fukushima, T. 1997, *Astron. J.*, **113(5)**, 45.
Gurzadyan, G. A. 1996, *Theory of Interplanetary Flights*, Gordon and Breach Publishers, Australia, Canada, United Kingdom.
Knudsen, N. W. 1953, Publ. Lund Obs., No. 12.
Ralston, A., Rabinowitz, P. 1978, *A First Course in Numerical Analysis*, McGraw-Hill Kogakusha Ltd., Tokyo.
Sharaf, M. A., Sharaf, A. A. 1998, *Celestial Mechanics and Dynamical Astronomy*, **69**, 331.

J. Astrophys. Astr. (2007) **28,** 17–27

Modeling Repulsive Gravity with Creation

R. G. Vishwakarma[1] & J. V. Narlikar[2]
[1]*Unidad Académica de Matemáticas, Universidad Autónoma de Zacatecas, C.P. 98068, Zacatecas, ZAC, Mexico.*
e-mail: rvishwa@mate.reduaz.mx
[2]*Inter-University Centre for Astronomy and Astrophysics, Post Bag 4, Ganeshkhind, Pune 411 007, India.*
e-mail: jvn@iucaa.ernet.in

Received 2006 May 8; accepted 2007 January 18

Abstract. There is a growing interest among cosmologists for theories with negative energy scalar fields and creation, in order to model a repulsive gravity. The classical steady state cosmology proposed by Bondi, Gold & Hoyle in 1948, was the first such theory which used a negative kinetic energy creation field to invoke creation of matter. We emphasize that creation plays a very crucial role in cosmology and provides a natural explanation to the various explosive phenomena occurring in local ($z < 0.1$) and extra galactic universe. We exemplify this point of view by considering the resurrected version of this theory – the quasi-steady state theory, which tries to relate creation events directly to the large scale dynamics of the universe and supplies more natural explanations of the observed phenomena.

Although the theory predicts a decelerating universe at the present era, it explains successfully the recent SNe Ia observations (which require an accelerating universe in the standard cosmology), as we show in this paper by performing a Bayesian analysis of the data.

Key words. Cosmology: theory, observation, creation—negative energy fields—SNe Ia.

1. Introduction

Remarkable progress has been made in various types of astrophysical and cosmological observations in recent years. Among these, the accurate measurements of the anisotropies in the CMB made by the WMAP experiment appear to offer the most promising determination of the cosmological parameters. The results of the WMAP experiment are however often quoted as providing a direct evidence for an accelerating universe, which however is not correct. The cosmological constraints as established by the WMAP team (Spergel *et al.* 2003) entirely rely on the power law spectrum assumption and could be erroneous (Kinney 2001; Hannestad 2001). Taken on their face value, the WMAP observations are fully consistent with the decelerating models like the CDM Einstein-de Sitter model (Vishwakarma 2003; Blanchard 2005).

The possibility of an accelerating universe in fact emerges from the measurements of distant SNe Ia, which look fainter than they are expected in the standard decelerating models. This observed faintness is generally explained by invoking some hypothetical source with negative pressure generally known as *'dark energy'*. This happens because the metric distance of an object, out to any redshift, can be increased by incorporating a *'fluid'* with negative pressure in Einstein's equations and hence the object looks fainter. The simplest and the most favoured candidate of dark energy is a positive cosmological constant Λ, which is however plagued with the horrible fine tuning problems – an issue amply discussed in the literature. This has led a number of cosmologists to resort to scalar field models called quintessence whose function is to cause the scale factor to accelerate at late times by violating the strong energy condition. While the scalar field models enjoy considerable popularity, they have not helped us to understand the nature of dark energy at a deeper level. By and large, the scalar field potentials used in the literature have no natural field theoretical justification and have to be interpreted as a low energy effective potential in an *ad hoc* manner. Moreover they also require fine tuning of the parameters in order to be viable (to find several other shortcomings, see for example, Padmanabhan 2005).

As desperate times call for desperate measures, the cosmologists, in order to model the dark energy, have now turned to *'phantom'* or *'ghost'* scalar field models with negative kinetic energy (Caldwell 2002; Carroll *et al.* 2003; Gibbons 2003; Singh *et al.* 2003; Sami & Toporensky 2004). The classical steady state cosmology proposed by Bondi & Gold (1948) and Hoyle (1948) was the first such theory which used a negative kinetic energy creation field to invoke creation of matter. It is interesting to note that, distinct from all the existing big bang models at that time, this model predicted an accelerating universe. However, it is unfortunate that the theory was not given any credit (which it deserved, despite the difficulties associated with it) when the SNe Ia observations started claiming an accelerating universe in 1998.

Once cosmologists have lost their inhibitions about negative energy fields, the time is ripe for considering the idea that the creation of matter plays an important role in cosmology. We exemplify this point of view by considering the resurrected version of the classical steady state theory, namely the quasi-steady state cosmology (QSSC) which has not been given proper attention as it deserves. This theory was proposed by Hoyle, Burbidge & Narlikar in 1993 (1993; 1995), wherein the introduction of negative kinetic energy scalar field is not *ad hoc* but is required to ensure that matter creation does not violate the law of conservation of matter and energy. However, first we emphasize that the idea of creation of matter is already present in general relativity, though hidden behind some simplifying assumptions.

With a suitable Lagrangian for the source terms, the Einstein field equations can be written as

$$R_{ik} - \frac{1}{2} g_{ik} R = -8\pi G[T_{ik}^{(\text{matter})} + T_{ik}^{(\Lambda)} + T_{ik}^{(\phi)} + \cdots], \tag{1}$$

where we have considered the speed of light $c = 1$. The only constraint on the source terms, which is imposed by this equation, is the conservation of the right-hand side through the Bianchi identities: $[R_{ij} - \frac{1}{2} R g_{ij}]^{;j} = 0 = [T_{ij}^{(\text{matter})} + T_{ij}^{(\Lambda)} + T_{ij}^{(\phi)} + \cdots]^{;j}$, implying that only the sum of all the energy-momentum tensors is conserved, individually they are not. If we take them conserved separately, as is in practice among the cosmologists, it can be done only through the additional assumption of no interaction

(minimal coupling) between different source fields, which though seems *ad hoc* and nothing more than a simplifying assumption. On the contrary, interaction is more natural and is a fundamental principle. Of course some ideal cases are consistent with the idea of minimal coupling, for example, $T_{ik}^{(\text{matter})}$ with a constant Λ. However, imposing this assumption on non-trivial cases would result in losing some important information. For example, taken on face value, a time-dependent Λ, with matter, implies matter creation and results in a Machian model (Vishwakarma 2002a). However, if one makes an additional assumption of no interaction between $T_{ik}^{(\text{matter})}$ and $\Lambda(t)$, these features are lost and $\Lambda(t)$ reduces to a constant. It is well known that even if one considers the Robertson–Walker spacetime (to avoid non-Machian Godel's solution of Einstein field equations), there still exists a non-Machian solution of Einstein field equations – the de Sitter solution. Creation has many more attractive features. It has been shown how the scalar creation field helps in resolving the problems of singularity, flatness and horizon in cosmology (Narlikar & Padmanabhan 1985). Such a negative energy creation field is responsible for a non-singular bounce from a high non-singular density state, as has been shown by Hoyle & Narlikar (1964). This idea has been recently used by Steinhardt and Turok in their oscillatory model (Steinhardt & Turok 2002). The quasi-steady state cosmology (QSSC) is also a Machian theory which is derived from an action principle based on Mach's Principle, and assumes that the inertia of matter owes its origin to other matter in the Universe. The stress-energy tensor for creation (corresponding to $T_{ik}^{(\phi)}$ in equation (1)) is given by

$$T_{ik}^{\text{creation}} = -f\left(C_i C_k + \frac{1}{4} C^l C_l \; g_{ik}\right), \tag{2}$$

where f is a positive coupling constant and the gradient $C_i \equiv \partial\phi/\partial x^i$ is the contribution from a trace-free zero rest mass scalar field ϕ of *negative* energy and stresses. The Λ in this theory (corresponding to a $T_{ik}^{(\Lambda)} \equiv -\Lambda g_{ik}/8\pi G$ of (1)) appears as a constant of nature with its value $\approx -2 \times 10^{-56}\ \text{cm}^{-2}$, which falls within the normally expected region of the magnitude of the cosmological constant. However, note that its sign is negative, which is a consequence of the Machian origin of the cosmological constant. The theory does not face the cosmological constant problem mentioned earlier. In fact, the Λ in the QSSC does not represent the energy density of the quantum fields, as this model does not experience the energy scales of quantum gravity except within the local centres of creation. The theory offers a purely stellar-based interpretation of all observed nuclei including the light ones (Burbidge *et al.* 1957; Burbidge & Hoyle 1998). In the following, we demonstrate in brief the main features of this cosmology and how it confronts the various observations (for more details, one can consult Sachs *et al.* 1996; Hoyle *et al.* 2000).

The QSSC represents a cyclic universe with its Robertson–Walker scale factor given by

$$S(t) = e^{t/P}\left[1 \; + \; \eta \cos\left(\frac{2\pi\tau}{Q}\right)\right], \tag{3}$$

where the timescales $P \approx 10^3\,\text{Gyr} \gg Q \approx 40$–$50\,\text{Gyr}$ are considerably greater than the Hubble time scale of 10–15 Gyr of the standard cosmology. The function $\tau(t)$ is very nearly like the cosmic time t, with significantly different behaviour for short

duration near the minima of the function $S(t)$. The parameter η has modulus less than unity, thus preventing the scale factor from reaching zero. Typically, $\eta \sim 0.8 - 0.9$. Hence there is no space-time singularity, or a violation of the law of conservation of matter and energy, as happens at the big bang epoch in the standard cosmology. The model has cycles of expansion and contraction (regulated respectively by the creation field and the negative Λ) of comparatively shorter period (Q) superposed on a long term (P) steady state-like expansion. Creation of matter, which occurs through explosive processes, is also periodic, being confined to pockets of strong gravitational fields around compact massive objects and the nuclei of existing galaxies. Such processes take place whenever the energy of the creation field quantum rises above a threshold energy, which is equal to the restmass energy of the created Planck particle.

The model provides a natural explanation to the various explosive phenomena occurring in local ($z < 0.1$) and extra galactic universe. By the early 1960s it had become clear that very large energy outbursts are taking place in the nuclei of galaxies. In the decades since then it has been found that many active nuclei are giving rise to X-rays, and relativistic jets, detected in the most detail as high frequency radio waves. A very large fraction of all of the energy which is detected in the compact sources is non-thermal in origin, and is likely to be incoherent synchrotron radiation or Compton radiation. In addition to this, we see several other explosive phenomena in the Universe, such as jets from radio sources, gamma-ray bursts, X-ray bursters, QSOs, etc. Generally it is assumed that a black hole plays the lead role in such an event by somehow converting a fraction of its huge gravitational energy into large kinetic energy of the 'burst' kind. In actuality however, we do not see infalling matter that is the signature of a black hole. Rather we see outgoing matter and radiation, which agree very well with the idea of creation events formulated in the framework of the QSSC.

There are several free parameters in the model which are estimated from the observations and provide a decelerating universe at the present cycle of expansion. It is then interesting to see how the model explains the SNe Ia and other observations! This is shown in the following section.

2. The high redshift supernovae Ia

It is generally accepted that metallic vapours are ejected from the SNe explosions which are subsequently pushed out of the galaxy through pressure of shock waves (Hoyle & Wickramasinghe 1988). Experiments have shown that metallic vapours on cooling, condense into elongated whiskers of $\approx 0.5 - 1$ mm length and $\approx 10^{-6}$ cm cross-sectional radius (Donn & Sears 1963; Nabarro & Jackson 1958). It can be shown that the extinction from the whisker dust adds an extra magnitude $\delta m(z)$ to the apparent magnitude $m(z)$ (arising from the cosmological evolution) of the SN light emitted at the epoch of redshift z, which is given by

$$\delta m(z) = 1.0857 \times \kappa \rho_{g0} \int_0^z (1+z')^2 \frac{\mathrm{d}z'}{H(z')}, \tag{4}$$

where κ is the mass absorption coefficient and ρ_{g0} is the whisker grain density at the present epoch. The net apparent magnitude is then given by

$$m^{\mathrm{net}}(z) = m(z) + \delta m(z). \tag{5}$$

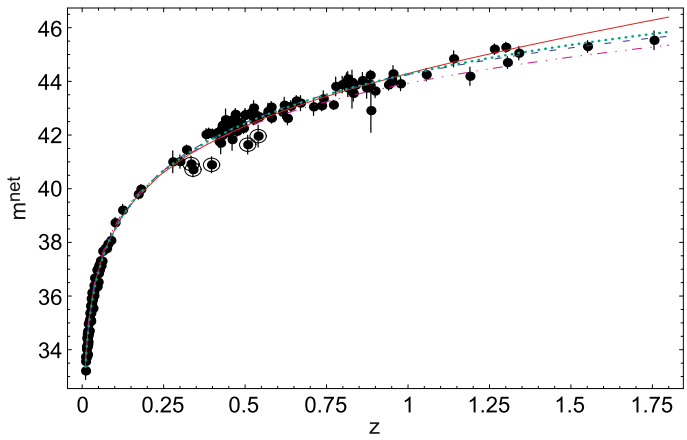

Figure 1. Some best-fitting models are compared with the 'gold sample' of SNe Ia data with 157 points as considered by Riess *et al.* (2004). The solid curve corresponds to the QSSC model with the whisker dust, the dotted curve corresponds to the flat ΛCDM model, the dashed curve corresponds to the spherical ΛCDM model, and the dashed-dotted curve corresponds to the Einstein-de Sitter model. The models differ significantly for $z > 1.2$. The encircled points seem to be general outliers which are missed by all the models.

By taking account of this effect, it has been shown that this kind of dust extinguishes radiation travelling over long distances and decelerating models without any dark energy (for example, the Einstein-de Sitter model) can also explain high redshift SNe Ia observations successfully (Vishwakarma 2002b, 2003, 2005). QSSC in fact resorts to this dust to explain not only the SNe Ia observations but also CMB, as we shall see in the following. It has been shown (Narlikar *et al.* 2002; Vishwakarma & Narlikar 2005) that by taking account of this effect, QSSC explains successfully the SNe Ia data from Perlmutter *et al.* (1999) and also shows an acceptable fit to the 'gold sample' of 157 SNe Ia recently published by Riess *et al.* (2004) which, in addition to having previously observed SNe, also includes some newly discovered highest-redshift SNe Ia by the Hubble Space Telescope. Though this sample is believed to have a 'high-confidence' quality of the spectroscopic and photometric record for individual supernovae, we note that there are some SNe (1997as, 1997bj, 2000eg, 2001iw, 2001iv) in this sample which do not seem to be consistent with any of the models generally considered in the fitting and appear as general outliers (see the encircled SNe in Figs. 1 and 2). By excluding these points, the fit to different models improves considerably. For example, the χ^2 value per degree of freedom (dof) for the best-fitting QSSC model reduces to 1.18 from the earlier $\chi^2/\text{dof} = 1.30$ obtained from the full sample of 157 points (Vishwakarma & Narlikar 2005). The fit to the standard (flat ΛCDM) cosmology improves tremendously from $\chi^2/\text{dof} = 1.14$ (from 157 points) to $\chi^2/\text{dof} = 0.99$ (from 152 points). The details of the fit (in the case of the frequentist approach) can be found in Vishwakarma & Narlikar (2005).

Though there is no clearly defined value of χ^2/dof for an acceptable fit, a *'rule of thumb'* for a *moderately* good fit is that χ^2 should be roughly equal to the number of dof. A more quantitative measure for the *goodness-of-fit* is given by the χ^2*-probability*. If the fitted model provides a typical value of χ^2 as x at n dof, this

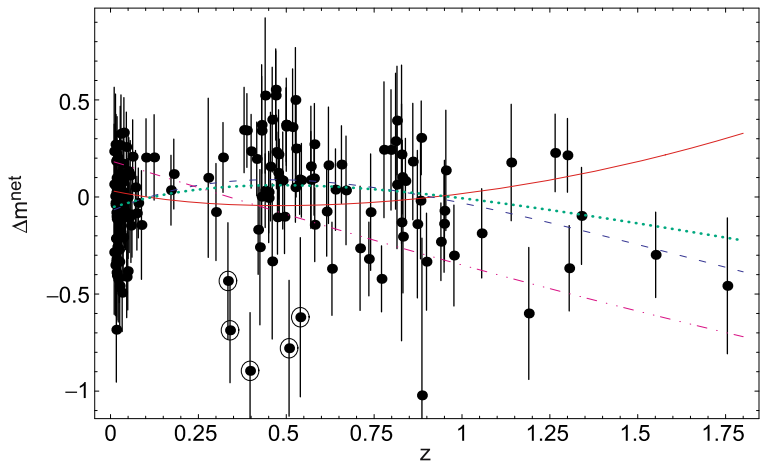

Figure 2. Modified Hubble diagram of the 'gold sample' of SNe Ia minus a fiducial model ($\Omega_{m0} = 0$, $\Omega_{\Lambda 0} = 0$). The relative magnitude ($\Delta m^{\rm net} \equiv m^{\rm net} - m_{\rm fiducial}$) is plotted for some best-fitting models, by using the original error bars. The solid curve corresponds to the QSSC model with the whisker dust, the dotted curve corresponds to the flat ΛCDM model, the dashed curve corresponds to the spherical ΛCDM model, and the dashed-dotted curve corresponds to the Einstein-de Sitter model. The encircled points seem to be general outliers which are missed by all the models.

probability is given by

$$Q(x, n) = \frac{1}{\Gamma(n/2)} \int_{x/2}^{\infty} e^{-u} u^{n/2-1} {\rm d}u. \tag{6}$$

Roughly speaking, it measures the probability that *the model does describe the data and any discrepancies are mere fluctuations which could have arisen by chance*. To be more precise, $Q(x, n)$ gives the probability that a model which does fit the data at n dof, would give a value of χ^2 as large or larger than x. If Q is very small, the apparent discrepancies are unlikely to be chance fluctuations and the model is ruled out. It may however, be noted that the χ^2-probability strictly holds only when the models are linear in their parameters and the measurement errors are normally distributed. It is though common, and usually not too wrong, to assume that the χ^2-distribution holds even for models which are not strictly linear in their parameters, and for this reason, the models with a probability as low as $Q > 0.001$ are usually deemed acceptable (Press *et al.* 1986). Models with vastly smaller values of Q, say, 10^{-18} are rejected. The probability Q for the best-fitting QSSC to the full sample is obtained as 0.007, which is though very small, but acceptable. By excluding the above-mentioned 5 outliers, Q improves to 0.062. The corresponding probabilities in the case of the standard ΛCDM cosmology are obtained as 0.109 and 0.534.

We note that the fit to the QSSC is considerably worse than those in the standard ΛCDM cosmology. However, one cannot compare the relative merits of the models on the basis of the χ^2-probability (frequentist approach), which uses the best-fitting parameter values and hence judges only the maximum likely performance of the models. The more appropriate theory for such comparisons is the Bayesian theory which does not hinge upon the best-fitting parameter values and evaluates the overall

performance of the models by using average likelihoods (rather than the maximum likelihoods), given by the *Bayes factor* B. The theory employs the premise that if we assume an equal prior probability for competing models, the probability for a given model is proportional to the marginalised likelihood called *evidence*. We have described this theory, in brief, in the Appendix (for more details, see Drell *et al.* 2000; John & Narlikar 2002).

In order to calculate the Bayes factor B for the two models QSSC and the standard ΛCDM, first we have to fix the prior probabilities for the free parameters. While the flat QSSC has four free parameters $\kappa\rho_{\rm g0}H_0^{-1}$, $\Omega_{\Lambda 0}$, $z_{\rm max}$ and $\mathcal{M}$ (see the Appendix of Vishwakarma & Narlikar (2005), the standard ΛCDM has only two free parameters Ω_{m0} and $\mathcal{M}$. We would like to mention that the whisker dust was already introduced in the QSSC in order to explain the CMB which put a constraint on the density of the dust $\rho_{\rm g0} \approx 10^{-34}$ g cm^{-3} (Narlikar *et al.* 2003). This value, taken together with the observational constraints on κ (Wickramasinghe & Wallis 1996) and H_0 (Freedman & Turner 2003) from other observations, supplies a value of the parameter $\kappa\rho_{\rm g0}H_0^{-1}$ close to the best-fitting value estimated from the SNe Ia observations. Hence we assign a prior on the parameter $\kappa\rho_{\rm g0}H_0^{-1}$ that it lies in the range $\kappa\rho_{\rm g0}H_0^{-1} \in [3.5, 6]$. We also note that $\Omega_{\Lambda 0}$ in the QSSC does not receive any significant contribution from $\Omega_{\Lambda 0} < -0.3$. Taking account of this and the theoretical constraint that Λ in the QSSC is negative, we assign a prior on the parameter $\Omega_{\Lambda 0} \in [-0.3, 0]$. To the rest two parameters in the QSSC, about which we do not have prior information, we assign liberal priors: $z_{\rm max} \in [5, 10]$ (to be consistent with the highest redshift ≈ 7 observed so far) and $\mathcal{M} \in [41, 45]$ (which is the common parameter). For the parameter Ω_{m0} in the flat ΛCDM model, we assume that $\Omega_{m0} \in [0, 1]$ (which is equivalent to assigning $\Omega_{\Lambda 0} \in [0, 1]$). It may be noted that the $\Omega_{\Lambda 0}$ in the two models are altogether different quantities (though they have been denoted by the same symbol in order to match the general convention) and there is no reason to assign the same probability for them in the two different models.

When calculated for the full 'gold sample' of 157 points, these prior probabilities give a Bayes factor favouring the standard ΛCDM over the QSSC as $B = 2.77$, which though indicates an evidence against the QSSC, however, the evidence is not definite and is not worth more than a bare mention (for the interpretation of B, see the Appendix). Our assigning $\Omega_{\Lambda 0} \in [-0.3, 0]$ in the QSSC can raise eyebrows, as this probability is very conservative compared to the one in the ΛCDM. However, assigning this probability is due to the reason that the likelihood for Λ in the QSSC does not receive any significant contribution from $\Omega_{\Lambda 0} < -0.3$, as mentioned earlier. For example, increasing the domain of Λ in its prior to $\Omega_{\Lambda 0} \in [-1, 0]$ in the QSSC, results in lowering the likelihood of the model, as expected. This gives the Bayes factor $B = 9.30$, which indicates that the evidence against the QSSC is definite, though not strong.

One should also note that a proper assessment of the probability for a model is given by $p = 1/(1+B)$. Thus for the above-mentioned two choices of the prior probabilities, the corresponding probabilities for the QSSC are 0.27 and 0.10, which are reasonably good probabilities.

It is interesting to note that the whisker dust, which is a vital ingredient of the QSSC, does not make any significant improvement in the fit to the ΛCDM cosmology. It may be argued that this kind of dust can create too much optical depth for the high redshift objects and they need to be excessively bright in order to be seen. However, from our

calculations, we find that the objects over the present cycle right up to the maximum redshift will be fainter, at the most, by ~ 6 magnitudes only and it is possible to see even sources from many previous cycles, though they will be very faint.

In Fig. 1, we have compared some best-fitting models with the actual data points. In order to have a better visual comparison of these models, we magnify their differences by plotting the relative magnitude with respect to a fiducial model $\Omega_{m0} = \Omega_{\Lambda 0} = 0$, without any whiskers (which has a reasonably good fit: $\chi^2/\mathrm{dof} = 1.2$). This has been shown in the 'modified' Hubble diagram in Fig. 2. In the following sections we describe in brief, the other important features of the QSSC.

3. The CMB

As far as the origin and nature of the CMB is concerned, the QSSC uses a fact that is always ignored by standard cosmologists. If we suppose that most of the ^{4}He found in our own and external galaxies (about 24% of the hydrogen by mass) was synthesized by hydrogen burning in stars, the energy released amounts to about 4.37×10^{-13} erg cm^{-3}. This is almost exactly equal to the energy density of the microwave background radiation with $T = 2.74$ K! In the standard cosmology, this has to be dismissed as a coincidence. Thus according to the QSSC, the CMB is the relic starlight left by the stars of the previous cycles which has been thermalized by the metallic whisker dust emitted by the supernovae. As a typical cycle proceeds from the maximum of scale factor towards the next minimum, the wavelengths of the starlight from the previous cycle are shortened since the universe contracts by a considerable factor. It has been shown (Narlikar *et al.* 1997) that at wavelengths 100 m–20 cm, sufficient optical depth exists for this radiation to thermalize in about twenty cycles. The production of microwaves in this fashion goes on in each cycle, and the process of frequent absorption and re-radiation by whiskers will eventually generate a uniform background, *except for the contribution from the latest generation of clusters*. These will stand out as inhomogeneities on the overall uniform background arising from certain intrinsic inhomogeneities of the process as well as from the cosmological model. A quantitative analysis shows that this process requires an intergalactic dust of density of $\approx 10^{-34}$ g cm^{-3}, which is very close to the best-fitting value estimated from the SNe Ia observations. The theory also explains the peaks at $l \sim 200$ and $l \sim 600$ which are related, in this cosmology, to the clusters and groups of clusters (Narlikar *et al.* 2003). Also, we have taken stock of the WMAP observations in (Narlikar *et al.* 2007).

Though these studies do not give predictions as sharp as those given by the standard Big Bang cosmology, however, one should note the attitudinal difference between this approach and the standard one. In the Big Bang cosmology, the inferences are related to the postulated initial conditions prevailing well beyond the range of direct observations (at redshifts 1100). Whereas the QSSC interpretation links the inhomogeneities of the radiation field to those of the matter field, on which we do not have very accurate data at present, but which may be observable one day.

It may also be worthwhile to mention that there are claims that like the dipole, the quadrupole and the octopole harmonics of the CMB spectrum also have their origin in the solar system (Starkman *et al.* 2004). If this is correct, then subtracting this foreground contribution from the rest of the signal (in order to have the temperature fluctuations only at the time of the Big Bang) would place the inflationary model in serious trouble.

4. The non-baryonic dark matter

Unlike the standard big bang cosmology, the QSSC allows the dark matter to be baryonic. It may be recalled that the standard cosmology predicts the existence of non-baryonic, though as yet undetected, particles to solve the problems of structure formation and of the missing mass in bound gravitational systems such as galaxies and clusters of galaxies. The most favoured candidate of non-baryonic dark matter postulated by many astrophysicists, cosmologists and particle physicists is a massive but very weakly interacting particle called WIMP (Weakly Interacting Massive Particle), a hypothetical elementary particle that was produced moments after the Big Bang. Currently there are a number of WIMP detection experiments underway. Among these, the DAMA experiment (Bernabei *et al.* 2003), which measures the annual modulation in WIMP interactions with the sodium-iodide detectors caused by the earth's rotation around the Sun, is the only one to have claimed a positive signal. However, the results of this experiment are controversial as other more sensitive searches have not detected nuclear recoils due to WIMP interactions (Akerib *et al.* 2004; Angloher *et al.* 2005) and concluded that almost all the events measured by DAMA were from neutrons, and should not be attributed to scattering events from dark-matter WIMPs. It is therefore fair to say that this scheme has still to demonstrate its viability.

However, in the framework of the QSSC, the dark matter need not be necessarily non-baryonic. It can be in the form of baryonic matter being the relic of very old stars of the previous cycles. A typical QSSC cycle has a lifetime long enough for most stars of masses exceeding ~ 0.5–$0.7\,M_\odot$ to have burnt out. Thus stars from previous cycles will be mostly extinct as radiators of energy. Their masses will continue however, to exert a gravitational influence on visible matter. The so-called dark matter seen in the outer reaches of galaxies and within clusters may very well be made up, at least in part, of these stellar remnants.

It may be timely to mention that the recent data on distant x-ray clusters obtained from XMM and Chandra projects indicate that the observed abundances of clusters at high redshift, taken at face value, give $0.9 < \Omega_{m0} < 1.07$ (at 1σ) (Blanchard 2005). This favours a matter-dominated model and is consistent with the value of Ω_{m0} in the QSSC estimated from the SNe Ia and CMB observations. However, it is hard to reconcile with the concordance model.

5. Conclusion

In order to explain the current observations in the framework of the standard cosmology, one has to trust a preposterous composition for the constituent of the Universe which defies any simple explanation, thereby posing probably the greatest challenge theoretical physics has ever faced. We think this is the right time to seriously consider alternative theories which present more natural explanations to the observed phenomena, especially when there is neither independent observational evidence for non-baryonic dark matter, dark energy and inflation, nor have they a firm basis in a well-established theory of particle physics. Furthermore, it is always necessary for healthy science to have an alternative model to the dominant paradigm.

Acknowledgement

RGV thanks the Abdus Salam ICTP, for providing hospitality and travel assistance (under his associateship programme), where a part of the work was done.

Appendix

The *Bayes factor* is a ratio of average likelihoods (rather than the maximum likelihoods used for model comparison in frequentist statistics) for two models M_i and M_j, and is given by

$$B_{ij} = \frac{\mathcal{L}(M_i)}{\mathcal{L}(M_j)} \equiv \frac{p(D|M_i)}{p(D|M_j)}, \tag{A.1}$$

where the likelihood for the model M_i, $\mathcal{L}(M_i)$ is the probability $p(D|M_i)$ to obtain the data D if the model M_i is the true one. For a model M_i with free parameter, say, α and β (generalization for the models with more parameters is straight forward), this probability is given by

$$\mathcal{L}(M_i) \equiv p(D|M_i) = \int \mathrm{d}\alpha \int \mathrm{d}\beta \; p(\alpha|M_i)p(\beta|M_i)\mathcal{L}_i(\alpha, \beta), \tag{A.2}$$

where $p(\alpha|M_i)$ and $p(\beta|M_i)$ are the prior probabilities for the parameters α and β respectively, assuming that the model M_i is true. $\mathcal{L}_i(\alpha, \beta)$ is the likelihood for α and β in the model M_i and is given by the usual χ^2-statistic:

$$\mathcal{L}_i(\alpha, \beta) = \exp\left[-\frac{\chi_i^2(\alpha, \beta)}{2}\right]. \tag{A.3}$$

For flat prior probabilities for the parameters α and β, i.e., assuming that we have no prior information regarding α and β except that they lie in some range $[\alpha, \alpha + \Delta\alpha]$ and $[\beta, \beta + \Delta\beta]$, we have $p(\alpha|M_i) = 1/\Delta\alpha$ and $p(\beta|M_i) = 1/\Delta\beta$. Hence the expression for the likelihood of the model M_i reduces to

$$\mathcal{L}(M_i) = \frac{1}{\Delta\alpha}\frac{1}{\Delta\beta}\int_{\alpha}^{\alpha+\Delta\alpha}\int_{\beta}^{\beta+\Delta\beta} \exp\left[-\frac{\chi_i^2(\alpha, \beta)}{2}\right]\mathrm{d}\beta \; \mathrm{d}\alpha. \tag{A.4}$$

The *Bayes factor* B_{ij}, given by (A.1), which measures the relative merits of model M_i over model M_j, is interpreted as follows (Drell *et al.* 2000; John & Narlikar 2002; and the references therein). If $1 < B_{ij} < 3$, there is an evidence against M_j when compared with M_i, but it is not worth more than a bare mention. If $3 < B_{ij} < 20$, the evidence against M_j is definite but not strong. For $20 < B_{ij} < 150$, this evidence is strong and for $B_{ij} > 150$, it is very strong.

References

Akerib, D. S. *et al.* 2004, *Phys. Rev. Lett.*, **93**, 211301.
Angloher, G. *et al.* 2005, *Astropart. Phys.* **23**, 325.
Bernabei, R. *et al.* 2003, *Riv. Nuovo Cim.*, **26**, 1.
Blanchard, A. 2005, astro-ph/0502220.
Bondi, H., Gold, T. 1948, *MNRAS*, **108**, 252.
Burbidge, E. M., Burbidge, G. R., Fowler, W. A., Hoyle, F. 1957, *Rev. Mod. Phys.*, **29**, 547.
Burbidge, G., Hoyle, F. 1998, *ApJ*, **509**, L1.
Caldwell, R. R. 2002, *Phys. Lett. B*, **545**, 23.
Carroll, S. M., Hoffman, M., Trodden, M. 2003, *Phys. Rev. D*, **68**, 023509.
Donn, B., Sears, G. W. 1963, *Science*, **140**, 1208.

Drell, P. S., Loredo, T. J., Wasserman, I. 2000, *ApJ*, **530**, 593.
Freedman, W. L., Turner, M. S. 2003, *Rev. Mod. Phys.*, **75**, 1433.
Gibbons, G. 2003, hep-th/0302199.
Hannestad, S. 2001, *Phys. Rev. D*, **63**, 043009.
Hoyle, F. 1948, *MNRAS*, **108**, 372.
Hoyle, F., Burbidge, G., Narlikar, J. V. 1993, *ApJ*, **410**, 437.
Hoyle, F., Burbidge, G., Narlikar, J. V. 1995, *Proc. Roy. Soc. A*, **448**, 191.
Hoyle, F., Burbidge, G., Narlikar, J. V. 2000, A Different Approach to Cosmology (Cambridge, Cambridge University Press); and the references therein.
Hoyle, F., Narlikar, J. V. 1964, *Proc. Roy. Soc. A*, **278**, 465.
Hoyle, F., Wickramasinghe, N. C. 1988, *Astrophys. Space Sci.*, **147**, 245.
John, M. V., Narlikar, J. V. 2002, *Phys. Rev. D*, **65**, 043506.
Kinney, W. H. 2001, *Phys. Rev. D*, **63**, 043001.
Nabarro, F. R. N., Jackson, P. J. 1958, In: *Growth and Perfection in Crystals* (eds) R. H. Duramus *et al.* (John Wiley: New York).
Narlikar, J. V., Burbidge, G., Vishwakarma, R. G. 2007, 'Cosmology and Cosmogony in a Cyclic Universe' (submitted to *J. Astrophys. Astr.*).
Narlikar, J. V., Padmanabhan, T. 1985, *Phys. Rev. D*, **32**, 1928.
Narlikar, J. V., Vishwakarma, R. G., Burbidge, G. 2002, *PASP*, **114**, 1092 (astro-ph/0205064).
Narlikar, J. V., Vishwakarma, R. G., Hajian, A., Souradeep, T., Burbidge, G., Hoyle, F. 2003, *ApJ*, **585**, 1 (astro-ph/0211036).
Narlikar, J. V., Wickramasinghe, N. C., Sachs, R., Hoyle, F. 1997, *Int. J. Mod. Phys. D*, **6**, 125.
Padmanabhan, T. 2005, In: *100 Years of Relativity – Spacetime Structure: Einstein and Beyond* (ed.) A. Ashtekar (World Scientific, Singapore), (gr-qc/0503107).
Perlmutter, S. *et al.* 1999, *ApJ*, **517**, 565.
Press, W. H., Teukolsky, S. A., Vetterling, W. T., Flannery, B. P. 1986, Numerical Recipes (Cambridge University Press).
Riess, A. G. *et al.* 2004, *ApJ*, **607**, 665.
Sachs, R., Narlikar, J. V., Hoyle, F. 1996, *A&A*, **313**, 703.
Sami, M., Toporensky, A. 2004, *Mod. Phys. Lett. A*, **19**, 1509.
Singh, P., Sami, M., Dadhich, N. 2003, *Phys. Rev. D*, **68**, 023522.
Spergel, D. N. *et al.* 2003, *ApJS*, **148**, 175.
Starkman, G. *et al.* 2004, *Phys. Rev. Lett.*, **93**, 221301.
Steinhardt, P. J., Turok, N. 2002, *Science*, **296**, 1436.
Vishwakarma, R. G. 2002a, *Class. Quant. Grav.*, **19**, 4747 (gr-qc/0205075).
Vishwakarma, R. G. 2002b, *MNRAS*, **331**, 776 (astro-ph/0108118).
Vishwakarma, R. G. 2003, *MNRAS*, **345**, 545 (astro-ph/0302357).
Vishwakarma, R. G. 2005, *MNRAS*, **361**, 1382 (astro-ph/0506217).
Vishwakarma, R. G., Narlikar, J. V. 2005, *Int. J. Mod. Phys. D*, **14**, 345 (astro-ph/0412048).
Wickramasinghe, N. C., Wallis, D. H. 1996, *Astrophys. Space Sci.*, **240**, 157.

J. Astrophys. Astr. (2007) **28,** 29–40

Kinematical Diagrams for Conical Relativistic Jets

Gopal-Krishna[1], Pronoy Sircar[2] & Samir Dhurde[3]
[1]*National Centre for Radio Astrophysics, Tata Institute of Fundamental Research, Pune University Campus, Post Bag No. 3, Pune 411 007, India.*
e-mail: krishna@ncra.tifr.res.in
[2]*Department of Physics, Indian Institute of Technology, Kanpur, India.*
e-mail: pronoys@iitk.ac.in
[3]*Inter-University Centre for Astronomy & Astrophysics, Pune University Campus, Post Bag No. 4, Pune 411 007, India.*
e-mail: samir@iucaa.ernet.in

Received 2006 August 7; accepted 2007 February 28

Abstract. We present diagrams depicting the expected inter-dependences of two key kinematical parameters of radio knots in the parsec-scale jets of blazars, deduced from VLBI observations. The two parameters are the apparent speed ($v_{\rm app} = c\beta_{\rm app}$) and the *effective* Doppler boosting factor ($\delta_{\rm eff}$) of the relativistically moving radio knot. A novel aspect of these analytical computations of β–δ diagrams is that they are made for parsec-scale jets having a *conical* shape, with modest opening angles (ω up to 10°), in accord with the VLBI observations of the nuclei of the nearest radio galaxies. Another motivating factor is the recent finding that consideration of a conical geometry can have important implications for the interpretation of a variety of radio observations of blazar jets. In addition to uniform jet flows (i.e., those having a uniform bulk Lorentz factor, Γ), computational results are also presented for stratified jets where an ultra-relativistic central spine along the jet axis is surrounded by a slower moving sheath, possibly arising from a velocity shear.

Key words. Blazars: general—galaxies: active, quasars, jets, nuclei, radio continuum.

1. Introduction

Ejection of a pair of jets of relativistic material is the key manifestation of the radio-loud phase of Active Galactic Nuclei (AGN) and micro-quasars. On parsec scale, two basic attributes of relativistic jets of AGN, when directed appropriately close towards the observer, are: **(i)** an apparent superluminal motion and **(ii)** a strong Doppler boosting of the radiation received. Attempts to use these effects for probing the physics of AGN have been reviewed extensively in the literature (e.g., Begelman *et al.* 1984; Phinney 1985; Barthel 1994; Urry & Padovani 1995; Gopal-Krishna 1995; Scheuer 1996; Zensus 1997; Livio 1997; Ghisellini 2004; Giovannini 2004; Wiita 2006; Collin-Souffrin 2006). Until recently, quantitative discussion of both these properties was

limited essentially to "collimated" jets (i.e., jets whose full opening angle, $\omega = 0$). However, recent VLBI observations provide good evidence for conical shape of AGN jets, with typical opening angle of several degrees on parsec scale. Salient examples are the nuclear radio jets of the nearest radio galaxies M87 (Biretta *et al.* 2002; Ly *et al.* 2004; Dodson *et al.* 2006) and Centaurus A (Horiuchi *et al.* 2006). Additional evidence comes from the studies by Tavecchio *et al.* (2004) and Jorstad *et al.* (2005). Conical geometry on parsec scale is also expected in certain theoretical models of jet formation (e.g., Meier *et al.* 2001). Thus, a fresh evaluation of the jet kinematics, taking into account the conical geometry of the jets, is warranted. This would also have a direct bearing on the schemes to relate the VLBI radio observations of blazar nuclei to their observed properties in the x-ray and γ-ray bands (e.g., Tavecchio *et al.* 2000; Celotti 2001; Kellermann *et al.* 2004) and to the energetics of the jet flow on parsec-scale (e.g., Sikora *et al.* 2005; Henri & Saugé 2006).

It has become evident from recent studies that considering a conical geometry of jets, even with modest opening angles, can have a major influence on the interpretation of several observed kinematical properties of the radio knots in the parsec-scale AGN jets, which usually remain largely unresolved in the VLBI observations (see Gopal-Krishna *et al.* 2004, 2006, hereafter referred to as Paper I and Paper II). This is particularly so when the jet's bulk speed is extremely relativistic ($\Gamma = 30$–100). Such extremely high bulk Lorentz factors of jets have in fact been inferred in a number of analyses of the rapid flux variability of blazars in the TeV range (e.g., Krawczynski *et al.* 2001; Konopelko *et al.* 2003; Henri & Saugé 2006) and at centimetre wavelengths (e.g., Qian *et al.* 1991; Wagner & Witzel 1995; Blandford 2001; Rickett *et al.* 2002; Macquart & de Bryun 2006). For the extensively observed jet of the quasar 3C273, the jet's bulk Lorentz factor, required to explain the radio and optical data is $\Gamma =$ 50–100 (Jester *et al.* 2006). Additional evidence for such ultra-relativistic jets comes from the detection of extremely superluminal radio knots from VLBI observations of several blazars, suggesting Γ to be above 20 to 60, or even more (e.g., Fujisawa *et al.* 1999; Jorstad *et al.* 2004, 2005; Piner & Edwards 2005). Note that as in Paper I & II, the conical jet flow has been assumed here to be *ballistic*, as also considered in many other studies (e.g., Falcke & Biermann 1996; Lobanov & Zensus 2001; Kaiser 2006).

Briefly, the main conclusions resulting from the consideration of modest opening angles of AGN jets ($\omega \sim 5$–10°, Papers I & II) are:

- The relatively slow (often subluminal) apparent motions of the VLBI radio knots of TeV blazars (e.g., Marscher & Marchenko 1999; Piner *et al.* 1999; Piner & Edwards 2004; Giroletti *et al.* 2004) can be reconciled with the ultra-relativistic jet speeds ($\Gamma = 30$–100) inferred from the flux variations mentioned above, without recourse to postulating phenomena like "pattern speed" (e.g., Lind & Blandford 1985; Vermeulen & Cohen 1994), or a spine-sheath jet flow (section 2).
- The viewing angle of an ultra-relativistic jet from the line-of-sight can be substantially larger than that inferred typically by combining the flux variability and proper motion measurements of the VLBI radio knots. This would normally imply not only an increased probability of the jet being observed, but also a smaller (and hence less uncomfortable) correction for foreshortening of the jet due to projection (e.g., see Schilizzi & de Bryun 1983).

2. Kinematical diagrams for uniform Lorentz factor jets

Very likely, the radio knots are manifestations of the relativistic shocks excited across the jet flow, e.g., due to instabilities in the flow speed, or some external perturbations (e.g., Marscher 1980; Laing 1980; Marscher & Gear 1985; Hughes *et al.* 1985; Laing 1993; Dermer & Chiang 1998; Spada *et al.* 2001). A useful conventional way to encapsulate the kinematical behaviour of the radio knots occuring in a fully collimated ($\omega = 0^\circ$) relativistic jet is to display the Doppler factor (δ) against the apparent speed ($v_{\rm app} = c\beta_{\rm app}$), for a range of bulk Lorentz factors, Γ, and viewing angle, θ, of the jet (e.g., Nesci *et al.* 2005). Here we present numerical computations of the kinematics of the radio knots occuring in *conical* relativistic jets; elsewhere we have discussed in detail some implications of these results concerning the effect of the *conical* jet geometry on various estimated jet parameters (Gopal-Krishna *et al.* 2007). In accordance with the canonical picture, we assume the knots to coincide with the (relatively thin) transverse shocks in the jet flow, such that each knot is taken to have the shape of a circular disk of uniform intrinsic emissivity. Further, each surface element of the knot is assumed to move ballistically, with the bulk speed of the jet, thus maintaining the *conical* shape of the jet. The two sets of computations reported here correspond to the cases of (i) a uniform Γ across the radio knot (approximated with a transverse shock disc) and (ii) a radially decreasing Γ from the centre of the radio knot (i.e., the jet axis). This latter version approximates a transverse shock formed in a jet described by the currently popular "fast spine + slow sheath" (i.e., stratified) jet configuration of AGN jets, as considered by numerous authors (e.g., Baan 1980; Bicknell 1985; Sol *et al.* 1989; Pelletier & Roland 1989; Komissarov 1990; Laing 1993; Begelman *et al.* 1994; Swain *et al.* 1998; Hardcastle *et al.* 1999; Attridge *et al.* 1999; Chiaberge *et al.* 2000; Giovannini *et al.* 2001; Meier 2003; Giroletti *et al.* 2004; Ghisellini *et al.* 2005). Clearly, due to velocity shear, such knots would soon begin to deviate from a laminar flow and would therefore be more prone to distortion/disruption. However, consideration of such stages is beyond the scope of the present exploratory study.

Details of the analytical prescription employed here are available in Papers I & II. The basic equations for a discrete component moving in a jet with a bulk speed v directed at an angle θ from the line of sight, are:

$$\delta = \frac{1}{\Gamma[1 - \beta\cos(\theta)]} \tag{1}$$

where,

$$\beta = \frac{v}{c} \quad \text{and} \quad \Gamma = (1 - \beta^2)^{-1/2}, \tag{2}$$

$$\beta_{\rm app} = \frac{\beta\sin(\theta)}{1 - \beta\cos(\theta)}, \tag{3}$$

$$S_{\rm obs} = \int_\Omega \delta^p(\Omega')S_{\rm em}(\Omega')d\Omega \equiv A(\theta)S_{\rm em}, \tag{4}$$

$$\delta_{\rm eff} = A^{1/p}(\theta), \tag{5}$$

$$\vec{\beta}_{\rm app,eff} = \frac{1}{S_{\rm obs}}\int_\Omega \vec{\beta}(\Omega')\delta^p(\Omega')S_{\rm em}(\Omega')d\Omega. \tag{6}$$

Here, $S_{\rm em}$ and $S_{\rm obs}$ are the emitted and the (Doppler boosted) observed flux density and $A(\theta)$ is the flux boosting factor averaged over the (circular) radio knot. The viewing angle, θ of the jet measures the angular offset of the radio disc's center (i.e., the jet axis) from the direction of the AGN core. The parameter $p = n - \alpha$, where α is the spectral index (flux density $\propto$ frequency$^{\alpha}$). The term n is equal to 2 if the emission arises from a continuous jet and $n = 3$ if discrete knots are the dominant emitters (e.g., Scheuer & Readhead 1979; Phinney 1985; Lind & Blandford 1985; also, Ryle & Longair 1966). For the present purpose of compact radio sources, we shall adopt $\alpha = 0$.

The numerical integration was carried out by dividing the (circular) radio knot of diameter ω into small pixels, each being 1% of the disc's diameter. For each pixel (i, j), angular offset from the AGN core was computed and combined with the jet's bulk Lorentz factor (see below), to find: the Doppler factor, $\delta_{i,j}$, the vector $\vec{\beta}_{{\rm app}(i,j)}$; the flux boosting factor, $A_{i,j}$, taken to be $\delta_{i,j}^{p}$; and the product of the last two terms (which amounts to apparent velocity of the pixel, weighted by its apparent flux). Finally, the average values of each of these parameters computed for all the pixels across the radio disc were determined (e.g., $A_{\rm ave}$) and these are taken to be the *effective* values for the entire radio knot. Note that the effective δ for the knot is then $\delta_{\rm eff} = A_{\rm ave}^{1/p}$, where we have taken two values for p (= 2 and 3), widely employed in the literature for compact radio sources (see above). Further details of the method are available in Papers I and II.

2.1 *Kinematical diagrams for the 'stratified' (spine-sheath) jets*

To quantify the *spine-sheath* type stratification in the jet, we approximate the cross-section of the jet flow with a radially decreasing bulk Lorentz factor, $\Gamma(r)$, according to an exponential law with e-folding length equal to $\omega/2q$ [$\Rightarrow \Gamma(r) = \Gamma_{(r=0)} e^{(-2rq/\omega)}$], where r is the transverse angular separation from the centre of the radio disc/knot (i.e., from the jet axis). Two representative values for q (= 1 and 2) were chosen. (The choice of minimum $\Gamma_{(r=0)}$ is constrained by the requirement that the corresponding Γ at the jet's surface ($r' = \omega/2$) does not drop below the physical limit of $\Gamma = 1$.) Note that while the choice of an exponential form for $\Gamma(r)$ is admittedly arbitrary, it should be reasonable enough for the illustrative calculations presented here, particularly since no specific form of $\Gamma(r)$ has yet been established.

3. Results

The first set of kinematical diagrams, shown in the left column of Figs. 1 & 2 display the *effective* values of δ versus $\beta_{\rm app}$ of a VLBI radio knot, computed for a range of Γ and θ and taking three representative values of ω (= 1°, 5° and 10°) and two values of p (= 2 and 3), as mentioned above. The plots in the right column have been extracted from the corresponding plots in the left column, by zooming on the limited (but more commonly used) range of Γ up to 20. Note that for all these plots, Γ was assumed to remain constant across the radio emitting disc/knot (i.e., $q = 0$ case). In Figs. 3–6 we show the corresponding diagrams for the cases of transverse Γ gradient (i.e., Γ decreasing away from the jet axis), taking two values for the (exponential) scale length ($q = 1$ and 2) (section 2.1). All these β–δ diagrams provide a generalization to *conical* jets and spine-sheath geometries of the simple kinematical diagrams presented

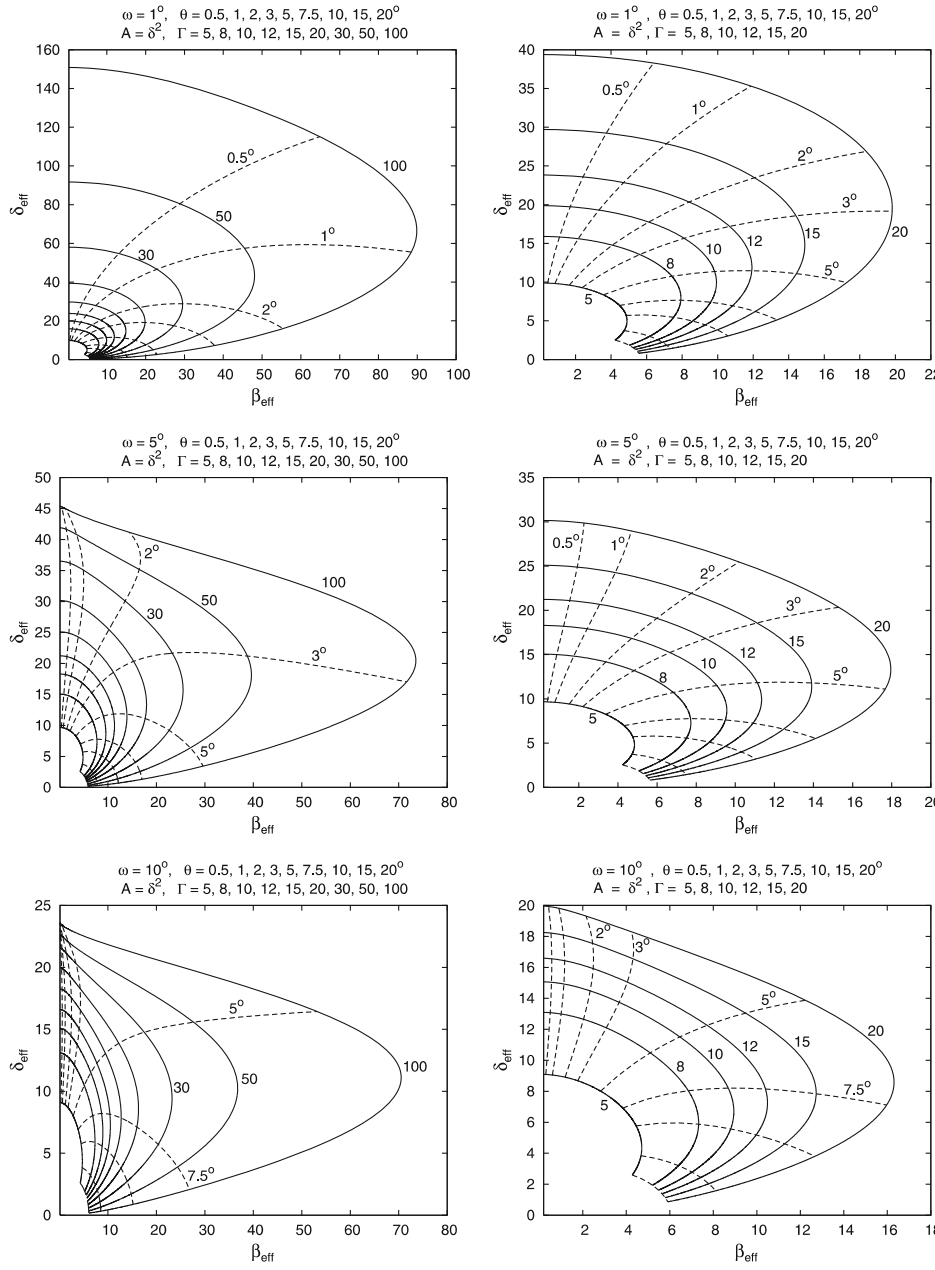

Figure 1. The curves show for a knot in a relativistic jet the computed *effective* values of Doppler factor (δ) and the apparent speed (in units of c), $\beta_{\rm app}$. Each solid curve is drawn for fixed bulk Lorentz factor (Γ) of the jet, whose value is shown besides the curve (when possible) and also at the top of each panel. The broken curves are for different values of the jet viewing angle (θ), as shown at the top of each panel. The panels on the right side show, for clarity, the zoomed versions of the inner parts of the corresponding panels to the left (section 3). All these panels refer to a Doppler boosting index $p = 2$ and a quotient $q = 0$ for the transverse gradient of the jet's bulk Lorentz factor (i.e., a uniform Γ case; section 2). The full opening angle of the jet, $\omega = 1^\circ$ for the top panels, 5° for the middle panels and 10° for the lower panels, as shown.

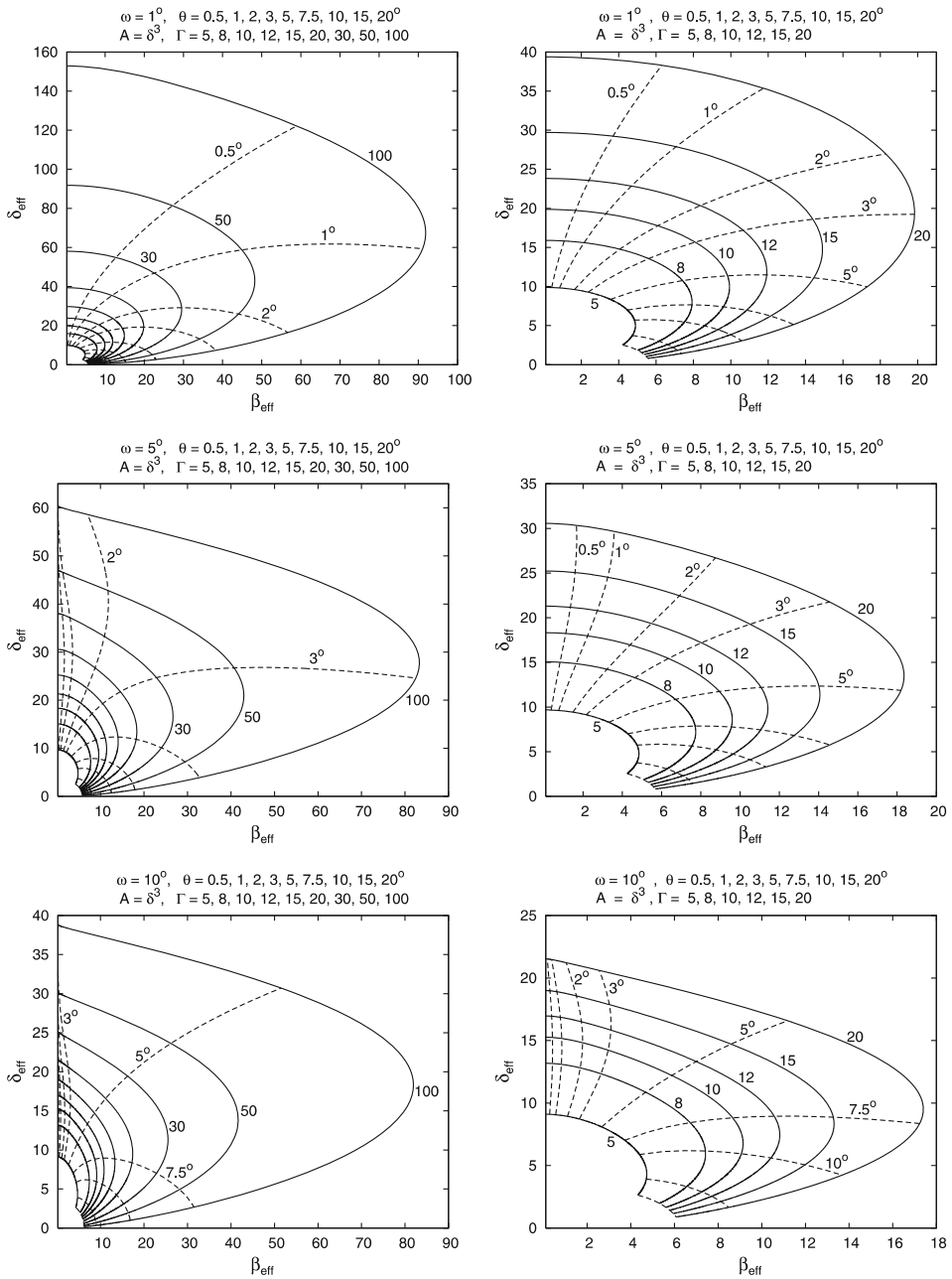

Figure 2. Same as Fig. 1, except that $p = 3$ and $q = 0$ (see the top of each panel).

by Nesci *et al.* (2005) for the case of fully collimated jets ($\omega = 0$). Hence, these generalized β–δ diagrams are likely to prove useful in quantitative interpretation of the VLBI and related multiband observations of blazars (see also Gopal-Krishna *et al.* 2007).

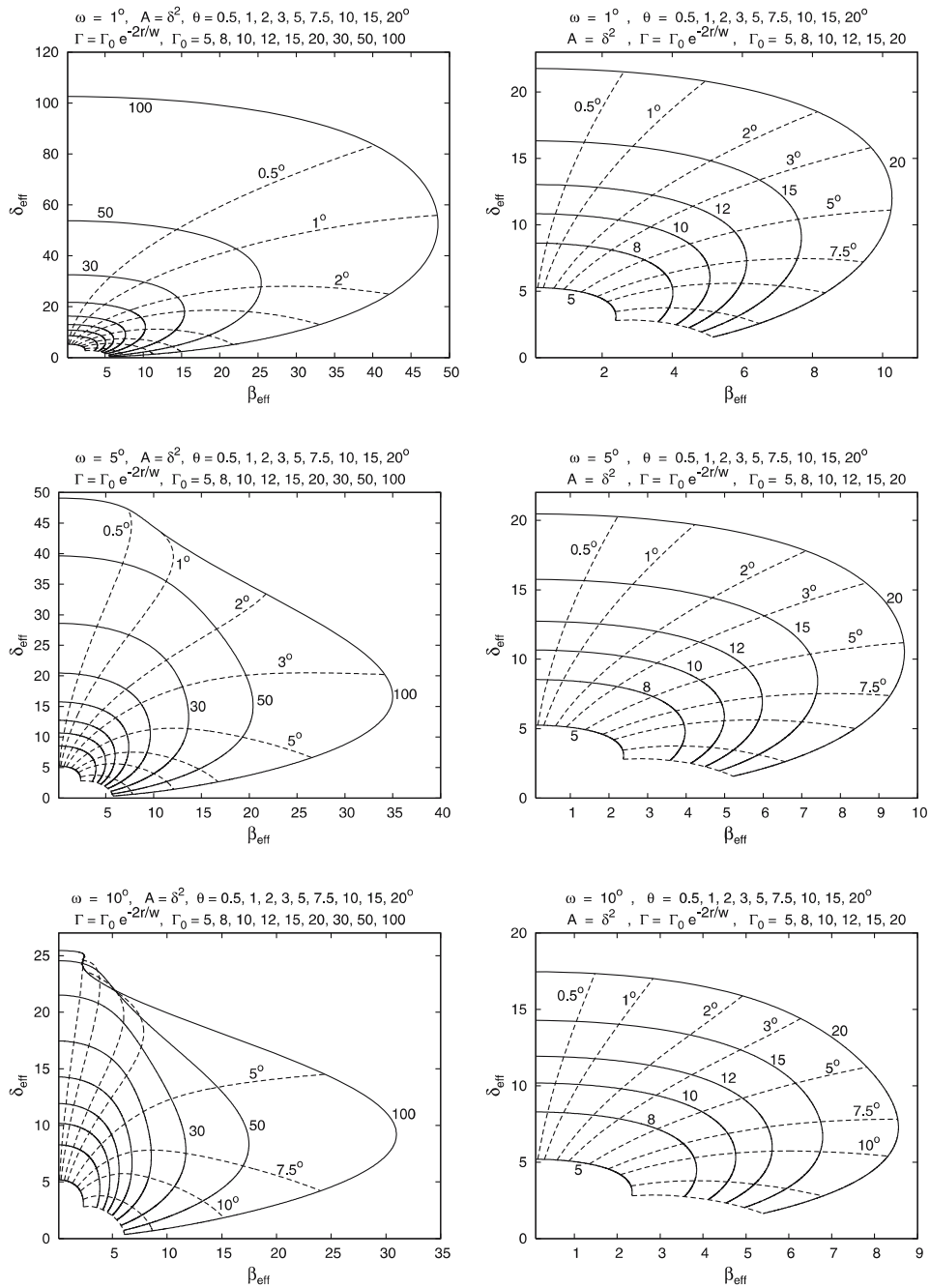

Figure 3. Same as Fig. 1, except that $p = 2$ and $q = 1$ (see the top of each panel). This is the case of a mildly stratified jet flow.

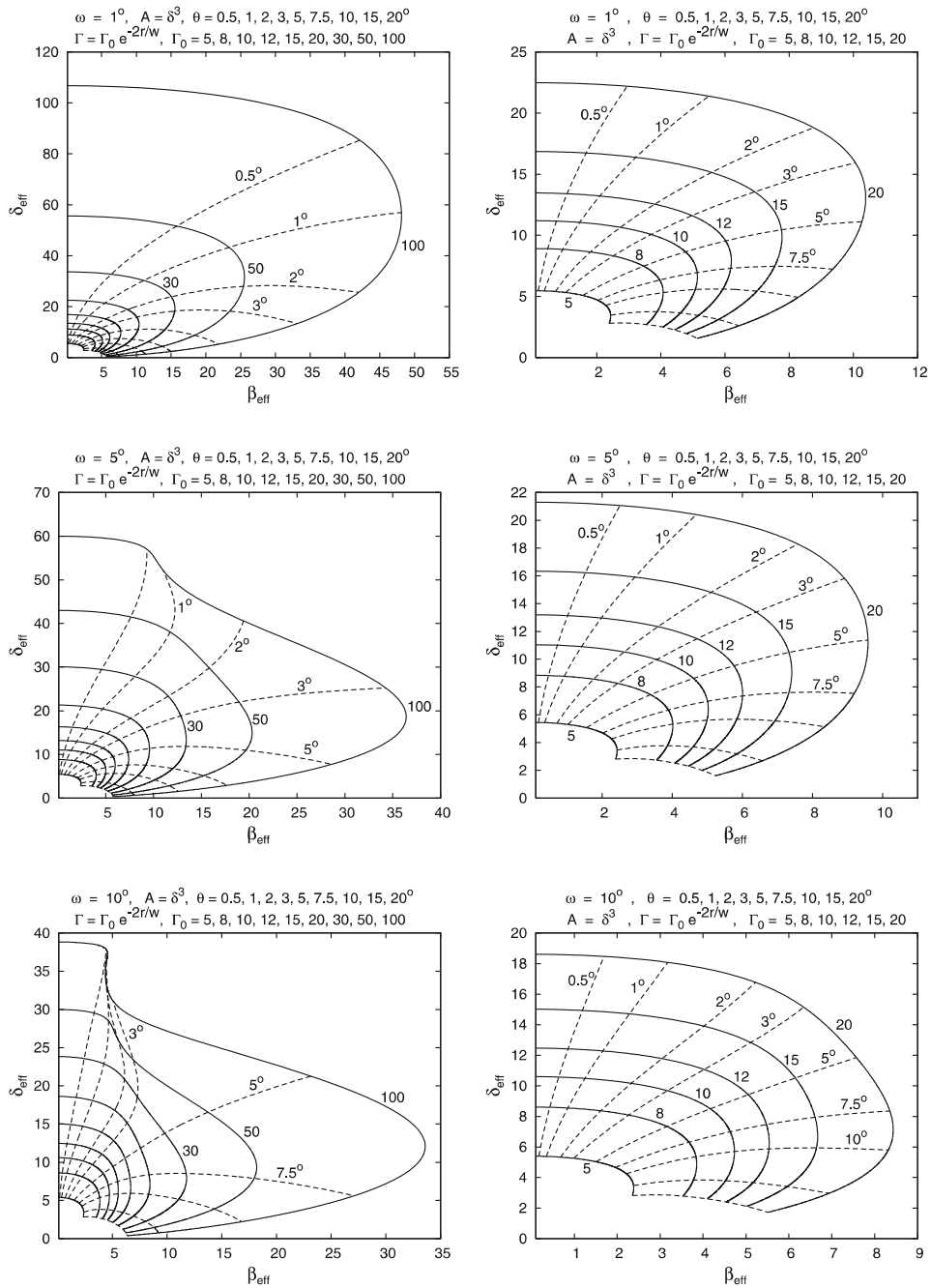

Figure 4. Same as Fig. 1, except that $p = 3$ and $q = 1$ (see the top of each panel). This is the case of a mildly stratified jet flow.

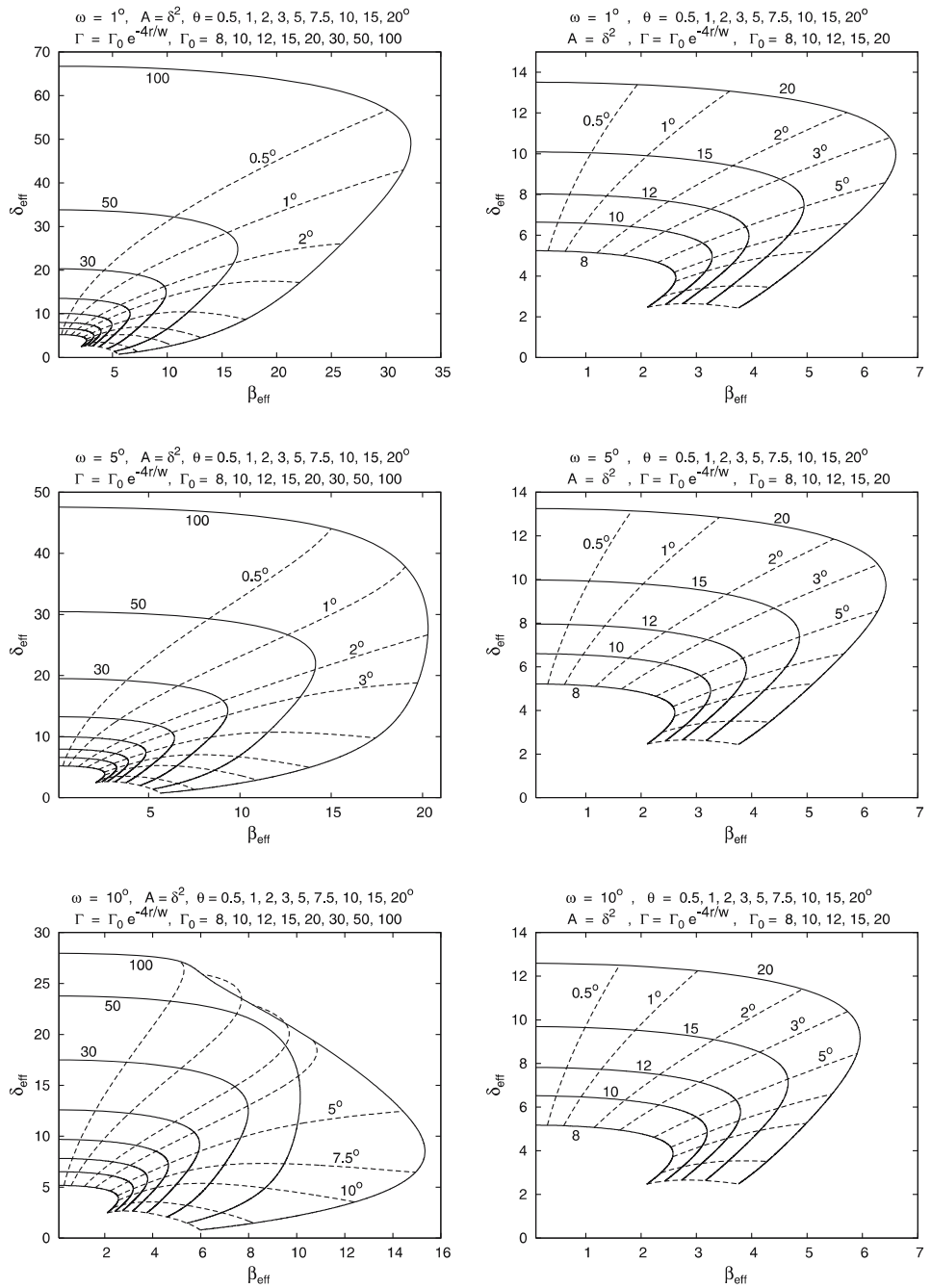

Figure 5. Same as Fig. 1, except that $p = 2$ and $q = 2$ (see the top of each panel). This is the case of a sharply stratified jet flow.

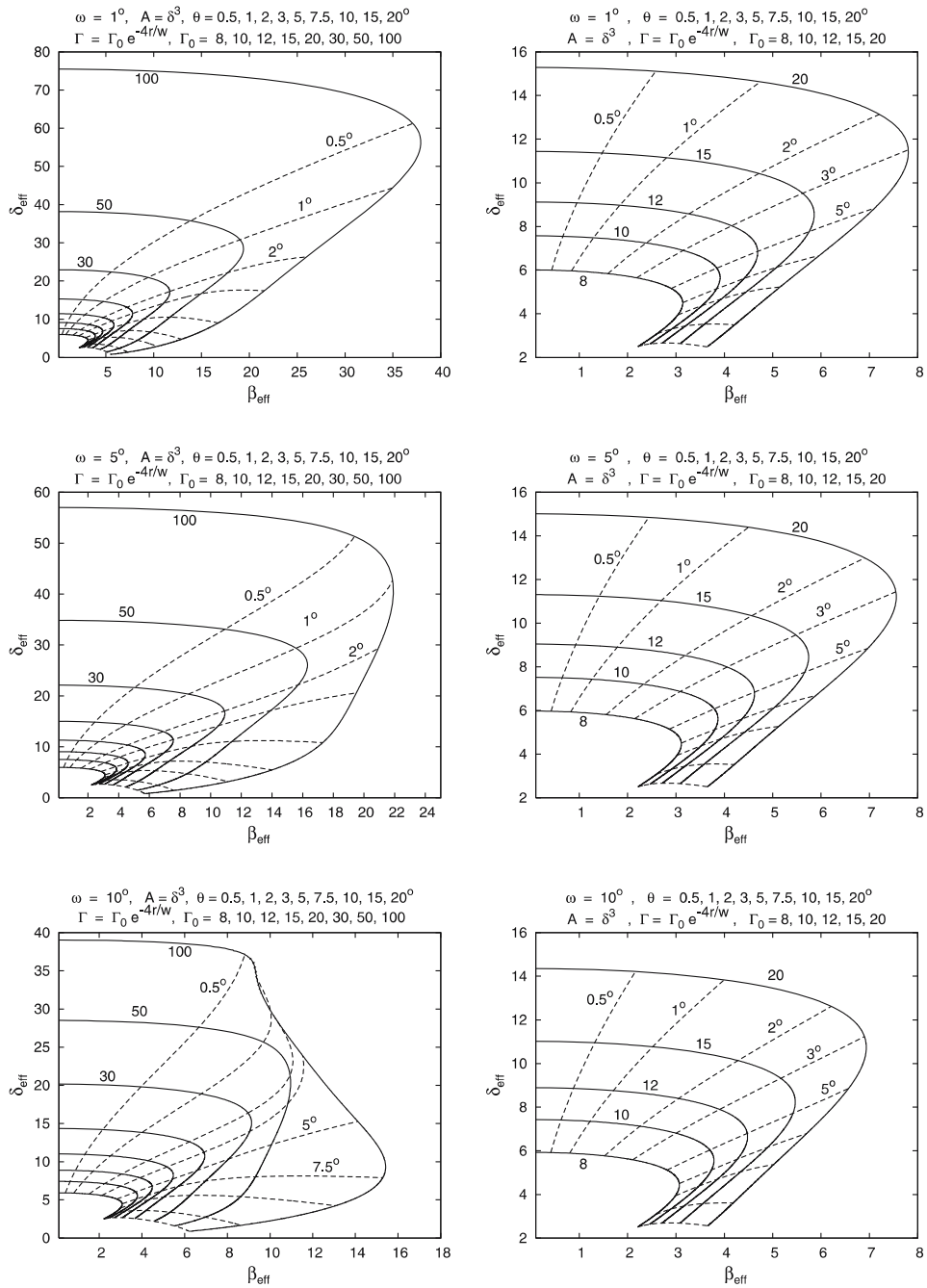

Figure 6. Same as Fig. 1, except that $p = 3$ and $q = 2$ (see the top of each panel). This is the case of a sharply stratified jet flow.

Acknowledgements

We thank the referee Prof. Ken Kellermann for his helpful advice on the manuscript and also Prof. Paul J. Wiita for the discussions. P. S. also thanks the Indian Academy of Sciences, Bangalore for a summer student fellowship and the National Centre for Radio Astrophysics (NCRA-TIFR), Pune for the facilities provided for his summer project, during which this work was carried out.

References

Attridge, J. M., Roberts, D. H., Wardle, J. F. C. 1999, *ApJ*, **518**, L87.
Baan, W. A. 1980, *ApJ*, **239**, 433B.
Barthel, P. D. 1994, *PASP*, **54**, 175.
Begelman, M. C., Blandford, R. D., Rees, M. J. 1984, *Reviews of Modern Physics*, **56**, 255.
Begelman, M. C., Rees, M. J., Sikora, M. 1994, *ApJ*, **429**, L57.
Bicknell, G. V. 1985, *PASAu*, **6**, 130.
Biretta, J., Junor, W., Livio, M. 2002, *New Astr. Rev.*, **46**, 239.
Blandford, R. D. 2001, *PASA*, **224**, 499.
Celotti, A. 2001, *ASPC*, **250**, 93.
Chiaberge, M., Celotti, A., Capetti, A., Ghisellini, G. 2000, *A&A*, **358**, 104.
Collin-Souffrin, S. 2006, *AIPC*, **861**, 587.
Dermer, C. D., Chiang, J. 1998, *New Astr.*, **3**, 157.
Dodson, C. D. 2006, *PASJ*, **58**, 243.
Dodson, R., Edwards, P. G., Hirabayashi, H. 2006, *PASJ*, **58**, 243.
Falcke, H., Biermann, P. L. 1996, *A&A*, **308**, 371.
Fujisawa, K., Kobayashi, H., Wajima, K., Hirabayashi, H., Kameno, S., Inoue, M. 1999, *PASJ*, **51**, 537.
Ghisellini, G. 2004, *New Astr. Rev.*, **48**, 375.
Ghisellini, G., Tavecchio, F., Chiaberge, M. 2005, *A&A*, **432**, 401.
Giovannini, G. 2004, *Ap&SS*, **293**, 1.
Giovannini, G., Cotton, W. D., Feretti, L., Lara, L., Venturi, T. 2001, *ApJ*, **552**, 508.
Giroletti, M. *et al.* 2004, *ApJ*, **600**, 127.
Gopal-Krishna 1995, *Proc. Natl. Acad. Sci.*, **92**, 11399.
Gopal-Krishna, Dhurde, S., Wiita, P. J. 2004, *ApJ*, **615**, L81 (Paper I).
Gopal-Krishna, Wiita, P. J., Dhurde, S. 2006, *MNRAS*, **369**, 1287 (Paper II).
Gopal-Krishna, Dhurde, S., Sircar, P., Wiita, P. J. 2007, *MNRAS*, in press.
Hardcastle, M. J., Alexander, P., Pooley, G. G., Riley, J. M. 1999, *MNRAS*, **304**, 135.
Henri, G., Saugé 2006, *ApJ*, **640**, 185.
Horiuchi, S., Meier, D. L., Preston, R. A., Tingay, S. J. 2006, *PASJ*, **58**, 211.
Hughes, P. A., Aller, H. D., Aller, M. F. 1985, *ApJ*, **298**, 296.
Jester, S., Harris, D. E., Marshall, H. L., Meisenheimer, K. 2006 *ApJ*, **648**, 900.
Jorstad, S. G., Marscher, A. P., Lister, M. L., Stirling, A. M., Cawthorne, T. V., Gómez, J.-L., Gear, W. K. 2004, *AJ*, **127**, 3115.
Jorstad, S. G. *et al.* 2005, *AJ*, **130**, 1418.
Kaiser, C. R. 2006, *MNRAS*, **367**, 1083.
Kellermann, K. I. *et al.* 2004, *ApJ*, **609**, 539.
Komissarov, S. S. 1990, *Soviet Astron. Lett.*, **16**, 284.
Konopelko, A. K., Mastichiadis, A., Kirk, J. G., de Jager, O. C., Stecker, F. W. 2003, *ApJ*, **559**, 851.
Krawczynski, H. *et al.* 2001, *ApJ*, **559**, 187.
Laing, R. A. 1980, *MNRAS*, **193**, 439.
Laing, R. A. 1993, "Radio observations of jets: large scales" In: *Space Telescope Sci. Inst. Symp. 6: Astrophysical Jets* (eds) Burgarella D., O'Dea C. (Cambridge: CUP), 95.
Lind, K. R., Blandford, R. D. 1985, *ApJ*, **295**, 358.
Livio, M. 1997, *ASPC*, **121**, 845.
Lobanov, A. P., Zensus, J. A. 2001, *Science*, **294**, 128.

Ly, C., Walker, R. C., Wrobel, J. M. 2004, *AJ*, **127**, 119.
Macquart, J.-P., de Bruyn A. G. 2006, *A&A*, **446**, 185.
Marscher, A. P. 1980, *ApJ*, **235**, 386.
Marscher, A. P., Marchenko, S. 1999, *ASPC*, **159**, 417.
Marscher, A. P., Gear, W. K. 1985, *ApJ*, **298**, 114.
Meier, D. L. 2003, *New Astr. Rev.*, **47**, 667.
Meier, D. L., Koide, S., Uchida, Y. 2001, *Science*, **291**, 84.
Nesci, R., Massaro, E., Rossi, C., Sclavi, S., Maesano, M., Montagni, F. 2005, *AJ*, **130**, 1466.
Phinney, S. 1985, In: *"Astrophysics of Active Galaxies and Quasi-stellar Objects", Mill Valley: University Science Books* (ed.) Miller J. S. (Oxford University Press) 453.
Piner, B. G., Unwin, S. C., Wehrle, A. E., Edwards, P. G., Fey, A., Kingham, K. A. 1999, *ApJ*, **525**, 176.
Piner, B. G., Edwards, P. G. 2004, *ApJ*, **600**, 115.
Piner, B. G., Edwards, P. G. 2005, *ApJ*, **622**, 168.
Pellettier, G., Roland, J. 1989, *A&A*, **224**, 24.
Qian, S. J., Quirrenbach, A., Witzel, A., Krichbaum, T. P., Hummet, C. A., Zensus, J. A. 1991, *A&A*, **241**, 15.
Rickett, B. J., Kedziora-Chudczer, L., Jauncey, D. L. 2002, *ApJ*, **581**, 103.
Ryle, M., Longair, M. S. 1966, *MNRAS*, **136**, 123.
Scheuer P. A. G. 1996, *ASPC*, **100**, 333.
Scheuer P. A. G., Readhead, A. C. S. 1979, *Nature*, **277**, 182.
Schilizzi, R. T., de Bruyn, A. G. 1983, *Nature*, **303**, 26.
Sikora, M., Begelman, M. C., Madejski, G. M., Lasota, J.-P. 2005, *ApJ*, **625**, 72.
Sol, H., Pelletier, G., Asseo, E. 1989, *MNRAS*, **237**, 411.
Spada, M., Ghisellini, G., Lazzati, D., Celotti, A. 2001, *MNRAS*, **325**, 1559.
Swain, M. R., Bridle, A. H., Baum, S. A. 1998, *ApJ*, **507**, L29.
Tavecchio, F., Maraschi, L., Sambruna, R. M., Urry, C. M. 2000, *ApJ*, **544**, 23.
Tavecchio, F., Maraschi, L., Sambruna, R. M., Urry, C. M., Cheung, C. C., Gambill, J. K., Scarpa, R. 2004, *ApJ*, **614**, 64.
Urry, C. M., Padovani, P. 1995, *PASP*, **107**, 803.
Vermeulen, R. C., Cohen, M. H. 1994, *ApJ*, **430**, 467.
Wagner, S., Witzel, A. 1995, *ARA&A*, **33**, 163.
Wiita, P. J. 2006, *J. Korean Physical Soc.*, **49**, 1753.
Zensus, J. A. 1997, *ARA&A*, **35**, 607.

J. Astrophys. Astr. (2007) **28,** 41–53

GMRT and VLA Observations at 49 cm and 20 cm of the HII Region near $l = 24.8°$, $b = 0.1°$

N. G. Kantharia[1], W. M. Goss[2], D. Anish Roshi[3], Niruj R. Mohan[4] & Francois Viallefond[5]

[1]*National Centre for Radio Astrophysics, Tata Institute of Fundemental Research, Post Bag 3, Ganeshkhind, Pune 411 007, India.*
e-mail: ngk@ncra.tifr.res.in

[2]*National Radio Astronomy Observatory, P. O. Box 0, Socorro, NM 87801, USA.*
e-mail: mgoss@aoc.nrao.edu

[3] *Raman Research Institute, Sadashivnagar, Bangalore 560 080, India.*
e-mail: anish@rri.res.in

[4] *Sterrewacht, Leiden, The Netherlands.*
e-mail: mohan@strw.leidenuniv.nl

[5]*Observatoire de Paris, Paris, France.*
e-mail: fviallef@maat.obspm.fr

Received 2006 September 5; accepted 2007 March 8

Abstract. We report multi-frequency radio continuum and hydrogen radio recombination line observations of HII regions near $l = 24.8°$, $b = 0.1°$ using the Giant Metrewave Radio Telescope (GMRT) at 1.28 GHz ($n = 172$), 0.61 GHz ($n = 220$) and the Very Large Array (VLA) at 1.42 GHz ($n = 166$). The region consists of a large number of resolved HII regions and a few compact HII regions as seen in our continuum maps, many of which have associated infrared (IR) point sources. The largest HII region at $l = 24.83°$ and $b = 0.1°$ is a few arcmins in size and has a shell-type morphology. It is a massive HII region enclosing $\sim 550\,M_\odot$ with a linear size of 7 pc and an rms electron density of $\sim 110\,cm^{-3}$ at a kinematic distance of 6 kpc. The required ionization can be provided by a single star of spectral type O5.5.

We also report detection of hydrogen recombination lines from the HII region at $l = 24.83°$ and $b = 0.1°$ at all observed frequencies near $V_{lsr} = 100\,km\,s^{-1}$. We model the observed integrated line flux density as arising in the diffuse HII region and find that the best fitting model has an electron density comparable to that derived from the continuum. We also report detection of hydrogen recombination lines from two other HII regions in the field.

Key words. Interstellar medium—radio recombination lines—HII regions—envelopes.

1. Introduction

Since the discovery of radio recombination lines (RRLs) (Dravskikh & Dravskikh 1964; Sorochenko & Borozich 1964; Palmer *et al.* 1967) these have been widely used as diagnostics of ionized media in the Galaxy. Theoretical studies of line formation have shown that the width of RRLs depends sensitively on the principal quantum number ($\Delta\nu \propto n^{4.4}$; $\Delta\nu$ in Hz) and is also proportional to the electron density (Shaver 1975). Thus low-frequency (< a few GHz) RRLs from dense ionized regions (> a few times 10^2 cm^{-3}) will be broadened, resulting in the reduction of the peak line intensity. This reduction in peak intensity makes it difficult to detect RRLs at low frequencies from high density gas. On the other hand, low frequency RRLs from low density ionized regions are not affected by this limitation. Hence these RRLs form an ideal probe to study lower density regions such as photodissociation regions, diffuse, extended HII regions (e.g., Lockman 1989; Lockman *et al.* 1996), envelopes of HII regions and the extended low density warm ionized medium (ELDWIM) (Anantharamaiah 1985; Heiles *et al.* 1996; Roshi & Anantharamaiah 2001). Typically such regions have densities $\leq 100\,\mathrm{cm}^{-3}$ and emission measures $\leq 10^5$ pc cm^{-6}.

In this paper, we present Giant Metrewave Radio Telescope (GMRT) and Very Large Array (VLA) observations of the HII region complex near $l = 24.8^\circ$ and $b = 0.1^\circ$ in RRLs and radio continuum at frequencies of 0.61, 1.28 and 1.42 GHz. This region consists of HII regions in various evolutionary stages, ranging from protostellar stage showing bipolar outflows detected in CO (Furuya *et al.* 2002) to the diffuse, extended HII region like G24.83 + 0.10. The primary objective of these observations was to image and determine the physical properties of the diffuse HII region from which low frequency hydrogen and carbon RRLs near 327 MHz have been observed (Anantharamaiah 1985). The details of the observations are given in section 2. The radio continuum morphology and physical properties of the diffuse HII regions derived from the continuum are described in section 3. In section 4, we present the RRL data. We end with a summary of the results. A near kinematic distance of 6 kpc was determined for G24.83 + 0.10 and is used in this paper.

2. Observations

Table 1 gives the details of the observations. The 0.61 and 1.28 GHz observations were carried out with the GMRT and the 1.42 GHz observations were carried out with the VLA.

2.1 *GMRT observations: 0.61 and 1.28 GHz*

The GMRT (Swarup *et al.* 1991; Ananthakrishnan & Rao 2002) consists of 30–45 m diameter dishes spread over a 25 km region and operating in five frequency bands. For our 1.28 GHz observations, we selected a 2 MHz bandwidth which includes both the H172α and the C172α lines. We followed an observing sequence of 20 min on-source followed by 6 min on the phase calibrator. The bandpass calibrator was observed thrice during the run for 30 min each. After editing corrupted data, we were left with 23 antennas on 13 April 2002 and 17 antennas on 14 April 2002, which were used in the subsequent analysis.

Table 1. Observation details.

Parameter	0.610 GHz	1.28 GHz	1.42 GHz
Telescope	GMRT	GMRT	VLA (C, DnC)
Field of view	$\sim 44'$	$\sim 25'$	$\sim 40'$
Phase centre			
α_{2000}	$18^h36^m11^s$	Same	Same
δ_{2000}	$-07^\circ10'40''$	Same	Same
Date of observation	17/1/2003	13/4/2002 14/4/2002	5/10/2002, 13/12/2002, 20/1/2003
Transitions	H220α C220α	H172α C172α	H166α C166α
Rest freq			
Hydrogen (GHz)	0.613405	1.28117	1.42473
Carbon (GHz)	0.613711	1.28181	1.42544
Phase cal	1822-096	1822-096	1831-126
Flux cal	3C286,3C48	3C286,3C48	1923+210
Bandpass cal	1822-096	3C286,3C48	3C48
Bandwidth (km s^{-1})	480	480	330
Channel width (km s^{-1})	3.8	3.7	2.6
Resolution continuum	$14'' \times 7''$	$8'' \times 6''$	$35'' \times 25''$
PA	-24°	38°	40°
Line	$22'' \times 22''$	$23'' \times 12''$	$35'' \times 25''$
PA	-57°	42°	40°
RMS noise			
Contm (mJy/beam)	2	2	4
Line (mJy/beam)	3	2	3

At 0.61 GHz, we selected a bandwidth of 1 MHz which includes both the hydrogen (H220α) and carbon (C220α) lines. The observing procedure was similar to that used at 1.28 GHz.

The GMRT data were converted to FITS format and imported into the NRAO AIPS package for further analysis. Single channel data on the bandpass calibrators were first examined, edited and gain calibrated. Bandpass solutions were then generated on these calibrators, which were applied to the data on all objects before averaging the line free channels to generate the continuum database. A first order spectral baseline was removed from the calibrated UV line data using the task UVLIN. The final continuum and line images were generated using IMAGR. The estimated flux density scale has errors $\leq 20\%$.

At the time of our observations, noise switching to measure the system temperature was not available. We estimated the difference in system temperature between the primary calibrator and the target source positions using the 0.408 GHz continuum map (Haslam 1982) and assuming a spectral index for the galactic background emission of -2.7. These correction factors were then applied to the data.

2.2 *VLA observations: 1.42 GHz*

The VLA observations were conducted with a bandwidth of 1.5 MHz which contained the observing frequencies for both the H166α and the C166α lines toward

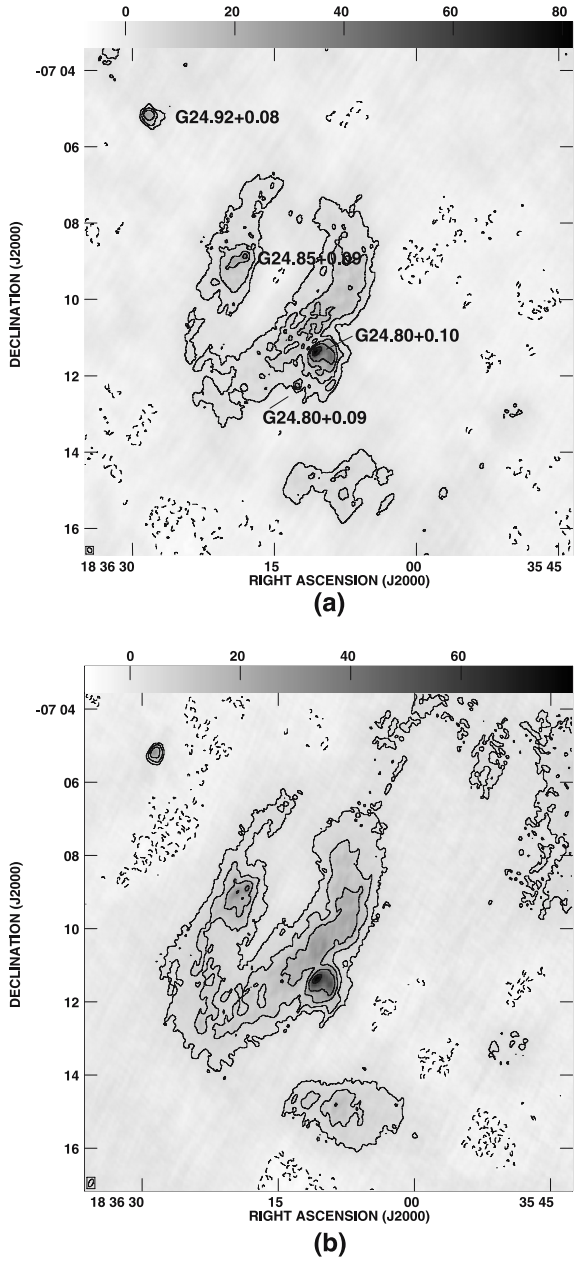

Figure 1. **(a)** GMRT 1.28 GHz continuum image of the region near $l = 24.8^\circ$, $b = 0.1^\circ$ with an angular resolution of $8.4'' \times 6.3''$, PA $= 38^\circ$. The RMS noise in the image is 1.5 mJy/beam. The grey scale ranges from -7 mJy/beam to 82 mJy/beam. The contour levels are -7.5, -3.75, 3.75, 7.5, 15, 30, 60, 120, 240, 480 mJy/beam. The compact sources detected in the field are labelled. **(b)** GMRT 0.61 GHz continuum image of the region near $l = 24.8^\circ$, $b = 0.1^\circ$. The angular resolution of the image is $13.8'' \times 6.7''$, PA $= -24^\circ$ and the RMS noise is 1.6 mJy/beam. The grey scale ranges from -8 mJy/beam to 80 mJy/beam. The contour levels are -8, -4, 4, 8, 16, 32, 64, 128, 256, 512 mJy/beam.

Table 2. Continuum flux densities and derived spectral indices using 0.61 and 1.42 GHz data of the diffuse HII regions marked in Fig. 2. Since some of the regions lie outside the 1.28 GHz primary beam, no flux density at this frequency is available.

Source	ν GHz	Flux density[1] Jy	$\alpha_{1.42\,\mathrm{GHz}}^{0.61\,\mathrm{GHz}}$	IRAS Pt src
G24.83 + 0.10	0.61	8.4	$-0.1^{-0.2}_{+0.3}$	IRAS 18335-0713A
	1.28	5.3		IRAS 18335-0711
	1.42	7.7		
	4.87[2]	7.1		
G24.74 + 0.08	0.61	1.29	$-0.6^{-0.2}_{+0.3}$	-
	1.28	0.65		
	1.42	0.80		
G24.74 + 0.16	0.61	0.21	$-0.4^{-0.2}_{+0.3}$	IRAS 18331-0715
	1.42	0.15		
G24.68 − 0.16	0.61	2.6	$-0.15^{-0.2}_{+0.25}$	IRAS 18341-0727
	1.42	2.3		
G24.71 − 0.13	0.61	3.9	$-0.3^{-0.2}_{+0.25}$	IRAS 18340-0724
	1.42	3.0		
G24.74 − 0.2	0.61	1.4	$-0.85^{-0.15}_{+0.2}$	IRAS 18344-0725
	1.42	0.68		

[1] The error in flux density at 0.61 GHz and 1.28 GHz is $\leq 20\%$. The error bars on the spectral index are estimated using a conservative 20% error on the 0.61 GHz data points.
[2] Data from Downes *et al.* (1980).

G24.83 + 0.10. Observing details are listed in Table 1. Data were obtained in the C and DnC configurations. The observations on the target source were interspersed with 5 min runs on the phase calibrator. The bandpass calibrator was observed at the beginning and end of the run for about 10 min. Data from all the three short runs were combined to obtain the final dataset. These data were online Hanning smoothed. The VLA data were analysed using standard tasks in AIPS. The continuum image was made by using the central 75% of the total number of channels. Gain and bandpass calibration were then carried out and applied to continuum and line data. The continuum emission was removed from the line UV data using UVLSF in AIPS. These data were then used to generate a cube using IMAGR. For further analysis, moment maps were also generated.

3. The continuum emission

Figures 1 and 2 show the continuum images. The 1.28 (Fig. 1a) and 0.61 (Figs. 1b and 2b) GHz continuum images were obtained after applying a UV taper of 30 kλ. At 1.28 GHz, the largest angular scale to which the GMRT is sensitive is $\sim 7'$, thus

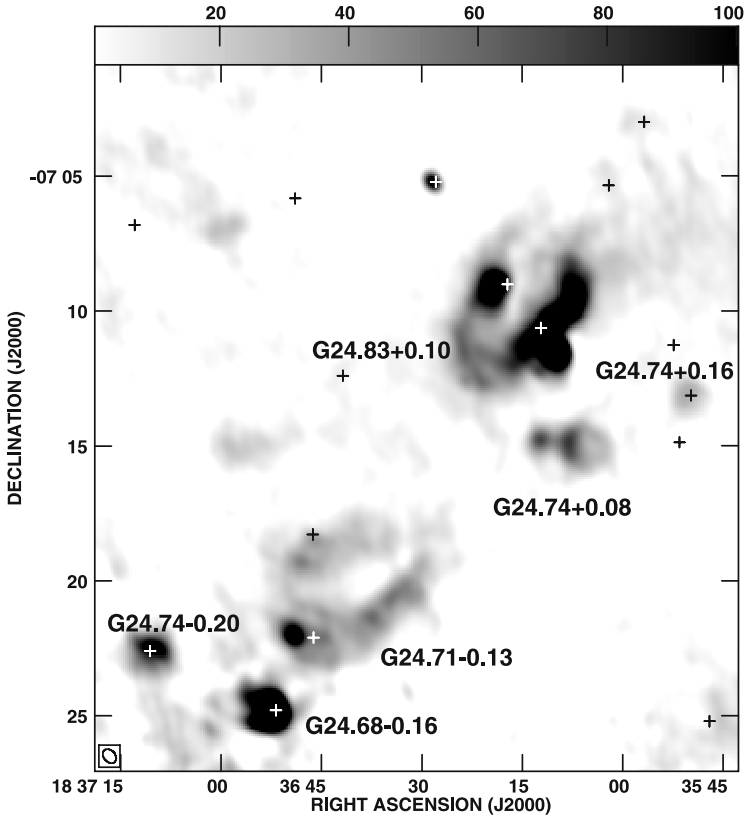

Figure 2(a). VLA radio continuum image at 1.42 GHz of the field near $l = 24.8^\circ$, $b = 0.1^\circ$. The image has been smoothed to a resolution of $35'' \times 25''$, PA $= 40^\circ$. The grey scale is plotted from 1 mJy/beam to 100 mJy/beam. The image has been corrected for primary beam attenuation. The diffuse HII regions (angular size $\geq 1'$) in the field are labelled. Crosses mark positions of IRAS point sources in the region obtained from the IRAS PSC.

the images are likely to be missing flux density for the large scale components. We, therefore, use the 0.61 and 1.42 GHz data to derive the spectral index listed in Table 2. The compact sources are marked in the 1.28 GHz image (see Fig. 1a). Resolved sources within the field of view of the observations are marked in the 1.42 GHz image shown in Fig. 2(a). Table 2 lists the estimated flux densities of these diffuse sources after correcting for the primary beam attenuation. IRAS point sources are marked by crosses in Fig. 2.

Infrared sources are associated with most of the diffuse regions shown in Fig. 2(a). To ascertain that these are not chance superpositions, we examined the 60 μm flux density of these sources as given in the IRAS PSC and these are labelled in Fig. 2. Codella *et al.* (1994) have shown that IRAS PSC sources which have $S_{60\,\mu\mathrm{m}} > 100$ Jy have more than 80% probability of being associated with HII regions and that for sources with $S_{60\,\mu\mathrm{m}} > 300$ Jy, chance superpositions are expected to be very rare. The $S_{60\,\mu\mathrm{m}}$ for all the sources are plotted in Fig. 2(b). As seen here, almost all the sources lying on the diffuse sources have a flux density > 100 Jy at 60 μm and most of them have flux densities > 300 Jy totally ruling out chance superpositions. The IRAS sources

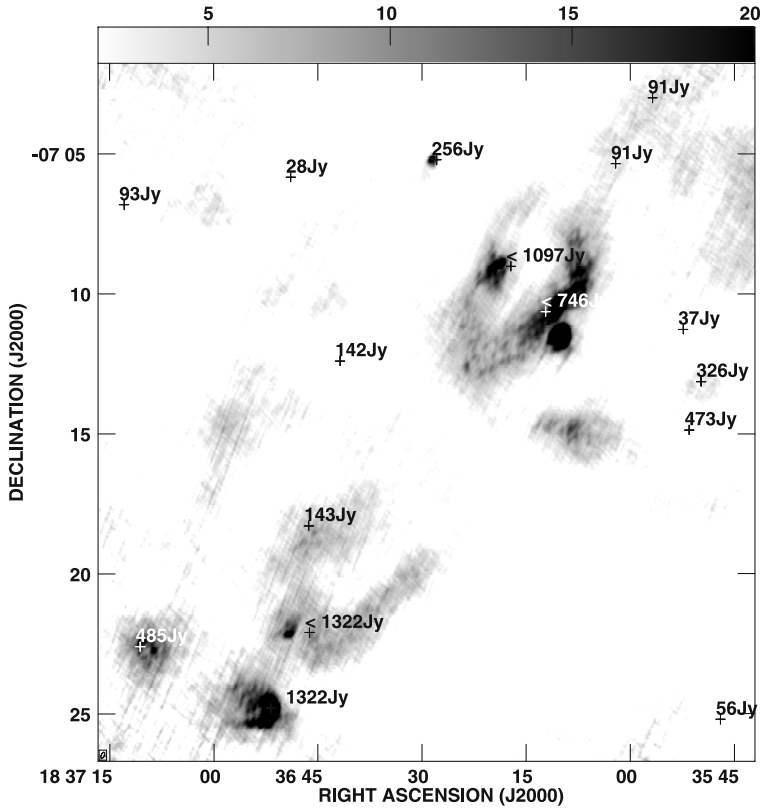

Figure 2(b). GMRT image of the same field as in (a) at 0.61 GHz. The image has an angular resolution of $14'' \times 7''$, PA $= -24^\circ$. The grey scale is plotted from 2 mJy/beam to 20 mJy/beam. The image has been corrected for primary beam attenuation. The crosses mark the positions of sources from the IRAS PSC. The strength of these sources at 60 μm are marked.

seen in rest of the field have lower flux densities at 60 μm and are hence unlikely to be associated with HII regions.

In the rest of this section, we discuss the continuum emission from these diffuse sources alongwith the associated IRAS point sources. The radio continuum peaks in the diffuse regions are most likely indicating the positions of the exciting stars.

G24.83+0.10: The brightest diffuse source in the field of view is G24.83+0.10. The integrated flux density of this source at 1.42 GHz is 7.7 Jy. Higher resolution images of this source are shown in Fig. 1. The diffuse source has a shell-like morphology with multiple radio peaks. The size of the shell is about 7 pc. The spectral index of this HII region is consistent with optically thin free-free emission. Two IRAS point sources are coincident with this object as shown in Fig. 2. The IR emission arises from the thermal dust located within the OB associations heated by the stars (Conti & Crowther 2004). At higher angular resolutions and frequencies, several compact/ultracompact HII regions are detected in G24.80 + 0.09. The ultracompact HII regions in G24.8 + 0.09 have been extensively studied at higher frequencies with high angular resolution (for example, Furuya *et al.* 2002). These HII regions are found to be in an earlier stage of evolution (Furuya *et al.* 2002). The UV photons from the stars exciting the UC HII regions

Table 3. The physical parameters derived from the measured continuum flux density of the source at 1.42 GHz are listed below. We assumed an electron temperature of 7000 K for the HII region and a spherical geometry. The exciting star is tabulated from Panagia (1973). All the regions are taken to be at a kinematic distance of 6 kpc.

Source	$\langle N_e \rangle$ cm^{-3}	d pc	$\langle EM \rangle$ 10^4 $pccm^{-6}$	Mass $M_\odot$	N_{Lyc} 10^{48} sec^{-1}	Exc star
G24.83 + 0.10	110	7	9.0	550	23	O5.5
G24.74 + 0.08	90	4	3.5	70	2	O8
G24.74 + 0.16	75	3	1.5	15	0.4	B0
G24.68 − 0.16	160	4	9.5	120	7	O6.5
G24.71 − 0.13	70	7	3.5	340	9	O6
G24.74 − 0.20	160	3	6.5	35	2	O8

might be absorbed by dust resulting in photon-bound UC HII regions. The other peak in the W arc, i.e., G24.80 + 0.10 appears to be a cometary type HII region with a tail extending to the N along the diffuse arc. The IRAS point source IRAS 18335-0713A is located in the W arc of G24.83 + 0.10, close to G24.80 + 0.10 and has a flux density of < 746 Jy at 60 μm. G24.80 + 0.10 (which appears as a single radio peak in the VLA 20 cm image (see Fig. 2a)) is resolved into a double-peaked structure in the higher resolution GMRT images at 20 cm and 49 cm. Another compact HII region G24.85 + 0.09 (see Fig. 1a) is located in the E arc of this diffuse HII region. This region also seems to display a cometary structure. The IRAS point source IRAS 18335-0711 with a flux density of < 1097 Jy at 60 μm is associated with this region.

We have used the VLA data at 1.42 GHz to estimate the average physical parameters of the diffuse HII region. The RMS electron density, emission measure, mass of the ionized gas and the number of Lyman continuum photons required to maintain ionization are listed in Table 3 (using Mezger and Henderson 1967). The exciting star types, assuming a single ionizing star are listed in the last column of Table 3 (Panagia 1973). G24.83 + 0.10 is the most massive HII region. In the next section, we discuss the observed recombination line emission from this region.

G24.74+0.08: This diffuse region located due S of G24.83 + 0.10 (see Fig. 2) does not have an associated IRAS point source. However, strong emission in the mid-infrared band E (MSX data) is observed, indicating its thermal nature and the predominance of hot dust. Moreover, MIR band A (MSX data), dominated by emission from polycylic aromatic hydrocarbons (PAH), is also detected from this region. The physical properties of this HII region estimated from its continuum strength are listed in Table 3.

G24.71−0.13: This is another shell-like HII region. The size of the shell is about 7 pc similar to that of G24.83 + 0.10. G24.71 − 0.13 (see Fig. 2) has a radio continuum peak in the middle of the two arcs forming the shell with diffuse emission along the rest of the shell. The IRAS point source IRAS 18340-0724 of strength < 1322 Jy at 60 μm is located close to the radio peak and is associated with the HII region. The properties of G24.71 − 0.13 derived from the continuum observations are listed in Table 3.

Other continuum sources: The HII regions G24.74 + 0.16, G24.74 − 0.2 and G24.68 − 0.16 are relatively compact with an angular extent less than ∼ 2'. All three have associated IRAS point sources. G24.68 − 0.16 is the brightest among these four

sources (2.3 Jy at 1.42 GHz). IRAS 18341-0727 with a flux density of 1322 Jy at 60 μm is associated with it. The emission is consistent with a free-free spectrum. This source appears to have a shell-like morphology with the shell opening to the E. G24.74 + 0.16 located to the W of G24.83 + 0.10 is the weakest continuum source in the field. The IRAS point source IRAS 18331-0715 of strength 326 Jy at 60 μm is associated with this HII region. The derived physical properties of these HII regions are also listed in Table 3. An unresolved HII region, G24.92 + 0.08 (see Fig. 1a) is located to the NE of G24.83 + 0.10. This HII region has an associated IRAS point source IRAS 18337-0707 with a strength of 256 Jy at 60 μm.

4. The recombination line emission

Hydrogen RRLs from G24.83 + 0.10 were detected at all three observed frequencies. Figure 3 shows the spectra integrated over the diffuse emission at frequencies 1.42 ($n = 166$), 1.28 ($n = 172$) and 0.61 ($n = 220$) GHz. The feature detected at ~ 100 km s^{-1} is the hydrogen line. Gaussians were fitted to the hydrogen features at the observed frequencies and the results are shown in Table 4. We did not detect carbon RRL from G24.83 + 0.10 at any of the observed frequencies. If the carbon lines were formed in an associated PDR with $T_e \leq 400$ K, $n_e > 3$ cm^{-3} and pathlengths ≥ 11 pc, we should have been able to detect these near 1 GHz, with the sensitivity of our present observations.

The integrated line intensity map of the H166α emission (in grey) is shown in Fig. 4 along with the radio continuum (in contours) at 1.42 GHz. Line emission from G24.83+0.10 follows the continuum morphology. Strongest line emission is observed near the compact objects G24.80 + 0.10 and G24.85 + 0.09. Line emission is also observed from the diffuse source (see Fig. 4). The line emission arising in G24.83+0.10 is observed at velocities ranging from ~ 100 to 110 km s^{-1}. No systematic gradient in the velocity is observed. Higher velocities $\sim 115 - 125$ km s^{-1} are observed for the RRL arising near the UCHII G24.80 + 0.09. From their high resolution CO observations, Furuya *et al.* (2002) find that the LSR velocity of the molecular cloud is about 111 km s^{-1} with velocity in the outflow ranging from 90 to 130 km s^{-1}. Thus, the RRL we detect near this UCHII region is most likely related to this star forming region. Interestingly, the lines appear to be narrow with widths at half maximum of $12 - 24$ km s^{-1}. The narrow lines seem to arise close to the HII region G24.80 + 0.10 whereas the widest lines are observed near the UC HII region G24.85 + 0.09.

Weak H RRL emission from diffuse HII regions G24.74 + 0.08 and G24.74 + 0.16 is also detected (see Fig. 4) at 102 km s^{-1} and 106 km s^{-1} respectively.

Due to the low signal-to-noise ratio for the H220α and H172α data, we were unable to obtain a detailed distribution of emission across G24.80 + 0.01. Thus, the integrated line flux densities of G24.83 + 0.10 at the three observed frequencies were used to model the line emission. The departure coefficients required in the models were calculated using the computer code of Salem & Brocklehurst (1979), modified by Walmsley & Watson (1982) and Payne *et al.* (1994). These were then used to calculate the expected line flux density at the different quantum numbers following Shaver (1975).

The best fitting model that we found had the following parameters, typical of a diffuse HII region: $T_e = 7000$ K and $n_e = 100$ cm^{-3} and a line of sight extent of 11 pc. The background radiation field for all the models was assumed to be a HII region with

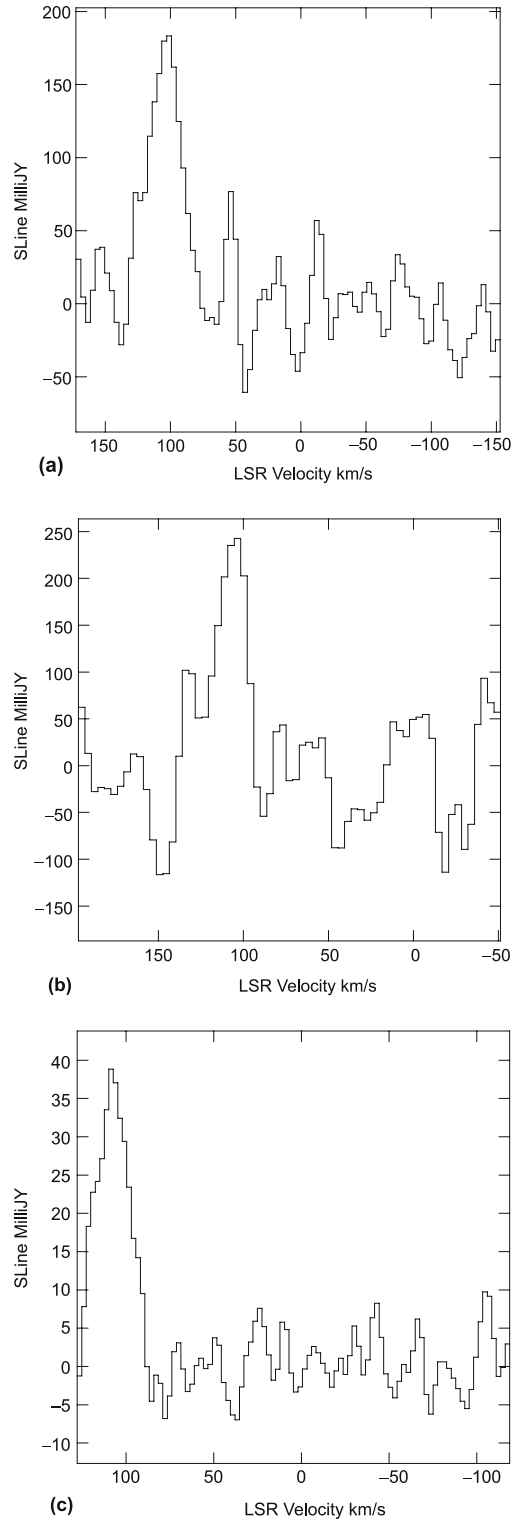

Figure 3.

Table 4. Gaussian fits to the source-integrated line emission from the sources in the G24 region. The 1σ errors on the fitted parameters are given.

Source	Line	S_{int} mJy	V_{LSR} km s^{-1}	FWHM km s^{-1}	Reference
G24.83 + 0.10	H166α	153 ± 15	109 ± 4	20 ± 8	A
	H172α	179 ± 13	104.5 ± 1	30 ± 1.5	A
	H220α	250 ± 39	107 ± 1.5	19 ± 2	A
	H87α[1]	370	108.6 ± 0.4	26.9 ± 0.8	B
	H110α[2]	164	107	23	C
	H272α[3]	235	83 ± 2	77 ± 4	D
	C272α[3]		46 ± 2	34 ± 5	D

[1] Integrated over 3′; [2] Integrated over 2.6′; [3] Integrated over $2^\circ \times 6'$.
A: Present work; B: Lockman (1989); C: Downes *et al.* (1980); D: Anantharamaiah (1985).

an emission measure of 5×10^4 pc cm^{-6}. However we note that this model overpredicts the line flux near 0.61 GHz.

Detailed modelling of the observed RRLs should include contribution from both the diffuse HII region and the compact HII regions observed embedded within the diffuse shell. With the signal to noise of the current data, such a detailed model is not warranted. High resolution sensitive low frequency RRL observations are required to model different regions of G24.80 + 0.10 independently.

5. Summary

Using the GMRT and VLA, we have carried out low frequency interferometric continuum and recombination line observations of the HII regions near $l = 24.8^\circ$, $b = 0.1^\circ$ at 0.61 GHz, 1.28 GHz and 1.42 GHz. We detect continuum emission from ultra-compact, compact and extended HII regions. In this paper, we discuss the morphology and average properties of the HII regions using our continuum observations. G24.83 + 0.10, the main region under study in this paper, is diffuse and massive encompassing $\sim 550\,M_\odot$ within a linear size of 7 pc with an rms electron density of $\sim 100\,cm^{-3}$. The source has a shell-like morphology with the shell opening to the N. Two IRAS point sources are clearly associated with this HII region. Compact HII regions are embedded in this diffuse region.

We detected H220α, H172α and H166α lines from G24.83 + 0.10. We also detected RRLs from two other HII regions in the field near 100 km s^{-1}. The model which

Figure 3 Caption

Figure 3. **(a)** The GMRT 172α spectrum integrated over G24.83 + 0.10. The spectrum has been Hanning smoothed to a velocity resolution of 7.2 km s^{-1}. **(b)** The GMRT 220α spectrum integrated over the continuum source G24.83 + 0.10. The spectrum has been Hanning smoothed to a velocity resolution of 7.6 km s^{-1}. **(c)** The VLA 166α spectrum integrated over a small region in G24.83 + 0.10 near G24.80 + 0.10 located in the W arc. The spectrum has been Hanning smoothed to a spectral resolution of 5.2 km s^{-1}. The feature near 100 km s^{-1} in all the three figures are the H RRLs.

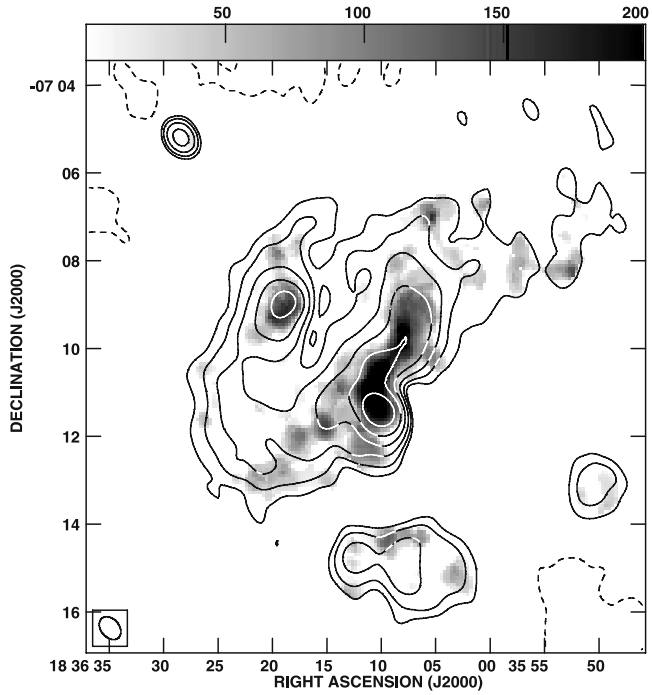

Figure 4. VLA moment 0 image of H166α (grey scale) of the HII region G24.83 + 0.10 superposed on a contour map of the continuum emission from the region at 1.42 GHz. The image has a resolution of 35″ × 25″ PA= 40°.

provides the best fit to the source-integrated RRL data towards G24.83 + 0.1 has the following parameters: $T_e = 7000$ K, $n_e \sim 100\,\mathrm{cm}^{-3}$ and path length of 11 pc. However this model overestimates the line strength at 0.61 GHz. We note that the electron density of the diffuse HII region that best fits the observed line data is similar to the rms electron density found from the continuum observations. More sensitive RRL observations of G24.83+0.10 are required to develop a detailed model of emission for this HII region.

Acknowledgments

We thank the anonymous referee for several useful comments which have especially helped to improve the IR part of the paper. We thank the staff of GMRT who made these observations possible. GMRT is run by the National Centre for Radio Astrophysics of the Tata Institute of Fundamental Research. The National Radio Astronomy Observatory is a facility of the National Science Foundation operated under cooperative agreement by Associated Universities, Inc. N. G. K. thanks Prasad Subramanian for useful comments on the manuscript. The IRAS point source catalog was obtained from VizieR.

References

Anantharamaiah, K. R. 1985, *J. Astrophys. Astr.*, **6**, 177.
Ananthakrishnan, S., Rao, A. P. 2002, In: *Multicolour Universe* (eds) Manchanda, R. K., Paul, B., p. 233.

Codella, C., Felli, M., Natale, V. 1994, *Astron. Astrophys.*, **284**, 233.
Conti, P., Crowther, P. A. 2004, *Mon. Not. R. Astron. Soc.*, **355**, 899.
Downes, D., Wilson, T. L., Bieging, J., Wink, J. 1980, *Astron. Astrophys. Suppl.*, **40**, 379.
Dravskikh, Z. V., Dravskikh, A. F. 1964, *Astron. Tsirk*, **158**, 2.
Furuya, R. S., Cesaroni, R., Codella, C. *et al.* 2002, *Astron. Astrophys.*, **390**, L1.
Haslam, C. G. T., Salter, C. J., Stoffel, H., Wilson, W. E. 1982, *Astron. Astrophys. Suppl.*, **47**, 1.
Heiles, C., Reach, W. T., Koo, B.-C. 1996, *Astrophys. J.*, **466**, 191.
Lockman, F. J. 1989, *Astrophys. J.*, **71**, 469.
Lockman, F. J., Pisano, D. J., Howard G. J. 1996, *Astrophys. J.*, **472**, 173.
Mezger, P. G., Henderson, A. P. 1967, *Astrophys. J.*, **147**, 471.
Palmer, P., Zuckermann, B., Penfield, H., Lilley, A. F., Mezger, P. G. 1967, *Astron. J.*, **72**, 821.
Panagia, N. 1973, *Astron. J.*, **78**, 929.
Payne, H. E., Anantharamaiah, K. R., Erickson, W. C. 1994, *Astrophys. J.*, **430**, 690.
Roshi, D. A., Anantharamaiah, K. R. 2001, *Astrophys. J.*, **557**, 226.
Salem, M., Brocklehurst, M. 1979, *Astrophys. J. Suppl.*, **39**, 633.
Shaver, P. A. 1975, *Pramana*, **5**, 1.
Sorochenko, R. L., Borozich, E. V. 1964, *Rep. Sov. Acad. Sci.*, **162**, 603.
Swarup, G., Ananthakrishnan, S., Kapahi, V. K. *et al.* 1991, *Curr. Sci.*, **60**, 95.
Walmsley, C. M., Watson, W. D. 1982, *Astrophys. J.*, **260**, 317.

J. Astrophys. Astr. (2007) **28,** 55–66

Evolution of the Distribution of Neutron Exposures in the Galaxy Disc: An Analytical Model

Wenyuan Cui[1,2,3], Weijuan Zhang[1] & Bo Zhang[1,2,*]
[1]*Department of Physics, Hebei Normal University, Shijiazhuang 050 016, P. R. China.*
[2]*National Astronomical Observatories, Chinese Academy of Sciences, Beijing 100 012, P. R. China.*
[3]*Graduate School of Chinese Academy of Sciences, Beijing 100 049, P. R. China.*
[*]*e-mail: zhangbo@hebtu.edu.cn*

Received 2006 January 5; accepted 2007 February 6

Abstract. In this work, based on the analytical model with delayed production approximation developed by Pagel & Tautvaišienė (1995) for the Galaxy, the analytic solutions of the distribution of neutron exposures of the Galaxy (hereafter NEG) are obtained. The present results appear to reasonably reproduce the distribution of neutron exposures of the solar system (hereafter NES). The strong component and the main component of the NES are built up in different epochs. Firstly, the strong component is produced by the s-process nucleosynthesis in the metal-poor AGB stars, starting from [Fe/H] ≈ -1.16 to [Fe/H] ≈ -0.66, corresponding to the time interval $1.06 < t < 2.6$ Gyr. Secondly, the main component is produced by the s-process in the galactic disk AGB stars, starting from [Fe/H] ≈ -0.66 to [Fe/H] ≈ 0, corresponding to the time interval $t > 2.6$ Gyr. The analytic solutions have the advantage of an understanding of the structure and the properties of the NEG. The NEG is believed to be an effective tool to study the s-process element abundance distributions in the Galaxy at different epochs and the galactic chemical evolution of the neutron-capture elements.

Key words. Nucleosynthesis—neutron exposures—abundances—stars: AGB stars: low-mass stars—galaxy: evolution.

1. Introduction

Many efforts have been made to explain the solar-system abundance of elements associated with the slow neutron-capture process (s-process). The s-process occurs mainly during hydrostatic He-burning phases of stellar evolution, which can be investigated either through nucleosynthesis computations in stellar models (Straniero *et al.* 1995; Gallino *et al.* 1998) or by the phenomenological models, mostly by the so-called classical model (Käppeler *et al.* 1989; Busso *et al.* 1999). The classical s-process was first outlined by Burbidge *et al.* (1957). Clayton *et al.* (1961) showed that the solar-system abundances of the s-only isotopes cannot be reproduced by a single neutron irradiation of an iron seed. Because the solar-system s-element composition was assumed to be

the result of a superposition of different distributions of neutron exposures, a satisfactory solution was found by Seeger *et al.* (1965) assuming an exponential decreasing distribution of neutron exposures. When sufficiently detailed input data became available, it turned out that three different exponential distributions of neutron exposures are required for a complete description of the observed s-process abundance (Clayton & Rassbach 1967; Clayton & Ward 1974; Beer & Macklin 1985; Käppeler *et al.* 1982). The current understanding of the s-process is supported by many observational and theoretical works, indicating that the heavier s-nuclei, from Sr to Pb, belonged to the so-called main component. Below A$\approx$ 90, this component fails to describe the steep increase of the $\sigma_i N_i$ (*i.e.*, the products of neutron cross section times s-process abundance) curve towards the iron seed. Therefore, a weak component is added which is characterized by a smaller mean neutron exposure. Finally, a strong component is to be postulated in order to account for the abundance maximum at lead. Each of these distributions is expressed as

$$\rho(\tau) = \frac{f_i N_{Fe}(\mathrm{Z}_\odot)}{\tau_{0,i}} \exp(-\tau/\tau_{0,i}), \tag{1}$$

where f_i is the initial solar fraction of ^{56}Fe that has been irradiated, τ means the time-integrated neutron flux (*i.e.*, $\tau = \int_0^t n_n v_T dt'$, where n_n is the neutron density and v_T is the thermal velocity at temperature T), and $\tau_{0,i}$ is the mean neutron exposure. These three components will be distinguished by adding indices 1, 2 and 3 to the parameters f_i and $\tau_{0,i}$, respectively. An exponential distribution for $\rho(\tau)$ appears to provide for an excellent fit of the main and strong s-process components.

A physical justification for the choice of $\rho(\tau)$ seemed to appear when Ulrich (1973) showed that an exponential distribution of exposures was the natural consequence of repeated He-shell flashes during the AGB phase. If $\Delta\tau$ is the neutron exposure per pulse, and if r denotes the overlap of the Nth and the $(N-1)$th convective shell, after N pulses the fraction of material having experienced an exposure $\tau = N\Delta\tau \sim r^N \equiv \exp(-\tau/\tau_0)$, where $\tau_0 = -\Delta\tau/lnr$ is the mean neutron exposure. During the AGB evolution, stars experience enough helium shell flashes to establish an exponential distribution. There, the s-process was assumed to occur in convective thermal pulses, the classical analysis was considered to yield 'effective' conditions characterizing the stellar scenarios (Käppeler *et al.* 1990). In order to test the reliability of the classical model, Goriely (1997) has developed a new s-process model based on the superposition of a large number of canonical astrophysical events, the abundance predictions of the multi-event model are in good agreement with those of the widely used exponential model.

The knowledge of the last stages of evolution for AGB stars was improved by a series of investigations originally based on the activation of the ^{13}C(α, n)^{16}O reaction in low-mass stars (Straniero *et al.* 1997; Gallino *et al.* 1998). The model suggested that the main s-process component results from low-mass stars with masses between 1.5 and 3.0 $\mathrm{M}_\odot$. According to the models, the ^{13}C neutron source is activated under radiative conditions during the intervals between subsequent He-shell burning episodes. The stellar models show that the distribution of neutron exposures is definitely non-exponential, and actually very difficult to be described analytically (Arlandini *et al.* 1995). The overabundances of elements heavier than iron observed at the surface of MS and S stars (Smith & Lambert 1990) clearly indicate that the s-process takes place during the AGB phase in the evolution of low- and intermediate-mass stars

($0.8 \leq M(\mathrm{M}_{\odot}) \leq 8$). Though the observations in the solar neighborhood exhibit a spread in the respective s-abundances, it is remarkable that the solar s-abundance distribution lays roughly at the center of the spread observed in MS and S stars (Busso *et al.* 1999). Most galactic disk AGB stars in the mass range $1.5 \leq M(\mathrm{M}_{\odot}) \leq 3.0$ can be considered as suitable sites for reproducing the main component. The mean physical conditions are found in AGB star models down to a metallicity slightly lower than $1/2\,\mathrm{Z}_{\odot}$ by Gallino *et al.* (1998).

In fact, the solar system chemical composition is the result of a complex galactic evolution mechanism, and depends upon the details of the stellar formation history, initial mass functions, chemical yields, etc. Since the s-process distribution and the neutron exposure distribution vary strongly for TP-AGB stars with different metallicity (Busso *et al.* 1999; Travaglio *et al.* 1999; Raiteri *et al.* 1999), the comparison with the solar distributions has to be complemented by the s-process distribution and the neutron exposure distribution for the TP-AGB stars with a fixed initial mass and metallicity. Recently, a quantitative calculation of the evolution of neutron exposure distribution in the Galaxy (NEG) has been carried out through solving a set of differential equations by Cui *et al.* (2007). The numerical models of the galactic chemical evolution attempts to combine ideas on the end-products of stellar evolution with ideas on the formation and evolution of galaxies in order to understand the NEG. As this involves many uncertainties, we prefer to adopt an analytical approach, which is to parameterize the problem as simply as possible, to study the characters of the NEG. This is somewhat unfashionable nowadays, but it has the advantage that one can immediately understand the structure and the properties of the NEG, which is not always the case in elaborate numerical models.

The distribution of neutron exposures plays a key role in the theory of s-process nucleosynthesis. It is a basis of qualitative and quantitative analysis as well as a good understanding of s-process abundance distribution. Therefore to investigate it more deeply is very necessary. In this work, we use the NEG $\rho_{\mathrm{Gal}}(\tau, t)$ to study the chemical evolution characteristic of the Galaxy for s-process elements. One of the main goals of this work was the development of an analytical model for the NEG, to provide stronger evolutionary support to the distribution of neutron exposures in the solar system. The paper is organized as follows: in section 2 we deduce the evolution equation of the NEG from the chemical evolution model developed by Pagel & Tautvaišienė (1995); in section 3 we discuss the distributions of neutron exposures in the AGB stars (hereafter the NEAGB); in section 4 we show the analytic solutions for the NEG and compare them with those of two parameterized models, namely the multi-event model (Goriely 1997) and the exponential model (Beer *et al.* 1997). Finally, in section 5 we summarize the main conclusions.

2. The basic equations and assumptions

In the chemical evolution model there are five variables: the total system mass M, the mass of 'gas' g, the mass existing in the form of stars (including compact remnants) s, the mass inflow rate F and the abundances X_i of the element(s). Instantaneous recycling cannot be used to describe the formation of elements, such as s-process elements, to which there is a significant contribution from stars that take a non-negligible time to complete their evolution. We adopt the analytical model with delayed production approximation developed by Pagel & Tautvaišienė (1995) for the disc of the Galaxy

assuming no galactic wind and the inflowing material to be unprocessed, which works by the simple device of assuming that the delayed element or component of stars to be released at a single time Δ after the onset of stars formation and is then instantaneously recycled. This model has been adopted to investigate different aspects of the chemical evolution of the Galaxy. In the delayed production approximation, each generation of the AGB stars ejects a newly synthesized s-process element after a fixed time delay Δ, then the abundance of s-element in the gas, with the dimensionless time-like variable u introduced, is governed by (Pagel & Tautvaišienė 1995)

$$\frac{dX_i}{du} + \frac{F}{\omega g} = 0 \qquad \text{if } u < \omega\Delta, \tag{2}$$

$$= p_i(u - \omega\Delta)\frac{g(u - \omega\Delta)}{g(u)} \qquad \text{if } u \geq \omega\Delta, \tag{3}$$

where time-like variable u is defined by

$$u \equiv \int_0^t \omega(t')dt', \tag{4}$$

$\omega(t)$ is the transition probability for diffuse material ('gas') to change into stars in unit time at time t. p_i is the yields of s-process element i ejected from each generation of the AGB stars, which is given by an expression of the form

$$p_i = \frac{1}{\alpha}\int_{m_l}^{m_u} m_{Dup} x_i \phi(m)dm, \tag{5}$$

where the limits of integration are taken from $m_{l,\min} = 0.1$ to $m_U = 62\,\mathrm{M}_\odot$ (Miller & Scalo 1979), $\phi(m)$ is initial mass function (IMF) adopted from Miller & Scalo (1979) too, α is the lock-up fraction (assumed constant), m_{Dug} is the total mass dredged up from the intershell of an AGB star and x_i the mass fraction of species i in the intershell.

We use $\rho_{\mathrm{Gal}}(\tau, t)d\tau$ to represent the number of iron seed nuclei (per 10^{12} H atoms) that has received a neutron exposure between τ and $\tau + d\tau$ in the Galaxy before time t. The abundance of an s-process nucleus i at time t can be written as (Cui *et al.* 2007)

$$N_i(t) = \frac{1}{\sigma_i}\int_0^\infty \Psi_i(\tau)\rho_{\mathrm{Gal}}(\tau, t)d\tau, \tag{6}$$

where $\Psi_i(\tau)$ is the solution of a single exposure (see Clayton 1961). In addition, the abundance of an s-process nucleus i in the AGB star with the initial metallicity Z is given by an expression of the form

$$n_i(Z) = \frac{1}{\sigma_i}\int_0^\infty \Psi_i(\tau)\rho_{AGB}(\tau, Z)d\tau, \tag{7}$$

where $\rho_{AGB}(\tau, Z)$ is the distribution of neutron exposures for an AGB star (NEAGB). Combining equations (2) and (3) with equations (6) and (7), the rate of change of the NEG can be written as

$$\frac{d\rho_{\mathrm{Gal}}(\tau, u)}{du} + \frac{F}{\omega g}\rho_{\mathrm{Gal}}(\tau, u) = 0 \qquad \text{if } u < \omega\Delta, \tag{8}$$

$$= p'\rho_{AGB}(\tau, Z_{u-\omega\Delta})\frac{g(u - \omega\Delta)}{g(u)} \qquad \text{if } u \geq \omega\Delta, \tag{9}$$

p' is the fraction of ejection mass dredged up from AGB stars, which is given by the formula

$$p' = \frac{1}{\alpha} \int_{m_l}^{m_u} m_{Dup} \phi(m) dm. \tag{10}$$

Given a functional form for $\rho_{(AGB)}(\tau, Z)$ and the constant parameters α and $\omega\Delta$, equations (8) and (9) can be solved to give the NEG as a function of u. For simplicity, we hold on to the assumption of constant mass dredged up from AGB stars for s-process elements despite the possibility raised by the work of Straniero *et al.* (2003) that the mass depends on the initial metallicity.

3. The adopted NEAGB

3.1 *Radiative model: Case A*

We adopt the 25th pulse computed by Gallino *et al.* (1998), corresponding to a core mass $M_C = 0.70\ M_\odot$, as typical, and assume that all the pulses are identical. Then the distribution of the neutron exposure in the AGB stars can be obtained by three separated steps (Cui *et al.* 2007). Firstly, considering that the seed nuclei previously entered the He intershell at the same pulse would suffer different irradiation times after experiencing a fixed number of pulses, we can investigate the total fraction of seed nuclei in the He intershell with different irradiation times. Secondly, taking into account that the seed nuclei in different layers of the ^{13}C-pocket would achieve different neutron exposures per irradiation, by introducing the conception of multi-order distribution of neutron exposures and using the Monte Carlo method, we can calculate the distribution of neutron exposures for the seed nuclei that can experience any irradiation times. Finally, we can obtain the final distribution of neutron exposures in the He intershell of low-mass AGB stars. There are four parameters in this model: they are the overlap factor r, the mass fraction q, the temperature T and the maximum of neutron exposure $\tau_{\max}$ in the q layer. In this paper, since we are particularly interested in the exposure distribution which can reproduce the solar s-abundance distribution, we adopt the ^{13}C abundance profile indicated as standard (ST) by Gallino *et al.* (1998) (see their Fig. 1), and take $r = 0.45$, $q = 0.05$. Figure 1 adopted from Cui *et al.* (2007) shows the resulting asymptotic distribution of exposures in the ^{13}C burning scenario. As Fig. 1 shows, even in the above simplified scenario, the distribution of neutron exposures is much more complex than the usually exponential form.

3.2 *Convective model: Case B*

There is another possibility for the synthesis of s-process elements in the AGB stars, *i.e.*, with nucleosynthesis taking place during thermal pulses (Aoki *et al.* 2001). In this case, the neutron irradiation is derived primarily by the reaction $^{13}C(\alpha, n)^{16}O$, with a minor contribution from the marginal burning of ^{22}Ne. During the AGB evolution, stars experience enough He shell flashes to establish an exponential distribution

$$\rho_{AGB}(\tau, Z) = \frac{N(Fe, Z)}{\tau_0(Z)} \exp(-\tau/\tau_0(Z)). \tag{11}$$

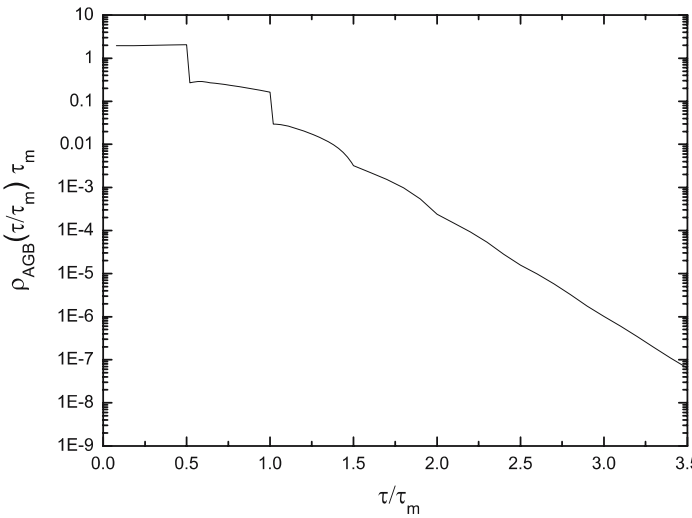

Figure 1. Distribution of the neutron exposures inside a single low-mass AGB star that can reproduce the main s-component in the solar system (adopted from Cui *et al.* 2007).

The resulting pattern of AGB nucleosynthesis, and its dependence on the initial metallicity of the star, have been discussed by Gallino *et al.* (1999). Since the ^{13}C neutron source is of primary nature, the typical neutron density in the nucleosynthesis zone scales roughly as $1/Z^{0.6}$ from $Z_\odot$ down to $1/50\ Z_\odot$. At lower metallicities, the effect of the primary poisons prevails (Busso *et al.* 1999).

4. Results of the evolution of NEG

The basic functions in our study, such as $F(u)$ and $g(u)$, are adopted from the analytical forms given by Pagel & Tautvaišienė (1995). In the present work, we concentrate on the influence of initial metallicity of the low-mass AGB stars to the NEG. More massive AGB stars, in the range 3–8 $M_\odot$, do not give a relevant contribution to the main and strong components (Travaglio *et al.* 1999). At the end of the TP–AGB phase, the s-process contributions to the ISM are determined by the amount of matter cumulatively dredged-up from the He shell to the surface and lost by stellar winds.

In general, the production factor of a given s-element varies with Z that is close to solar, the strong increase of the s-process yields with increasing neutron exposure dominates and the secondary nature of these nuclei is overcompensated, so that they actually increase for decreasing Z, instead of constant (Travaglio 1999, see their Fig. 2). Starting from AGB stars with nearly solar metallicity, first the s-process builds up the s-elements belonging to the Zr-peak. Then the Zr-peak yields decrease while the Ba-peak production increases, reaching a maximum at [Fe/H] $\simeq -0.6$. Despite the dependence of s-process yield on the initial metallicity of the AGB stars being very complex and non linear (Busso *et al.* 1999), the galactic disk AGB stars in the mass range $1.5 \leq M_\odot \leq 3.0$ can be considered as suitable sites for reproducing the main component, and the mean physical conditions of the main component are found in AGB star models down to a metallicity slightly lower than $1/2\ Z_\odot$ by Gallino *et al.* (1998). At lower metallicities, the neutron-flux skips Ba and feeds Pb, which reaches a maximum

yields at [Fe/H] $= -1$ (Gallino *et al.* 1999). Eventually, yields of Pb also decrease which makes the secondary nature of the s-process evident. At very low metallicities, essentially all the Fe group seeds are converted to Pb. The origin of Pb and the site of the strong component of the s-process, should be attributed to these low-mass, metal-poor AGB stars. Below [Fe/H] $\simeq -2$ only a small s-signature is presented on heavy elements.

Based on the above discussion, we assume that the strong and the main components in the Galaxy are built up in different epochs. Firstly, the strong component is produced by the s-process nucleosynthesis in the metal-poor AGB stars, starting from [Fe/H] ≈ -1.16 to [Fe/H] ≈ -0.66, corresponding to the time interval $\omega\Delta < u < u_c$ where $u_c = 2.6$ is the time at which the contribution of the galactic disk AGB stars to the gas began. Secondly, the main component is produced by the s-process nucleosynthesis in the galactic disk AGB stars, starting from [Fe/H] ≈ -0.66 to [Fe/H] ≈ 0. The solutions of NEG are given by

$$\rho_{\rm Gal}(\tau, u) = 0 \qquad \text{if } u < \Delta, \tag{12}$$

$$= p' e^{\omega\Delta} \left(\frac{u_1 + u_0}{u + u_0}\right)^3 (u - \omega\Delta) \frac{N_{Fe1}}{\tau_{01}} e^{-\frac{\tau}{\tau_{01}}} \qquad \text{if } \omega\Delta \le u \le u_1 + \omega\Delta, \tag{13}$$

$$= p' \frac{e^{\omega\Delta}}{(u + u_0)^3} \left[(u_1 + u_0)^3 u_1 + \frac{(u + u_0 - \omega\Delta)^4 - (u_1 + u_0)^4}{4}\right] \frac{N_{Fel}}{\tau_{01}} e^{-\frac{\tau}{\tau_{01}}}$$

$$\text{if } u_1 + \omega\Delta \le u < u_c, \tag{14}$$

$$= p' \frac{e^{\omega\Delta}}{(u_c + u_0)^3} \left[(u_1 + u_0)^3 u_1 + \frac{(u_c + u_0 - \omega\Delta)^4 - (u_1 + u_0)^4}{4}\right] \frac{N_{Fel}}{\tau_{01}} e^{-\frac{\tau}{\tau_{01}}}$$

$$+ p' \frac{e^{\omega\Delta}}{(u + u_0)^3} \left[\frac{(u + u_0 - \omega\Delta)^4}{4} - \frac{(u_c + u_0 - \omega\Delta)^4}{4}\right] \frac{N_{Fe2}}{\tau_{02}} e^{-\frac{\tau}{\tau_{02}}} \quad \text{if } u \ge u_c. \tag{15}$$

The first term on the right of the solution (14) represents the contribution of the metal-poor AGB stars that formed during the early time of the Galaxy. In this case, the characteristics of s-process nucleosynthesis in the AGB stars are that the abundance of Fe seeds is lower but the neutron exposure is larger. The second term represents the contribution of the galactic disk AGB stars that formed during the later time of the Galaxy. The characteristics are that the abundance of Fe seeds is higher but the neutron exposure is lower. Obviously, the analytical model has the advantage that one can immediately understand the structure and the properties of the NEG, which is not the case in elaborate numerical models (Cui *et al.* 2007).

Gallino *et al.* (1998) have pointed out that the neutron density is relatively low, reaching $\sim 10^7$ cm^{-3} in the q layer at solar metallicity and the temperature $T_9 = 0.1$ (in units of 10^9 K). According to the standard ^{13}C profile with solar metallicity given by Gallino *et al.* (1998), we take the maximum neutron exposure $\tau_{\rm max} \approx 1.79$ mbarn^{-1} (kT $= 30$ keV) for the low-mass AGB stars with [Fe/H] $= -0.3$ in Case A. We choose $\tau_0 = 0.296(T_9/0.348)^{1/2}$ mbarn^{-1} for low-mass AGB stars with [Fe/H] $= -0.3$ in Case B, which corresponds to a mean neutron exposure of the solar system. We use the primary nature of ^{13}C neutron source to calculate the neutron exposure per thermal

Table 1. The parameters of NEAGB.

	$\omega\Delta \leq u < u_c$	$u \geq u_c$
Case A	$\Delta\tau_{\max} = 10mb^{-1}$	$\Delta\tau_{\max} = 1.79mb^{-1}$
	[Fe/H] $= -1$	[Fe/H] $= -0.3$
Case B	$\tau_0 = 8mb^{-1}$	$\tau_0 = 0.296mb^{-1}$
	[Fe/H] $= -1$	[Fe/H] $= -0.3$

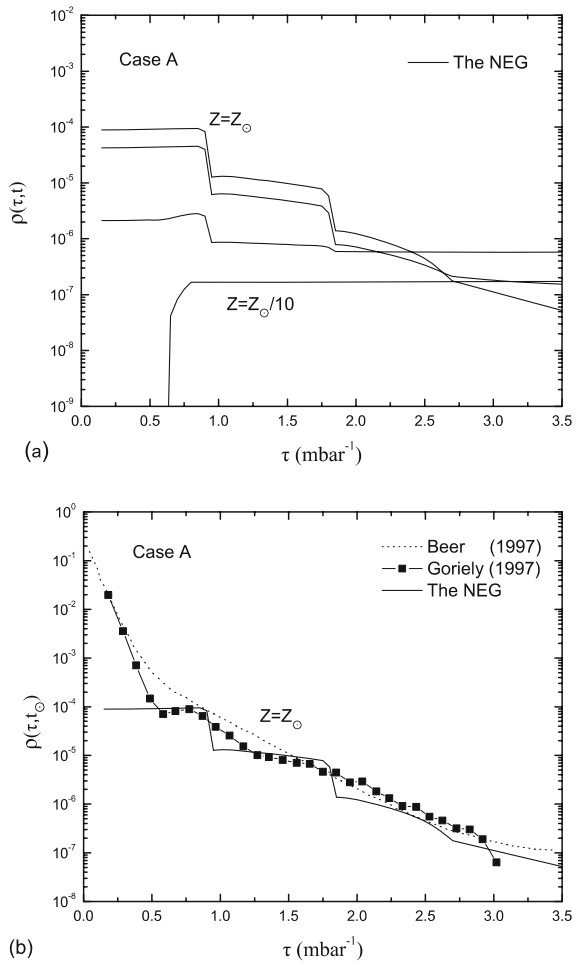

Figure 2. The evolution of the NEG as a function of neutron exposures (τ) for Case A. **(a)** The solid lines show respectively, the resulting NEG at $Z = 1/10,\ 1/5,\ 1/2,\ Z_\odot$. **(b)** The NEG at $Z = Z_\odot$. The analytical exposure distribution predicted by the exponential model of Beer *et al.* (1997) and the exposure distribution predicted by the canonical multi-event s-process model of Goriely (1997) are also shown for comparison.

pulse $\Delta\tau_{\max}$ and τ_0 at other metallicities. The basic parameters of the NGAGB adopted in our model are given in Table 1. Taking the total mass dredged up into the envelope of the AGB stars as $\Delta m_{Dup} = 10^{-2}\ M_\odot$ (Busso *et al.* 1992), the ejection mass fraction

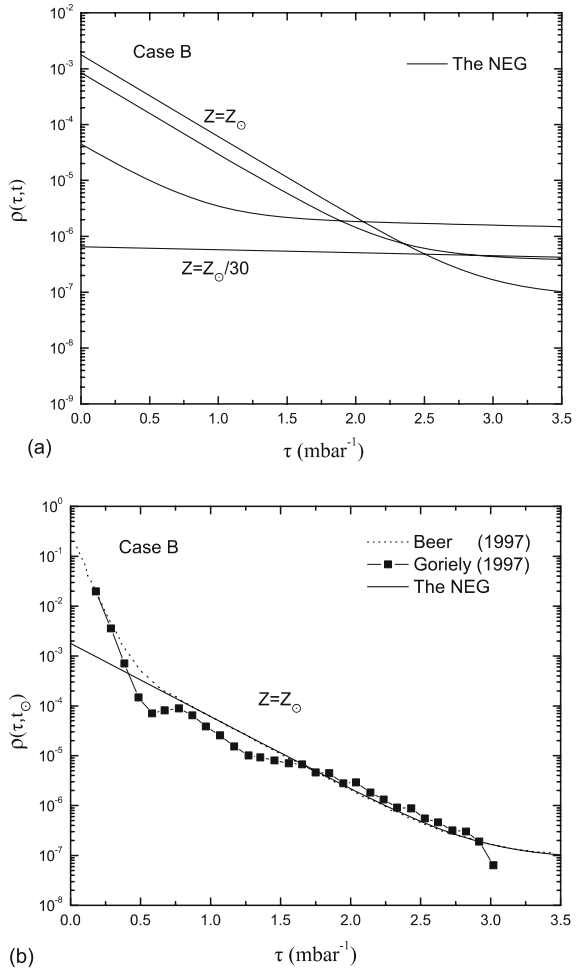

Figure 3. The evolution of the NEG as a function of neutron exposures (τ) for Case B. **(a)** The solid lines show respectively, the resulting NEG at $Z = 1/30,\ 1/5,\ 1/2,\ Z_{\odot}$. **(b)** The NEG at $Z = Z_{\odot}$. The analytical exposure distribution predicted by the exponential model of Beer *et al.* (1997) and the exposure distribution predicted by the canonical multi-event s-process model of Goriely (1997) are also shown for comparison.

$p' = 3.66 \times 10^{-4}$ is expected. The large parameter $\omega\Delta = 1.06$ for s-process elements is expected to reflect the time-scale for low-mass AGB stars.

In Figs. 2(a) and 3(a) we present our results for the evolution of the NEG as a function of metallicity for Case A and Case B. The [Fe/H] scale here is the indication of a time scale, albeit a non-linear one. From Fig. 2, we can see that there is a strong dependence of NEG on the metallicity or time of the Galaxy, the s-process contribution becomes important starting from [Fe/H] $= -1.5$ at lower values of [Fe/H], the contribution of s-process nucleosynthesis rapidly decreases due to the strong dependence of stellar yields on metallicity. The low-mass and metal-poor AGB stars are the dominant producers for the larger neutron exposures of the NEG, with a complex dependence on metallicity and a maximum efficiency at [Fe/H] $= -1$. The galactic

disk AGB stars are the dominant producers for the lower neutron exposures of the NEG, with a maximum efficiency at [Fe/H] $= -0.6$. We also compare in Figs. 2(b) and 3(b), our results of NEG at $Z = \mathrm{Z}_{\odot}$ with the previous studies by Beer *et al.* (1997), obtained with the so-called classical approach and the results obtained by Goriely (1997) with multi-events model. Notice that our results are in agreement with the other results.

From the solutions of NEG and its definition, *i.e.*, $f(u)N_{Fe}(0) = \int_0^{\infty} \rho_{\mathrm{Gal}}(\tau, t)d\tau$, we can obtain the total number of iron seed nuclei in the Galaxy, which have been irradiated in all AGB stars, before time t

$$f(u)N_{Fe}(0) = 0 \quad \text{if } u < \Delta, \tag{16}$$

$$= p' \frac{e^{\omega\Delta}}{(u+u_0)^3}\left[(u_1+u_0)^3 u_1 + \frac{(u+u_0-\omega\Delta)^4 - (u_1+u_0)^4}{4}\right] N_{Fe1}$$

$$\text{if } u_1 + \omega\Delta \le u < u_c, \tag{17}$$

$$= p' \frac{e^{\omega\Delta}}{(u_c+u_0)^3}\left[(u_1+u_0)^3 u_1 + \frac{(u_c+u_0-\omega\Delta)^4 - (u_1+u_0)^4}{4}\right] N_{Fe1}$$

$$+ p' \frac{e^{\omega\Delta}}{(u+u_0)^3}\left[\frac{(u+u_0-\omega\Delta)^4}{4} - \frac{(u_c+u_0-\omega\Delta)^4}{4}\right] N_{Fe2} \quad \text{if } u \ge u_c. \tag{18}$$

In Fig. 4 we present the calculated results for the evolution of the $f(u)N_{Fe}(0)$ as a function of time t. From the definition of the NEG, we can know that the factor $f(u_{\odot})N_{Fe}(0)$ in the NES is the number of Fe seeds for the corresponding components that has been irradiated before the formation of the solar system. Our results are in agreement with the previous studies by Beer *et al.* (1997), obtained with the classical approach.

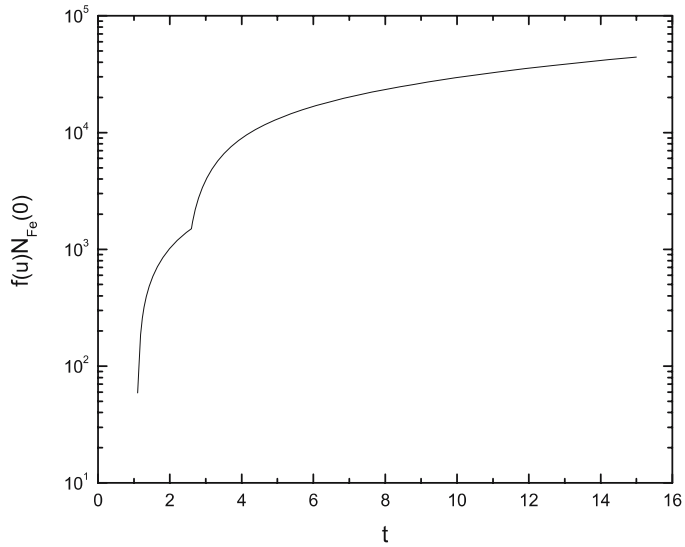

Figure 4. The evolution of the $f(u)N_{Fe}(0)$ as a function of time t.

It should be noticed that, for the lower neutron exposures, the NEG of Case A is slightly underestimated as compared to the results given by Beer *et al.* (1997), which is also obtained by the numerical model (Cui *et al.* 2007). The values of NEG for the lower neutron exposures are much more model dependent. The origin of discrepancies might be attributed to specific characteristics of the galactic evolution model (e.g., the age-metallicity relation). The uncertainty in this evaluation may depend on the set of prescriptions adopted to estimate the s-process yields from the AGB stars with varying metallicity and on the general prescriptions adopted in the galactic chemical evolution model.

5. Conclusions

Based on the analytical model with delayed production approximation developed by Pagel & Tautvaišienė (1995) for the Galaxy, the analytic solutions of the NEG are obtained. The results of the galactic evolution model, compared with the exposure distribution in the solar system, confirm the basic nucleosynthesis scenario outlined by Straniero *et al.* (1997) and Gallino *et al.* (1998). The ^{13}C neutron source, activated during the interpulse phases of low-mass TP-AGB stars, accounts for the exposure distribution in the solar system. From a comparison of the best fit by the classical analysis to the main component and the strong component by Beer *et al.* (1997) with the results of our NEG, one can obtain a clear indication that not a unique AGB stellar model, nor the classical analysis, was able to explain the main and strong component in the solar system, which must be considered as the outcome of different generations of AGB stars which are prior to the solar system formation.

The present results appear to reasonably reproduce the distribution of neutron exposures of the solar system. The strong component and the main component in the NES are built up in different epochs. Firstly, the strong component is produced by the s-process nucleosynthesis in the metal-poor AGB stars, starting from [Fe/H] ≈ -1.16 to [Fe/H] ≈ -0.66, corresponding to the time interval $1.06 < t < 2.6$ Gyr. Secondly, the main component is produced by the s-process nucleosynthesis in the galactic disk AGB stars, starting from [Fe/H] ≈ -0.66 to [Fe/H] ≈ 0, corresponding to the time interval $t > 2.6$ Gyr. As a matter of fact, our NEG calculations confirm, in a quantitative way, what was anticipated by Gallino *et al.* (1998), *i.e.*, that the role previously attributed to the strong component is actually played by the low-mass AGB stars with low metallicity. It is evident that the analytic solutions obtained in this work have the advantage of an understanding of the structure and the properties of the NEG, whereas numerical calculations can only supply evidence of the special case considered (Cui *et al.* 2007). From the definition of the NEG, we can know that the factor $f(u)N_{Fe}(0)$ in the NES is the number of Fe seeds for the corresponding component that has been irradiated before the formation of the solar system.

Despite the large number of approximations (the total amount of dredged-up material, the parameterization on the ^{13}C pocket and the dependence on the initial mass of low-mass AGB stars for the s-process nucleosynthesis, etc.) for the analytical model, the agreement of the model results with the NES provides a strong support to the validity of the analytical solutions of the NEG and the NEAGB prescriptions adopted in this work. The NEG is therefore believed to be an effective tool to study the s-process element abundance distributions in the Galaxy at different epochs and the galactic chemical evolution of the neutron-capture elements. It presents new features

in comparison to the classical exponential model. Hopefully, the NEG can help us to improve our understanding of the exposure distribution in the solar system and open new perspectives in this direction.

Acknowledgements

We thank the referee for an extensive and helpful review, containing very relevant scientific advice which improved this paper greatly. This work has been supported by the National Natural Science Foundation of China under Grant 10373005 and 10673002.

References

Aoki, W., Ryan, S. G., Norris, J. E., Beers, T. C., Ando, H., Iwamoto, N., Kajino, T., Mathews, G. J., Fujimoto, M. Y. 2001, *ApJ*, **561**, 346.
Arlandini, C., Gallino, R., Busso, M., Straniero, O. 1995, In: Stellar Evolution: what should be done (eds) Noels, A., Fraipont-Caro, D., Gabriel, M., Grevesse, N., Demarque, P. (Liège: Univ. de Liège), 447.
Beer, H., Corvi, F., Mutti, P. 1997, *ApJ*, **474**, 843.
Beer, H., Macklin, R. L. 1985, *Phys. Rev. C*, **32**, 738.
Burbidge, E. M., Burbidge, G. R., Fowler, W. A., Hoyle, F. 1957, *Rev. Mod. Phys.*, **29**, 547.
Busso, M., Gallino, R., Lambert, D. L., Raiteri, C. M., Smith, V. V. 1992, *ApJ*, **399**, 218.
Busso, M., Gallino, R., Wasserburg, G. J. 1999, *ARA&A*, **37**, 239.
Clayton, D. D., Fowler, W. A., Hull, T. E., Zimmerman, B. A. 1961, *Ann. Phys.*, **12**, 331.
Clayton, D. D., Rassbach, M. E. 1967, *ApJ*, **168**, 69.
Clayton, D. D., Ward, R. A. 1974, *ApJ*, **193**, 397.
Cui, W. Y., Zhang, F. H., Zhang, W. J., Zhang, L., Zhang, B. 2007, *Chinese J. Astron. Astrophys.* (ChJAA), in press.
Gallino, R., Arlandini, C., Busso, M., Lugaro, M., Travaglio, C., Straniero, O., Chieffi, A., Limongi, M. 1998, *ApJ*, **497**, 388.
Gallino, R., Busso, M., Lugaro, M., Travaglio, C., Arlandini, C., Vaglio, P. 1999, In: *Nuclei in the Cosmos V* (eds) Prantzos, N., Harissopulos, S. (Paris: Edition Frontières), 216.
Goriely, S. 1997, *A&A*, **327**, 845.
Käppeler, F., Beer, H., Wisshak, K., Clayton, D. D., Macklin, R. L., Ward, R. A. 1982, *ApJ*, **257**, 821.
Käppeler, F., Beer, H., Wisshak, K. 1989, *Rep. Progr. Phys.*, **52**, 945.
Käppeler, F., Gallino, R., Busso, M., Picchio, G., Raiteri, C. M. 1990, *ApJ*, **354**, 630.
Miller, G. E., Scalo, J. M. 1979, *ApJS*, **41**, 513.
Pagel, B. E. J., Tautvaišiene, G. 1995, *MNRAS*, **276**, 505.
Raiteri, C. M., Villata, M., Gallino, R., Busso, M., Gravanzola, A. 1999, *ApJ*, **518**, L91.
Seeger, P. A., Fowler, W. A., Clayton, D. D. 1965, *ApJS*, **11**, 121.
Smith, V. V., Lambert, D. L. 1990, *ApJS*, **72**, 387.
Straniero, O., Chieffi, A., Limongi, M., Busso, M., Gallino, R., Arlandini, C. 1997, *ApJ*, **478**, 332.
Straniero, O., Domnguez, I., Cristallo, S., Gallino, R. 2003, *PASA*, **20**, 389.
Straniero, O., Gallino, R., Busso, M., Chieffi, A., Raiteri, C. M., Salaris, M., Limongi, M. 1995, *ApJ*, **440**, L85.
Travaglio, C., Galli, D., Gallino, R., Busso, M., Ferrini, F., Straniero, O. 1999, *ApJ*, **521**, 691.
Ulrich, R. K. 1973, In: Explosive Nucleosynthesis (eds) Schramm, D. N., Arnett, W. D. (Austin: University of Texas Press), p. 139.

J. Astrophys. Astr. (2007) **28,** 67–99

Cosmology and Cosmogony in a Cyclic Universe

Jayant V. Narlikar[1,*], Geoffrey Burbidge[2] & R. G. Vishwakarma[3]
[1]*Inter-University Centre for Astronomy and Astrophysics, Pune 411 007, India.*
[2]*Center for Astrophysics and Space Sciences, University of California, San Diego, CA 92093-0424, USA.*
[3]*Department of Mathematics, Autonomous University of Zacatecas, Zacatecas, ZAC C.P. 98060, Mexico.*
[*]*e-mail: jvn@iucaa.ernet.in*

Received 2006 August 8; accepted 2007 April 26

Abstract. In this paper we discuss the properties of the quasi-steady state cosmological model (QSSC) developed in 1993 in its role as a cyclic model of the universe driven by a negative energy scalar field. We discuss the origin of such a scalar field in the primary creation process first described by F. Hoyle & J. V. Narlikar forty years ago. It is shown that the creation processes which take place in the nuclei of galaxies are closely linked to the high energy and explosive phenomena, which are commonly observed in galaxies at all redshifts.

The cyclic nature of the universe provides a natural link between the places of origin of the microwave background radiation (arising in hydrogen burning in stars), and the origin of the lightest nuclei (H, D, He^3 and He^4). It also allows us to relate the large scale cyclic properties of the universe to events taking place in the nuclei of galaxies. Observational evidence shows that ejection of matter and energy from these centers in the form of compact objects, gas and relativistic particles is responsible for the population of quasi-stellar objects (QSOs) and gamma-ray burst sources in the universe.

In the later parts of the paper we briefly discuss the major unsolved problems of this integrated cosmological and cosmogonical scheme – the understanding of the origin of the intrinsic redshifts, and the periodicities in the redshift distribution of the QSOs.

Key words. Cosmology—cosmogony—high energy phenomena.

1. Introduction

1.1 *Cosmological models*

The standard cosmological model accepted by the majority at present is centered about the big bang which involves the creation of matter and energy in an initial explosion. Since we have overwhelming evidence that the universe is expanding, the only alternative to this picture appears to be the classical steady-state cosmology of

Bondi, Gold & Hoyle (Bondi & Gold 1948; Hoyle 1948) or a model in which the universe is cyclic with an oscillation period which can be estimated from observation. In this latter class of model the bounce at a finite minimum of the scale factor is produced by a negative energy scalar field. Long ago, Hoyle & Narlikar (1964) emphasized the fact that such a scalar field will produce models which oscillate between finite ranges of scale. In the 1960s theoretical physicists shied away from scalar fields, and more so those involving negative energy. Later Narlikar & Padmanabhan (1985) discussed how the scalar creation field helps resolve the problems of singularity, flatness and horizon in cosmology. It now appears that the popularity of inflation and the so-called new physics of the 1980s have changed the 1960s' mind-set. Thus Steinhardt & Turok (2002) introduced a negative potential energy field and used it to cause a bounce from a non-singular high density state. It is unfortunate that they did not cite the earlier work of Hoyle & Narlikar which had pioneered the concept of non-singular bounce through the agency of a negative energy field, at a time when the physics community was hostile to these ideas. Such a field is required to ensure that matter creation does not violate the conservation of matter and energy.

Following the discovery of the expansion of the universe by Hubble in 1929, practically all of the theoretical models considered were of the Friedmann type, until the proposal by Bondi, Gold & Hoyle in 1948 of the classical steady state model which first invoked the creation of matter. A classical test of this model lay in the fact that, as distinct from all of the big bang models, it predicted that the universe must be accelerating (cf Hoyle & Sandage 1956). For many years it was claimed that the observations indicated that the universe is decelerating, and that this finding disproved the steady state model. Not until much later was it conceded that it was really not possible to determine the deceleration parameter by the classical methods then being used. Gunn & Oke (1975) were the first to highlight the observational uncertainties associated with this test. Of course many other arguments were used against the classical steady state model (for a discussion of the history, see Hoyle *et al.* 2000, chapters 7 and 8). But starting in 1998 studies of the redshift-apparent magnitude relation for supernovae of Type 1A showed that the universe *is* apparently accelerating (Riess *et al.* 1998; Perlmutter *et al.* 1999). The normal and indeed the proper way to proceed, after this result was obtained, should have been at least to acknowledge that, despite the difficulties associated with the steady state model, this model had all along been advocating an accelerating universe.

It is worth mentioning that McCrea (1951) was the first to introduce vacuum related stresses with equation of state $p = -\rho$ in the context of the steady state theory. Later Gliner (1970) discussed how vacuum-like state of the medium can serve as original (non-singular) state of a Friedmann model.

The introduction of dark energy is typical of the way the standard cosmology has developed; viz., a new assumption is introduced specifically to sustain the model against some new observation. Thus, when the amount of dark matter proved to be too high to sustain the primordial origin of deuterium, the assumption was introduced that most of the dark matter has to be non-baryonic. Further assumptions about this dark matter became necessary, e.g., cold, hot, warm, to sustain the structure formation scenarios. The assumption of inflation was introduced to get rid of the horizon and flatness problems and to do away with an embarrassingly high density of relic magnetic monopoles. As far as the dark energy is concerned, until 1998 the general attitude towards the cosmological constant was typically as summarized by Longair

in the Beijing cosmology symposium: "None of the observations to date require the cosmological constant" (Longair 1987). Yet, when the supernovae observations could not be fitted without this constant, it came back with a vengeance as dark energy.

Although the popularity of the cosmological constant and dark energy picked up in the late 1990s, there had been earlier attempts at extending the Friedmann models to include effects of vacuum energy. A review of these models, vis-a-vis observations may be found in the article by Carroll & Press (1992).

We concede that with the assumptions of dark energy, non-baryonic dark matter, inflation, etc., an overall self consistent picture has been provided within the framework of the standard model. One demonstration of this convergence to self consistency is seen from a comparison of a review of the values of cosmological parameters of the standard model by Bagla *et al.* (1996), with the present values. Except for the evidence from high redshift supernovae, in favour of an accelerating universe which came 2–3 years later than the above review, there is an overall consistency of the picture within the last decade or so, including a firmer belief in the flat ($\Omega = 1$) model with narrower error bars.

Nevertheless we also like to emphasize that the inputs required in fundamental physics through these assumptions have so far no experimental checks from laboratory physics. Moreover an epoch dependent scenario providing self-consistency checks, e.g., CMB anisotropies, cluster baryon fraction as a function of redshift does not meet the criterion of 'repeatability of scientific experiment'. We contrast this situation with that in stellar evolution where stars of different masses constitute repeated experimental checks on the theoretical stellar models thus improving their credibility.

Given the speculative nature of our understanding of the universe, a sceptic of the standard model is justified in exploring an alternative avenue wherein the observed features of the universe are explained with fewer speculative assumptions. We review here the progress of such an alternative model.

In this model creation of matter is brought in as a physical phenomenon and a negative kinetic energy scalar field is required to ensure that it does not violate the law of conservation of matter and energy. A simple approach based on Mach's principle leads naturally to such a field within the curved spacetime of general relativity described briefly in section 2. The resulting field equations have the two simplest types of solutions for a homogeneous and isotropic universe:

- those in which the universe oscillates but there is no creation of matter, and
- those in which the universe steadily expands with a constant value of H_o being driven by continuous creation of matter.

The simplest model including features of both these solutions is the *Quasi-Steady State Cosmology* (QSSC), first proposed by Hoyle *et al.* (1993). It has the scale factor in the form:

$$S(t) = \exp\left(\frac{t}{P}\right)\{1 + \eta\cos\theta(t)\}, \quad \theta(t) \approx \frac{2\pi t}{Q}, \tag{1}$$

where P is the long term 'steady state' timescale of expansion while Q is the period of a single oscillation.

Note that it is essential for the universe to have a long term expansion; for a universe that has only oscillations without long term expansion would run into problems like the Olbers paradox. It is also a challenge in such a model to avoid running into 'heat death'

through a steady increase of entropy from one cycle to the next. These difficulties are avoided if there is a creation of new matter at the start of each oscillation as happens in the QSSC, and also, if the universe has a steady long term expansion in addition to the oscillations. New matter in such a case is of low entropy and the event horizon ensures a constant entropy within as the universe expands.

The QSSC has an additional attractive feature if one uses the criterion of the Wheeler and Feynman, absorber theory of electromagnetic radiation (Wheeler & Feynman 1945, 1949). This theory provided a very natural explanation of why in actuality the electromagnetic signals propagate into the future, i.e., via retarded solutions, despite the time-symmetry of the basic equations. By writing the theory in a relativistically invariant action-at-a-distance form, Wheeler and Feynman showed that suitable absorptive properties of the universe can lead to the breaking of time-symmetry. As was discussed by Hogarth (1962) and later by Hoyle & Narlikar (1963, 1969, 1971) who also extended the argument to quantum electrodynamics, the Wheeler–Feynman theory gives results consistent with observations only if the past absorber is imperfect and the future absorber is perfect. This requirement is *not* satisfied by a simply cyclic universe or by an ever-expanding big bang universe but is satisfied by the QSSC because of expansion being coupled with cyclicity.

One may question as to why one needs to have the Wheeler–Feynman approach to electrodynamics in preference to field theory. The advantages are many, including

- a satisfactory explanation of the Dirac formula of radiative reaction,
- the unambiguous deduction of why one uses retarded solutions in preference to advanced ones and
- a resolution of the ultraviolet divergences in quantum electrodynamics.

Rather than go into these aspects in detail we refer the reader to a recent review by Hoyle & Narlikar (1995).

Since cosmology seeks to deal with the large-scale properties of the universe, it inevitably requires a strong connection with fundamental physics. In the big bang cosmology particle physics at very high energy is considered very relevant towards understanding cosmology. In the same spirit we believe that the action at a distance approach to fundamental physics brings about an intimate link of microphysics with cosmology. The Wheeler–Feynman approach is an excellent demonstration of such a connection.

1.2 *Cosmogony*

In this paper, we shall discuss this cosmological model, but first we want to indicate the importance of the observed behavior of the galaxies (the observed cosmogony) in this approach.

Now that theoretical cosmologists have begun to look with favor on the concepts of scalar negative energy fields, and the creation process, they have taken the position that this subject can only be investigated by working out models based on classical approaches of high energy physics and their effects on the global scale. In all of the discussions of what is called precision cosmology there is no discussion of the remarkable phenomena which have been found in the comparatively nearby universe showing that galaxies themselves can eject what may become, new galaxies. We believe that only when we really understand how individual galaxies and clusters, etc., have formed,

evolve, and die (if they ever do) shall we really understand the overall cosmology of the universe. As was mentioned earlier, the method currently used in the standard model is to suppose that initial quantum fluctuations were present at an unobservable epoch in the early universe, and then try to mimic the building of galaxies using numerical methods, invoking the dominance of non-baryonic matter and dark energy for which there is no independent evidence.

In one sense we believe that the deficiency of the current standard approach is already obvious. The model is based on only some parts of the observational data. These are: all of the details of the microwave background, the abundances of the light elements, the observed dimming of distant supernovae, and the large-scale distribution of the observed galaxies. This has led to the conclusion that most of the mass-energy making up the universe has properties which are completely unknown to physics. This is hardly a rational position, since it depends heavily on the belief that all of the laws of physics known to us today can be extrapolated back to scales and epochs where nothing is really testable; and that there is nothing new to be learned.

In spite of this, a very persuasive case has been made that all of the observational parameters can be fitted together to develop what is now becoming widely accepted as a new standard model, the so-called ΛCDM model (Spergel *et al.* 2003). There have been some publications casting doubt on this model, particularly as far as the reality of dark energy and cold, dark matter are concerned (Meyers *et al.* 2004; Blanchard *et al.* 2003). It is usual to dismiss them as controversial and to argue that a few dissenting ideas on the periphery of a generally accepted paradigm are but natural. However, it is unfortunately the case that a large fraction of our understanding of the extragalactic universe is being based on the belief that there *was* a beginning and an inflationary phase, and that the seeds of galaxies all originate from that very early phase.

We believe that an alternative approach should be considered and tested by observers and theorists alike. In this scheme the major themes are:

- that the universe is cyclic and there was no initial big bang, and
- *all* of the observational evidence should be used to test the model.

As we shall show, this not only includes the observations which are used in the current standard model, but also the properties and interactions of galaxies and QSOs which are present in the local ($z < 0.1$) universe.

Possibly the most perceptive astronomer in recent history was Viktor Ambartsumian the famous Armenian theorist. Starting in the 1950s and 1960s he (Ambartsumian 1965) stressed the role of explosions in the universe arguing that the associations of galaxies (groups, clusters, etc.) showed a tendency to expand with far larger kinetic energy than is expected by assuming that the gravitational virial condition holds.

We shall discuss the implications of the cluster dynamics in section 6. Here we take up the issue emphasized by Ambartsumian that there apparently exist phenomena in nuclei of galaxies where matter seems to appear with large kinetic energy of motion directed outwards. In section 6, we will also include other phenomena that share the same property, namely explosive creation of matter and energy. We shall refer to such events as mini-creation events.

Since these phenomena appear on the extragalactic scale and involve quasi-stellar objects, active galaxies, powerful radio sources and clusters and groups of galaxies at all redshifts, we believe they must have an intimate connection with cosmology. Indeed, if one looks at standard cosmology, there too the paradigm centres around the ‘big bang’

which is itself an explosive creation of matter and energy. In the big bang scenario, the origin of all of the phenomena is ultimately attributed to a single origin in the very early universe. No connection has been considered by the standard cosmologists between this primordial event and the *mini-creation events* (MCEs, hereafter) that Ambartsumian talked about. In fact, the QSOs and AGN are commonly ascribed to supermassive black holes as 'prime movers'. In this interpretation the only connection with cosmology is that it must be argued that the central black holes are a result of the processes of galaxy formation in the early universe.

In the QSSC we have been trying to relate such mini-creation events (MCEs) directly to the large-scale dynamics of the universe. We show in sections 2–4 that the dynamics of the universe is governed by the frequency and power of the MCEs, and there is a two-way feedback between the two. That is, the universe expands when there is a large MCE activity and contracts when the activity is switched off. Likewise, the MCE activity is large when the density of the universe is relatively large and negligible when the density is relatively small. In short, the universe oscillates between states of finite maximum and minimum densities as do the creation phases in the MCEs.

This was the model proposed by Hoyle *et al.* (1993) and called the *quasi-steady state cosmology* (QSSC in brief). The model was motivated partly by Ambartsumian's ideas and partly by the growing number of explosive phenomena that are being discovered in extragalactic astronomy. In the following sections we discuss the cosmological model and then turn to the various phenomena which are beginning to help us understand the basic cosmogony. Then we discuss and look at the phenomena themselves in the framework of this cosmology. Finally, we discuss some of the basic problems that have been uncovered by the new observations for which no theoretical explanation has so far been proposed.

2. Gravitational equations with creation of matter

The mathematical framework for our cosmological model has been discussed by Hoyle *et al.* (1995; HBN hereafter), and we outline briefly its salient features. To begin with, it is a theory that is derived from an action principle based on Mach's Principle, and assumes that the inertia of matter owes its origin to other matter in the universe. This leads to a theoretical framework wider than general relativity as it includes terms relating to inertia and creation of matter. These are explained in the Appendix, and we use the results derived there in the following discussion.

Thus the equations of general relativity are replaced in the theory by

$$R_{ik} - \frac{1}{2}g_{ik}R + \lambda g_{ik} = 8\pi G\left[T_{ik} - f\left(C_iC_k - \frac{1}{4}g_{ik}C^lC_l\right)\right], \tag{2}$$

with the coupling constant f defined as

$$f = \frac{2}{3\tau^2}. \tag{3}$$

[We have taken the speed of light $c = 1$.] Here $\tau = \hbar/m_P$ is the characteristic lifetime of a Planck particle with mass $m_P = \sqrt{3\hbar/8\pi G}$. The gradient of C with respect to space-time coordinates $x^i (i = 0, 1, 2, 3)$ is denoted by C_i. Although the above equation defines f in terms of the fundamental constants it is convenient to keep its

identity on the right hand side of Einstein's equations since there we can compare the C-field energy tensor directly with the matter tensor. Note that because of positive f, the C-field has $negative$ kinetic energy. Also, as pointed out in the Appendix, the constant λ is $negative$ in this theory.

The question now arises of why astrophysical observation suggests that the creation of matter occurs in some places but not in others. For creation to occur at the points $A_0, B_0, \ldots$ it is necessary classically that the action should not change (i.e., it should remain stationary) with respect to small changes in the space-time positions of these points, which can be shown to require

$$C_i(A_0)C^i(A_0) = C_i(B_0)C^i(B_0) = \cdots = m_P^2. \tag{4}$$

This is in general not the case: in general the magnitude of $C_i(X)C^i(X)$ is much less that m_P^2. However, as one approaches closer and closer to the surface of a massive compact body $C_i(X)C^i(X)$ is increased by a general relativistic time dilatation factor, whereas m_P stays fixed.

This suggests that we should look for regions of strong gravitational field such as those near collapsed massive objects. In general relativistic astrophysics such objects are none other than black holes, formed from gravitational collapse. Theorems by Penrose, Hawking and others (see Hawking & Ellis 1973) have shown that provided certain positive energy conditions are met, a compact object undergoes gravitational collapse to a space-time singularity. Such objects become black holes before the singularity is reached. However, in the present case, the negative energy of the C-field intervenes in such a way as to violate the above energy conditions. What happens to such a collapsing object containing a C-field apart from ordinary matter? We argue that such an object does not become a black hole. Instead, the collapse of the object is halted and the object bounces back, thanks to the effect of the C-field. We will refer to such an object as a compact massive object (CMO) or a near-black hole (NBH). In the following section we discuss the problem of gravitational collapse of a dust ball with and without the C-field to illustrate this difference.

3. Gravitational collapse and bounce

Consider how the classical problem of gravitational collapse is changed under the influence of the negative energy C-field. First we describe the classical problem which was first discussed by B. Datt (1938). We write the space-time metric inside a collapsing homogeneous dust ball in comoving coordinates (t, r, θ, ϕ) as

$$\mathrm{d}s^2 = \mathrm{d}t^2 - a^2(t)\left[\frac{\mathrm{d}r^2}{1-\alpha r^2} + r^2(\mathrm{d}\theta^2 + \sin^2\theta \mathrm{d}\phi^2)\right], \tag{5}$$

where r, θ, ϕ are constant for a typical dust particle and t is its proper time. Let the dust ball be limited by $r \leq r_b$.

In the above problem we may describe the onset of collapse at $t = 0$ with $a(0) = 1$ and $\dot{a}(0) = 0$. The starting density ρ_0 is related to the constant α by

$$\alpha = \frac{8\pi G\rho_0}{3}. \tag{6}$$

The field equations (2) *without* the C-field and the cosmological constant then tell us that the equation of collapse is given by

$$\dot{a}^2 = \alpha\left(\frac{1-a}{a}\right), \tag{7}$$

and the space-time singularity is attained when $a(t) \to 0$ as $t \to t_S$, where

$$t_S = \frac{\pi}{2\sqrt{\alpha}}. \tag{8}$$

Note that we have ignored the λ-term as it turns out to have a negligible effect on objects of size small compared to the characteristic size of the universe.

The collapsing ball enters the event horizon at a time $t = t_H$ when

$$r_b a(t_H) = 2GM, \tag{9}$$

where the gravitational mass of the dust ball is given by

$$M = \frac{4\pi}{3} r_b^3 \rho_0 = \frac{\alpha r_b^3}{2G}. \tag{10}$$

This is the stage when the ball becomes a black hole.

When we introduce an ambient C-field into this problem, it gets modified as follows. In the homogeneous situation under discussion, C is a function of t only. Let, as before $a(0) = 1$, $\dot{a}(0) = 0$ and let $\dot{C}$ at $t = 0$, be given by β. Then it can be easily seen that the equation (7) is modified to

$$\dot{a}^2 = \alpha\left(\frac{1-a}{a}\right) - \gamma\left(\frac{1-a}{a^2}\right), \tag{11}$$

where $\gamma = 2\pi G f \beta^2 > 0$. Also the earlier relation (6) is modified to

$$\alpha = \frac{8\pi G\rho_0}{3} - \gamma. \tag{12}$$

It is immediately clear that in these modified circumstances $a(t)$ cannot reach zero, the space-time singularity is averted and the ball bounces at a minimum value $a_{\min} > 0$, of the function $a(t)$.

Writing $\mu = \gamma/\alpha$, we see that the second zero of $\dot{a}(t)$ occurs at $a_{\min} = \mu$. Thus even for an initially weak C-field, we get a bounce at a finite value of $a(t)$.

But what about the development of a black hole? The gravitational mass of the black hole at any epoch t is estimated by its energy content, i.e., by,

$$M(t) = \frac{4\pi}{3} r_b^3 a^3(t)\left\{\rho - \frac{3}{4} f\dot{C}^2\right\}$$

$$= \frac{\alpha r_b^3}{2G}\left(1 + \mu - \frac{\mu}{a}\right). \tag{13}$$

Thus the gravitational mass of the dust ball *decreases* as it contracts and consequently its effective Schwarzschild radius decreases. This happens because of the

reservoir of negative energy whose intensity rises faster than that of dust density. Such a result is markedly different from that for a collapsing object with positive energy fields only. From (13) we have the ratio

$$F \equiv \frac{2GM(t)}{r_b a(t)} = \alpha r_b^2 \left\{ \frac{1+\mu}{a} - \frac{\mu}{a^2} \right\}. \tag{14}$$

Hence,

$$\frac{dF}{da} = \frac{\alpha r_b^2}{a^2} \left\{ \frac{2\mu}{a} - (1+\mu) \right\}. \tag{15}$$

We anticipate that $\mu \ll 1$, i.e., the ambient C-field energy density is much less than the initial density of the collapsing ball. Thus F increases as a decreases and it reaches its maximum value at $a \cong 2\mu$. This value is attainable, being larger than $a_{\min}$. Denoting this with $F_{\max}$, we get

$$F_{\max} \cong \frac{\alpha r_b^2}{4\mu}. \tag{16}$$

In general $\alpha r_b^2 \ll 1$ for most astrophysical objects. For the Sun, $\alpha r_b^2 \cong 4 \times 10^{-8}$, while for a white dwarf it is $\sim 4 \times 10^{-6}$. We assume that μ, although small compared to unity, exceeds such values, thus making $F_{\max} < 1$. *In such circumstances black holes do not form.*

We consider scenarios in which the object soon after bounce picks up high outward velocity. From (11) we see that maximum outward velocity is attained at $a = 2\mu$ and it is given by

$$\dot{a}_{\max}^2 \approx \frac{\alpha}{4\mu}. \tag{17}$$

As $\mu \ll 1$, we expect $\dot{a}_{\max}$ to attain high values. Likewise the C-field gradient ($\dot{C}$ in this case) will attain high values in such cases.

Thus, such objects after bouncing at $a_{\min}$ will expand and as $a(t)$ increases the strength of the C-field falls while for small $a(t)$ $\dot{a}$ increases rapidly as per equation (11). This expansion therefore resembles an explosion. Further, the high local value of the C-field gradient will trigger off creation of Planck particles. We will return to this explosive phase in section 7 to illustrate its relevance to high energy phenomena.

It is worth stressing here that even in classical general relativity, the external observer never lives long enough to observe the collapsing object enter the horizon. Thus all claims to have observed black holes in X-ray sources or galactic nuclei really establish the existence of compact massive objects, and as such they are consistent with the NBH concept. A spinning NBH, for example can be approximated by the Kerr solution limited to region outside the horizon (in an NBH there is no horizon). In cases where $\dot{C}$ has not gone to the level of creation of matter, an NBH will behave very much like a Kerr black hole.

The theory would profit most from a quantum description of the creation process. The difficulty, however, is that Planck particles are defined as those for which the Compton wavelength and the gravitational radius are essentially the same, which means

that unlike other quantum processes, flat space-time cannot be used in the formulation of the theory. A gravitational disturbance is necessarily involved and the ideal location for triggering creation is that near a CMO. The C-field boson far away from a compact object of mass M may not be energetic enough to trigger the creation of a Planck particle. On falling into the strong gravitational field of a sufficiently compact object, however, the boson energy is multiplied by a factor $(1 - 2GM/r)^{-1/2}$ for a local Schwarzschild metric.

Bosons then multiply up in a cascade, one makes two, two make four, ..., as in the discharge of a laser, with particle production multiplying up similarly and with negative pressure effects ultimately blowing the system apart. This is the explosive event that we earlier referred to as a *mini-creation event* (MCE). Unlike the big bang, however, the dynamics of this phenomenon is *well defined and non-singular*. For a detailed discussion of the role of a NBH as well as the mode of its formation, see Hoyle *et al.* (2000), (HBN hereafter) p. 244–249.

While still qualitative, we shall show that this view agrees well with the empirical facts of observational astrophysics. For, as mentioned in the previous section, we do see several explosive phenomena in the universe, such as jets from radio sources, gamma-ray bursts, X-ray bursters, QSOs and active galactic nuclei, etc. Generally it is assumed that a black hole plays the lead role in such an event by somehow converting a fraction of its huge gravitational energy into large kinetic energy of the 'burst' kind. In actuality, we do not see infalling matter that is the signature of a black hole. Rather one sees outgoing matter and radiation, which agrees very well with the explosive picture presented above.

4. Cosmological models

The qualitative picture described above is too difficult and complex to admit an exact solution of the field equations (2). The problem is analogous to that in standard cosmology where a universe with inhomogeneity on the scale of galaxies, clusters, superclusters, etc., as well as containing dark matter and radiation is impossible to describe exactly by a general relativistic solution. In such a case one starts with simplified approximations as in models of Friedmann and Lemaitre and then puts in specific details as perturbation. The two phases of radiation-dominated and matter-dominated universe likewise reflect approximations implying that in the early stages the relativistic particles and photons dominated the expansion of the universe whereas in the later stages it was the non-relativistic matter or dust, that played the major role in the dynamics of the universe.

In the same spirit we approach the above cosmology by a mathematical idealization of a homogeneous and isotropic universe in which there are regularly phased epochs when the MCEs were active and matter creation took place while between two consecutive epochs there was no creation (the MCEs lying dormant). We will refer to these two situations as creative and non-creative modes. In the homogeneous universe assumed here, the C-field will be a function of cosmic time only. We will be interested in the matter-dominated analogues of the standard models since, as we shall see, the analogue of the radiation-dominated state never arises except locally in each MCE where, however, it remains less intense than the C-field. In this approximation, the increase or decrease of the scale factor $S(t)$ of the universe indicates an average

smoothed out effect of the MCEs as they are turned on or off. The following discussion is based on the work of Sachs *et al.* (1996).

We write the field equations (2) for the Robertson–Walker line element with $S(t)$ as scale factor and k as curvature parameter and for matter in the form of dust, when they reduce to essentially two independent equations:

$$2\frac{\ddot{S}}{S} + \frac{\dot{S}^2 + k}{S^2} = 3\lambda + 2\pi G f \dot{C}^2, \tag{18}$$

$$\frac{3(\dot{S}^2 + k)}{S^2} = 3\lambda + 8\pi G\rho - 6\pi G f \dot{C}^2, \tag{19}$$

where we have set the speed of light $c = 1$ and the density of dust is given by ρ. From these equations we get the conservation law in the form of an identity:

$$\frac{d}{dS}\{S^3(3\lambda + 8\pi G\rho - 6\pi G f \dot{C}^2)\} = 3S^2\{3\lambda + 2\pi G f \dot{C}^2\}. \tag{20}$$

This law incorporates "creative" as well as "non-creative" modes. We will discuss both in that order.

4.1 *The creative mode*

This has

$$T^{ik}_{;k} \neq 0 \tag{21}$$

which, in terms of our simplified model becomes

$$\frac{d}{dS}(S^3\rho) \neq 0. \tag{22}$$

For the case $k = 0$, we get a simple steady-state de Sitter type solution with

$$\dot{C} = m, \qquad S = \exp(t/P), \tag{23}$$

and from (18) and (19) we get

$$\rho = fm^2, \qquad \frac{1}{P^2} = \frac{2\pi G\rho}{3} + \lambda. \tag{24}$$

Since $\lambda < 0$, we expect that

$$\lambda \approx -\frac{2\pi G\rho}{3}, \qquad \frac{1}{P^2} \ll |\lambda|, \tag{25}$$

but will defer the determination of P to after we have looked at the non-creative solutions. Although Sachs *et al.* (1996) have discussed all cases, we will concentrate on the simplest one of flat space $k = 0$.

The rate of creation of matter is given by

$$J = \frac{3\rho}{P}. \tag{26}$$

As will be seen in the quasi-steady state case, this rate of creation is an overall average made of a large number of small events. Further, since the creation activity has ups and downs, we expect J to denote some sort of temporal average. This will become clearer after we consider the non-creative mode and then link it to the creative one.

4.2 *The non-creative mode*

In this case $T^{ik}_{;k} = 0$ and we get a different set of solutions. The conservation of matter alone gives

$$\rho \propto \frac{1}{S^3}, \tag{27}$$

while for (27) and a constant λ, (20) leads to

$$\dot{C} \propto \frac{1}{S^2}. \tag{28}$$

Therefore, equation (19) gives

$$\frac{\dot{S}^2 + k}{S^2} = \lambda + \frac{A}{S^3} - \frac{B}{S^4}, \tag{29}$$

where A and B are positive constants arising from the constants of proportionality in (27) and (28). We now find that the exact solution of (29) in the case $k = 0$, is given by

$$S = \bar{S}[1 + \eta \cos\theta(t)] \tag{30}$$

where η is a parameter and the function $\theta(t)$ is given by

$$\dot{\theta}^2 = -\lambda(1 + \eta\cos\theta)^{-2}\{6 + 4\eta\cos\theta + \eta^2(1 + \cos^2\theta)\}. \tag{31}$$

Here, $\bar{S}$ is a constant and the parameter η satisfies the condition: $|\eta| < 1$. Thus the scale factor never becomes zero and the model oscillates between finite scale limits

$$S_{\min} \equiv \bar{S}(1 - \eta) \leq S \leq \bar{S}(1 + \eta) \equiv S_{\max}. \tag{32}$$

The density of matter and the C-field energy density are given by

$$\bar{\rho} = -\frac{3\lambda}{2\pi G}(1 + \eta^2), \tag{33}$$

$$f\dot{C}^2 = -\frac{\lambda}{2\pi G}(1 - \eta^2)(3 + \eta^2), \tag{34}$$

while the period of oscillation is given by

$$Q = \frac{1}{\sqrt{-\lambda}} \int_0^{2\pi} \frac{(1 + \eta\cos\theta)d\theta}{\{6 + 4\eta\cos\theta + \eta^2(1 + \cos^2\theta)\}^{1/2}}. \tag{35}$$

The oscillatory solution can be approximated by a simpler sinusoidal solution with the same period:

$$S \approx 1 + \eta\cos\frac{2\pi t}{Q}. \tag{36}$$

Thus the function $\theta(t)$ is approximately proportional to t.

Notice that there is considerable similarity between the oscillatory solution obtained here and that discussed by Steinhardt & Turok (2002) in the context of a scalar field arising from phase transition. The bounce at finite minimum of scale factor is produced in both cosmologies through a negative energy scalar field. As we pointed out in the introduction, Hoyle & Narlikar (1964) [*see also* Narlikar (1973)] have emphasized the fact that such a scalar field can produce models which oscillate between finite ranges of scale. In the Hoyle–Narlikar paper cited above $\dot{C} \propto 1/S^3$, as opposed to (28), exactly as assumed by Steinhardt & Turok (2002) 38 years later. This is because instead of the trace-free energy tensor of equation (2) here, Hoyle and Narlikar had used the standard scalar field tensor given by

$$-f\left(C_i C_k - \frac{1}{2} g_{ik} C_l C^l\right). \tag{37}$$

Far from being dismissed as physically unrealistic, negative kinetic energy fields like the C-field are gaining popularity. Recent works by Rubano & Seudellaro (2004); Sami & Toporensky (2004); Singh *et al.* (2003) who refer to the earlier work by Hoyle & Narlikar (1964) have adapted the same ideas to describe phantom matter and the cosmological constant. In these works, solutions of vacuum field equations with a cosmological constant are interpreted as a steady state in which matter or entropy is being continuously created. Barrow *et al.* (2004) who obtain bouncing models similar to ours refer to the paper by Hoyle & Narlikar (1963) where C-field idea was proposed in the context of the steady state theory.

4.3 *The quasi-steady state solution*

The quasi-steady state cosmology is described by a combination of the creative and the non-creative modes. For this the general procedure to be followed is to look for a composite solution of the form

$$S(t) = \exp\left(\frac{t}{P}\right)\{1 + \eta \cos\theta(t)\} \tag{38}$$

wherein $P \gg Q$. Thus over a period Q as given by (35), the universe is essentially in a non-creative mode. However, at regular instances separated by the period Q it has injection of new matter at such a rate as to preserve an average rate of creation over period P as given by J in (26). It is most likely that these epochs of creation are those of the minimum value of the scale factor during oscillation when the level of the C-field background is the highest. There is a sharp drop at a typical minimum but the $S(t)$ is a continuous curve with a zero derivative at $S = S_{\min}$.

Suppose that matter creation takes place at the minimum value of $S = S_{\min}$, and that N particles are created per unit volume with mass m_0. Then the extra density added at this epoch in the creative mode is

$$\Delta\rho = m_0 N. \tag{39}$$

After one cycle the volume of the space expands by a factor $\exp(3Q/P)$ and to restore the density to its original value we should have

$$(\rho + \Delta\rho)e^{-3Q/P} = \rho, \quad \text{i.e.,} \quad \Delta\rho/\rho \cong 3Q/P. \tag{40}$$

The C-field strength likewise takes a jump at creation and declines over the following cycle by the factor $\exp(-4Q/P)$. Thus the requirement of "steady state" from cycle to cycle tells us that the change in the strength of $\dot{C}^2$ must be

$$\Delta\dot{C}^2 = \frac{4Q}{P}\dot{C}^2. \tag{41}$$

The above result is seen to be consistent with (40) when we take note of the conservation law (20). A little manipulation of this equation gives us

$$\frac{3}{4}\frac{1}{S^4}\frac{d}{dS}(f\dot{C}^2S^4) = \frac{1}{S^3}\frac{d}{dS}(\rho S^3). \tag{42}$$

However, the right-hand side is the rate of creation of matter per unit volume. Since from (40) and (41) we have

$$\frac{\Delta\dot{C}^2}{\dot{C}^2} = \frac{4}{3}\frac{\Delta\rho}{\rho}, \tag{43}$$

and from (23) and (24) we have $\rho = f\dot{C}^2$, we see that (42) is deducible from (40) and (41).

To summarize, we find that the composite solution properly reflects the quasi-steady state character of the cosmology in that while each cycle of duration Q is exactly a repeat of the preceding one, over a long timescale the universe expands with the de Sitter expansion factor $\exp(t/P)$. The two timescales P and Q of the model thus turn out to be related to the coupling constants and the parameters λ, f, G, η of the field equations. Further progress in the theoretical problem can be made after we understand the quantum theory of creation by the C-field.

These solutions contain sufficient number of arbitrary constants to assure us that they are generic, once we make the simplification that the universe obeys the Weyl postulate and the cosmological principle. The composite solution can be seen as an illustration of how a non-creative mode can be joined with the creative mode. More possibilities may exist of combining the two within the given framework. We have, however, followed the simplicity argument (also used in the standard big bang cosmology) to limit our present choice to the composite solution described here. HBN have used (38), or its approximation

$$S(t) = \exp\left(\frac{t}{P}\right)\left\{1 + \eta\cos\frac{2\pi t}{Q}\right\} \tag{44}$$

to work out the observable features of the QSSC, which we shall highlight next.

5. The astrophysical picture

5.1 *Cosmological parameters*

Coming next to a physical interpretation of these mathematical solutions, we can visualize the above model in terms of the following values of its parameters:

$$P = 20Q, \quad Q = 5 \times 10^{10}\,\text{yrs}, \quad \eta = 0.811,$$
$$\lambda = -0.358 \times 10^{-56}(\text{cm})^{-2}. \tag{45}$$

To fix ideas, we have taken the maximum redshift $z_{\max} = 5$ so that the scale factor at the present epoch S_0 is determined from the relation $S_0 = \bar{S}(1 - \eta)(1 + z_{\max})$. This set of parameters has been used in recent papers on the QSSC (Narlikar *et al.* 2002, 2003). For this model the ratio of maximum to minimum scale factor in any oscillation is around 9.6.

These parametric values are not uniquely chosen; they are rather indicative of the magnitudes that may describe the real universe. For example, $z_{\max}$ could be as high as 10 without placing any strain on the model. The various observational tests seek to place constraints on these values. Can the above model quantified by the above parameters cope with such tests? If it does we will know that the QSSC provides a realistic and viable alternative to the big bang.

5.2 *The radiation background*

As far as the origin and nature of the CMBR is concerned we use a fact that is always ignored by standard cosmologists. If we suppose that most of the ^{4}He found in our own and external galaxies (about 24% of the total by mass) was synthesized by hydrogen burning in stars, the energy released amounts to about 4.37×10^{-13} erg cm^{-3}. This is almost exactly equal to the energy density of the microwave background radiation with $T = 2.74$ K. For standard cosmologists this has to be dismissed as a coincidence, but for us it is a powerful argument in favor of the hypothesis that the microwave radiation at the level detected is relic starlight from previous oscillations in the QSSC which has been thermalized (Hoyle *et al.* 1994). Of course, this coincidence loses its significance in the standard big bang cosmology where the CMBR temperature is epoch-dependent.

It is then natural to suppose that the other light isotopes, namely D, ^{3}He, ^{6}Li, ^{7}Li, ^{9}Be, ^{10}B and ^{11}B were produced by stellar processes. It has been shown (cf. Burbidge & Hoyle 1998) that both spallation and stellar flares (for ^{2}D) on the surfaces of stars can explain the measured abundances. Thus *all* of the isotopes are ultimately a result of stellar nucleosynthesis (Burbidge *et al.* 1957; Burbidge & Hoyle 1998).

This option raises a problem, however. If we simply extrapolate our understanding of stellar nucleosynthesis, we will find it hard to explain the relatively low metallicity of stars in our Galaxy. This is still an unsolved problem. We believe but have not yet established that it may be that the initial mass function of the stars where the elements are made is dominated by stars which are only able to eject the outer shells while all of the heavy elements are contained in the cores which simply collapse into black holes. Using theory, we can construct a mass function which will lead to the right answer (we think) but it has not yet been done. But of course, our handwaving in this area is no better than all of the speculations that are being made in the conventional approach when it comes to the "first" stars.

The theory succeeds in relating the intensity and temperature of CMBR to the stellar burning activity in each cycle, the result emphasizing the causal relationship between the radiation background and nuclear abundances. But, how is the background thermalized? The metallic whisker shaped grains condensed from supernova ejecta have been shown to effectively thermalize the relic starlight (Hoyle *et al.* 1994, 2000). It has also been demonstrated that inhomogeneities on the observed scale result from the thermalized radiation from clusters, groups of galaxies, etc., thermalized at the minimum of the last oscillation (Narlikar *et al.* 2003). By using a toy model for these

sources, it has been shown that the resulting angular power spectrum has a satisfactory fit to the data compiled by Podariu *et al.* (2001) for the band power spectrum of the CMBR temperature inhomogeneities. Extending that work further we show, in the following, that the model is also consistent with the first- and third-year observations of the Wilkinson Microwave Anisotropy Probe (WMAP) (Page *et al.* 2003; Spergel *et al.* 2007).

Following Narlikar *et al.* (2003) we model the inhomogeneity of the CMBR temperature as a set of small disc-shaped spots, randomly distributed on a unit sphere. The spots may be either 'top hat' type or 'Gaussian' type. In the former case they have sharp boundaries whereas in the latter case they taper outwards. We assume the former for clusters, and the latter for the galaxies, or groups of galaxies, and also for the curvature effect. This is because the clusters will tend to have rather sharp boundaries whereas in the other cases such sharp limits do not exist. The resultant inhomogeneity of the CMBR thus arises from a superposition of random spots of three characteristic sizes corresponding to the three effects – the curvature effects at the last minimum of the the scale factor, clusters, and groups of galaxies. This is given by a seven-parameter model of the angular power spectrum (for more details, see Narlikar *et al.* 2003):

$$\mathcal{C}_l = A_1\ \ l(l+1)e^{-l^2\alpha_1^2} + A_2\ \frac{l^{\gamma-2}}{l+1}[\ \cos\alpha_2 P_l(\cos\alpha_2) - P_{l-1}(\cos\alpha_2)]^2$$
$$+ A_3\ l(l+1)e^{-l^2\alpha_3^2}, \tag{46}$$

where the parameters A_1, A_2, A_3 depend on the number density as well as the typical temperature fluctuation of each kind of spot, the parameters α_1, α_2, α_3 correspond to the multipole value l_p at which the $\mathcal{C}_l$ from each component peaks, and the parameter γ refers to the correlation of the hot spots due to clusters. These parameters are determined by fitting the model to the observations by following the method we have used in (Narlikar *et al.* 2003). We find that the observations favour a constant in place of the first gaussian profile in equation (46), resulting in a six-parameter model with A_1, A_2, A_3, α_2, α_3 and γ as the remaining free parameters. We should mention that the first gaussian profile of equation (46) had been conjectured by Narlikar *et al.* (2003) to be related to the signature of space-time curvature at the last minimum scale of oscillation. This conjecture was analogous to the particle horizon in the standard cosmology. In the QSSC, there is no particle horizon and the current observations suggest that the curvature effect on CMBR inhomogeneity is negligible.

For the actual fitting, we consider the WMAP-three year data release (Spergel *et al.* 2007). The data for the mean value of TT power spectrum have been binned into 39 bins in multipole space. We find that the earlier fit (Narlikar *et al.* 2003) of the model is worsened when we consider the new data, giving $\chi^2 = 129.6$ at 33 degrees of freedom. However, we should note that while the new dataset (WMAP-three year) has generally increased its accuracy, compared with the WMAP-one year observations, for $l \leq 700$, the observations for higher l do not seem to agree. This is clear from Fig. 1 where we have shown these two observations simultaneously. If we exclude the last three points from the fit, we can have a satisfactory fit giving $\chi^2 = 83.6$ for the best-fitting parameters $A_1 = 890.439 \pm 26.270$, $A_2 = 2402543.93 \pm 3110688.86$, $A_3 = 0.123 \pm 0.033$, $\alpha_2 = 0.010 \pm 0.0001$, $\alpha_3 = 0.004 \pm 0.000004$ and $\gamma = 3.645 \pm 0.206$, we shall see in the following that the standard cosmology also supplies a similar fit to

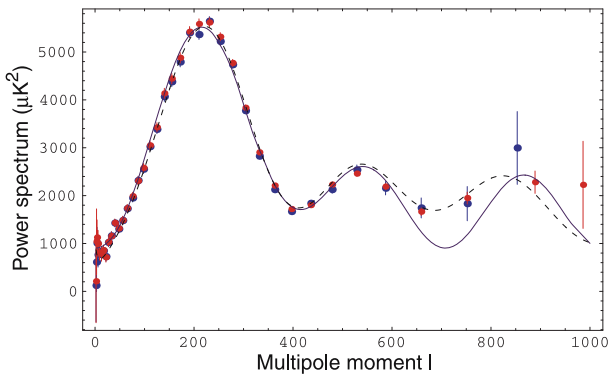

Figure 1. We plot the best-fitting angular power spectrum curves to the WMAP-three year data averaged into 39 bins. The continuous curve corresponds to the QSSC with 6 parameters and the dashed one to the big bang model with $\Omega_m = 0.2$, $\Omega_\Lambda = 0.8$. We notice that the highest part of contribution to χ^2 is from the last three points and the first four points of the data, on which the observations have not settled yet, as becomes clear from the comparison of these data with the WMAP-one year data. The rest of the points have reasonable fits with the theoretical curves.

the data. It should be noted that the above-mentioned parameters in the QSSC can be related to the physical dimensions of the sources of inhomogeneities along the lines of Narlikar *et al.* (2003) and are within the broad range of values expected from the physics of the processes.

For comparison, we fitted the same binned data, to the anisotropy spectrum prediction of a grid of open-CDM and Λ-CDM models within the standard big bang cosmology. We varied the matter density, $\Omega_m = 0.1$ to 1 in steps of 0.1; the baryon density, $\Omega_b h^2$ from 0.005 to 0.03 in steps of 0.004 where h is the Hubble constant in units of 100 km s^{-1} Mpc^{-1}; and the age of the universe, t_0 from 10 Gyr to 20 Gyr in steps of 2 Gyr. For each value of Ω_m we considered an open model and one flat where a compensating Ω_Λ was added. For the same binned dataset, we find that the minimum value of χ^2 is obtained for the flat model ($\Omega_m = 0.2 = 1 - \Omega_\Lambda$, $\Omega_b h^2 = 0.021$, $t_0 = 14$ Gyr and $h = 0.75$) with $\chi^2 = 95.9$ for the full data and $\chi^2 = 92.7$ from the first 36 points. Though the fit can be improved marginally by fine tuning the parameters further. However, it should be noted that the error bars (we have used) provided by the WMAP team provide only a rough estimate of the errors, not the exact error bars. For a proper assignment of errors, it is suggested to use the complete Fisher matrix. However, one should note that some components that go into making the Fisher matrix, depend on the particular models. This makes the errors model-dependent which prohibits an independent assessment of the viability of the model. Hence until the model-independent errors are available from the observations, we have to do with our above procedures and qualities of fit for both theories.

Figure 1 shows the best-fitting angular power spectrum curve obtained for QSSC by using the six-parameter model. For comparison, we have also drawn the best-fitting big bang model.

We mention in passing that recent work (Wickramasinghe 2005) indicates that small traces of polarization would be expected in the CMBR wherever it passes through optically thin clouds of iron whiskers. These whiskers being partially aligned along

the intracluster magnetic fields will yield a weak signal of polarization on the scale of clusters or smaller objects.

It should be noted that the small scale anisotropies do not constitute as crucial a test for our model as they do for standard cosmology. Our general belief is that the universe is inhomogeneous on the scales of galaxy-cluster-supercluster and the QSSC model cannot make detailed claims of how these would result in the anisotropy of CMBR. In this respect, the standard model subject to all its assumptions (dark matter, inflation, dark energy, etc.) makes much more focussed predictions of CMBR anisotropy.

It is worth commenting on another issue of an astrophysical nature. The typical QSSC cycle has a lifetime long enough for most stars of masses exceeding ~ 0.5–$0.7\,M_\odot$ to have burnt out. Thus stars from previous cycles will be mostly extinct as radiators of energy. Their masses will continue, however, to exert a gravitational influence on visible matter. The so-called dark matter seen in the outer reaches of galaxies and within clusters may very well be made up, at least in part, of these stellar remnants.

To what extent does this interpretation tally with observations? Clearly, in the big bang cosmology the timescales are not long enough to allow such an interpretation. Nor does that cosmology permit dark matter to be baryonic to such an extent. The constraints on baryonic dark matter in standard cosmology come from

- the origin and abundance of deuterium and
- considerations of large scale structure.

The latter constraint further requires the nonbaryonic matter to be cold. In the QSSC, as has been shown before, these constraints are not relevant. For other observational issues completely handled by the QSSC, see Hoyle *et al.* (2000).

The QSSC envisages stars from previous cycles to have burnt out and remained in and around their parent galaxies as dark matter. These may be very faint white dwarfs, neutron stars and even more massive remnants of supernovae, like near black holes. Their masses may be in the neighbourhood of $M_\odot$, or more, i.e., much larger than planetary or brown dwarf masses. Thus one form of baryonic dark matter could be in such remnants. In this connection results from surveys like MACHO or OGLE would provide possible constraints on this hypothesis. We should mention here that unlike the standard cosmology, the QSSC does not have limits on the density of baryonic matter from considerations of deuterium production or formation of large-scale structure.

6. Explosive cosmogony

6.1 *Groups and clusters of galaxies*

We have already stated that it was Ambartsumian (1965) who first pointed out that the simplest interpretation of many physical systems of galaxies ranging from very small groups to large clusters is that they are expanding and coming apart. Since most of the observations are of systems at comparatively small redshifts it is clear that this takes place at the current epoch, and while we do not have direct evidence of the situation at large redshifts, it is most likely a general phenomenon.

Why has this effect been so widely ignored? The answer to this is clearly related to the beliefs of earlier generations of cosmologists. From a historical point of view, the

first physical clusters were identified in the 1920s, and it was Zwicky, and later others who supposed that they must be stable systems. By measuring individual redshifts of a number of the galaxies in such a cluster it is possible to get a measurement of the line-of-sight random motions. For stability the virial condition $2E_K + \Omega = O$ needs to be satisfied where E_K and Ω are the average values of the kinetic energy and potential energy of the cluster members. Extensive spectroscopic studies from the 1950s onward showed that nearly always, the kinetic energy of the visible matter far exceeds the potential energy apparent from the visible parts of the galaxies. Many clusters have structures which suggest they are stable and relaxed. Thus it was deduced that in these clusters there must be enough dark matter present to stabilize them. This was, originally, one of the first pieces of evidence for the existence of dark matter.

The other argument was concerned with the ages of the galaxies. Until fairly recently it has been argued that all galaxies have stellar populations which include stars which are very old, with ages of the order of H_o^{-1}, i.e., that they are all as old as the classic big bang universe. However we now know that young galaxies with ages $\ll H_o^{-1}$ do exist. But the major point made by Ambartsumian was, and is, that there are large numbers of clusters of galaxies, and many small groups, which are physically connected but clearly from their forms and their relative velocities, appear to be unstable.

In this situation the use of the virial theorem is totally inappropriate. It is worthwhile pointing out that if the virial theorem holds, the random motions of the galaxies should follow a steady state distribution such as

$$F(\mathbf{v}) \propto \exp\left[-\frac{\mathbf{v}^2}{2\sigma^2}\right]. \tag{47}$$

So far there is no observational demonstration that this is indeed the case. The conclusion drawn from $2E_K + \Omega > O$ as based on visible components only should rather be that the clusters are manifestly *not* in dynamical equilibrium.

Unfortunately, over the last thirty years the virial approach has been wedded to the idea that all galaxies are old, and it is this misreading of the data that led to the view that most galaxies were formed in the early universe and cannot be forming now. For example, Ostriker *et al.* (1974) argued in a very influential paper that the masses of physical systems of galaxies increase linearly with their sizes. As one of us pointed out at the time (Burbidge 1975) this result was obtained completely by assuming that at every scale, for binary galaxies, very small groups, larger groups, and rich clusters, the virial condition of stability holds. Thus it was argued that more and more dark matter is present as the systems get bigger.

Modern evidence concerning the masses of clusters has been obtained from X-ray studies, the Sunyaev–Zeldovich effect, and gravitational lensing (cf. Fabian 1994; Carlstrom *et al.* 2002; Fort & Mellier 1994 and many other papers). All of these studies of rich clusters of galaxies show that large amounts of matter in the form of hot gas and/or dark matter must be present. However, evidence of enough matter to bind small or irregular clusters has not been found in general, and these are the types of configurations which Ambartsumian was originally considering. A system such as the Hercules Cluster is in this category. Also the very compact groups of galaxies (cf. Hickson 1997) have been a subject of debate for many years since a significant fraction of them ($\sim$ 40%) contain one galaxy with a redshift very different from the others. Many statistical studies of these have been made, the orthodox view being that such

galaxies must be "interlopers"; foreground or background galaxies. Otherwise they either have anomalous redshifts, or are exploding away from the other galaxies.

We also have the problem of interacting galaxies, briefly referred to earlier in section 1. In modern times it has been generally supposed that when two galaxies are clearly in interaction they must be coming together (merging) and never coming apart. There are valid ways of deciding whether or not mergers are, or have occurred. The clearest way to show that they are coming together is to look for tidal tails (Toomre & Toomre 1972), or, if they are very closely interwoven, to look for two centers, or two counter rotating systems. For some objects this evidence does exist, and mergers are well established. But to assume that merging is occurring in all cases is unreasonable: there may well be systems where we are seeing the ejection of one galaxy from another as Ambartsumian proposed. Thus when the virial condition is not satisfied, and the systems are highly irregular and appear to be coming apart, then perhaps they *are* coming apart, and never have been separate. Here we are clearly departing from the standard point of view.

If one assumes that clusters may not be bound, their overall astrophysics changes from that of bound 'steady' clusters. Issues like the nature of intracluster medium, the role of the halo, generation of X-rays will require a new approach in the case where clusters are expanding. Further, the ejection of new matter provides additional inputs to the dynamics of the system. For example, the energy of ejection will play a role in heating the intracluster gas. This important investigation still needs to be carried out. However, a preliminary discussion may be found in Hoyle *et al.* (2000), chapter 20.

6.2 *Explosions in individual galaxies*

By the early 1960s, it had become clear that very large energy outbursts are taking place in the nuclei of galaxies.

The first evidence came from the discovery of powerful radio sources and the realization that the nuclei of the galaxies which they were identified with, had given rise to at least 10^{59}–10^{61} ergs largely in the form of relativistic (Gev) particles and magnetic flux which had been ejected to distances of ≥ 100 kpc from the region of production.

A second line of evidence comes from the classical Seyfert galaxies which have very bright star-like nuclei which show very blue continua, and highly excited gas which has random motions $\gtrsim 3000$ km sec^{-1}, and must be escaping from the nucleus. We know that the gas is being ejected because we see it through absorption in optical and X-ray spectra of Seyfert nuclei, and the wavelengths of the absorption lines are shifted to the blue of the main emission. The speeds observed are very large compared with the escape velocity. Early data were described by Burbidge *et al.* (1963).

In the decades since then it has been shown that many active nuclei are giving rise to X-rays, and to relativistic jets, detected in the most detail as high frequency radio waves. A very large fraction of all of the energy which is detected in the compact sources is non-thermal in origin, and is likely to be incoherent synchrotron radiation or Compton radiation.

Early in the discussion of the origin of these very large energies it was concluded that the only possible energy sources are gravitational energy associated with the collapse of a large mass, and the ejection of a small fraction of the energy, or we are indeed seeing mass and energy being created in the nuclei (cf. Hoyle *et al.* 1964).

Of course the most conservative explanation is that the energy arises from matter falling into massive black holes with an efficiency of conversion of gravitational energy to whatever is seen, of order 10%. This is the argument that has been generally advanced and widely accepted (cf. Rees 1984).

Why do we believe that this is not the correct explanation? After all, there is good evidence that many nearby galaxies (most of which are not active) contain collapsed supermassive objects in their centers with masses in the range 10^6–$10^8\,M_\odot$.

The major difficulty is associated with the efficiency with which gravitational energy can be converted into very fast moving gas and relativistic particles, a problem that has haunted us for more than forty years (Burbidge & Burbidge 1965). In our view the efficiency factor is not 10% but close to 0.1%–1%. The reasons why the efficiency factor is very small are the following. If the energy could be converted directly the efficiency might be as high as $\sim 8\%$, or even higher from a Kerr rotating black hole. But this energy will appear outside the Schwarzschild radius as the classical equivalent of gravitons. This energy has to be used to heat an accretion disk or generate a corona in a classical AGN, or generate very high energy particles which can propagate outward in a radio source, then heat gas which gives rise to shock waves, which accelerate particles, which in turn radiate by the synchrotron process. Thermodynamics tells us that the efficiency at each of these stages is $\lesssim 10\%$. If there are 3 to 4 stages the overall efficiency is $\sim 10^{-3}$–10^{-4}. This is borne out by the measured efficiency by which relativistic beams are generated in particle accelerators on earth, and by the efficiency associated with the activity in the center of M87 (cf. Churasov *et al.* 2002).

If these arguments are not accepted, and gravitational energy is still claimed to be the only reasonable source, another problem appears.

For the most luminous sources, powerful radio sources and distant QSOs the masses involved must be much greater than the typical values used by the black hole-accretion disk theorists. If one uses the formula for Eddington luminosity (cf. for details pages 109–111, 408–409 of Kembhavi & Narlikar 1999) one arrives at black hole masses of the order $10^8\,M_\odot$ on the basis of perfect efficiency of energy conversion. An efficiency of ≤ 0.01 would drive the mass up a hundred fold at least, i.e., to $10^{10}\,M_\odot$ or greater. So far there is no direct evidence in any galaxy for such large dark masses. The largest masses which have been reliably estimated are about $10^9\,M_\odot$.

In general it is necessary to explain where the bulk of the energy released which is not in the relativistic particle beams, is to be found. A possible explanation is that it is much of this energy which heats the diffuse gas in active galaxies giving rise to the extended X-ray emission in clusters and galaxies.

An even harder problem is to explain how the massive black holes in galaxies were formed in the first place. Were they formed before the galaxies or later? In the standard model both scenarios have been tried, but no satisfactory answer has been found.

In our model the energy comes with creation in the very strong gravitational fields very close to the central NBH, where the process can be much more efficient than can be expected in the tortuous chain envisaged in the classical gravitational picture. We shall discuss this in section 7.

Would very massive galaxies result if the universe allows indefinitely large time for galaxy formation? Earlier ideas (Hoyle 1953; Binney 1977; Rees & Ostriker 1977; Silk 1977) seemed to suggest so. In the present case two effects intervene to make massive galaxies rather rare. The first one is geometrical. Because of steady long-term expansion, the distance between two galaxies formed, say, n cycles ago, would

have increased by a factor $\sim \exp(nQ/P)$, and their density decreased by the factor $\sim \exp(-3nQ/P)$. For $n \gg 1$, we expect the chance of finding such galaxies very small.

The second reason working against the growth of mass in a region comes from the negative energy and pressure of the C-field. As the mass grows through creation, the C-field also mounts and its repulsive effect ultimately causes enough instability for the mass to break up. Thus the large mass grows smaller by ejecting its broken parts.

What is ejected in an MCE? Are the ejecta more in the form of particles or radiation or coherent objects? All three are produced. For a discussion of the mechanism leading to ejection of coherent objects, see Hoyle *et al.* (2000), chapter 18.

6.3 *Quasi-stellar objects*

In the early 1960s QSOs were discovered (Matthews & Sandage 1963; Schmidt 1963; cf. Burbidge & Burbidge 1967 for an extensive discussion) as star-like objects with large redshifts. Very early on, continuity arguments led to the general conclusion that they are very similar to the classical Seyfert galaxies, i.e., they are the nuclei of galaxies at much greater distances. However, also quite early in the investigations, it became clear that a good case could also be made for supposing that they are more likely to be compact objects *ejected* from comparatively local, low redshift active galaxies (Hoyle & Burbidge 1966).

After more than thirty years of controversy this issue has not yet been settled, but a very strong case for this latter hypothesis based on the observations of the clustering of many QSOs about active galaxies has been made (Burbidge *et al.* 1971; Arp 1987; Burbidge 1996).

If this is accepted, it provides direct evidence that in the creation process active galaxies are able to eject compact sources with large intrinsic redshifts. What was not predicted was the existence of intrinsic redshifts. They present us with an unsolved problem, but one which must be closely connected to the creation process. A remarkable aspect of this problem is that the intrinsic redshifts show very clear peaks in their distribution with the first peak at $z = 0.061$ and with a periodicity of the form $\triangle \log(1 + z) = 0.089$ (cf. Karlsson 1971; Burbidge & Napier 2001). The periodicity is in the intrinsic redshift component (z_i), and in order to single out that component, either the cosmological redshift component z_c must be very small, i.e., the sources must be very close to us, or it must be known and corrected for by using the relation $(1 + z_{\rm obs}) = (1 + z_c)(1 + z_i)$. Thus a recent claim that the periodicity is not confirmed (Hawkins *et al.* 2002) has been shown to be in error (Napier & Burbidge 2003).

It is admitted that the evidence from gravitational lensing provides an overall consistent picture for the standard cosmological hypothesis. The evidence on quasars of larger redshift being lensed by a galaxy of lower redshift, together with the time delay in the radiation found in the two lensed images can be explained by this hypothesis. This type of evidence needs to be looked at afresh if the claim is made that quasars are much closer than their redshift-distances. In such cases, the lensing models can be 'scaled' down but the time-delay will have to be checked for lower values. To our knowledge no such exercise has been carried out to date. We hope to examine this issue in a later paper.

6.4 *Gamma-ray bursts*

One of the most remarkable phenomena discovered in recent years relate to very short-lived ($\lesssim$ minutes) bursts of high energy photons (γ-ray and X-ray) which can apparently occur anywhere in the sky, and which sometimes can be identified with a very faint optical and/or radio source, an afterglow, which may fade with time. Sometimes a very faint object remains. The first optical observation in which a redshift could be measured led to the conclusion that those sources are extragalactic. Using the redshifts as distance indicators this has led to the conclusion that the energies emitted lie in the range 10^{50}–10^{54} ergs, with most of them $\gtrsim 10^{53}$ ergs, if the explosions take place isotropically. If energies involving single stars are invoked the energies can be reduced if beaming is present. The most recent observations have suggested that the events are due to forms of supernovae which are beamed. In the usual interpretation it is assumed that the redshifts which have been measured for the gamma-ray bursts are cosmological (cf. Bloom 2003). However in a recent study using all (more than 30) gamma-ray bursts (GRBs) with measured redshifts it was shown that the redshift distribution strongly suggests that they are closely related to QSOs with the same intrinsic redshift peaks (Burbidge 2003, 2004). Also an analysis of the positions of all of the GRBs for which we have positions (about 150) shows that a number of them are very near to already identified QSOs (Burbidge 2003). All of this suggests that the GRBs are due to explosions of objects (perhaps *in* QSOs) which have themselves been ejected following a creation process from active galaxies. In general they have slightly greater cosmological redshifts and thus are further away ($\leq$ 500 Mpc) than the galaxies from which most of bright QSOs are ejected. While we do not claim that this hypothesis is generally accepted, Bloom (2003) has shown that there are peculiarities in the redshift distribution interpreted in the conventional way. More observations may clarify this situation.

7. Dynamics and spectrum of radiation from an MCE

A discussion of how a minicreation event arises can be found in section 3. Thus we took the modified problem of a collapsing dust ball in the presence of the C-field as a toy-model of how a realistic massive object would behave. In the classic Oppenheimer-Snyder case the dust ball collapses to become a black hole, eventually ending in space-time singularity. In the modified problem, as we saw in section 3, the dust ball need not become a black hole. It certainly does not attain singularity, but bounces at a finite radius. We saw that after bounce its outward speed rapidly rises before it ultimately slows down to a halt. In the phase of rapid expansion it resembles the classical white hole, which is the reverse of the classical collapse without the C-field. The white hole solution can be used to approximate the behavior of an MCE as seen by an external observer, because the former can be handled exactly in an analytic way. In essence, we use the notation of section 3 with slight modification.

We begin with a discussion of a white hole as considered by Narlikar *et al.* (1974) within the framework of standard general relativity. Consider a massive object emerging from a space-time singularity in the form of an explosion. To simplify matters Narlikar, Apparao & Dadhich (*op. cit.*) considered the object as a homogeneous dust

ball, for which one can use comoving coordinates. As described in section 3, the line element within the object is given by

$$\mathrm{d}s^2 = \mathrm{d}t^2 - a^2(t)\left[\frac{\mathrm{d}r^2}{(1-\alpha r^2)} + r^2(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2)\right], \tag{48}$$

where c = speed of light is taken as unity, $a(t)$ is the expansion factor and α is a parameter related to the mass M and the comoving radius r_b of the object by

$$2GM = \alpha r_b^3. \tag{49}$$

The similarity of equation (48) to the Robertson–Walker line element of cosmology is well known. Also, if we change t to $-t$, equation (48) represents a freely collapsing ball of dust. The parameter α is related to the dust density ρ_0 at $a = 1$, by the relation

$$\alpha = \frac{8\pi G\rho_0}{3}. \tag{50}$$

The formulae (48)–(50) are the same as (5), (10) and (6) of section 3. However, in section 3 we were discussing the contracting phase, while here we are interested in the expanding mode. For convenience therefore, we will measure t from the instant of explosion so that $a(0) = 0$. For $t > 0$, $a(t)$ satisfies the equation

$$\dot{a}^2 = \frac{\alpha(1-a)}{a}, \tag{51}$$

so that it attains its maximum value $a = 1$ at

$$t = t_0 = \frac{\pi}{(2\sqrt{\alpha})}. \tag{52}$$

We will investigate light emission from the white hole in the interval $0 < t < t_0$. The equation (51) can be solved in a parametric form by defining

$$a = \sin^2\xi, \quad 0 \le \xi \le \pi/2. \tag{53}$$

The ξ is related to the comoving time coordinate t by

$$t = \frac{2t_0}{\pi}(\xi - \sin\xi\cos\xi). \tag{54}$$

The white hole bursts out of the Schwarzschild radius at $t = t_c$, $\xi = \xi_c$, where

$$\sin\xi_c = (\alpha r_b^2)^{1/2}. \tag{55}$$

The space exterior to the white hole is described by the Schwarzschild line element

$$\mathrm{d}s^2 = [1 - (2GM/R)]\mathrm{d}T^2 - \frac{\mathrm{d}R^2}{1-(2GM/R)} - R^2(\mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\phi^2). \tag{56}$$

A typical Schwarzschild observer has R = constant, θ = constant, ϕ = constant. We wish to calculate the spectrum of radiation from the white hole as seen by a

Schwarzschild observer with $R = R_1 \gg 2GM$. To simplify matters further, we will take the luminosity spectrum of the white hole as $L\delta(\nu - \nu_0)$, where $L =$ constant.

Suppose two successive light signals are sent out from the surface at comoving instants t and $t + \mathrm{d}t$ and are received by the observer at R_1 at instants T and $T + \mathrm{d}T$ measured in the Schwarzschild coordinates. Then a straightforward calculation shows that

$$\frac{\mathrm{d}T}{\mathrm{d}t} = \frac{\sin\xi}{\sin(\xi + \xi_c)}. \tag{57}$$

So an electromagnetic wave of frequency ν_0 emitted from the surface appears to the receiver to have the frequency

$$\nu = \nu_0 \left[\frac{\sin(\xi + \xi_c)}{\sin\xi}\right]. \tag{58}$$

A result of this type is suitable for working out the spectrum of the radiation as seen by the Schwarzschild observer. Under our assumption $L/h\nu_0$ photons of frequency ν_0 are being emitted per unit $t-$ time from the surface. The number emitted in the interval $[t,\ t + \mathrm{d}t]$ is therefore $L\mathrm{d}t/h\nu_0$. The same number must be received in the interval $[T,\ T + \mathrm{d}T]$, but with frequencies in the range $(\nu,\ \nu + \mathrm{d}\nu)$ where $\mathrm{d}\nu$ is related to $\mathrm{d}t$ through equations (54) and (58). A simple calculation gives

$$\mathrm{d}t = \frac{(4t_0\nu_0^3 \sin^3\xi_c\ \mathrm{d}\nu)}{\pi(\nu^2 + \nu_0^2 - 2\nu\nu_0\cos\xi_c)^2}. \tag{59}$$

Writing $E = h\nu,\ E_0 = h\nu_0$, the number of photons in the range $[E,\ E + \mathrm{d}E]$ received from the white hole per unit area at $R = R_1$ is given by

$$N(E)\,\mathrm{d}E = \frac{Lt_0}{\pi^2 R_1^2} \times \frac{E_0^2 \sin^3\xi_c\,\mathrm{d}E}{(E^2 + E_0^2 - 2EE_0\cos\xi_c)^2}. \tag{60}$$

For $E \gg E_0$

$$N(E)\,\mathrm{d}E \cong Lt_0E_0^2 \times \frac{\sin^3\xi_c}{\pi^2 R_1^2}\frac{\mathrm{d}E}{E^4}. \tag{61}$$

The energy spectrum $I(E)$ is given by

$$I(E) = EN(E) \propto E^{-3}. \tag{62}$$

This is the spectrum at the high energy end under the simplifying assumptions made here. More general (and perhaps more realistic) assumptions can lead to different types of spectra which can also be worked out. Following Narlikar *et al.* (1974) possible fields in high energy astrophysics where MCEs might find applications are as follows:

- The hard electromagnetic radiation from the MCEs situated at the centres of, say Seyfert galaxies, can be a source of background X and gamma radiation. The energy spectrum (60) seems, at first sight to be too steep compared to the observed spectrum $\propto E^{-1.2}$. But absorption effects in the gas present in the nuclei

surrounding the MCE tend to flatten the spectrum given by equation (60). Detailed calculation with available data shows that these absorption effects can in fact flatten the E^{-3} spectrum to $\sim E^{-1}$ form in the range 0.2 keV to 1 keV. At lower energies, the ultraviolet radiation seems to be of the right order of magnitude to account for the infrared emission of $\sim 10^{45}$ erg s^{-1} through the dust grain heating mechanism.

- The transient nature of X-ray and gamma-ray bursts suggests an MCE origin. The shape of the spectrum at the emitting end is likely to be more complicated than the very simple form assumed in the above example. *In general, however, the spectrum should soften with time.*
- Although Narlikar *et al.* (1974) had worked out the spectrum of photons, it is not difficult to see that similar conclusions will apply to particles of non-zero rest mass provided they have very high energy, with the relativistic γ-factor $\gg 1$. It is possible therefore to think of MCEs in the Galaxy on the scale of supernovae, yielding high energy cosmic rays right up to the highest energy observed.

This picture of a white hole gives a quantitative but approximate description of radiation coming out of an MCE, which is without a singular origin and without an event horizon to emerge out of. Ideally we should have used the modified C-field solution described in section 3 to calculate the exact result. This, however has proved to be an intractable problem analytically as an explicit exterior solution is not known.

The collapse problem with an earlier version of the C-field was discussed by Hoyle & Narlikar (1964) in which proof was given that an exterior solution matching the homogeneous dust ball oscillation exists. However an explicit solution could not be given. The same difficulty exists with this solution also and further work, possibly using numerical general relativity, may be required. We mention in passing, that a similar matching problem exists in inflationary models where a Friedmann bubble emerges within an external de Sitter type universe.

The above type of expansion has one signature. Its explosive nature will generate strong blueshifts, thus making the radiation of high frequency, which softens to that at lower frequencies as the expansion slows down. This model therefore has the general features associated with gamma-ray bursts and transient X-ray bursters.

A further generalization of this idea at a qualitative level corresponds to the introduction of spin so as to correspond to the Kerr solution in classical general relativity. If we consider an MCE to have axial symmetry because of spin, the tendency to go round the axis is strong in a region close to the 'equator' and not so strong away from it. In classical general relativity the ergosphere identifies such a region: it shrinks to zero at the poles. At the poles therefore we expect that the ejection outwards will be preferentially directed along the axis and so we may see jets issuing in opposite directions.

In the very first paper on the QSSC, Hoyle *et al.* (1993) had pointed to the similarity between an MCE and the standard early universe. In particular they had shown that the creation of matter in the form of Planck particles leads to their subsequent decay into baryons together with release of very high energy. These 'Planck fireballs' have a density temperature relationship of the form $\rho \propto T^3$ which permits the synthesis of light nuclei just as in the classical big bang model. However, these authors drew attention to the circumstance that the relevant (ρ, T) domain for this purpose in the QSSC is very different from the (ρ, T) domain in the primordial nucleosynthesis of standard cosmology.

8. Concluding remarks

The oscillating universe in the QSSC, together with a long-term expansion, driven by a population of mini-creation events provides the missing dynamical connection between cosmology and the 'local' explosive phenomena. The QSSC additionally fulfills the roles normally expected of a cosmological theory, namely:

- it provides an explanation of the cosmic microwave background with temperature, spectrum and inhomogeneities related to astrophysical processes (Narlikar *et al.* 2003),
- it offers a purely stellar-based interpretation of all observed nuclei (*including* light ones) (Burbidge *et al.* 1957; Burbidge & Hoyle 1998),
- it generates baryonic dark matter as part of stellar evolution (Hoyle *et al.* 1994),
- it accounts for the extra dimming of distant supernovae *without* having recourse to dark energy (Narlikar *et al.* 2002; Vishwakarma & Narlikar 2005), and it also suggests a possible role of MCEs in the overall scenario of structure formation (Nayeri *et al.* 1999).

The last mentioned work shows that preferential creation of new matter near existing concentrations of mass can lead to growth of clustering. A toy model based on million-body simulations demonstrates this effect and leads to clustering with a 2-point correlation function with index close to -1.8. Because of the repulsive effect of the C-field, it is felt that this process may be more important than gravitational clustering. However, we need to demonstrate this through simulations like those in our toy model, *together with* gravitational clustering.

There are two challenges that still remain, namely understanding the *origin* of anomalous redshifts and the observed *periodicities* in the redshifts. Given the QSSC framework, one needs to find a scenario in which the hitherto classical interpretation of redshifts is enriched further with inputs of quantum theory. These are huge problems which we continue to wrestle with.

Acknowledgements

One of us, (JVN) thanks College de France, Paris for hospitality when this work was in process. RGV is grateful to IUCAA for hospitality which facilitated this collaboration.

Appendix

Field theory underlying the QSSC

Following Mach's principle, we begin with the hypothesis that inertia of any particle of matter owes its origin to the existence of all other particles of matter in the universe. If the particles are labelled $a, b, c, \ldots$ and the element of proper time of ath particle in Riemannian space-time is denoted by ds_a, then we express the inertia of particle a by the sum

$$M_a(A) = \sum_{b \neq a} \int \lambda_b \tilde{G}(A, B) ds_b = \sum_{b \neq a} M^{(b)}(A), \tag{A1}$$

where A is a typical point on the world line of particle a. $\tilde{G}(A, B)$ is a scalar propagator communicating the inertial effect from B to A. The coupling constant λ_b denotes the intensity of the effect and without loss of generality may be set equal to unity. Likewise we may replace $M_a(A)$ by a scalar mass function $M(X)$ of a general space-time point X, denoting the mass acquired by a particle at that point. As in Riemannian geometry we will denote by R_{ik} the Ricci tensor and by R the scalar curvature.

The individual contributors to $M(X)$ are the scalar functions $M^{(b)}(X)$, which are determined by the propagators $\tilde{G}(X, B)$. The simplest theory results from choice of a conformally invariant wave equation for $M^{(b)}(X)$,

$$\Box M^{(b)}(X) + \frac{1}{6}RM^{(b)}(X) + M^{(b)}(X)^3 = \int \frac{\delta_4(X, B)}{\sqrt{-g(B)}} ds_b. \tag{A2}$$

The expression on the right-hand side identifies the worldline of b as the source. Why conformal invariance? In a theory of long range interactions influences travel along light cones and light cones are entities which are globally invariant under a conformal transformation. Thus a theory which picks out light cones for global communication is naturally expected to be conformally invariant. (A comparison may be made with special relativity. The *local* invariance of speed of light for all moving observers leads to the requirement of local Lorentz invariance of a physical theory.)

Although the above equation is non-linear, a simplification results in the smooth fluid approximation describing a universe containing a larger number of particles. Thus $M(X) = \sum_b M^{(b)}(X)$ satisfies an equation

$$\Box M + \frac{1}{6}RM + \Lambda M^3 = \sum_b \int \frac{\delta_4(X, B)}{\sqrt{-g(B)}} d^4 s_b. \tag{A3}$$

What is Λ? Assuming that there are N contributing particles in a cosmological horizon size sphere, we will get

$$\Lambda \approx N^{-2}, \tag{A4}$$

since adding N equations of the kind (A2) leads to the cube term having a reduced coefficient by this factor, because of the absence of cross products $M^{(b)}M^{(c)}$ type $(b \neq c)$. Typically the observable mass in the universe is $\sim 10^{22}\mathrm{M}_\odot$ within such a sphere, giving $N \sim 2 \times 10^{60}$ if the mass is typically that of a Planck particle. We shall return to this aspect shortly. With this value for N, we have

$$\Lambda \approx 2.5 \times 10^{-121}. \tag{A5}$$

With these definitions we now introduce the action principle from which the field equations can be derived. In particle–particle interaction form it is simply

$$\mathcal{A} = -\sum_a \int M_a(A) ds_a. \tag{A6}$$

Expressed in terms of a scalar field function $M(X)$, it becomes

$$\mathcal{A} = -\frac{1}{2}\int \left(M_i M^i - \frac{1}{6}RM^2\right)\sqrt{-g}\; d^4x + \frac{1}{4}\Lambda \int M^4 \sqrt{-g} d^4x$$

$$- \sum_a \int \frac{\delta_4(X, A)}{\sqrt{-g(A)}} M(X) ds_a. \tag{A7}$$

For example, the variation $M \to M + \delta M$ leads to the wave equation (A2). The variation of space-time metric gives rise to gravitational equations. The variation of particle world lines gives rise to another scalar field, however, if we assume the worldlines to have finite beginnings. This is where creation of matter explicitly enters the picture. The characteristic mass of a typical particle that can be constructed in the theory using the available fundamental constants c, G and $\hbar$ is the Planck mass

$$m_P = \left(\frac{3\hbar c}{4\pi G}\right)^{1/2}. \tag{A8}$$

We shall assume therefore that the typical basic particle created is the Planck particle with the above mass. We shall take $\hbar = 1$ in what follows. Imagine now the worldline of such a particle beginning at a world-point A_0.

A typical Planck particle a exists from A_0 to $A_0 + \delta A_0$, in the neighborhood of which it decays into n stable secondaries, $n \simeq 6 \cdot 10^{18}$, denoted by $a_1, a_2, \ldots, a_n$. Each such secondary contributes a mass field $m^{(a_r)}(X)$, say, which is the fundamental solution of the wave equation

$$\Box m^{(a_r)} + \frac{1}{6} R m^{(a_r)} + n^2 m^{(a_r)^3} = \frac{1}{n} \int_{A_0+\delta A_0} \frac{\delta_4(X, A)}{\sqrt{-g(A)}} da, \tag{A9}$$

while the brief existence of a contributes $c^{(a)}(X)$, say, which satisfies

$$\Box c^{(a)} + \frac{1}{6} R c^{(a)} + c^{(a)^3} = \int_{A_0}^{A_0+\delta A_0} \frac{\delta_4(X, A)}{\sqrt{-g(A)}} da. \tag{A10}$$

Summing $c^{(a)}$ with respect to $a, b, \ldots$ gives

$$c(X) = \sum_a c^{(a)}(X), \tag{A11}$$

the contribution to the total mass $M(X)$ from the Planck particles during their brief existence, while

$$\sum_a \sum_{r=1}^{n} m^{(a_r)}(X) = m(X) \tag{A12}$$

gives the contribution of the stable secondary particles.

Although $c(X)$ makes a contribution to the total mass function

$$M(X) = c(X) + m(X) \tag{A13}$$

that is generally small compared to $M(X)$, there is the difference that, whereas $m(X)$ is an essentially smooth field, $c(X)$ contains small exceedingly rapid fluctuations and so can contribute significantly to the derivatives of $c(X)$. The contribution to $c(X)$ from Planck particles a, for example, is largely contained between two light cones, one from A_0, the other from $A_0 + \delta A_0$. Along a timelike line cutting these two cones the contribution to $c(X)$ rises from zero as the line crosses the light cone from A_0, attains some maximum value and then falls back effectively to zero as the line crosses the second light cone from $A_0 + \delta A_0$. The time derivative of $c^{(a)}(X)$ therefore involves the

reciprocal of the time difference between the two light cones. This reciprocal cancels the short duration of the source term on the right-hand side of (A10). The factor in question is of the order of the decay time τ of the Planck particles, $\sim 10^{-43}$ seconds. No matter how small τ may be, the reduction in the source strength of $c^{(a)}(X)$ is recovered in the derivatives of $c^{(a)}(X)$, which therefore cannot be omitted from the gravitational equations.

The derivatives of $c^{(a)}(X)$, $c^{(b)}(X)$, ... can as well be negative as positive, so that in averaging many Planck particles, linear terms in the derivatives do disappear. It is therefore not hard to show that after such an averaging the gravitational equations become

$$R_{ik} - \frac{1}{2}g_{ik}R - 3\Lambda m^2 g_{ik} = \frac{6}{m^2}\left[-T_{ik} + \frac{1}{6}(g_{ik}\Box m^2 - m^2_{;ik}) + \left(m_i m_k - \frac{1}{2}g_{ik}m_l m^l\right) + \frac{2}{3}\left(c_i c_k - \frac{1}{4}g_{ik}c_l c^l\right)\right]. \tag{A14}$$

Since the same wave equation is being used for $c(X)$ as for $m(X)$, the theory remains scale invariant. A scale change can therefore be introduced that reduces $M(X) = m(X) + c(X)$ to a constant, or one that reduces $m(X)$ to a constant. Only that which reduces $m(X)$ to a constant, viz.,

$$\Omega = \frac{m(X)}{m_P} \tag{A15}$$

has the virtue of not introducing small very rapidly varying ripples into the metric tensor. Although small in amplitude such ripples produce non-negligible contributions to the derivatives of the metric tensor, causing difficulties in the evaluation of the Riemann tensor, and so are better avoided. Simplifying with (A14) does not bring in this difficulty, which is why separating of the main smooth part of $M(X)$ now proves an advantage, with the gravitational equations simplifying to

$$8\pi G = \frac{6}{m_P^2}, \quad m_P \text{ a constant}, \tag{A16}$$

$$R_{ik} - \frac{1}{2}g_{ik}R + \lambda g_{ik} = -8\pi G\left[T_{ik} - \frac{2}{3}\left(c_i c_k - \frac{1}{4}g_{ik}c_l c^l\right)\right]. \tag{A17}$$

We define the cosmological constant λ by

$$\lambda = -3\Lambda m_P^2 \approx -2 \times 10^{56}\ \text{cm}^{-2}. \tag{A18}$$

This value falls within the normally expected region of the magnitude of the cosmological constant. Note, however, that its sign is negative! This has been the consequence of the Machian origin of the cosmological constant through the non-linear equations (A2), (A3).

It has been on (A17) that the discussion of what is called the quasi-steady state cosmological model (QSSC) has been based. A connection with the C-field of the earlier steady state cosmology can also be given. Writing

$$C(X) = \tau c(X), \tag{A19}$$

where τ is the decay lifetime of the Planck particle, the action contributed by Planck particles $a, b, \ldots,$

$$-\sum_a \int_{A_0}^{A_0+\delta A_0} c(A)da \tag{A20}$$

can be approximated as

$$-C(A_0) - C(B_0) - \cdots, \tag{A21}$$

which form corresponds to the C-field used in the steady state cosmology.

Thus the equation (A17) is replaced by

$$R_{ik} - \frac{1}{2}g_{ik}R + \lambda g_{ik} = -8\pi G\left[T_{ik} - f\left(C_iC_k - \frac{1}{4}g_{ik}C_lC^l\right)\right], \tag{A22}$$

with the earlier coupling constant f defined as

$$f = \frac{2}{3\tau^2}. \tag{A23}$$

[We remind the reader that we have taken the speed of light $c = 1$.]

The question now arises of why astrophysical observation suggests that the creation of matter occurs in some places but not in others. For creation to occur at the points $A_0, B_0, \ldots$ it is necessary classically that the action should not vary with respect to small changes in the space-time positions of these points, which was shown earlier to require

$$C_i(A_0)C^i(A_0) = C_i(B_0)C^i(B_0) = \cdots = m_P^2. \tag{A24}$$

More precisely, the field $c(X)$ is required to be equal to m_P at $A_0, B_0, \ldots,$

$$c(A_0) = c(B_0) = \cdots = m_P. \tag{A25}$$

(For, equation (A19) tells us that connection between c and C is through the lifetime τ of Planck particle.)

As already remarked in the main text, this is in general not the case: in general the magnitude of C^iC_i is much less than m_P. However, close to the event horizon of a massive compact body $C_i(A_0)C^i(A_0)$ is increased by a relativistic time dilatation factor, whereas m_P^2 stays fixed. Hence, near enough to an event horizon the required conservation conditions can be satisfied, which has the consequence that creation events occur only in compact regions, agreeing closely with the condensed regions of high excitation observed so widely in astrophysics.

References

Ambartsumian, V. A. 1965, *Structure and Evolution of Galaxies, Proc. 13th Solvay Conf. on Physics, University of Brussels* (New York: Wiley Interscience), 24.

Arp, H. C. 1987, *Quasars, Redshifts and Controversies* (Interstellar Media, Berkeley, California).

Bagla, J. S., Padmanabhan, T., Narlikar, J. V. 1996, *Comm. Astrophys.*, **18**, 289.

Barrow, J., Kimberly, D., Magueijo, J. 2004, *Class. Quant. Grav.*, **21**, 4289.

Binney, J. 1977, *ApJ*, **215**, 483.
Blanchard, A., Souspis, B., Rowan-Robinson, M., Sarkar, S. 2003, *A&A*, **412**, 35.
Bloom, J. S. 2003, *A.J.*, **125**, 2865.
Bloom, J. S., Kulkarni, S. R., Djorgovsky, S. G. 2001, *A.J.*, **123**, 1111.
Bondi, H., Gold, T. 1948, *MNRAS*, **108**, 252.
Burbidge, E. M., Burbidge, G. R., Fowler, W. A., Hoyle, F. 1957, *Rev. Mod. Phys.*, **29**, 547.
Burbidge, E. M., Burbidge, G., Solomon, P., Strittmatter, P. A. 1971, *ApJ*, **170**, 223.
Burbidge, G. 1975, *ApJ*, **106**, L7.
Burbidge, G. 1996, *A&A*, **309**, 9
Burbidge, G. 2003, *ApJ*, **585**, 112.
Burbidge, G. 2004, "The Restless High Energy Universe", *Conf. Proc. Nuclear Physics B.*, **305**, 132.
Burbidge, G., Burbidge, E. M. 1965, *The Structure and Evolution of Galaxies, Proc. of 13th Solvay Conference on Physics, University of Brussels* (New York: Wiley Interscience), 137.
Burbidge, G., Burbidge, E. M. 1967, *Quasi-Stellar Objects* (San Francisco, W. H. Freeman).
Burbidge, G., Hoyle, F. 1998, *ApJ*, **509**, L1.
Burbidge, G., Napier, W. M. 2001, *A.J.*, **121**, 21.
Burbidge, G., Burbidge, E. M., Sandage, A. 1963, *Rev. Mod. Phys*., **35**, 947.
Carlstrom, J., Holder, G., Reese, E. 2002, *ARAA*, **40**, 643.
Carroll, S. M., Press, W. H. 1992, *ARAA*, **30**, 499.
Churasov, E., Sunyaev, R., Forman, W., Bohringer, H. 2002, *MNRAS*, **332**, 729.
Datt, B. 1938, *Z. Phys.*, **108**, 314.
Fabian, A. C. 1994, *ARAA*, **32**, 277.
Fort, B., Mellier, Y. 1994, *A&A Rev.*, **4**, 239.
Gliner, E. B. 1970, *Soviet Physics-Doklady*, **15**, 559.
Gunn, J. B., Oke, J. B. 1975, *ApJ*, **195**, 255.
Hawking, S. W., Ellis, G. F. R. 1973, *The Large Scale Structure of Space-time*, Cambridge.
Hawkins, E., Maddox, S. J., Merrifield, M. R. 2002, *MNRAS*, **336**, L13.
Hickson, P. 1997, *ARAA*, **35**, 377.
Hogarth, J. E. 1962, *Proc. Roy. Soc.*, **A267**, 365.
Hoyle, F. 1948, *MNRAS*, **108**, 372.
Hoyle, F. 1953, *ApJ*, **118**, 513.
Hoyle, F., Burbidge, G. 1966, *ApJ*, **144**, 534.
Hoyle, F., Narlikar, J. V. 1963, *Proc. Roy. Soc.*, **A277**, 1.
Hoyle, F., Narlikar, J. V. 1964, *Proc. Roy. Soc.*, **A278**, 465.
Hoyle, F., Narlikar, J. V. 1969, *Ann. Phys. (N.Y.)*, **54**, 207.
Hoyle, F., Narlikar, J. V. 1971, *Ann. Phys. (N.Y.)*, **62**, 44.
Hoyle, F., Narlikar, J. V. 1995, *Rev. Mod. Phys.*, **61**, 113.
Hoyle, F., Burbidge, G., Narlikar, J. V. 1993, *ApJ*, **410**, 437.
Hoyle, F., Burbidge, G., Narlikar, J. V. 1994, *MNRAS*, **267**, 1007.
Hoyle, F., Burbidge, G., Narlikar, J. V. 1995, *Proc. Roy. Soc.*, **A448**, 191.
Hoyle, F., Burbidge, G., Narlikar, J. V. 2000, *A Different Approach to Cosmology* (Cambridge: Cambridge University Press).
Hoyle, F., Fowler, W. A., Burbidge, E. M., Burbidge, G. 1964, *ApJ*, **139**, 909.
Hoyle, F., Sandage, A. 1956, *PASP*, **68**, 301.
Karlsson, K. G. 1971, *A&A*, **13**, 333.
Kembhavi, A. K., Narlikar, J. V. 1999, *Quasars and Active Galactic Nuclei* (Cambridge: Cambridge University Press).
Longair, M. S. 1987, *IAU Symposium 124, "Observational Cosmology"* (eds) Hewitt, A., Burbidge, G., Fang, L. Z., Reidel, D. Dordrecht, p. 823.
Matthews, T. A., Sandage, A. R. 1963, *ApJ*, **138**, 30.
McCrea, W. H. 1951, *Proc. Roy. Soc.*, **A206**, 562.
Meyers, A. D., Shanks, T., Outram, J. J., Srith, W. J., Wolfendale, A. W. 2004, *MNRAS*, **347**, L67.
Napier, W., Burbidge, G. 2003, *MNRAS*, **342**, 601.
Narlikar, J. V. 1973, *Nature*, **242**, 35.
Narlikar, J. V., Padmanabhan, T. 1985, *Phys. Rev.*, **D32**, 1928.
Narlikar, J. V., Apparao, M. V. K., Dadhich, N. K. 1974, *Nature*, **251**, 590.

Narlikar, J. V., Vishwakarma, R. G., Burbidge, G. 2002, *PASP*, **114**, 1092.
Narlikar, J. V., Vishwakarma, R. G., Hajian, A., Souradeep, T., Burbidge, G., Hoyle, F. 2003, *ApJ*, **585**, 1.
Nayeri, A., Engineer, S., Narlikar, J. V., Hoyle, F. 1999, *ApJ*, **525**, 10.
Ostriker, J. P., Peebles, P. J. E., Yahil, A. 1974, *ApJ*, **193**, L1.
Page, L. *et al.* 2003, *Astrophys. J. Suppl.*, **148**, 233.
Perlmutter, S. *et al.* 1999, *ApJ*, **517**, 565.
Podariu, S., Souradeep, T., Gott III, J. R., Ratra, B., Vogeley, M. S. 2001, *ApJS*, **559**, 9.
Rees, M. J. 1984, *ARAA*, **22**, 471.
Rees, M. J., Ostriker, J. P. 1977, *MNRAS*, **179**, 541.
Riess, A. *et al.* 1998, *AJ*, **116**, 1009.
Rubano, C., Seudellaro, P. 2004, astro-ph/0410260.
Sachs, R., Narlikar, J. V., Hoyle, F. 1996, *A&A*, **313**, 703.
Sami, M., Toporensky, A. 2004, *Mod. Phys. Lett. A*, **19**, 1509.
Schmidt, M. 1963, *Nature*, **197**, 1040.
Silk, J. 1977, *ApJ*, **211**, 638.
Singh, P., Sami, M., Dadhich, N. 2003, *Phys. Rev.*, **D68**, 023522.
Spergel, D. *et al.* 2003, *ApJS*, **148**, 175.
Spergel, D. N. *et al.* 2006, astro-ph/0603449.
Steinhardt, P. J., Turok, N. 2002, *Science*, **296**, 1436.
Toomre, A., Toomre, J. 1972, *ApJ*, **178**, 623.
Vishwakarma, R. G., Narlikar, J. V. 2005, *Int. J. Mod. Phys. D*, **14(2)**, 345.
Wheeler, J. A., Feynman, R. P. 1945, *Rev. Mod. Phys.*, **17**, 157.
Wheeler, J. A., Feynman, R. P. 1949, *Rev. Mod. Phys.*, **21**, 425.
Wickramasinghe, N. C. 2005, *Current Issues in Cosmology*, Proceedings of the Colloquium on 'Cosmology: Facts and Problems', Paris (Cambridge: Cambridge University Press), 152.

J. Astrophys. Astr. (2007) **28,** 101–116

Some Doubts on the Validity of the Foreground Galactic Contribution Subtraction from Microwave Anisotropies

Martín López-Corredoira
Instituto de Astrofísica de Canarias, C/.Vía Láctea, s/n, E-38200 La Laguna (S/C de Tenerife), Spain.
e-mail: martinlc@iac.es

Received 2007 April 5; accepted 2007 August 30

Abstract. The Galactic foreground contamination in CMBR anisotropies, especially from the dust component, is not easily separable from the cosmological or extragalactic component. In this paper, some doubts will be raised concerning the validity of the methods used until now to remove Galactic dust emission and will show that none of them achieves its goal.

First, I review the recent bibliography on the topic and discuss critically the methods of foreground subtraction: the cross-correlation with templates, analysis assuming the spectral shape of the Galactic components, the "maximum entropy method", "internal linear combination", and "wavelet-based high resolution fitting of internal templates". Second, I analyse the Galactic latitude dependence from WMAP data. The frequency dependence is discussed with data in the available literature. The result is that all methods of subtracting the Galactic contamination are inaccurate. The Galactic latitude dependence analysis or the frequency dependence of the anisotropies in the range 50–250 GHz put a constraint on the maximum Galactic contribution in the power spectrum to be less than $\sim 10\%$ (68% C. L.) for an ~ 1 degree scale, and possibly higher for larger scales.

The origin of most of the signals in the CMBR anisotropies is not Galactic. In any case, the subtraction of the galaxy is not accurate enough to allow a "precision Cosmology"; other sources of contamination (extragalactic, solar system) are also present.

Key words. Cosmic microwave background—ISM: clouds—dust, extinction—ISM: structure.

1. Introduction

The cosmic microwave background radiation (CMBR) has been interpreted as the relict radiation of an early stage of the Universe. Its black-body spectrum (Mather *et al.* 1994) of 2.75 K reveals a very small dependence on sky position. Measurements of the anisotropies, carried out by several teams of researchers over the last two decades, have claimed to provide information on the structural formation of the Universe, inflation in its early stages, quantum gravity, topological defects (strings, etc.), dark matter type

and abundance, the determination of cosmological parameters (H_0, Ω_m, Ω_Λ, etc.), the geometry and dynamics of the Universe, the thermal history of the Universe at the recombination epoch, etc. (e.g., Bennett *et al.* 2003a). There are even some researchers who claim that the golden era of cosmology, the age of the precision cosmology, has arrived with new experiments to measure these anisotropies and other optimistic interpretations of observational cosmology. However, all that glitters is not gold and what is claimed as a source of cosmological information is often something else. One should be very careful to ensure that there are no contaminants along the light path of this microwave radiation, Galactic contamination being one important factor. If one wants to do "precision cosmology", then even greater care must be taken. Many authors (see the brief review in section 2) have studied the different components of the Galactic microwave foreground radiation, but in my view these efforts are still insufficient to separate the components appropriately. The foreground emission is small; there is no "*Galactic foreground contamination which is 1000 times more intense than the desired signal*" (as claimed by Robitaille 2007) in off-plane regions; this is completely wrong, but neither is the foreground contamination entirely negligible. At least, some doubts on the validity of the foreground Galactic subtraction from microwave anisotropies can be expressed, and this is precisely the topic of this paper, together with some analysis to constrain Galactic emission. In section 2, I review the recent bibliography on the topic and discuss critically the methods of foreground subtraction. I analyse the Galactic latitude dependence (section 3) from WMAP first year data. The frequency dependence is discussed (section 4) with data in the available literature. Other possible sources of contamination are also discussed in section 5.

2. Some comments on the present status of CMBR foreground analysis

The consideration of the Galactic foregrounds has been present for a long time now. However, an analysis of the papers analysing the problem reveals that it was sometimes underestimated, and that the problem of galaxy subtraction has become more complex in recent years.

For instance, Gutiérrez de la Cruz *et al.* (1995); Davies *et al.* (1996) thought that Tenerife data at 15 GHz (see also Gutiérrez *et al.* 2000) were likely to be dominated by cosmological fluctuations, and that 33 GHz data were almost unaffected by the Galaxy ($\sim 4\,\mu$K in scales of 5–15 degrees). More recent analyses (WMAP, Bennett *et al.* 2003b, Fig. 10), however, have claimed that for comparable angular scales the galaxy is dominant in the anisotropies at 22.8 GHz (and the Galactic contamination at 15 GHz is not much smaller than at 22.8 GHz), and at 33 GHz the Galactic anisotropies are of the same order as those from the CMBR. Also, at small scales, less than a degree, the foreground Galactic emission dominates (Leitch *et al.* 2000). Halverson *et al.* (2002) and Fernández-Cerezo *et al.* (2006) claim that the cosmological signal is still dominant at $|b| > 40^\circ$ around 15 GHz and 31 GHz respectively for intermediate and small angular scales (around a degree and down to ≈ 12 arcminutes respectively), but their analysis is based on correlations with templates, which, as we will see below, might not be appropriate. An explanation for this is that it is mainly due to the existence of a new kind of emission correlated with the dust that was not discovered previously; the synchrotron and or free–free emission might also have been underestimated. Positive correlations between the microwave anisotropies, including the region around 15 GHz, and far-infrared maps, which trace Galactic dust, were found (Kogut *et al.*

1996; Leitch *et al.* 1997, 2000; de Oliveira-Costa *et al.* 1997, 1998; Finkbeiner *et al.* 1999, 2004; Fernández-Cerezo *et al.* 2006). Casassus *et al.* (2004) report that in the Helix region the emission at 31 GHz and 100 μm are well correlated; they estimate that the 100 μm-correlated radio emission, presumably due to dust, accounts for at least 20% of the 31 GHz emission in the Helix (so the total dust emission is higher than 20% because there is a non-correlated component too). Watson *et al.* (2005) show that the anomalous emission at around 23 GHz of the Perseus molecular cloud (Temp $\sim$ 1 mK) is an order of magnitude larger than the emission expected from synchrotron + free–free + thermal dust. There is also some correlation of 10, 15 GHz maps with H_α maps, but very low, so the free-free emission is detected at levels far lower than the dust correlation (de Oliveira-Costa *et al.* 1999, 2002). The most likely current explanation for this emission correlated with dust around 15–50 GHz is spinning dust grain emission (Draine & Lazarian 1998) and/or magnetic dipole emission from ferromagnetic grains (Draine & Lazarian 1999); although some other authors (e.g., Mukherjee *et al.* 2001) do not agree. Whatever it is, it is now clear that the 15–40 GHz range is dominated by the galaxy.

A question might arise as to whether the contamination in the remaining frequencies (40–200 GHz) is being correctly accounted for. Typical calculations of dust contamination claim that it should not be predominant, and that its contribution can be subtracted accurately, but how sure can we be of this statement? How accurate is the subtraction of the dust foreground signal?

2.1 *Dust foreground removal with templates*

The first difficulty in the subtraction of the dust component is to know exactly how much emission there is in each line of sight. First approximations came with extrapolations from the IRAS far-infrared and DIRBE data (with the zodiacal light subtracted, as well as the cosmic infrared background) used to model the dust thermal emission in microwaves. Templates were taken from these infrared maps and extrapolated in amplitude by a common factor for all pixels. The problem is that, as said by Finkbeiner *et al.* (1999), *'a template approach is often carelessly used to compare observations with expected contaminants, with the correlation amplitude indicating the level of contamination. (. . .) These templates ignore well-measured variation in dust temperature and variations in dust/gas ratio.'* According to López-Corredoira (1999), the growing contrast of colder clouds in the background of the diffuse interstellar medium will produce much higher microwave anisotropies than the product of the template extrapolation. Neither is it a good strategy to subtract a scaled IRAS template to remove the spinning dust in multifrequency data since it produces large residual differences (Leitch *et al.* 2000).

A better approximation to this dust emission in the microwave region came with the adoption of an extrapolation with colour corrections in each pixel. This method assigns a different temperature to each pixel. This colour correction, together with a ν^2 emission emissivity (Schlegel *et al.* 1998), gave a much tighter agreement with FIRAS data. However, it was still inconsistent with the FIRAS data below 800 GHz in amplitude (Finkbeiner *et al.* 1999): a 14% error in the mean amplitude at 500 GHz. Indeed, no power-law emissivity function fits the FIRAS data in the 200–2100 GHz region (Finkbeiner *et al.* 1999). Furthermore, laboratory measurements suggest that the universality of ν^2 emissivity is an oversimplification, with different species of grains

having different emissivity laws (Finkbeiner *et al.* 1999). A better approximation is an extrapolation with two components ($\nu^{1.7}$, $\langle T \rangle = 9.5$ K and $\nu^{2.7}$, $\langle T \rangle = 16$ K); each pixel has an assignation of two temperatures (Finkbeiner *et al.* 1999) with correction factors that are a function of these temperatures. This still fails in the predictions of FIRAS from the IRAS–DIRBE extrapolation by 15% in zones dominated by atomic gas (Finkbeiner *et al.* 1999).

Finkbeiner *et al.*'s (1999) approach, although much better than a direct extrapolation of the template, is insufficient. The problem is difficult to solve because we are trying to do an extrapolation by a factor $\sim$15–60 in frequency (from 1250 GHz [240 μm] or 3000 GHz [100 μm]) to 50–90 GHz). This is equivalent, for instance, to the attempt to derive a map of stellar emission in 12 μm as an extrapolation of the emission in optical B-filter. Frankly, when I see IRAS-12 μm map of point sources and the Palomar plates in blue filters, I observe huge differences, and I do not know how we can extrapolate the second map to obtain the first one. Each star has a different colour, and stars which are very bright in blue may be very faint at mid-infrared and *vice versa*. The same thing happens with the diffuse + cloud emission: there are hot regions, cold regions, different kinds of emitters (molecular gas, atomic gas) and we have to integrate all this into each line of sight. The assumption that with only two temperatures we can extrapolate the average flux, and that a colour term can correct the pixel-to-pixel differences is comparable to the assumption that a model with only two kinds of stars and the knowledge of $(B - V)$ for each star we can extrapolate the star counts from optical to mid-infrared for the whole galaxy.

The cold clouds, invisible or very faint in far-infrared surveys, are potential elements to produce a very significant emission at 40–200 GHz at the amplitude level of the observed anisotropies. Bernard *et al.* (1999) have shown that some cirrus features with a high value of $I_{100\,\mu\mathrm{m}}/I_{60\,\mu\mathrm{m}}$ of 29.5 (the average is 3.2) are cold dust regions ($T \approx 13$ K instead of the average of 17.5 K). Colder clouds may also exist. Lawrence (2001) claims that part of the SCUBA-850 μm sources might be local sources at 7 K. And who knows whether there are even colder clouds that might emit significantly only in the range of microwaves? Paladini *et al.* (2007) show that the microwave emission in the outer Galactic plane ($R > 8.9$ kpc) is much higher than expected, suggesting the possible presence of very cold dust.

It is also very common (e.g., Masi *et al.* 2001; Fernández-Cerezo *et al.* 2006) to calculate the Galactic dust contribution in some microwave data by just making a cross-correlation between these data and some far-infrared map of the sky. This is simply wrong and not even valid for ascertaining the order of magnitude of such contamination. Since the templates in infrared cannot be used as maps of the microwave Galactic emission, there will be non-correlated emission coming from the dust too. That is, there will be many clouds that emit substantially in microwaves, and that are not detected in far-infrared maps.

Therefore, any method of foreground subtraction which uses templates will have serious credibility problems with regard to the goodness of the subtraction. This applies not only to those methods that use templates directly with coupling coefficients derived from cross-correlation but also to MEM (maximum entropy method; Bennett *et al.* 2003b), which uses in the initial stage templates for the dominant foreground components and also establishes some *a priori* conditions of their spectral behaviour. Moreover, any calculation of the limits of such contamination based on cross-correlations will not be totally accurate.

2.2 *Dust foreground removal without templates*

If removal of foreground contamination is applied without the use of templates, we might need a knowledge of all of the physical components of this foreground emission and their spectral behaviour.

With ISO-90,180 μm, a power law down to scale of $3'$ is found in the power spectrum (Herbstmeier *et al.* 1998), with similar shape as that derived from the 21 cm line for Galactic cirrus. At 410 GHz, the dust emission is also roughly a power law in off-plane regions (Masi *et al.* 2001), although I see in the plot for $b = -17^\circ$ from Fig. 2 of Masi *et al.* (2001) that there would be a better fit if we admitted that there was a maximum in the dust emission power spectrum around $l = 200$. In any case, even if a power law is present in the far infrared or at 410 GHz, the extrapolation to 50–90 GHz is not direct and, moreover, the spectral index of this power law is highly variable. The presence of different dust-emitting components causes the spectral indices of the foregrounds to vary with position. Kiss *et al.* (2003) have measured spectral index changes from -5.1 to -2.1, depending on the region, and also changes with wavelength; they point out the existence of dust at different temperatures, in particular of a cold, extended component. Dupac *et al.* (2004) also show with the PRONAOS balloon-borne experiment that the submillimetre/millimetre spectral index is found to vary between roughly 1 and 2.5 on small scales ($3.5'$ resolution). Even small spectral index variations as small as $\Delta\alpha \sim 0.1$ can have a substantial impact on how channels should be combined and on the attainable accuracy (Tegmark 1998), so the errors of the subtraction assuming certain power spectrum for the dust are serious.

Another technique that does not use templates is (Bennett *et al.* 2003b): ILC (internal linear combination). It assumes nothing about the particular frequency dependencies or morphologies of the foregrounds and tries to minimize the variance in twelve different regions of the sky with the combination of the five available WMAP frequencies. There is degeneracy of solutions, an infinite number of maps can be generated from the five basis sets, there is no way to test whether the maximum likelihood solution is the correct one (Robitaille 2007), and it is not enough to divide the sky into twelve regions. Bennett *et al.* (2003b) themselves warn against its use for cosmological analysis, and Eriksen *et al.* (2004b) show that it is not effective to remove all residual foregrounds. The ILC method performs quite badly, especially for dust (Eriksen *et al.* 2004b), in part because of the variability of the spectral indices. Indeed, a tremendous coefficient variability in the twelve sections was obtained (Robitaille 2007). Therefore, ILC maps are not clean enough to allow cosmological conclusions to be arrived at (Eriksen *et al.* 2004b): '*[The] ILC map, which by eye looks almost free of foreground residuals, has been extensively used for scientific purposes—despite the fact that there are strong (and difficult to quantify) residual foregrounds present in the map. (. . .) the ILC map is indeed highly contaminated by residual foregrounds, and in particular, that the low-l components, which have received the most attention so far, are highly unstable under the ILC cleaning operation*' (Eriksen *et al.* 2005).

The WI-FIT method ('Wavelet based hIgh resolution Fitting of Internal Templates', Hansen *et al.* 2006) does not require *a priori* templates, but takes the information about the foregrounds by taking differences of temperature maps at different frequencies. However, for the application in presently available maps, it requires the assumption that the spectral indices are constant in space, which, as said, is a very inaccurate approximation. Their assumption that the Galactic emission in each pixel is proportional to

the difference in temperature maps ($(T_i^{\nu} - T_i^{\nu'}) \propto T_i^{\nu}$) is in general incorrect because $T_i^{\nu'}$ is not proportional to T_i^{ν}, the temperature in each pixel being the superposition of many different emissions with different temperatures. Again, we have here the same problem as with the use of templates: the assumption that there is only one temperature along each line of sight, and that the intensity of this emission is describable with a simple average fixed power law multiplying a black body emission with an average temperature. Hansen *et al.* (2006) claim that their method is good because they obtain similar results to Bennett *et al.* (2003b), but this may be due to their similar assumptions.

Foreground contamination residuals are found even for the best available supposed clean CMB maps. Naselsky *et al.* (2006) found some correlation between $\Delta l = 4n$ and $n = 1,\ 2$ spherical harmonic multipole domain, which is caused by a symmetric signal in the Galactic coordinate system. de Oliveira-Costa & Tegmark (2006) find that the alignment of low-l multipoles appears to be rather robust to Galactic cut and different foreground contaminations. It is also relevant that Bernui *et al.* (2007a) found a preferred direction in the fluctuations of the WMAP data while there should be no preferential direction in the sky (isotropy); but the significance of this detection is only at 95% C.L. so it could be just a chance detection, or it could also indicate some kind of contamination. The statistical anisotropy in different circles of the sky is also found by Then (2006). Liu & Zhang (2006) have also studied the cross-correlation between WMAP and Egret γ-ray data and concluded that an unknown source of radiation, most likely of Galactic origin, is implied by their analysis. Verschuur (2007) speculates that this unknown radiation of Galactic origin might take place in the surface of Galactic HI structures moving through interstellar space and/or interacting with one another.

The spatial association found by Verschuur (2007) on scales of 1–2 degrees between interstellar neutral hydrogen, integrated in maps over ranges of 10 km/s, and WMAP-ILC maps, which should be clean of foreground contamination through the ILC methods, is especially significant for the present discussion too. Several extended areas of excess emission at high Galactic latitudes ($b > 30^{\circ}$) are present in both maps. According to Verschuur (2007), these structures have typical distances from the Sun of the order of 100 pc.

2.3 *Non-gaussianity*

Another aspect to think about is the non-gaussianity of the anisotropies distribution. Standard theories involving inflation generally predict a pattern of gaussian noise; whereas non-standard theories based on symmetry breaking and the generation of defects have more distinctive signatures (Coulson *et al.* 1994). The fact is that a non-gaussian distribution was discovered in the anisotropies against all predictions of current inflationary cosmological models by many different authors with different methods (Ferreira *et al.* 1998; Pando *et al.* 1998; Magueijo 2000a, b; Chiang *et al.* 2003; Coles *et al.* 2004; Eriksen *et al.* 2004a; Park 2004; Vielva *et al.* 2004; Liu & Zhang 2005; Raeth *et al.* 2007; Bernui *et al.* 2007b) and yet the leading cosmologists still claim inflationary cosmology to be valid and to be able to constrain cosmological parameters from the anisotropies. This might mean that the analyses of non-gaussianity are wrong, or that they are not significant enough, or that the inflation model is incorrect, or that some contamination is present in the maps, which are supposed to be clear of any

contamination. Among all these possibilities, maybe more than one applies, and I suspect that one of these is the incorrect subtraction of the Galactic contamination although other factors may also be important. As a matter of fact, Liu & Zhang (2005) show with certain tests that residual foreground contamination in WMAP-cleaned maps may contribute to this non-gaussian features significantly, and Tojeiro *et al.* (2006) have argued that the non-gaussianity is associated with cold spots of unsubtracted foregrounds. The lowest spherical harmonic modes in the map are significantly contaminated with foreground radiation (Chiang *et al.* 2007). However, Galactic contamination may not be the only reason: using the WMAP data and 2MASS galaxy catalog, Cao *et al.* (2006) show that the non-gaussianity of the 2MASS galaxies is imprinted on WMAP maps too. Nonetheless, Rubiño-Martín *et al.* (2006) and McEwen *et al.* (2006) claim that only few regions have such non-gaussian anisotropies due to contamination, e.g., the Corona Borealis supercluster region, and that most regions of the sky are gaussian.

3. Galactic latitude dependence

A simple way to test whether the Galactic anisotropies are somewhat significant would be to examine their variation with the position in the sky, for example with Galactic latitude dependence (López-Corredoira 1999). Higher Galactic anisotropies are expected towards the plane. Moreover, the lines of sight in the southern Galactic hemisphere should give more Galactic anisotropies than the northern one, because the Sun is approximately 15 pc above the Galactic plane (Hammersley *et al.* 1995; Chyzy *et al.* 2005) and the scale height of the cold dust component is not very high, although unknown in principle. For dust observed in infrared, the scale-height is 40 pc (Unavane *et al.* 1998)[1]. If the cold dust responsible for the microwave fluctuation in the range 60–100 GHz had a scale-height of 40 pc as well, the excess column density in the south over the north $\left(\frac{2-\exp(-15/40)}{\exp(-15/40)}-1\right)$ would be 91%, for any Galactic latitude. If we assume a cosecant law for Galactic latitude dependence, this absolute value of the asymmetry would be ≈ 3 times higher at $|b| = 20^\circ$ than at $|b| = 90^\circ$. Of course, this is a rough calculation because the scale-height might change in microwaves with respect to the value inferred from the extinction in infrared, there is longitude dependence, flares, the accuracy of the cosecant law for low latitude regions is not good. Cloud sizes and other factors should be taken into account too.

Nonetheless, we are limited by the variations of the anisotropies produced by the own cosmological anisotropies, especially those at large scale, which have a range of possible gradients that could even mimic the Galactic gradients. Multipoles of low-l with $l \geq 2$ are part of the fluctuations, and one might suspect that they are responsible for the gradients and the difference north/south. One way to check whether these variations with Galactic position exist and whether they are within the expected cosmological variation or not is to measure them with real data (WMAP) and compare with realizations.

[1]Note that, although Unavane *et al.* use near-infrared data, the scale-height of 40 pc does not refer to the old stellar population (the scale-height of the old stellar population is 285 pc; López-Corredoira *et al.* 2002) but to the dust distribution inferred through the extinction maps.

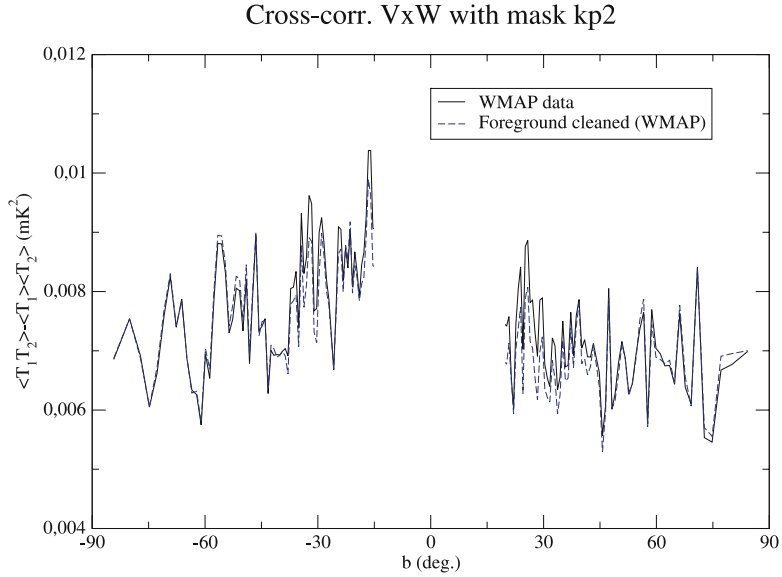

Figure 1. Variance as a function of Galactic latitude in WMAP first-year data (only points with less than 10% of pixels in the mask kp2). Each point in the plot represents the pixels in the ring for all Galactic longitudes and $\Delta(\sin b) = 0.01$.

3.1 *WMAP data analysis*

Figure 1 shows the variation with latitude of the variance ($\langle T_1 T_2\rangle - \langle T_1\rangle\langle T_2\rangle$) and with T_1, T_2, the antenna temperature at bands V and W of WMAP (61 and 94 GHz) (Bennett *et al.* 2003a). This variance is free of instrumental noise because it corresponds to the cross-correlation of two independent maps. Each point in the plot represents the cross-variance in the ring for all Galactic longitudes and $\Delta(\sin b) = 0.01$ (which comprises around 15,000 pixels). We used the data with no foreground subtraction with the kp2 mask (Bennett *et al.* 2003a), which removes 15% of the sky, mainly from the Galactic plane, the same mask that was used in the derivation of their WMAP power spectrum. The figure shows two features:

- there is a significant decrease of fluctuations toward the polar caps, especially for the southern cap;
- there is a significant excess in the southern fluctuations ($b < 0$) with respect to the north fluctuations (this fact was already present in COBE-DMR data; López-Corredoira 1999, subsection 3.4).

The least squares linear fits ($y = a + c|b|$) to the ranges $b < -20^\circ$ and $b > 20^\circ$ are respectively:

$$\sigma^2_{\mathrm{cross},b<-20^\circ}(|b|) = (8.99 \pm 0.28) \times 10^{-3} - (27.8 \pm 5.9) \times 10^{-6}|b|(\mathrm{deg})\ \mathrm{mK}^2$$

$$\sigma^2_{\mathrm{cross},b>20^\circ}(|b|) = (7.74 \pm 0.24) \times 10^{-3} - (16.2 \pm 5.1) \times 10^{-6}|b|(\mathrm{deg})\ \mathrm{mK}^2. \quad (1)$$

That is, $c = 0$ is excluded within a $4.7 - \sigma$ level for the south and $3.2 - \sigma$ for the north for the first feature, and $\sigma^2_{\rm cross}(b = -20^\circ) - \sigma^2_{\rm cross}(b = 20^\circ) = (1.01 \pm 0.40) \times 10^{-3}\,{\rm mK}^2$, i.e., the second feature. Equality between north and south is excluded to within $2.5 - \sigma$.

Bennett *et al.* (2003b, Fig. 10) point out that in the W bands the Galactic contribution constitutes around 10% of the power spectrum. If we examine the same dependence in the same maps once the foregrounds have been cleaned according to their estimations, the results are (see also Fig. 1):

$$\sigma^2_{{\rm cross},b<-20^\circ}(|b|) = (8.59 \pm 0.27) \times 10^{-3} - (20.6 \pm 5.6) \times 10^{-6}|b|({\rm deg})\,{\rm mK}^2$$

$$\sigma^2_{{\rm cross},b>20^\circ}(|b|) = (6.88 \pm 0.23) \times 10^{-3} + (2.2 \pm 4.7) \times 10^{-6}|b|({\rm deg})\,{\rm mK}^2, \quad (2)$$

and $\sigma^2_{\rm cross}(b = -20^\circ) - \sigma^2_{\rm cross}(b = 20^\circ) = (1.25 \pm 0.38) \times 10^{-3}\,{\rm mK}^2$. The latitude dependence on the south is less significant ($c = 0$ excluded at $3.7 - \sigma$) and disappears in the north, but the south-north difference is higher than before ($3.3 - \sigma$).

These higher fluctuations in the southern Galactic hemisphere have also been observed by Eriksen *et al.* (2004a), Hansen *et al.* (2004a, b), Chyzy *et al.* (2005) or Bernui *et al.* (2006) with higher asymmetry for lower latitudes (Chyzy *et al.* 2005). The higher asymmetry for lower values of $|b|$ is expected from the Galactic contamination because the dust emission is roughly proportional to $1/\sin(|b|)$.

3.2 *Simulations*

We now do the same measures in 20 random realizations of the expected WMAP sky (Eriksen *et al.* 2004b, 2005) (i.e., 20 measures of the excess south/north, and 40 measures of the slope).

The average slope c (of a linear law $\sigma^2(|b|) = a - c|b|$) is $+11 \times 10^{-6}\,{\rm mK}^2/{\rm deg}$ and the r.m.s. is $21 \times 10^{-6}\,{\rm mK}^2/{\rm deg}$, so our measured values of $+27.8 \times 10^{-6}\,{\rm mK}^2/{\rm deg}$ and $+16.2 \times 10^{-6}\,{\rm mK}^2/{\rm deg}$ for the south and the north respectively are indeed within the expected values only for cosmological fluctuations. The fact that we have a non-null value for the average c ($\langle c \rangle = +11.3 \pm 3.3 \times 10^{-6}$) is due to the way the fluctuations are measured in different rings with different latitudes; all rings have the same number of points but different shapes; for instance, toward the polar caps the areas are different from the rings near the plane.

The difference south/north in the same realizations is on average $0.33 \pm 0.28 \times 10^{-3}\,{\rm mK}^2$ (compatible with zero, which is the exact expected value, since there are no reasons for an asymmetry north/south) and the r.m.s. is $1.3 \times 10^{-3}\,{\rm mK}^2$. Again, our measured value in WMAP of $1.25 \times 10^{-3}\,{\rm mK}^2$ is within the expected statistical values.

3.3 *Latitude dependence on small angular scales*

We might think that the variations among zones could be reduced if we used only a high-l (small angular scale) range of the power spectrum with the same shape for all selected regions instead of circular rings of variable thickness and height parallel to the Galactic equator. If we do the analysis of $[l(l+1)C_l/(2\pi)]$ for $316 \leq l \leq 354$ with 28 points distributed with different Galactic longitudes and latitudes $\pm 36^\circ$, $\pm 54^\circ$, $\pm 72^\circ$, $\pm 90^\circ$

in circles of radius of 9° which are not overlapping to each other, we get

$$[l(l+1)C_l/(2\pi)](|b| = 90^\circ) = 3.0(\pm 0.6) \times 10^{-3}\,\mathrm{mK}^2/\mathrm{deg} \quad (3)$$

$$\frac{d[(l(l+1)C_l/(2\pi)]}{d|b|} = +3(\pm 10) \times 10^{-6}\,\mathrm{mK}^2/\mathrm{deg}, \quad (4)$$

that is, the error bar is larger than the measured slope.

3.4 *Constraints on the maximum Galactic contamination*

All angles: We are therefore very limited in ascertaining the Galactic contamination through this kind of analysis. The values of the present simulations allow us to put a higher limit but these numbers permit a very wide range of Galactic contamination amplitudes. On average, for both hemispheres, it should be less than $+43 \times 10^{-6}\,\mathrm{mK}^2/\,\mathrm{deg}(1\sigma)[+64 \times 10^{-6}\,\mathrm{mK}^2/\,\mathrm{deg}(2\sigma)]$. Since Bennett *et al.*'s (2003b) foreground corrections reduce the value of the slope c by an average of $13 \times 10^{-6}\,\mathrm{mK}^2/\,\mathrm{deg}$ and they claim that this is representative of 10% of contamination by the galaxy, an estimate of the maximum limit of contamination by the galaxy is less than 33% (1σ) [49% (2σ)].

Another way to analyse it would be to suppose that the Galactic contamination follows a cosecant law in the dependence on Galactic latitude. Then $c_{\mathrm{gal}}/a_{\mathrm{gal}} = 9.5 \times 10^{-3}$ (derived from a linear fit in the same conditions as for the analysis of the WMAP data); so the excess over the average in the simulation ($\langle c\rangle = +11.3 \times 10^{-6}$) in the south of $c_{\mathrm{gal}} = \Delta c = 27.8 \times 10^{-6} - 11.3 \times 10^{-6} \pm 21 \times 10^{-6}\,\mathrm{mK}^2/\,\mathrm{deg} = 16.5 \pm 21 \times 10^{-6}\,\mathrm{mK}^2/\,\mathrm{deg}$ implies that $a_{\mathrm{gal,\ south}} = 1.7 \pm 2.2 \times 10^{-3}$, i.e., 19 ± 25% of the total emission ($a = 8.9 \times 10^{-6}$). This means that the Galactic emission in the south is less than 44% (1σ) [69% (2σ)]. The same analysis for the northern hemisphere gives: less than 35% (1σ) [64% (2σ)]. Given that, as said above with regard to the asymmetry south/north, the contamination in the south should be 1.94 times the contamination in the north, we can further reduce the constraint in the north: less than 26% (1σ) [41% (2σ)]. Therefore, on average in the north and the south, the contamination is less than 35% (1σ) [55% (2σ)].

Small angular scales: With the numbers at small angular scales given in subsection 3.3, we can set the constraint: $\delta[(l(l+1)C_l/(2\pi)](|b| = 20^\circ) - [l(l+1)C_l/(2\pi)](|b| = 90^\circ)] < 0.49 \times 10^{-3}\,\mathrm{mK}^2/\mathrm{deg}(1\sigma)$ [$< 1.2 \times 10^{-3}\,\mathrm{mK}^2/\,\mathrm{deg}(2\sigma)$], which implies, assuming a dust emission proportional to $1/\sin(|b|)$, that $[l(l+1)C_l/(2\pi)](|b| = 90^\circ) < 2.5 \times 10^{-4}(1\sigma)[< 6.2 \times 10^{-4}(2\sigma)]$. Comparing with equation (3), we get the constraint that the Galactic emission must be lower than 8% (1σ) [20% (2σ)]. We see, then, that the restriction of values of l is more efficient for constraining the Galactic contamination.

4. The frequency dependence

In my opinion, the most important proof presented to claim that most of the off-plane microwave anisotropies have a cosmological origin is the frequency dependence. The near independence in the range of 50–90 GHz is not conclusive because this could

be produced by the galaxy too, mainly as a combination of thermal dust and rotational dust emission (López-Corredoira 1999). However, the surveys with a larger range of frequencies, such as that at 170 GHz by Ganga *et al.* (1993), BOOMERANG (Netterfield *et al.* 2002) or ARCHEOPS (Tristam *et al.* 2005), have shown that the power spectrum is practically the same in amplitude and shape with respect to WMAP measures and is constant over the frequency range 50–250 GHz. The spectral analysis of FIRAS + WMAP by Fixsen (2003) also shows that the anisotropies are Planckian, although Fixsen (2003) assumes that the frequency and spatial dependences of the galaxy are separable and modelled by templates, which is not a trustworthy assumption. Since Galactic dust thermal emission of 7–20 K temperature grains should have some frequency dependence, we must conclude that they are not predominant. We have, within the error ranges of the calibration and within the error bars of the power spectrum (around a 10% in total), a constant amplitude over a factor 5 in frequency.

The amplitude of the anisotropies depends on two factors (López-Corredoira 1999): the variation of the mean flux with frequency, and the ratio of diffuse and cloud emission as a function of frequency; but in any case, the variation of $(\Delta T)^2$ should be at least twice as high at 250 GHz than at 50 GHz. This minimum factor roughly equals to two stems from the multiplication of the first two factors specified in López-Corredoira (1999, equation 2):

$$\frac{(\Delta T)^2(\nu = 250\,\mathrm{GHz})}{(\Delta T)^2(\nu = 50\,\mathrm{GHz})} = \left\langle \frac{T(\nu = 250\,\mathrm{GHz})}{T(\nu = 50\,\mathrm{GHz})} \right\rangle^2 \left\langle \frac{\frac{\delta T}{T}(\nu = 250\,\mathrm{GHz})}{\frac{\delta T}{T}(\nu = 50\,\mathrm{GHz})} \right\rangle^2 . \quad (5)$$

$\left\langle T(\nu = 250\,\mathrm{GHz})/T(\nu = 50\,\mathrm{GHz}) \right\rangle \approx 5$ (López-Corredoira 1999, Fig. 5), and $\left\langle \frac{\delta T}{T}(\nu = 50\,\mathrm{GHz})/\frac{\delta T}{T}(\nu = 250\,\mathrm{GHz}) \right\rangle <\approx 3.3$ (López-Corredoira 1999, equation 12) so $(\Delta T)^2(\nu = 250\,\mathrm{GHz})/(\Delta T)^2(\nu = 50\,\mathrm{GHz}) >\approx 2.3$.

Assuming this roughly minimum factor of two in the increase in Galactic anisotropy at 250 GHz with respect to 50 GHz, if we observe a maximum of 10% excess in the total contribution, it means that the Galactic emission should be lower than a 10% of the total contribution of the anisotropies in average in the range of frequencies.

5. Other contaminants

Although it is not the matter for this paper, it is important to point out that, apart from the cosmological and the Galactic signal, there may be other contaminants: either from closer sources (in the solar system or the solar neighbourhood), or extragalactic sources much closer than $z = 1000 - 1500$ (as is supposed for the cosmological origin).

Copi *et al.* (2006, 2007) show that the two lowest cosmologically interesting multipoles, $l = 2$ and 3, are not statistically isotropic. The planes of the quadrupole and the octopole are unexpectedly aligned (de Oliveira-Costa 2004; Copi *et al.* 2006, 2007; Land & Magueijo 2007; Rakić & Schwarz 2007). Indeed, the combined quadrupole plus octopole is surprisingly aligned with the geometry and direction of motion of the solar system: the plane they define is perpendicular to the ecliptic plane and to the plane defined by the dipole direction, and the ecliptic plane carefully separates stronger from weaker extrema, running within a couple of degrees of the null-contour between a maximum and a minimum over more than 120° of the sky. The axis of maximum

asymmetry of the WMAP data tends to lie close to the ecliptic axis (Eriksen *et al.* 2004a). There is alignment of the quadrupole and octopole with each other and they are correlated with each other: 99.6% C.L. (Copi *et al.* 2007). There are also statistically significant correlations with local geometry, especially that of the solar system ($>$ 99.9% C.L.; Copi *et al.* 2006). Moreover, the angular two-point correlation function at scales $>$ 60 degrees in the regions outside the Galactic cut is approximately zero in all wavebands and is discrepant with the best fit ΛCDM inflationary model: 99.97% C.L. for the discrepancy (Copi *et al.* 2007; Rakić & Schwarz 2007). Eriksen *et al.* (2004a) also found that the ratio of the large-scale fluctuation amplitudes in the southern ecliptic hemisphere is high at the level 98–99% , with an absence of large-scale power in the vicinity of the north ecliptic pole. This asymmetry is stable with respect to frequency and sky coverage. This seems to point to the presence of some contamination in relation with the solar system or neighbourhood, although part of the effect may be due to non-uniform observational time sky coverage (Chyzy *et al.* 2005). Abramo *et al.* (2006) claim that if a hypothetic foreground produced by a cold spot in the Local Supercluster is subtracted from the CMB data, the amplitude of the quadrupole is substantially increased, and the statistically improbable alignment of the quadrupole with the octopole is substantially weakened, but this does not explain the coincidence of the alignment with the ecliptic.

The extragalactic contamination is still not very clear. In this respect, the discovery that the hard X-ray background measured by the HEAO-1 satellite and the number counts of radio galaxies in the NVSS survey are correlated with the WMAP microwave fluxes (Boughn & Crittenden 2004) should be pointed out. One could raise the objection that microwave anisotropies are produced over a very wide range of angles, which include structures of several degrees and largest structures while galaxies and clusters are generally smaller. However, if one considers the very large structures of filaments, walls, hyperclusters, etc., one can find in the sky structures of even several tens of degrees; that is, excess of densities of galaxies over regions with several tens of degrees. As a matter of fact, Narlikar *et al.* (2003) explain the peak at $l = 200$ and other peaks too in the power spectrum in terms of rich clusters of galaxies. The galaxy counts in López-Corredoira & Betancort-Rijo (2004, Fig. 1) also suggest structures of this kind and size. I did the cross-correlation of these near-infrared galaxy counts with COBE-DMR-90 GHz and found nothing significant (cross-corr. for $|b| > 20^\circ$: $5.5(\pm 11.2) \times 10^{-7}$ K). However, as said in subsection 2.3, using the WMAP data and 2MASS galaxy catalog, Cao *et al.* (2006) show that the non-gaussianity of the 2MASS galaxies is imprinted on WMAP maps, and Cabre *et al.* (2006) cross-correlate the third-year-WMAP data with optical galaxy samples extracted from the SDSS-DR4 giving a positive signal-to-noise of about 4.7.

Another consideration that could point to the importance of the non-cosmological extragalactic contamination over photons of cosmological origin is that the simulations of gravitational lensing of the microwave background by galaxy clusters at $z < 1$ under any plausible Big Bang model variation produces far more dispersion in the angular size of the primary acoustic peaks than WMAP observations allow (Lieu & Mittaz 2005). When all the effects are taken together, it is difficult to understand how WMAP could reveal no evidence whatsoever of lensing by groups and clusters. Cool spots in the microwave background are too uniform in size to have travelled from $z = 1000$–1500 to us. There should be a spread of sizes around the average, with some of these cool spots noticeably larger and others noticeably smaller. But this dispersion of sizes

is not seen in the data. Too many cool spots have the same size. Moreover, the observed WMAP Sunyaev–Zel'dovich effect caused by the clusters only accounts for about 1/4 of the expected decrement (Lieu *et al.* 2006); although the level of Sunyaev–Zel'dovich effect is observed as in the predictions in radio (Bonamente *et al.* 2006). Effects like a central cooling flow in clusters, the abundance of hot cluster gas, large scale radial decline in the temperature, uncertainties in the β-model and the role played by cluster radio sources are too weak to change the estimation (Lieu *et al.* 2006). Under Big Bang premises, this implies that the cosmological parameters (including the Hubble constant, the amount of dark matter, etc.) used to predict the original, pre-lensed sizes of the cool and hot spots in the microwave background might be wrong or some of these cool spot structures are caused by nearby physical processes and are not really remnants of the creation of the Universe; or there is some other, unknown factor damping the effects of dispersion and focusing. It was speculated that the large-scale curvature of space may not entirely be an initial value problem related to inflation. The absence of gravitational lensing of the CMB points to the possibility that even effects on light caused by wrinkles in the space of the late (nearby) Universe have been compensated for, beyond some distance scale, by a mechanism that maintains a flat geometry over such scales. Or high energy electrons may synchrotron radiate in the intracluster magnetic field of strength B $<\sim 1\mu$G to produce cluster microwave emissions in the WMAP passbands that account for the missing Sunyaev–Zel'dovich effect flux (Lieu & Quenby 2006). However, one is also tempted to interpret this in terms of the importance of extragalactic non-cosmological contribution in the microwave anisotropies.

6. Summary and conclusions

Summing up, CMBR anisotropies are not totally free from solar system, Galactic or extragalactic contamination, and an accurate way of correcting all these contributions is still to be devised since we do not have accurate information on the microwave emission of any of these contributions.

Considering only the Galactic dust component, the main topic of this paper, it was shown that the methods used to remove it (templates, cross-correlations, assumption of a Galactic power spectrum, MEM, ILC, etc.) are all inaccurate and one should not expect to produce maps clean of Galactic dust contamination by applying them. The analysis of Galactic latitude dependence of these anisotropies and the fact that the power spectrum is nearly independent of the frequency over a range 50–250 GHz can be considered at least as a proof that the Galactic dust emission is lower than a 10% on the $\sim$ 1 degree scale [or double of this value considered to 2σ], possibly higher for lower l multipoles. This uncertainty in the Galactic contamination may produce important systematic errors in some cosmological parameters. The determination of the error bars in the cosmological parameters taking into account the possible foreground contamination values is beyond the scope of this paper. Further research is needed in this respect. In any case, one thing is clear: the present error bars calculated for the cosmological parameters (e.g., Spergel *et al.* 2007) are very significantly underestimated and the range of possible values is not as small as indicated by the "precision cosmology" claim.

Therefore, I conclude that the most pessimistic possibilities among those claimed in López-Corredoira (1999) (that Galactic contamination could be dominant) is discarded, but I keep my position that we are far from achieving a "precision cosmology"

with CMBR data. My suggestion with regard to the problem would be that rather than performing an incorrect subtraction of the foreground and claiming that a clean map is obtained that is purely cosmological, we might use better the map without corrections or only with a first-order approximation subtraction, and calculate the cosmological parameters with the appropriate error bars taking into account the uncertainties in the foregrounds (Galactic or any other), to be constrained only by the frequency dependence.

Acknowledgements

Thanks are given to J. A. Rubiño-Martín (IAC, Tenerife, Spain) for providing some numerical codes to read HEALPix format data and the calculation of the power spectrum; to H. K. Eriksen (Inst. Theor. Astroph., Oslo, Norway) for providing me the maps of the simulations derived in his paper Eriksen *et al.* (2005); to C. M. Gutiérrez (IAC, Tenerife, Spain) and the anonymous referee for helpful comments on this paper; and to T. J. Mahoney (IAC, Tenerife, Spain) for proof-reading this paper.

References

Abramo, L. R., Sodre, L. Jr., Wuensche, C. A. 2006, *Phys. Rev. D*, **74(8)**, 3515.
Bennett, C. L., Halpern, M., Hinshaw, G. *et al.* 2003a, *ApJS*, **148**, 1.
Bennett, C. L., Hill, R. S., Hinshaw, G. *et al.* 2003b, *ApJS*, **148**, 97.
Bernard, J. P., Abergel, A., Ristorcelli, I. *et al.* 1999, *A&A*, **347**, 640.
Bernui, A., Mota, B., Reboucas, M. J., Tavakol, R. 2007a, *A&A*, **464**, 479.
Bernui, A., Tsallis, C., Villela, T. 2007b, astro-ph/0703708.
Bernui, A., Villela, T., Wuensche, C. A., Leonardi, R., Ferreira, I. 2006, *A&A*, **454**, 409.
Bonamente, M., Joy, M. K., LaRoque, S. J., Carlstrom, J. E., Reese, E. D., Dawson, Kyle S. 2006, *ApJ*, **647**, 25.
Boughn, S., Crittenden, R. 2004, *Nature*, **427**, 45.
Cabré, A., Gaztañaga, E., Manera, M., Fosalba, P., Castander, F. 2006, *MNRAS*, **372**, L23.
Cao, L., Chu, Y.-Q., Fang, L.-Z. 2006, *MNRAS*, **369**, 645.
Casassus, S., Readhead, A. C. S., Pearson, T. J., Nyman, L.-A., Shepherd, M. C., Bronfman, L. 2004, *ApJ*, **603**, 599.
Chiang, L.-Y., Naselsky, P. D., Coles, P., 2007, *ApJ*, **664**, 8.
Chiang, L.-Y., Naselsky, P. D., Verkhodanov, O. V., Way, M. J. 2003, *ApJ*, **590**, 65.
Chyzy, K. T., Novosyadlyj, B., Ostrowski, M. 2005, astro-ph/0512020.
Coles, P., Dineen, P., Earl, J., Wright, D. 2004, *MNRAS*, **350**, 983.
Copi, C. J., Huterer, D., Schwarz, D. J., Starkman, G. D. 2006, *MNRAS*, **367**, 79.
Copi, C. J., Huterer, D., Schwarz, D. J., Starkman, G. D. 2007, *Phys. Rev. D*, **75(2)**, 3507.
Coulson, D., Ferreira, P., Graham, P., Turok, N. 1994, *Nature*, **368**, 27.
Davies, R. D., Gutiérrez, C. M., Hopkins, J. *et al.* 1996, *MNRAS*, **278**, 883.
de Oliveira-Costa, A., Kogut, A., Devlin, M. J. *et al.* 1997, *ApJ*, **482**, L17.
de Oliveira-Costa, A., Tegmark, M., Page, L. A., Boughn, S. P. 1998, *ApJ*, **509**, L9.
de Oliveira-Costa, A., Tegmark, M., Gutiérrez, C. M., Jones, A. W., Davies, R. D., Lasenby, A. N., Rebolo, R., Watson, R. A. 1999, *ApJ*, **527**, L9.
de Oliveira-Costa, A., Tegmark, M., Finkbeiner, D. P. *et al.* 2002, *ApJ*, **567**, 363.
de Oliveira-Costa, A., Tegmark, M., Zaldarriaga, M., Hamilton, A. 2004, *Phys. Rev. D*, **69(6)**, 3516.
de Oliveira-Costa, A., Tegmark, M. 2006, *Phys. Rev. D*, **74(2)**, 3005.
Draine, B. T., Lazarian, A. 1998, *ApJ*, **494**, L19.
Draine, B. T., Lazarian, A. 1999, *ApJ*, **512**, 740.
Dupac, X., Bernard, J. P., Boudet, N., Giard, M., Lamarre, J. M., Mény, C., Pajot, F., Ristorcelli, I. 2004, In: Multiwavelength Cosmology (Astroph. Space Sci. library, vol. 301), Kluwer Academic Publishers, Dordrecht (The Netherlands), p. 89

Eriksen, H. K., Hansen, F. K., Banday, A. J., Górski, K. M., Lilje, P. B. 2004a, *ApJ*, **605**, 14.
Eriksen, H. K., Banday, A. J., Górski, K. M., Lilje, P. B. 2004b, *ApJ*, **612**, 633.
Eriksen, H. K., Banday, A. J., Górski, K. M., Lilje, P. B. 2005, astro-ph/0508196.
Fernández-Cerezo, S., Gutiérrez, C. M., Rebolo, R. *et al.* 2006, *MNRAS*, **370**, 15.
Ferreira, P. G., Górski, K. M., Magueijo, J. 1998, *ApJ*, **503**, L1.
Finkbeiner, D. P., Davis, M., Schlegel, D. J. 1999, *ApJ*, **524**, 867.
Finkbeiner, D. P., Langston, G. I., Minter, A. H. 2004, *ApJ*, **617**, 350.
Fixsen, D. J. 2003, *ApJ*, **594**, L67.
Ganga, K., Cheng, E., Meyer, S., Page, L. 1993, *ApJ*, **410**, L57.
Gutierrez de La Cruz, C. M., Davies, R. D., Rebolo, R., Watson, R. A., Hancock, S., Lasenby, A. N. 1995, *MNRAS*, **442**, 10.
Gutiérrez, C. M., Rebolo, R., Watson, R. A., Davies, R. D., Jones, A. W., Lasenby, A. N. 2000, *ApJ*, **529**, 47.
Halverson, N. W., Leitch, E. M., Pryke, C. *et al.* 2002, *ApJ*, **568**, 38.
Hammersley, P. L., Garzón, F., Mahoney, T., Calbet, X. 1995, *MNRAS*, **273**, 206.
Hansen, F. K., Banday, A. J., Górski, K. M. 2004a, *MNRAS*, **354**, 641.
Hansen, F. K., Balbi, A., Banday, A. J., Górski, K. M. 2004b, *MNRAS*, **354**, 905.
Hansen, F. K., Banday, A. J., Eriksen, H. K., Górski, K. M., Lilje, P. B. 2006, *ApJ*, **648**, 784.
Herbstmeier, U., Ábrahám, P., Lemke, D. *et al.* 1998, *A&A*, **332**, 739
Kiss, C., Ábrahám, P., Klaas, U., Lemke, D., Héraudeau, P., del Burgo, C., Herbstmeier, U. 2003, *A&A*, **399**, 177.
Kogut, A., Banday, A. J., Bennett, C. L. *et al.* 1996, *ApJ*, **460**, 1.
Land, K., Magueijo, J. 2007, *MNRAS*, **378**, 153.
Lawrence, A. 2001, *MNRAS*, **323**, 147.
Leitch, E. M., Readhead, A. C. S., Pearson, T. J., Myers, S. T. 1997, *ApJ*, **486**, L23.
Leitch, E. M., Readhead, A. C. S., Pearson, T. J., Myers, S. T., Gulkis, S., Lawrence, C. R. 2000, *ApJ*, **532**, 37.
Lieu, R., Mittaz, J. P. D. 2005, *ApJ*, **628**, 583.
Lieu, R., Mittaz, J. P. D., Zhang, S.-N. 2006, *ApJ*, **648**, 176.
Lieu, R., Quenby, J. 2006, astro-ph/0607304.
Liu, X., Zhang, S. N. 2005, *ApJ*, **633**, 542.
Liu, X., Zhang, S. N. 2006, *ApJ*, **636**, L1.
López-Corredoira, M. 1999, *A&A*, **346**, 369.
López-Corredoira, M., Cabrera-Lavers, A., Garzón, F., Hammersley, P. L. 2002, *A&A*, **394**, 883.
López-Corredoira, M., Betancort-Rijo, J. 2004, *A&A*, **416**, 1.
Magueijo, J. 2000a, *ApJ*, **528**, L57.
Magueijo, J. 2000b, *ApJ*, **532**, L157.
Masi, S., Ade, P. A. R., Bock, J. J. *et al.* 2001, *ApJ*, **553**, L93.
McEwen, J. D., Hobson, M. P., Lasenby, A. N., Mortlock, D. J. 2006, *MNRAS*, **371**, L50.
Mukherjee, P., Jones, A. W., Kneissl, R., Lasenby, A. N. 2001, *MNRAS*, **320**, 224.
Narlikar, J. V., Vishwakarma, R. G., Hajian, A., Souradeep, T., Burbidge, G., Hoyle, F. 2003, *ApJ*, **585**, 1.
Naselsky, P. D., Novikov, I. G., Chiang, L.-Y. 2006, *ApJ*, **642**, 617.
Netterfield, C. B., Ade, P. A. R., Bock, J. J. *et al.* 2002, *ApJ*, **571**, 604.
Paladini, R., Montier, L., Giard, M., Bernard, J. P., Dame, T. M., Ito, S., Macías-Pérez, J. F. 2007, *A&A*, **465**, 839.
Pando, J., Valls-Gabaud, D., Fang, L.-Z. 1998, *Phys. Rev. Lett.*, **81(21)**, 4568.
Park, C.-G. 2004, *MNRAS*, **349**, 313.
Raeth, C., Schuecker, P., Banday, A. J. 2007, MNRAS, **380**, 466.
Rakić, A., Schwarz, D. J. 2007, *Phys. Rev. D*, **75(10)**, 3002.
Robitaille, P.-M. 2007, Progress in Physics 1/2007, 3.
Rubiño-Martín, J. A., Aliaga, A. M., Barreiro, R. B. *et al.* 2006, *MNRAS*, **369**, 909.
Schlegel, D. J., Finkbeiner, D. P., Davis, M. 1998, *ApJ*, **500**, 525.
Spergel, D. N., Bean, R., Dore, O. *et al.* 2007, *ApJS*, **170**, 377.
Tegmark, M. 1998, *ApJ*, **502**, 1.
Then, H. 2006, *MNRAS*, **373**, 139.
Tojeiro, R., Castro, P. G., Heavens, A. F., Gupta, S. 2006, *MNRAS*, **365**, 265.

Tristam, M., Patanchon, G., Macías-Pérez, J. F. *et al.* 2005, *A&A*, **436**, 785.
Unavane, M., Gilmore, G., Epchtein, N., Simon, G., Tiphène, D., Batz, B. de 1998, *MNRAS*, **295**, 119.
Verschuur, G. L. 2007, arXiv:0704.1125.
Vielva, P., Martínez-González, E., Barreiro, R. B., Sanz, J. L., Cayón, L. 2004, *ApJ*, **609**, 22.
Watson, R. A., Rebolo, R., Rubiño-Martín, J. A., Hildebrandt, S., Gutiérrez, C. M., Fernández-Cerezo, S., Hoyland, R. J., Battistelli, E. S. 2005, *ApJ*, **624**, L89.

J. Astrophys. Astr. (2007) **28,** 117–131

Effects of the Size of Cosmological N-body Simulations on Physical Quantities – II: Halo Formation and Destruction Rate

Jayanti Prasad
Harish-Chandra Research Institute, Allahabad 211 019, India.
e-mail: jayanti@mri.ernet.in

Received 2007 February 22; accepted 2007 September 5

Abstract. In this study we show how errors due to finite box size affect formation and the destruction rate for haloes in cosmological N-body simulations. In an earlier study we gave an analytic prescription of finding the corrections in the mass function. Following the same approach, in this paper we give analytical expressions for corrections in the formation rate, destruction rate and the rate of change in comoving number density, and compute their expected values for the power law ($n = -2$) and LCDM models.

Key words. Large scale structure of Universe—N-body simulations—galaxies.

1. Introduction

In prevalent models of structure formation in the Universe, galaxies, clusters of galaxies and other large-scale structures are believed to have formed due to gravitational amplification of small perturbations (Peebles 1980; Padmanabhan 1993, 2002; Peacock 1999; Bernardeau *et al.* 2002). In cold dark matter dominated models these perturbations collapse hierarchically, i.e., smaller structures form first and then they merge and form larger structures.

In models of structure formation the main goal is to understand the growth of perturbations at various scales. It has been found that the linear perturbation theory closely follows the actual growth of perturbation at any scale as long as the amplitude of perturbation at that scale is small. However, the linear approximation breaks down once the amplitude becomes large because perturbations at various scales couple with each other and the system becomes nonlinear. Many approximations have been proposed (Zel'dovich 1970; Gurbatov *et al.* 1989; Saichev & Shandarin 1989; Matarrese *et al.* 1992; Brainerd *et al.* 1993; Hui & Bertschinger 1996; Bagla & Padmanabhan 1994; Sahni & Coles 1995) to explain nonlinear gravitational clustering but their validity has been limited to special cases. Cosmological N-body simulations are the main tools to understand gravitational clustering in nonlinear regime (Efstathiou *et al.* 1985; Bertschinger 1998; Bagla & Padmanabhan 1997a; Bagla 2005).

In N-body simulations we simulate a representative region of the universe which is in general large but finite. This restricts us from incorporating perturbations at scales greater than the size of the simulation box. We cannot also consider perturbations at

scales smaller than the grid length. This means that there is a natural truncation of power spectrum at large as well as small wave numbers. It has been shown in earlier studies (Peebles 1974, 1985; Little *et al.* 1991; Bagla & Padmanabhan 1997b) that the truncation of power spectrum at small scales does not affect large scale clustering in any significant way. However, significant effects of the truncation of power spectrum at large scales on clustering at small scales have been found (Gelb & Bertschinger 1994a, b; Bagla & Ray 2005; Power & Knebe 2006; Bagla & Prasad 2006, hereafter BP06).

Apart from N-body simulations, an analytical prescription given by Press & Schechter (1974) has been used to find mass function and related quantities. In this prescription, the fraction of mass in the collapsed objects having mass greater than a certain value, is identified with the fraction of volume in the initial density field which had density contrast, filtered over an appropriate scale, greater than a critical density contrast. The Press–Schechter formalism has been used in many studies to compute merging and mass function (Bond *et al.* 1991; Bower 1991; Lacey & Cole 1993, 1994; Sasaki 1994; Kitayama & Suto 1996; Cohn *et al.* 2001). In many of these studies, the main goal has been to test the results based on the Press–Schechter formalism against N-body simulations. Here our goal is to find the analytic corrections in the formation and destruction rate of haloes in N-body simulations due to finite box size of the simulation box.

The effects of the box size on the physical quantities in N-body simulations have been studied before. Gelb & Bertschinger (1994a) showed that the rms fluctuation in the mass and pairwise velocity dispersion at a given scale are underestimated when the size of the simulation box is reduced. Bagla & Ray (2005) computed the finite box effects on the average two-point correlation function, power spectrum and cumulative mass function in N-body simulations. They found that the required box size for simulating the LCDM model at high redshift is much larger than what people generally use. Power & Knebe (2006) studied the effects of box size on kinematical properties of cold dark matter haloes and concluded that the distribution of internal properties of haloes (like spin parameter) is not very sensitive to the box size. So far, most of the studies used N-body simulations for drawing their conclusions. In one of our earlier studies (BP06) we proposed an analytic prescription for estimating the corrections due to box size in N-body simulations for rms fluctuations in mass, two-point correlation function and the mass function.

The mass function is an important physical quantity in nonlinear gravitational clustering. It has been used to compare various theoretical models with observations (White *et al.* 1993; Ma & Bertschinger 1994; Ostriker & Steinhardt 1995; Bagla *et al.* 1996; White 2002). However, it does not contain all the information that is needed for various physical processes. For example, if we want to know the comoving number density of ionizing sources at any redshift then the mass function contains insufficient information since not all bound objects become the source of ionization, only those objects become ionizing sources which form during a particular period of time (Chiu & Ostriker 2000). In this situation we need to know the formation and destruction rate of collapsed objects as a function of redshift. The rate at which quasars form in the early universe depend in an important way on the merger rate which is another manifestation of the formation rate (Carlberg 1990). Apart from these there are many other cases like understanding the epoch of cluster formation etc., for which the formation and destruction rate of haloes are important (Bower 1991). In the present study we

carry forward our program and give analytic expressions for the corrections in the formation and destruction rate and estimate their values for the power law ($n = -2$) and LCDM models.

The plan of this paper is as follows. In section 2, we give the basic equations which are needed for our analysis. We give the expression for the clustering amplitude (rms fluctuations in mass), mass function and the rate of change in the comoving number density due to formation and destruction of haloes. In section 2.3, we show that the rate of change of comoving number density around a mass M can be written in terms of the rate at which the haloes of that mass are formed and the rate at which the haloes of that mass are destroyed due to merging. We present our results in section 3 and discuss their implications for the power law and LCDM models. In section 4 we summarize our results.

2. Basic equations

2.1 *Clustering amplitude*

In order to estimate the effects of box size on physical quantities in the linear regime, we use the mass variance $\sigma^2(r)$ (Peebles 1993; Peacock 1998; Padmanabhan 2002) as the base quantity.

$$\sigma^2(r) = 9 \int \frac{k^3 P(k)}{2\pi^2} \left(\frac{\sin kr - kr \cos kr}{k^3 r^3} \right)^2 \frac{dk}{k}. \tag{1}$$

On the basis of the correction in $\sigma^2(r)$, we find the corrections in other physical quantities. In BP06, we showed that the variance in mass is suppressed at all scales when we reduce the size of the simulation box. For example, if its actual value at scale r is $\sigma_0^2(r)$ then in the initial conditions of a simulation we obtain $\sigma^2(r, L_{\rm Box})$ where:

$$\sigma^2(r, L_{\rm Box}) = \sigma_0^2(r) - \sigma_1^2(r, L_{\rm Box}). \tag{2}$$

Here $L_{\rm Box}$ is the size of the simulation box and $\sigma_1^2(r, L_{\rm Box})$ is the correction due to the finite box size. Note that here $\sigma_1^2(r, L_{\rm Box})$ is a positive quantity so clustering amplitude $\sigma^2(r)$ is always underestimated when we reduce the size of the simulation box.

2.2 *Mass function and number density*

In the Press–Schechter formalism, we consider the initial density field as Gaussian random and smooth it over a filter of size r (or mass M), then the fraction of mass in the collapsed objects $F(>M)$, having mass greater than M, at the final epoch, can be identified with the fraction of volume in the initial density field which had smoothed density contrast greater than some critical density contrast δ_c which is computed on the basis of the spherical collapse model.

$$F(>M, t) = \frac{2}{\sqrt{\pi}} \int_{\frac{\delta_c(t)}{\sqrt{2}\sigma(M)}}^{\infty} e^{-x^2} dx = erfc\left(\frac{\delta_c(t)}{\sqrt{2}\sigma(M)} \right). \tag{3}$$

The comoving number density $N_{PS}(M,t)dM$ of objects which have mass in the range $[M, M+dM]$ at time t is given by

$$N_{PS}(M,t)dM = \frac{\rho_0}{M} \times \frac{dF(>M)}{dM} dM$$
$$= \sqrt{\frac{2}{\pi}} \frac{\rho_0}{M} \left(-\frac{\delta_c(t)}{\sigma^2(M)} \frac{d\sigma(M)}{dM} \right) \exp\left(-\frac{\delta_c^2(t)}{2\sigma^2(M)} \right) dM. \quad (4)$$

Here we can use $\delta_c(t) = \delta_c / D(t)$ where $D(t)$ is the linear growth factor which depends on the cosmological model being considered and $\delta_c(t)$ is taken to be 1.68 at the present epoch.

2.3 *The rate of change of number density*

We can find the rate of change in the comoving number density per unit time for objects which have mass in the range $[M, M+dM]$ from equation (4).

$$\left(\frac{dN_{PS}(M,t)}{dt} \right) dM = \sqrt{\frac{2}{\pi}} \frac{\rho_0}{M} \left(\frac{1}{D^2(t)} \frac{dD(t)}{dt} \right) \left(\frac{\delta_c}{\sigma^2(M)} \frac{d\sigma(M)}{dt} \right)$$
$$\times \left[1 - \frac{\delta_c^2}{\sigma^2(M) D^2(t)} \right] \exp\left(-\frac{\delta_c^2}{2\sigma^2(M) D^2(t)} \right) dM. \quad (5)$$

We can identify the first and second terms of the right-hand side with the destruction and formation rate respectively (Sasaki 1994; Kitayama & Suto 1996).

$$\left(\frac{dN_{PS}(M,t)}{dt} \right) dM = -\frac{1}{D(t)} \frac{dD(t)}{dt} \left[1 - \frac{\delta_c^2}{\sigma^2(M) D^2(t)} \right] N_{PS}(M,t) dM$$
$$= -\left(\frac{dN_{\rm Dest}(M,t)}{dt} \right) dM + \left(\frac{dN_{\rm Form}(M,t)}{dt} \right) dM. \quad (6)$$

The formation rate $(dN_{\rm Form}(M,t)/dt)dM$ quantifies the change in the comoving number density of objects around mass M, per unit time, due to the formation of objects in that mass range when objects of mass smaller than M merge together.

$$\left(\frac{dN_{\rm Form}(M,t)}{dt} \right) dM = \frac{1}{D(t)} \frac{dD(t)}{dt} \left[\frac{\delta_c^2}{\sigma^2(M) D^2(t)} \right] N_{PS}(M,t) dM$$
$$= \sqrt{\frac{2}{\pi}} \frac{\rho_0}{M} \left(\frac{1}{D^4(t)} \frac{dD(t)}{dt} \right) \left(-\frac{\delta_c^3}{\sigma^4(M)} \frac{d\sigma(M)}{dt} \right)$$
$$\times \exp\left(-\frac{\delta_c^2}{2\sigma^2(M) D^2(t)} \right) dM. \quad (7)$$

The destruction rate $(dN_{\rm Dest}(M,t)/dt)dM$ quantifies the rate of change of comoving number density of haloes in the mass range $[M, M+dM]$ when the haloes in that mass range merge together and form bigger haloes.

$$\left(\frac{dN_{\rm Dest}(M,t)}{dt}\right)dM = \frac{1}{D(t)}\frac{dD(t)}{dt}N_{PS}(M,t)dM$$

$$= \sqrt{\frac{2}{\pi}}\frac{\rho_0}{M}\left(\frac{1}{D^2(t)}\frac{dD(t)}{dt}\right)\left(-\frac{\delta_c}{\sigma^2(M)}\frac{d\sigma(M)}{dt}\right)dM$$

$$\times \exp\left(-\frac{\delta_c^2}{2\sigma^2(M)D^2(t)}\right)dM. \tag{8}$$

3. Box corrections

On the basis of the correction due to finite box size in $\sigma^2(r)$ (equation (2)) we can find the corrections in the number density, the rate of change in the number density, the rate of formation and the rate of destruction, i.e., equations (4), (5), (7), and (8) respectively.

3.1 *Corrections in comoving number density*

On the basis of the Press–Schechter formalism the comoving number density of objects in a mass range around a mass scale can be related to the rms fluctuations in the mass at that scale. Using the technique which we have developed to compute the corrections due to finite box size in the rms fluctuations in mass (see equation (2)), we can find the corrections in the comoving number density also.

The comoving number density which we expect in cosmological N-body simulations is given by

$$N_{PS}(M,t)dM = \sqrt{\frac{2}{\pi}}\frac{\rho_0}{M}\frac{\delta_c}{D(t)}\left(-\frac{1}{\sigma^2(M)}\frac{d\sigma(M)}{dM}\right)$$

$$\times \exp\left(-\frac{\delta_c^2}{2\sigma^2(M)D^2(t)}\right)dM$$

$$= N_{PS,0}(M,t)dM - N_{PS,1}(M,t)dM. \tag{9}$$

Here $N_{PS,0}(M,t)$ and $N_{PS}(M,t)$ are the theoretical (box size infinite) and the actual (box size finite) comoving number densities respectively.

$$N_{PS,0}(M,t)dM = \sqrt{\frac{2}{\pi}}\frac{\rho_0}{M}\frac{\delta_c}{D(t)}\left(\frac{-1}{\sigma_0^2(M)}\frac{d\sigma_0(M)}{dM}\right)$$

$$\times \exp\left(\frac{-\delta_c^2}{2\sigma_0^2(M)D^2(t)}\right)dM \tag{10}$$

and

$$N_{PS,1}(M,t)dM = \frac{1}{2}\left[1-\left(1-\frac{\sigma_1^2}{\sigma_0^2}\right)^{-3/2}\left(1-\frac{d\sigma_1^2}{d\sigma_0^2}\right)\right.$$
$$\left.\times\ \exp\left\{-\frac{\delta_c^2}{2D^2(t)}\left(\frac{1}{\sigma^2}-\frac{1}{\sigma_0^2}\right)\right\}\right] N_{PS,0}(M,t)dM \qquad (11)$$

or

$$N_{PS,1}(M,t)dM \approx \frac{1}{2}\left[\frac{d\sigma_1^2}{d\sigma_0^2}-\frac{3\sigma_1^2}{2\sigma_0^2}+\frac{\delta_c^2\sigma_1^2}{2D^2(t)\sigma_0^4}\right] N_{PS,0}(M,t)dM,$$
$$\text{if } \sigma_1^2/\sigma_0^2 < 1. \qquad (12)$$

In equation (12) the coefficient of $N_{PS,0}(M,t)$ changes sign at the scale for which $\sigma_0 \approx \delta_c/D(t)\sqrt{3} = \delta_c(t)/\sqrt{3}$ so the number density of objects below this scale is overestimated and the number density above this scale is underestimated. This is in accordance with our earlier results (BP06).

Figure (1) shows the theoretical comoving number density $N_{PS,0}$ and the actual comoving number density N_{PS}, which we expect in an N-body simulation for the power law model ($n = -2$) at the present epoch ($z = 0$). In this case we consider a simulation box of size $128\,h^{-1}$Mpc and normalize the initial power spectrum such that

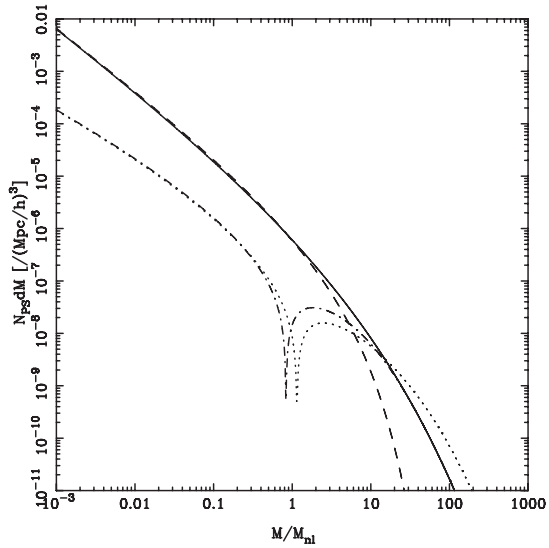

Figure 1. This figure shows the actual comoving number density $N_{PS,0}dM$ (solid line) and the comoving number density $N_{PS}\ dM$ (dashed line) which we expect in a cosmological N-body simulation at the present epoch ($z = 0$) for a power law ($n = -2$) model (see equations (9, 10)). Here we consider the size of the simulation box $128\,h^{-1}$Mpc and normalize the initial power spectrum such that the scale of non-linearity (r_{nl}) at $z = 0$ is $8\,h^{-1}$Mpc. In the figure, the exact (dot-dashed line) and approximate (dotted line) corrections in the comoving number density due to finite box size are also shown (see equations (11, 12)).

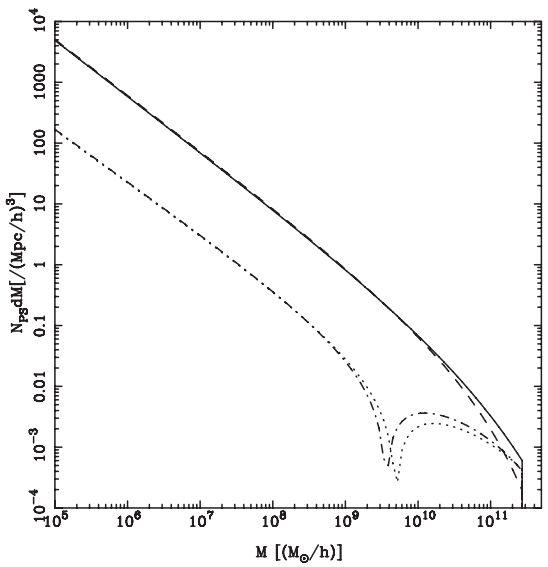

Figure 2. This figure shows the actual comoving number density $N_{PS,0}\,dM$ (solid line) and the comoving number density $N_{PS}\,dM$ (dashed line) which we expect in a cosmological N-body simulation for the LCDM model at $z = 6$ (see equations (9, 10)). Here the size of the simulation is taken to be $10\,h^{-1}$Mpc. We have also shown the exact (dot-dashed line) and approximate (dotted line) corrections in the comoving number density due to finite box size (see equations (11, 12)).

the scale of non-linearity at $z = 0$ is $8\,h^{-1}$Mpc. The actual and approximate corrections in the comoving number density are also shown in the figure. This figure shows that in N-body simulations the number density of large mass haloes is underestimated. However, it is overestimated for small mass haloes. This feature is more evident from the actual and approximate corrections (difference between the theoretical and actual values). In Fig. 2 we show the theoretical and the actual comoving number densities and corrections due to finite box size for the LCDM model. In this case we consider a simulation box of size $10\,h^{-1}$Mpc and compute physical quantities at $z = 6$. From Figs. 1 and 2 it is clear that the overall trend remains the same. This result is in agreement with the results of our earlier study (BP06).

3.2 *Formation rate*

The rate of formation of haloes of mass M which we expect in N-body simulation is given by equation (7)

$$\left(\frac{dN_{\rm Form}(M,t)}{dt}\right)dM = \sqrt{\frac{2}{\pi}}\frac{\rho_0}{M}\left(\frac{1}{D^4(t)}\frac{dD(t)}{dt}\right)\left(-\frac{\delta_c^3}{\sigma^4(M)}\frac{d\sigma(M)}{dt}\right) \times \exp\left(-\frac{\delta_c^2}{2\sigma^2(M)D^2(t)}\right)dM$$

$$= \left(\frac{dN_{\rm Form,0}(M,t)}{dt}\right)dM - \left(\frac{dN_{\rm Form,1}(M,t)}{dt}\right)dM. \quad (13)$$

Here $(dN_{\rm Form,0}(M,t)/dt)dM$ and $(dN_{\rm Form,1}(M,t)/dt)dM$ are the theoretical formation rate and the correction term respectively.

$$\left(\frac{dN_{\rm Form,0}(M,t)}{dt}\right)dM = \sqrt{\frac{2}{\pi}}\frac{\rho_0}{M}\left(\frac{1}{D^4(t)}\frac{dD(t)}{dt}\right)dM \times \left(-\frac{\delta_c^3}{\sigma_0^4(M)}\frac{d\sigma_0(M)}{dt}\right)\exp\left(-\frac{\delta_c^2}{2\sigma_0^2(M)D^2(t)}\right) \quad (14)$$

and

$$\left(\frac{dN_{\rm Form,1}(M,t)}{dt}\right)dM = \frac{1}{2}\left[1-\left(1-\frac{\sigma_1^2}{\sigma_0^2}\right)^{-5/2}\left(1-\frac{d\sigma_1^2}{d\sigma_0^2}\right) \times \exp\left\{-\frac{\delta_c^2}{2D^2(t)}\left(\frac{1}{\sigma^2}-\frac{1}{\sigma_0^2}\right)\right\}\right]\left(\frac{dN_{\rm Form,0}(M,t)}{dt}\right)dM \quad (15)$$

or

$$\left(\frac{dN_{\rm Form,1}(M,t)}{dt}\right)dM \approx \frac{1}{2}\left[\frac{d\sigma_1^2}{d\sigma_0^2}-\frac{5\sigma_1^2}{2\sigma_0^2}+\frac{\delta_c^2\sigma_1^2}{2D^2(t)\sigma_0^4}\right] \times \left(\frac{dN_{\rm Form,0}(M,t)}{dt}\right)dM \quad \text{if } \sigma_1^2/\sigma_0^2 < 1. \quad (16)$$

Figures 3 and 4 show the theoretical formation rate $(dN_{\rm Form,0}/dt)dM$ and the actual formation rate $(dN_{\rm Form}/dt)dM$ for the power law and LCDM models respectively. The parameters for the power law and LCDM model are the same as in Figs. 1 and 2 respectively. Since the formation rate is directly proportional to the comoving number density, it follows the same trend as the comoving number density. However, in this case the scale at which the correction term changes sign is different from the scale at which the correction term for the comoving number density changes, i.e., here it is the scale for which $\sigma_0 = \delta_c/D(t)\sqrt{5} = \delta_c(t)/\sqrt{5}$. From these figures we see that the formation rate of massive haloes is suppressed. However, that of low mass haloes is enhanced in N-body simulations when the box size is reduced. The main reason behind the suppression of the formation of large mass haloes is the absence of fluctuations in the initial density field at large scales due to truncation of power.

3.3 *Destruction rate*

Following the same approach as we have applied for the formation rate, we can find the corrections due to finite box size for destruction rate, i.e., equation (8) also.

$$\left(\frac{dN_{\rm Dest}(M,t)}{dt}\right)dM = \left(\frac{dN_{\rm Dest,0}(M,t)}{dt}\right)dM - \left(\frac{dN_{\rm Dest,1}(M,t)}{dt}\right)dM \quad (17)$$

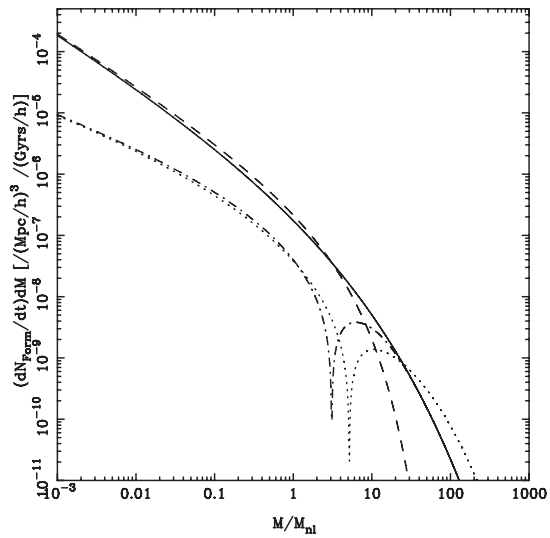

Figure 3. This figure shows the actual formation rate $(dN_{\rm Form,0}/dt)dM$ (solid line) and the formation rate $(dN_{\rm Form}/dt)dM$ (dashed line) which we expect in a cosmological N-body simulation for the power law model (see equations (13, 14)). All the parameters for this figure are the same as for Fig. 1. We have also shown the exact (dot-dashed line) and approximate (dotted line) corrections in the formation rate due to finite box size (see equations (15, 16)).

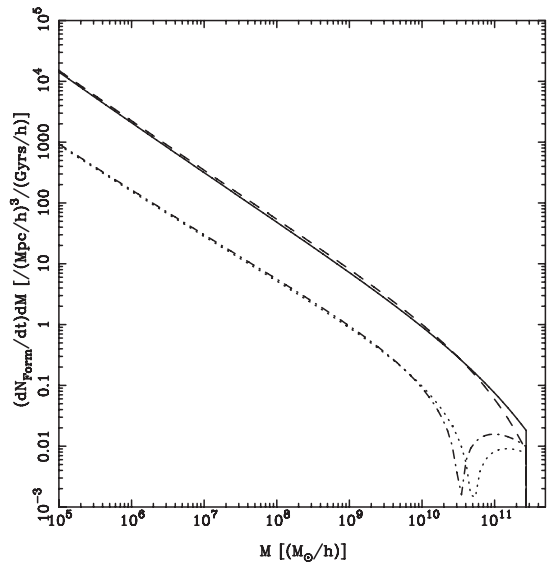

Figure 4. This figure shows the actual formation rate $(dN_{\rm Form,0}/dt)dM$ (solid line) and the formation rate $(dN_{\rm Form}/dt)dM$ (dashed line) which we expect in a cosmological N-body simulation for the LCDM model (see equations (13, 14)). For this figure all the parameters are identical to that of Fig. 2. The exact (dot-dashed line) and approximate (dotted line) corrections in the formation rate due to finite box size are also shown (see equations (15, 16)).

where

$$\left(\frac{dN_{\rm Dest,0}(M,t)}{dt}\right)dM = \sqrt{\frac{2}{\pi}}\frac{\rho_0}{M}\left(\frac{1}{D^2(t)}\frac{dD(t)}{dt}\right)\left(-\frac{\delta_c}{\sigma_0^2(M)}\frac{d\sigma_0(M)}{dt}\right) \times dM\exp\left(-\frac{\delta_c^2}{2\sigma_0^2(M)D^2(t)}\right) \tag{18}$$

and

$$\left(\frac{dN_{\rm Dest,1}(M,t)}{dt}\right)dM = \frac{1}{2}\left[1-\left(1-\frac{\sigma_1^2}{\sigma_0^2}\right)^{-3/2}\left(1-\frac{d\sigma_1^2}{d\sigma_0^2}\right) \times \exp\left\{-\frac{\delta_c^2}{2D^2(t)}\left(\frac{1}{\sigma^2}-\frac{1}{\sigma_0^2}\right)\right\}\right]\left(\frac{dN_{\rm Dest,0}(M,t)}{dt}\right) \tag{19}$$

or

$$\left(\frac{dN_{\rm Dest,1}(M,t)}{dt}\right)dM \approx \frac{1}{2}\left[\frac{d\sigma_1^2}{d\sigma_0^2}-\frac{3\sigma_1^2}{2\sigma_0^2}+\frac{\delta_c^2\sigma_1^2}{2D^2(t)\sigma_0^4}\right] \times \left(\frac{dN_{\rm Dest,0}(M,t)}{dt}\right) \quad \text{if } \sigma_1^2/\sigma_0^2 < 1. \tag{20}$$

Figures 5 and 6 show the destruction rates for the power law and LCDM model respectively. The parameters for the power law and LCDM models are the same as in Figs. 1 and 2 respectively. In this case the correction term changes sign at the same scale at which the correction term for the comoving number density changes. This is because $N_{PS,1}/N_{PS,0}$ and $N_{\rm Dest,1}/N_{\rm Dest,0}$ are equal and so $N_{PS,1}$ and $N_{\rm Dest,1}$ have the same scale of zero crossing. Here also we find that the destruction rate for the massive haloes are suppressed. However, they are enhanced for the low mass haloes when we reduce the size of the simulation box.

3.4 *Rate of change on number density*

The rate of change of the number density is defined as (see equation (7))

$$\left(\frac{N_{PS}(M,t)}{dt}\right)dM = \frac{-1}{D(t)}\frac{dD(t)}{dt}\left[1-\frac{\delta_c^2}{\sigma^2(M)D^2(t)}\right]N_{PS}(M,t)dM$$
$$= -\left(\frac{dN_{\rm Dest}(M,t)}{dt}\right)dM + \left(\frac{dN_{\rm Form}(M,t)}{dt}\right)dM. \tag{21}$$

We have already given the corrections for the formation rate $(dN_{\rm Form}/dt)dM$ and the destruction rate $(dN_{\rm Dest}/dt)dM$ in the last two sections. The correction in the rate of change of number density $(dN_{PS}(M,t)/dt)dM$ can be written in terms of the correction in the formation and destruction rates. In Figs. 7 and 8 we show the rate

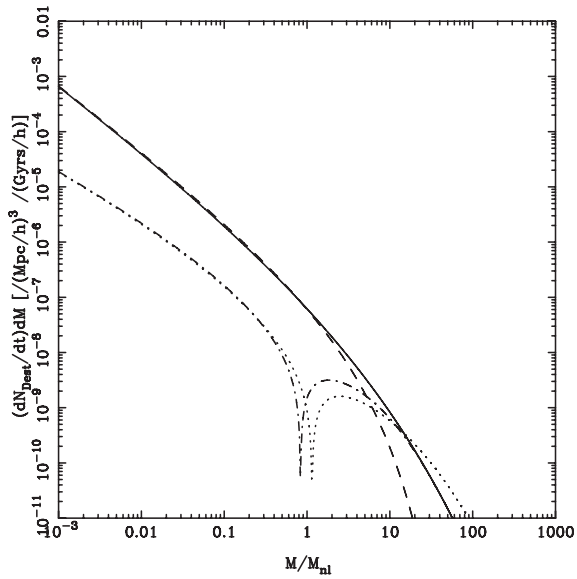

Figure 5. This figure shows the actual destruction rate $(dN_{\rm Dest,0}/dt)dM$ (solid line) and the formation rate $(dN_{\rm Dest}/dt)dM$ (dashed line) which we expect in a cosmological N-body simulation (see equations (17, 18)) for the power law model. All the parameters for this figure are identical to that of Fig. 1. We have also shown the exact (dot-dashed line) and approximate (dotted line) corrections in the destruction rates due to finite box size (see equations (19, 20)).

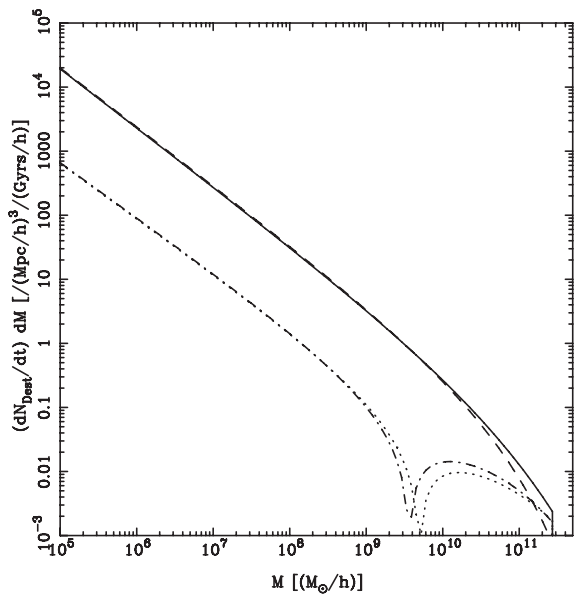

Figure 6. This figure shows the actual destruction rate $(dN_{\rm Dest,0}/dt)dM$ (solid line) and the formation rate $(dN_{\rm Dest}/dt)dM$ (dashed line) which we expect in a cosmological N-body simulation (see equations (17, 18)) for the LCDM simulation. All the parameters for this figure are identical to that of Fig. 2. The exact (dot-dashed line) and approximate (dotted line) corrections in the destruction rate due to finite box size are also plotted (see equations (19, 20)).

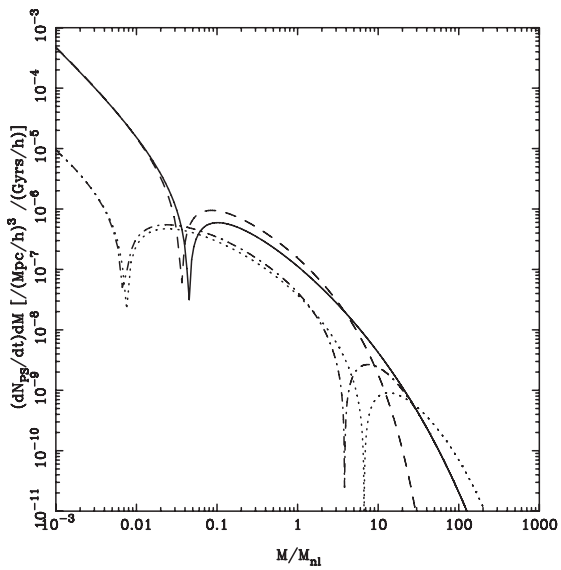

Figure 7. This figure shows the actual merger rate $(dN_{PS,0}/dt)dM$ (solid line) and the formation rate $(dN_{PS}/dt)dM$ (dashed line) which we expect in a cosmological N-body simulation for the power law ($n = -2$) (see equation (21)). All the parameters for this figure are the same as for Fig. 1. The exact (dot-dashed line) and approximate (dotted line) corrections in the rate of change in the comoving number density are also plotted.

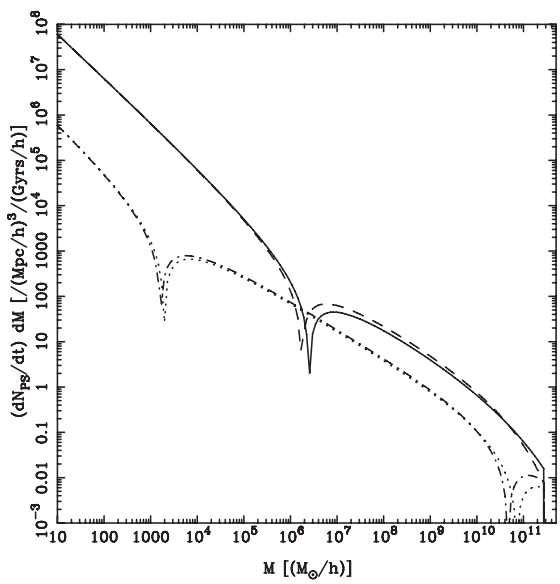

Figure 8. This figure shows the actual merger rate $(dN_{PS,0}/dt)dM$ (solid line) and the formation rate $(dN_{PS}/dt)dM$ (dashed line) which we expect in a cosmological N-body simulation for the LCDM model. All the parameters for this figure are the same as for Fig. 2. The exact (dot-dashed line) and approximate (dotted line) corrections in the rate of change in the comoving number density are also plotted.

of change of the comoving number density and the exact and approximate correction terms in it for the power law and LCDM model respectively. From equation (21) it is clear that for $\sigma(M) < \delta_c/D(t)$ the rate of change of number density $\dot{N}_{PS}$ is dominated by the formation rate, and for $\sigma(M) > \delta_c/D(t)$ by the destruction rate. For any time t, we can find a mass scale M_c for which $\sigma(M_c) = \delta_c/D(t)$, i.e., the formation and the destruction rate are equal and so there is no net change in the comoving number density of objects at that scale. In hierarchical clustering models, i.e., $\sigma(M)$ is a decreasing function of mass, comoving number density at large scales mainly changes due to the formation of massive haloes and at small scales due to destruction of smaller haloes.

The rate of change in the number density is underestimated at large and small scales, however, it is overestimated at intermediate scales. This feature is clear from Fig. 7 in which the actual and approximate error terms are positive at large and small scales but they are negative at intermediate scales and we have two zeroes crossing.

4. Discussion

In the hierarchical clustering models of structure formation, formation and destruction of haloes is a common process. In the present study we have shown that the formation and destruction rate of haloes due to gravitational clustering are affected significantly if the size of the simulation box is not sufficiently large. On the basis of the Press–Schechter formalism we have given the analytic expressions for the corrections in the comoving number density, formation rate, destruction rate and the rate of change in the number density of haloes at a given mass scale. We have considered the implications of our analysis for the power law ($n = -2$) and LCDM models. Since the box corrections are more important for models which have significant power at large scales, most of the models in which there is none or very less power at large scales are not affected by the size of the simulation box. However, models in which there is a lot of power at large scales (n is large and negative) the box effects can be quite large. In both the cases, i.e., power law and LCDM models, the scales at which we have shown the corrections are far below the size of the simulation box.

The main conclusions of the present study are as follows:

- If the size of the simulation box in N-body simulations is not large enough then the clustering amplitude is underestimated at all scales.
- At any given time, there is a scale above which merging is dominated by the formation rate and below which it is dominated by the destruction rate.
- The comoving number density of haloes is underestimated at large scales and overestimated at small scales when we reduce the size of the simulation box.
- The formation and destruction rate also get modified by reducing the size of the simulation box. Particularly, they are underestimated at large scales and overestimated at small scales.
- The suppression of the formation rate as well as destruction rate at large scales is mainly due to the absence of fluctuations in the initial density field at large scales due to limitation of box size.
- In N-body simulations which have small power at large scales the corrections due to box size can be ignored.

Observations suggest that cosmological perturbations were present at all scales in the initial density field that have been probed. Particularly in dark matter models

(Diemand *et al*. 2005, 2006) the index of the power spectrum at small scales becomes close to -3. So in order to simulate these models one has to be very careful in choosing the size of the simulation box. This is because if the index of power spectrum is close to -3 at the box scale also or if the scale of non-linearity is close to the size of the simulation box then the results can be significantly affected by the finite box effects.

In this and Bagla & Prasad (2006) we have presented our analytic results for taking into account the effects of finite box size on the mass function, formation and destruction rate and other physical quantities. In the next paper of this series we will present a detailed comparison of our analytic results with a set of cosmological N-body simulations of different box sizes.

Acknowledgements

I would like to thank Jasjeet Bagla for insightful comments and discussions. Numerical work for this study was carried out at cluster computing facilities at the Harish-Chandra Research Institute (http://www.cluster.mri.ernet.in). This research has made use of NASA's Astrophysical Data System.

References

Bagla, J. S., Padmanabhan, T. 1994, *MNRAS*, **266**, 227.
Bagla, J. S., Padmanabhan, T. 1997a, *Pramana*, **49** 161.
Bagla, J. S., Padmanabhan, T. 1997b, *MNRAS*, **286**, 1023.
Bagla, J. S., Padmanabhan, T., Narlikar, J. V. 1996, *Comments on Astrophysics*, **18**, 275.
Bagla, J. S., Ray, S. 2005, *MNRAS*, **358**, 1076.
Bagla, J. S., Prasad, J. 2006, *MNRAS*, **370**, 993.
Bagla, J. S. 2005, *CSci*, **88**, 1088.
Bernardeau, F., Colombi, S., Gaztanaga, R., Scoccimarro, R. 2002, *Phys. Rep.*, **367**, 1.
Bertschinger, E. 1998, *ARA&A*, **36**, 599.
Bond, J. R., Cole, S., Efstathiou, G., Kaiser, N. 1991, *ApJ*, **379**, 440.
Bower, R. G. 1991, *MNRAS*, **248**, 332.
Brainerd, T. G., Scherre R. J., Villumsen J. V. 1993, *ApJ*, **418**, 570.
Carlberg, R. G. 1993, *ApJ*, **350**, 505.
Chiu, W. A., Ostriker, J. P. 2000, *ApJ*, **534**, 507.
Cohn, J. D., Bagla, J. S., White, M. 2001, *MNRAS*, **325**, 1053.
Diemand, J., Moore, B., Stadel, J. 2005, *Nature*, **433**, 389.
Diemand, J., Kuhlen, M., Madau, P. 2006, *ApJ*, **649**, 1.
Efstathiou, G.; Davis, M.; White, S. D. M.; Frenk, C. S. 1985, *ApJ*, **57**, 241.
Gelb, J. M., Bertschinger, E. 1994a, *ApJ*, **436**, 467.
Gelb, J. M., Bertschinger, E. 1994b, *ApJ*, **436**, 491.
Gurbatov, S. N., Saichev, A. I., Shandarin, S.F. 1989, *MNRAS*, **236**, 385.
Hui, L., Bertschinger, E. 1996, *ApJ*, **471**, 1.
Kitayama, T., Suto, Y. 1996, *MNRAS*, **280**, 638.
Lacey, C., Cole, S. 1993, *MNRAS*, **262**, 627.
Lacey, C., Cole, S. 1994, *MNRAS*, **271**, 676.
Little, B., Weinberg, D. H., Park, C. 1991, *MNRAS*, **253**, 295.
Ma, C.-P., Bertschinger, E. 1994, *ApJ*, **434L**, 5.
Matarrese, S., Lucchin, F., Moscardini, L., Saez, D. 1992, *MNRAS*, **259**, 437.
Ostriker, J. P., Steinhardt, P. J. 1995, *Nature*, **377**, 600.
Padmanabhan, T. 1993, *Structure Formation in the Universe*, Cambridge University Press.
Padmanabhan, T. 2002, *Theoretical Astrophysics, Volume III: Galaxies and Cosmology*, Cambridge University Press.
Peacock, J. A. 1999, *Cosmological Physics*, Cambridge University Press.

Peebles, P. J. E. 1974, *A&A*, **32**, 391.
Peebles, P. J. E. 1980, *The large-scale structure of the universe*, Princeton University Press.
Peebles, P. J. E. 1985, *Apj*, **297**, 350.
Power, C., Knebe, A. 2006 *MNRAS*, **370**, 691.
Press, W. H., Schechter, P. 1974, *ApJ*, **187**, 425.
Sahni, V., Coles, P. 1995, *Phys. Rep.*, **262**, 1.
Sasaki, S. 1994, *Publ. Astron. Soc. Japan*, **46**, 427.
White, M. 2002, *ApJS*, **143**, 241.
White, S. D. M., Efstathiou, G., Frenk, C. S. 1993, *MNRAS*, **262**, 1023.
Zel'dovich, Ya B. 1970, *A&A*, **5**, 84.

J. Astrophys. Astr. (2007) **28,** 133–138

On Surface Tension for Compact Stars

R. Sharma* & S. D. Maharaj†
Astrophysics and Cosmology Research Unit, School of Mathematical Sciences, University of KwaZulu–Natal, Private Bag X54001, Durban 4000, South Africa.
**e-mail: 206526115@ukzn.ac.za*
†*e-mail: maharaj@ukzn.ac.za*

Received 2007 February 22; accepted 2007 May 15

Abstract. In an earlier analysis it was demonstrated that general relativity gives higher values of surface tension in strange stars with quark matter than neutron stars. We generate the modified Tolman–Oppenheimer–Volkoff equation to incorporate anisotropic matter and use this to show that pressure anisotropy provides for a wide range of behaviour in the surface tension than is the case with isotropic pressures. In particular, it is possible that anisotropy drastically decreases the value of the surface tension.

Key words. Relativity—pulsars—equation of state.

1. Introduction

Stars that are more compact than neutron stars, at present, have become a subject of considerable interest as they provide us natural laboratories for testing QCD. Over the last couple of decades, various models have been proposed to explain the compactness and properties of some of the observed compact objects. Pioneering works in this field have put forward new concepts of compact matter, namely strange stars (Witten 1984; Farhi & Jaffe 1984) and boson stars (Kaup 1968; Ruffini & Bonazzola 1969; Colpi *et al.* 1986). Due to the high matter densities within such stars one expects pressure to be anisotropic in general, *i.e.*, in the interior of such stars the radial pressure and tangential pressure are different. An anisotropic energy momentum is a topic which is often ignored in the calculations of compact stars. However, since the pioneering work of Bowers & Liang (1974) there has been extensive research in the study of anisotropic relativistic matter in general relativity. The analysis of static spherically symmetric anisotropic fluid spheres is important in relativistic astrophysics. Ruderman (1972) showed that nuclear matter may be anisotropic in the high density ranges of order 10^{15} gm cm^{-3} where nuclear interactions have to be treated relativistically. Anisotropy in compact objects may occur due to the existence of a solid core or the presence of type 3A superfluid (Kippenhahm & Weigert 1990), phase transition (Sokolov 1980), pion condensation (Sawyer 1972), slow rotation (Herrera & Santos 1997), mixture of two gases (Letelier 1980) or strong magnetic fields (Weber 1999). Also objects made up of self-interacting scalar particles known as boson stars are naturally anisotropic in their configurations. Anisotropic models for compact self gravitating objects have been studied by Herrera & Santos (1997); Rao *et al.* (2000); Corchero (2001); Mak & Harko

(2003); Ivanov (2002); Dev & Gleiser (2003); Hernández & Núñez (2004); Chaisi & Maharaj (2005), and many others. Anisotropic models for compact objects have been shown to achieve high red-shift values (Bowers & Liang 1974; Herrera & Santos 1997; Ivanov 2002; Mak & Harko 2003), and they are stable (Herrera & Santos 1997; Dev & Gleiser 2003). In this article, we show that pressure anisotropy may also affect the surface tension of compact stars. We believe that this aspect has not been considered yet in the context of anisotropic stellar models.

2. Surface tension of strange stars

In a recent paper by Bagchi *et al.* (2005), it has been shown that objects composed of u, d and s quarks popularly known as 'strange stars' give higher values of surface tension than neutron stars, a necessary criterion for the existence of stable strange stars in the Universe. This calculation is based on equations of state (EOS) for strange matter formulated by Dey *et al.* (1998). In an approximated linearized form, the EOS may be written as (Zdunik 2000; Gondek-Rosińska *et al.* 2000)

$$p = a(\rho - \rho_b), \tag{1}$$

where ρ is the energy density, ρ_b is the density at the surface, p is the isotropic pressure, and a is a parameter related to the velocity of sound ($a = dp/d\rho$).

To calculate the surface tension, one assumes that the star is a huge spherical ball composed of strange matter which is self-bound and non-rotating. The excess pressure on the surface of the star can be expressed as

$$|\Delta p|_{r=R} = \frac{2S}{R}, \tag{2}$$

where S is the surface tension of the star and R is the radius of curvature. At the surface

$$|\Delta p|_{r=R} = r_n \frac{dp}{dr}|_{r=R}, \tag{3}$$

where r_n is the radius of the quark particle given by $r_n = (1/\pi n)^{1/3}$ where n is the baryon number density. As strange stars are very compact, a relativistic treatment is necessary to find their configurations and other physical parameters. Thus for a given EOS, one uses the Tolman–Oppenheimer–Volkoff (TOV) equation (Oppenheimer & Volkoff 1939)

$$\frac{dp}{dr} = -\frac{G(\rho + p)\left[\frac{m(r)}{c^2 r} + \frac{4\pi r^2 p}{c^4}\right]}{c^2 r\left(1 - \frac{2Gm(r)}{r}\right)} \tag{4}$$

to find the surface tension of the star, making use of equations (2) and (3). This method helps to yield higher values of surface tension as compared to neutron stars including the possible explanation for the existence of strange stars in the Universe and other related phenomena like delayed γ-ray bursts (Bagchi *et al.* 2005).

However, at very high densities, anisotropy may be significant in such stars which may contribute to the surface tension. If we assume that pressure within such a star is anisotropic in general then the TOV equation (4) gets modified yielding different results as obtained by Bagchi *et al.* (2005). In the following sections, we derive the modified TOV equation with anisotropic pressure and perform some numerical calculations to show the effects of pressure anisotropy on the surface tension of compact stars.

3. Anisotropic TOV equation

We first formulate the modified TOV equation with anisotropic pressure. We assume the line element for a static spherical object in the standard form

$$ds^2 = -e^{\gamma(r)}c^2dt^2 + e^{\mu(r)}dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2), \tag{5}$$

where $\gamma(r)$ and $\mu(r)$ are the two unknown metric functions. Without any loss of generality, the energy momentum tensor for an anisotropic star may be written as

$$T_{ij} = (\rho c^2 + p_r)u_i u_j + p_r g_{ij} + (p_r - p_\perp)n_i n_j, \tag{6}$$

where u_i is the fluid four-velocity, n_i is a radially directed unit space-like vector. We assume that $p_r \neq p_\perp$ and $p_\perp - p_r = \Delta$ gives the measure of pressure anisotropy in this model.

The Einstein's field equations are then given by

$$\frac{8\pi G}{c^4}\rho = \frac{\left(1 - e^{-\mu}\right)}{r^2} + \frac{\mu' e^{-\mu}}{r}, \tag{7}$$

$$\frac{8\pi G}{c^4}p_r = \frac{\gamma' e^{-\mu}}{r} - \frac{\left(1 - e^{-\mu}\right)}{r^2}, \tag{8}$$

$$\frac{8\pi G}{c^4}p_\perp = \frac{e^{-\mu}}{4}\left(2\gamma'' + \gamma'^2 - \gamma'\mu' + \frac{2\gamma'}{r} - \frac{2\mu'}{r}\right), \tag{9}$$

where primes denote differentiation with respect to the radial coordinate r. Equations (7)–(9) may be combined together to yield

$$(\rho + p_r)\gamma' + 2p_r' + \frac{4}{r}(p_r - p_\perp) = 0 \tag{10}$$

which is a conservation equation.

If we write the metric function μ in terms of mass function $m(r)$ as

$$e^{-\mu} = 1 - \frac{Gm(r)}{c^2 r} \tag{11}$$

then equation (10) becomes

$$\frac{dp_r}{dr} = -(\rho + p_r)\frac{\left(\frac{Gm(r)}{c^2 r} + \frac{4\pi G r^2 p_r}{c^4}\right)}{r\left(1 - \frac{2Gm(r)}{c^2 r}\right)} + \frac{2}{r}(p_\perp - p_r). \tag{12}$$

Equation (12) is the the modified TOV equation in the presence of pressure anisotropy. For a given central density ρ_c or central pressure p_r^c and anisotropic parameter Δ, equation (12) may be integrated to find the mass $M = m(R)$ and radius R of the star provided the EOS $p_r = p_r(\rho)$ is known. Local anisotropy thus effects the geometry of the star.

At the surface of the star $r = R$, the radial pressure p_r vanishes. However, the tangential pressure $p_\perp$ is not necessarily zero at the surface. The two pressure profiles within the star should satisfy the following conditions: $p_r > 0$ and $p_\perp > 0$. The

maximum value of the anisotropic parameter Δ *vis-a-vis* the tangential pressure $p_\perp$ is constrained by the physical requirement that the radial pressure gradient dp_r/dr should be negative in the stellar interior; other physical requirements may, however, put a more stringent restriction on the values of Δ. Thus for finite values of $p_\perp$ at the boundary $\Delta(r = R) = p_\perp^b$, equation (12) becomes

$$\left.\frac{dp_r}{dr}\right|_{r=R} = -\frac{\rho_b \frac{GM}{c^2 R}}{R\left(1 - \frac{2GM}{c^2 R}\right)} + \frac{2p_\perp^b}{R}. \tag{13}$$

If $p_\perp^b$ is not negligible at the boundary, equation (13) shows that it is possible to get different sets of values of surface tension as obtained by Bagchi *et al.* (2005) for isotropic matter. Thus it is possible to generate a wide range of behaviour in the surface tension for anisotropic matter than is the case for isotropic pressures.

4. Numerical results

To get an estimate of the effects of pressure anisotropy on the surface tension, we consider the strange matter EOS given by equation (1). We consider two particular cases as discussed by Gondek-Rosińska *et al.* (2000):

- EOS SS1: where, $a = 0.463$, $\rho_b = 1.15 \times 10^{15}$ gm cm^{-3}, $\rho_c = 4.68 \times 10^{15}$ gm cm^{-3}, $n(r = R) = 0.725\,\text{fm}^{-3}$, $n(r = 0) = 2.35\,\text{fm}^{-3}$, $M = 1.435\ \text{M}_\odot$, $R = 7.07$ km.
- EOS SS2: where, $a = 0.455$, $\rho_b = 1.33 \times 10^{15}$ gm cm^{-3}, $\rho_c = 5.5 \times 10^{15}$ gm cm^{-3}, $n(r = R) = 0.805\,\text{fm}^{-3}$, $n(r = 0) = 2.638\,\text{fm}^{-3}$, $M = 1.323\ \text{M}_\odot$, and $R = 6.55$ km.

Numerical calculations show that for a given mass and radius, if we gradually introduce anisotropy, the absolute value of the surface tension decreases as can be seen in Fig. 1. For example, it is observed that even if we consider a tangential pressure of 100 MeV fm^{-3} at the surface, the surface tension decreases drastically. It is to be noted here that the anisotropy parameter should be so chosen that all the regularity conditions (Delgaty & Lake 1998) are satisfied. Thus, although in Fig. 1, the surface tension increases beyond a certain value of the anisotropic parameter, we ignore this region as the radial pressure gradient becomes positive in this region. The results are given in Table 1.

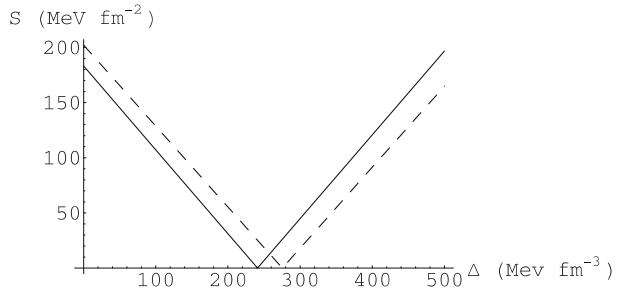

Figure 1. Surface tension S plotted against Δ. The solid line is for EOS SS1 and the dotted line is for EOS SS2.

Table 1. Anisotropic effect on the surface tension of strange stars.

		$\frac{dp_r}{dr}\|_{r=R}$ (MeV fm^{-3} km^{-1})		S (MeV fm^{-2})	
EOS	r_n (fm)	$p_\perp = 0$	$p_\perp = 100$ (MeV fm^{-3})	$p_\perp = 0$	$p_\perp = 100$ (MeV fm^{-3})
SS1	0.76	68.18	39.90	183.19	107.19
SS2	0.73	84.09	53.56	202.14	128.75

5. Discussions

We have shown that anisotropy plays an important role in the calculation of surface tension of compact stars. The origin of such anisotropies within compact objects may be different for different objects. We may, however, ask whether it is necessary at all to consider anisotropic effects on the surface tension of strange stars. The answer is in the affirmative since one possibility for the origin of anisotropies within strange stars could be the presence of charged particles at the surface. It has recently been reported that in strange stars, the electric field could be as high as 10^{19} eV/cm (Usov 2004), which indicates the possibility of a large charge distribution within such objects. Therefore we need to consider the effect of charge while deriving the the gross features of such stars. It can be shown that in the presence of charge, the TOV equation is modified to

$$\frac{dp}{dr} = -(\rho + p)\frac{\left(\frac{Gm(r)}{c^2 r} + \frac{4\pi G r^2 p}{c^4}\right)}{r\left(1 - \frac{2Gm(r)}{c^2 r}\right)} + \frac{Q(r)}{4\pi r^4}\frac{dQ(r)}{dr}, \tag{14}$$

where, $Q(r)$ is the total charge confined within a sphere of radius r. Note that the Einstein–Maxwell system is always anisotropic which is often treated as an isotropic system of field equations for mathematical simplicity (see for example, Ray *et al.* 2004). Also recent works (Schmitt 2005) suggest that a natural mechanism to explain the strong pulsar kicks in neutron stars could be the existence of asymmetric phases in quark matter.

It is to be noted here that, for simplicity, we ignored the effect of rotation in the present work although pulsars are magnetized rotators and a strong magnetic field ($\sim 10^{12}$ G) is observed at the surface of such stars. Pulsars known as magnetars may even have a magnetic field as strong as $\sim 10^{14-15}$ G. Though we do not have an established theory for the microscopic origin of such a strong magnetic field, it is agreed that Ferro-magnetization may occur in the high density quark matter which, in turn, may modify the EOS for strange matter. The derivation and the form of the modified EOS in the presence of strong magnetic field or superfluidity (responsible for anisotropy) is a complex issue and a more detailed analysis is required to see the effect of the modified EOS on the overall configuration *vis-a-vis* surface tension of compact stars.

To conclude, without going into the microscopic details of a star, it can be shown that surface tension is affected in the presence of anisotropy. For the very existence of strange stars in our Universe a crucial condition put forward was a large value of S by Alcock & Olinto (1989) which according to Bagchi *et al.* (2005) can be achieved by a general relativistic treatment of strange stars. However, in this article we have shown that a wide range of values of S are possible if we consider anisotropy in the energy momentum tensor; an issue ignored in the previous calculation (Bagchi *et al.* 2005).

Therefore, on the basis of surface tension for compact stars, no conclusive remarks at this moment can perhaps be made on the possible existence of strange stars. There could be, however, some other means to justify the existence of such stars which will be taken up elsewhere.

Acknowledgements

RS acknowledges the financial support (grant no. SFP2005070600007) from the National Research Foundation (NRF), South Africa. SDM acknowledges that this work is based upon research supported by the South African Research Chair Initiative of the Department of Science and Technology and the National Research Foundation.

References

Alcock, C., Olinto, A. 1989, *Phys. Rev. D*, **39**, 1233.
Bagchi, M., Sinha, M., Dey, M., Dey, J., Bhowmick, S. 2005, *Astron. & Astrophys.*, **440**, L33.
Bowers, R. L., Liang, E. P. T. 1974, *J. Astrophys.*, **188**, 657.
Chaisi, M., Maharaj, S. D. 2005, *Gen. Relat. Grav.*, **37**, 1177.
Colpi, M., Shapiro, S. L., Wasserman, I. 1986, *Phys. Rev. Lett.*, **57**, 2485.
Corchero, E. S. 2001, *Astrophysics and Space Science*, **275**, 259.
Delgaty, M. S. R., Lake, K. 1998, *Comput. Phys. Commun.*, **115**, 395.
Dev, K., Gleiser, M. 2003, *Gen. Relat. Grav.*, **35**, 1435.
Dey, M., Bombaci, I., Dey, J., Ray, S., Samanta, B. C. 1998, *Phys. Lett. B*, **438**, 123.
Farhi, E., Jaffe, R. L. 1984, *Phys. Rev. D*, **30**, 2379.
Gondek-Rosińska, D., Bulik, T., Zdunik, J. L., Gourgoulhon, E., Ray, S., Dey, J., Dey, M. 2000, *Astron. & Astrophys.*, **363**, 1005.
Hernández, H., Núñez, L. A. 2004, *Can. J. Phys.*, **82**, 29.
Herrera, L., Santos, N. O. 1997, *Phys. Rep.*, **286**, 53.
Ivanov, B. V. 2002, *Phys. Rev. D*, **65**, 104011.
Kaup, D. J. 1968, *Phys. Rev.*, **172**, 1331.
Kippenhahm, R., Weigert, A. 1990, *Stellar Structure and Evolution*, Springer, Berlin.
Letelier, P. 1980, *Phys. Rev. D*, **22**, 807.
Mak, M. K., Harko, T. 2003, *Proc. Roy. Soc. London A*, **459**, 393.
Oppenheimer, J. R., Volkoff, G. M. 1939, *Phys. Rev.*, **55**, 374.
Rao, J. K., Annapurna, M., Trivedi, M. M. 2000, *Pramana – J. Phys.*, **54**, 215.
Ray, S., Malheiro, M., Lemos, J. P. S., Zanchin, V. T. 2004, *Braz. J. Phys.*, **34**, 310.
Ruderman, R. 1972, *Ann. Rev. Astron. Astrophys.*, **10**, 427.
Ruffini, R., Bonazzola, S. 1969, *Phys. Rev.*, **187**, 1767.
Sawyer, R. F. 1972, *Phys. Rev. Lett.*, **29**, 382.
Schmitt, A., Shovkovy, I. A., Wang, Q. 2005, *Phys. Rev. Lett.*, **94**, 211101.
Sokolov, A. I. 1980, *JETP*, **79**, 1137.
Usov, V. V. 2004, *Phys. Rev. D*, **70**, 067301.
Weber, F. 1999, *Pulsars as astrophysical observatories for nuclear and particle physics*, Institute of Physics, Bristol.
Witten, E. 1984, *Phys. Rev. D*, **30**, 272.
Zdunik, J. L. 2000, *Astron. & Astrophys.*, **359**, 311.

J. Astrophys. Astr. (2007) **28,** 139–145

Does Si Play a Role in the Formation of Extrasolar Planet Systems?

C. Huang[1,3], G. Zhao[2,*], H. W. Zhang[1] & Y. Q. Chen[2]
[1]*Department of Astronomy, Peking University, Beijing 100 871, China.*
[2]*National Astronomical Observatories, Chinese Academy of Sciences, Beijing 100 012, China.*
[3]*National Center for Space Weather, China Meteorological Administration, Beijing 100 081, China.*
[*]*e-mail: gzhao@bao.ac.cn*

Received 2006 August 31; accepted 2007 August 13

Abstract. With the high signal-to-noise ratio spectra, we obtained Si abundances of 22 extrasolar planet host stars, and discussed some constraints on the planet formation. Using our silicon abundance results and other authors' Si abundance studies about planets-harboring stars, we investigated the correlation between the dynamical properties and the silicon abundance. We propose a hypothesis that higher primordial metallicity in the host stars' birth cloud with higher abundance of Si will make the cloud more sticky to bypass the time scale restriction in planet formation and easier to form the planets.

Key words. Planetary systems: formation—stars: abundances—stars: late-type.

1. Introduction

From the first discovery of extrasolar planet host system (Mayor & Queloz 1995), there are about 200 planet-harboring systems which have been found. Most of them are discovered by using the Doppler technique. One of the key factors which is relevant to the mechanism of planetary formation is that many such systems are really metal-rich (e.g., Gonzalez *et al.* 2001; Smith *et al.* 2001; Sadakane *et al.* 2003) compared with the metallicity distribution of nearby field F, G, or K stars which are known to hold non-planets. Gonzalez (1997, 1998) proposed two hypotheses to account for the correlation between metallicity and the presence of the giant planets: the first view shows that higher primordial metallicity in the host stars' birth cloud helps planets to be more easily or efficiently formed (e.g., Pollack *et al.* 1996). The second view proposed that the observed high metallicity may be the result of a "self-enrichment" process. Accretion of high-Z material after the outer convection zone of the host star has thinned to a certain minimum mass, elevates the apparent metallicity above its primordial value. Some studies (see, e.g., Takeda *et al.* 2001; Sadakane *et al.* 2002, 2003; Fischer & Valenti 2005; Huang *et al.* 2005a, 2005b; Shen *et al.* 2005) show that there exists no significant evidence to prove the "self-enrichment" hypothesis.

This work discusses the time-scale constraints in planet formation and presents the correlation between dynamical properties and silicon abundances. We try to examine

the "high primordial metallicity" hypothesis from a special point of view that metal-silicon relationship leads to fast planet formation.

2. The constraints in planet formation

According to the standard model of planet formation, when the dust settles to the midplane of the disk, grains coagulate and grow to small (1–10 km diameter) planetesimals. The planetesimals accumulate and accrete the gas to form terrestrial or giant planets. So there exist two stages in planet formation. During the first stage, dust to planetesimals, there exist interactions with the gas dominating. The gravity influence is neglectable. In the second stage, planetesimals to planets, gravity plays the bigger role.

2.1 *The stage of dust to planetesimals*

For the early solar-like nebula, the most abundant condensates are silicates and iron compounds. These initial micron-sized grains are believed to form the planetesimals by inelastic collisions. The grain growth process is determined by the sticking efficiency upon collisions (Poppe *et al.* 2000). The principal problem (Wetherill 1990) in this stage appears to be the need to grow fairly large ($>$ 10 m) planetesimals before the time the nebular gas is removed (3–10 Myr, Hartmann 1998). The characteristic settling time of a 1-μm grain will be 2.9×10^6 yr (Wetherill 1990). This time is too long compared to the time scale for the removal of the gas in the solar-like nebula and the turbulences will lengthen the settling time. Another problem is that the large particles will encounter a headwind which will cause their angular momentum loss and they will spiral inward towards the host star (Weidenschilling 1977). These particles should complete the transition from cm size to km size rather rapidly, especially for m-size particles, otherwise the material will not survive to form the planets (Lissauer 1993). However, the growth of solid bodies from mm-size to km-size still presents particular problems. There must be some way to increase the efficiency of the planetesimal formation.

2.2 *The stage of planetesimals to planets*

The early version of gas giant formation model provides a constraint that 10–20 $M_\oplus$ cores are needed to initiate efficient gas accretion (Pollack *et al.* 1996). The gas giant formation time scale is critical because the gas giant almost consists of hydrogen and helium, the gaseous solar-like nebula still being present when the gas giant began to form (Wetherill 1990). According to the present studies, the observed disk depletion time scale is 3–10 Myr (Walter 1986; Strom *et al.* 1989; Hartmann 1998). For the reasons above, the massive core must form rapidly in order to initiate gas accretion while the nebular gas is still available. So the planetesimals should find an efficient way to accumulate.

3. Results and discussion

3.1 *Data*

We obtained the silicon abundances of 22 samples from our new article (Huang *et al.* 2005a) and chose an other 59 samples from previous studies (Gonzalez *et al.* 2001;

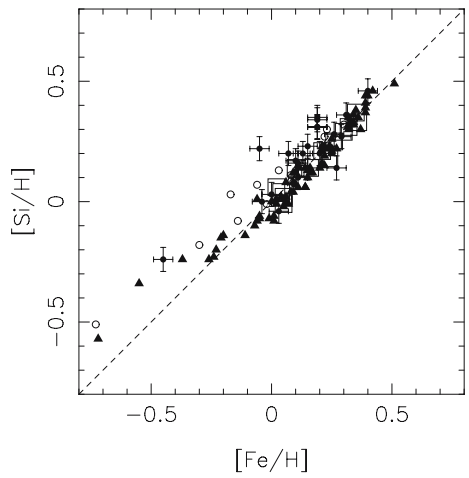

Figure 1. Plot of [Si/H] *vs.* [Fe/H]; our results of 22 samples with error bars are denoted by black points, the results from Gonzalez *et al.* (2001) are designed by hollow squares, data of the results from Bodaghee *et al.* (2003) are shown by black triangles, results from Sadakane *et al.* (2002) are shown by open circles.

Sadakane *et al.* 2002; Bodaghee *et al.* 2003). In Fig. 1, we plot [Si/H] *versus* [Fe/H], they have a good one-to-one correlation. It means the initial cloud with high primordial metallicity will have high primordial silicon abundance and more silicates. Our results of [Si/H] show the average of 0.06 dex higher than those that overlap in Bodaghee *et al.* (2003), the [Si/H] results from Sadakane *et al.* (2002) and Gonzalez *et al.* (2001) show a slight higher and a bit lower than the same samples in Bodaghee *et al.* (2003), respectively. The average values of [Si/H] are 0.12, 0.20, 0.19, 0.05, which are respective to 78 samples of Bodaghee *et al.* (2003), 22 samples of ours, 13 samples of Gonzalez *et al.* (2001), 12 samples of Sadakane *et al.* (2002). Si is a dominant element in the component of the Earth and silicates are good adhesive material in some circumstances.

3.2 *Discussion on the constraints*

Robinson *et al.* (2006) suggested that silicon and oxygen are both α elements and [Si/Fe] is a tracer of [O/Fe], so a silicon-enrich star implies an oxygen-rich star. The oxygen-rich star will increase the solid surface density and make more ice cores to form the planets. We discuss this correlation of Si-rich and planets from another point of view. We suppose that the Si-rich primordial cloud implies an Si-rich and O-rich cloud, so this type of cloud will make more silicates to increase the "stickness" of the cloud and be easy to form planets.

In section 2.1, if the dust settling proceeds more rapidly, the grain should grow larger. Grossman (1988) found chondrules of $\sim$ 1-mm radius in most stony meteorites. To mm-size particles, the corresponding settling time will be shortened to $\sim 3 \times 10^3$ yr (Wetherill 1990). According to their calculations, Kornet *et al.* (2004) reported that if the main process responsible for particle growth is collisional coagulation, then its efficiency should be high. The recent observations by Chakraborty

et al. (2004) show that asteroid-size bodies can exist in disks as young as 0.1 Myr, that means if their growth is due to the coagulation, the sticking efficiency must be very high. In accordance with the contact model (Johnson *et al.* 1971), the adhesive surface forces will increase the contact radius and the sticking probability. We suppose the cloud with high primordial silicon abundance could increase the sticking efficiency of particles that makes the km-size planetesimals form in the dust settling time $\sim 10^4$ yr.

In the early stage of planetesimal accumulation, the rate of growth of planetesimal increases as $\sim R^2$, where R is the radius of the growing body (see details in Lissauer 1993), and the gravity influence is weak. Hence, the runaway growth starts slowly and the surface property is important to the encounters. It appears that the "stickiness" is necessary in colliding cases during this stage. We assume that more silicates could influence the sticking efficiency of colliding bodies in the early stage of planetesimal accumulation and change the size of planetesimals in the early phase of runaway growth. It is possible that the "initial" size distribution of planetesimals significantly influences the growth times of the giant planets (Lissauer 1993). For example, it will cost the protoplanet $\sim 6 \times 10^5$ years to attain 15 $M\oplus$ (Lissauer 1987).

3.3 *Dynamical properties discussion*

The disk gas will be evaporated either through photo-evaporation or driven by central stars from the inner disk region (Hollenbach & Adams 2004) during a period of relative fixed time. So if the initial cloud has an efficient way to form planets, the situation of the disk gas will give a reflection of the giant planets' orbital properties. In Fig. 2, we plot 81 samples' orbital properties with their silicon abundance. We suppose that if the giant planet forms rapidly, it will increase the possibility to present itself near the central star or migrate to the central star when the gas at inner region is still available. The top diagram shows a slight decreasing trend of [Si/H] with increasing semimajor for our 22 samples when semimajor a < 3 AU. However, there is a large scatter in other samples. Maybe the scatter is caused by the stochastic events in the planet formation and the different results used by investigators. The mass of giant planet is dominated by gas, so [Si/H] should not be relevant to the mass. The middle diagram shows no obvious trend of [Si/H] with the mass. The bottom diagram shows a decreasing trend of [Si/H] with increasing eccentricity for our 22 samples and 13 samples from Gonzalez (2001). Despite the large scatter, the results from Bodaghee *et al.* (2003) also show the mimic trend. We obtained the slope by a linear least-squares fit for [Si/H] *versus* the eccentricities of our 22 samples in the bottom panel of Fig. 2, and the slope value is near -0.24. Though the straight-line fit to [Si/H] *versus* the eccentricities is not meant to imply that there is a real linear relation between [Si/H] and eccentricities, it is used to characterize the [Si/H] distribution by the slope with the value of -0.24. Small eccentricities are a requirement of the standard model for the formation of a giant-planet in the scenario of gradual accretion of solid particles in a disk and followed by gravitational accretion of gas (Smith *et al.* 2001). So the bottom diagram implies that maybe the initial cloud with high primordial silicon abundance adapts to forming giant planets. We suppose the inner disk gas will exist to influence the planets' eccentricities in the planet migration because the planet can form efficiently in the initial Si-rich cloud. However, it is difficult to quantify this scenario for lots of details in this process are obscure.

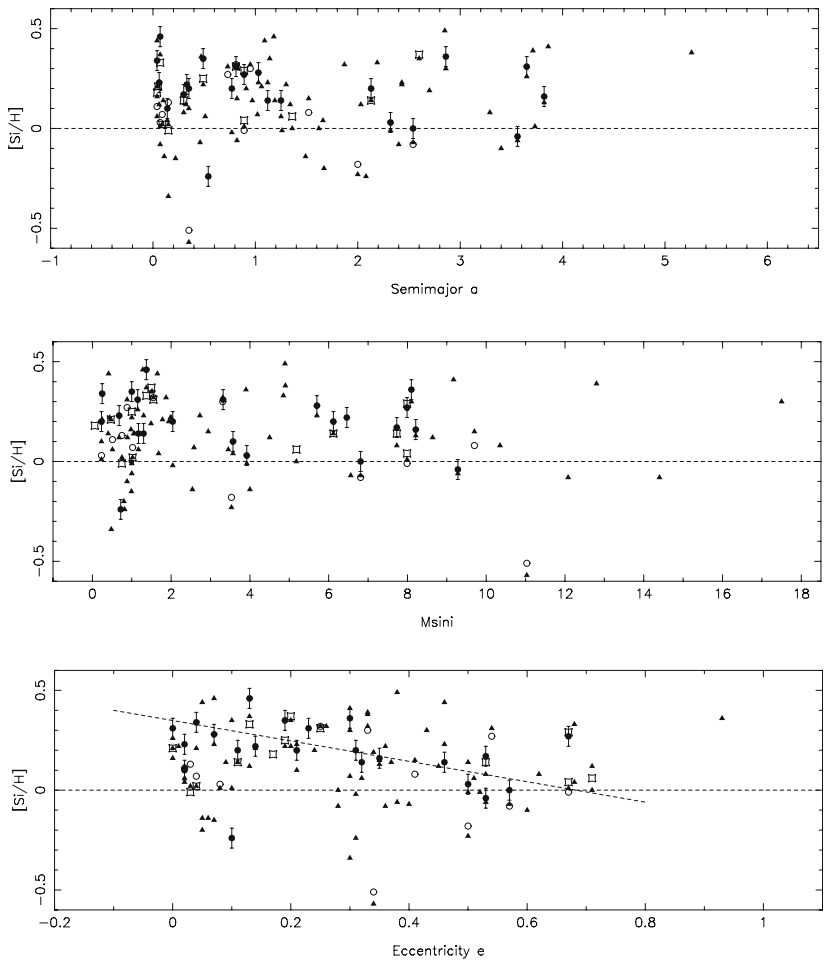

Figure 2. The trend of dynamical properties: (Semimajor a, Msini, eccentricity e) with [Si/H], symbols in figures are the same as in Fig. 1.

To investigate the two-order effects, we plot the relative abundances of [Si/Fe] *versus* the various planetary parameters in Fig. 3. The top diagram and the middle diagram in Fig. 3 show that the planet-host stars with high [Si/Fe] have planets of short-period orbits and small companion masses. The whys about the lack of short-period massive planets is that massive planets finish forming too late to experience the Type II migration (after opening the gap in the disk) which would cause the massive planets to get near to their central stars (Udry *et al.* 2003). The facts imply that maybe the [Si/Fe]-high cloud would not be easy to form massive planets. The extrasolar planet systems' discoveries may be influenced by the selection effects. However, most of the [Si/Fe]-rich planet hosts gather inside 1 AU of semimajor and their companies' masses are less than 4 Jupiter-masses. The semimajors and the companies' masses ranges are beyond the selection effect influence, so the results should not be caused by the selection effects of observation. In the bottom panel in Fig. 3, our results show that the planet-host stars with high [Si/Fe] are likely to have planets of small eccentricities,

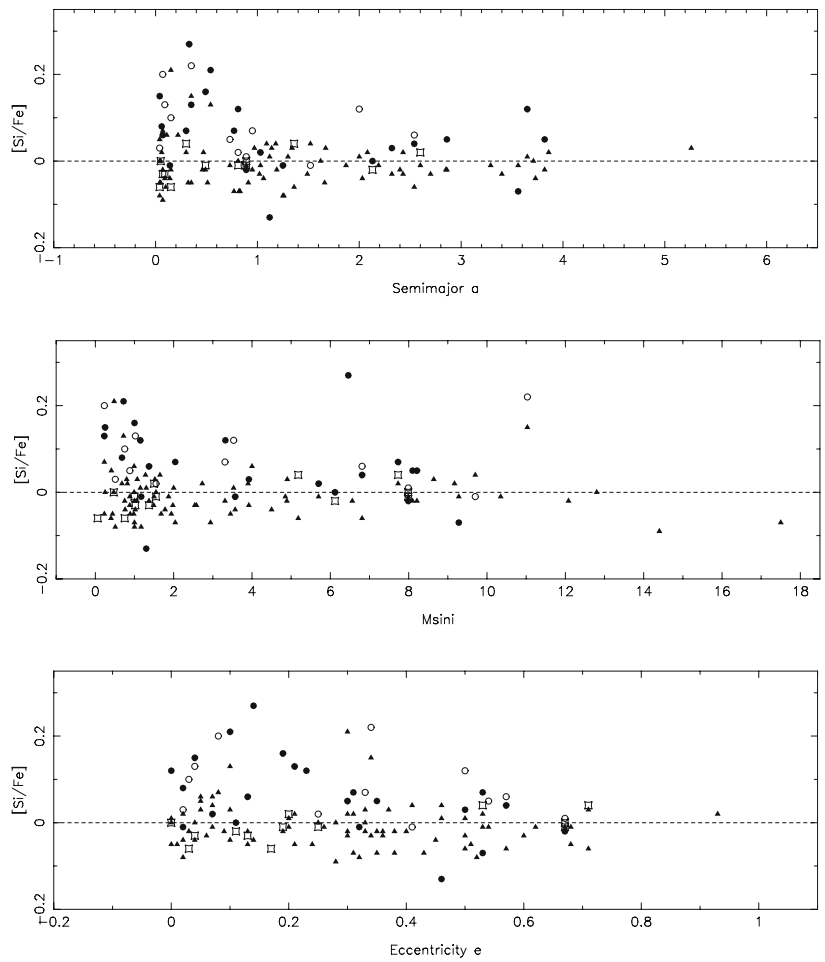

Figure 3. The trend of dynamical properties: (Semimajor a, Msini, eccentricity e) with [Si/Fe], symbols in figures are the same as in Fig. 1.

but there is a large scatter in [Si/Fe] *versus* eccentricity in other studies. So we cannot draw a firm conclusion.

4. Concluding remarks

Calculations reported by Kornet *et al.* (2005) show that the disk viscous coefficient will influence the rate of the planet occurrence. A too high viscous coefficient will reduce the rate of the planet occurrence. Maybe that can explain why the relative frequency of planet hosts has a sharp cutoff when the metallicity reaches 0.4 dex (Israelian 2004). As one might expect, in the disk with [Fe/H] > 0.5, there are more silicates which make the disk viscous coefficient too high to let the star associated with planets. Robinson *et al.* (2006) reported that the planet-host stars exhibit statistically significant silicon enrichment in their observations, their simulations predict the important relationship of the silicon enrichment with the core-accretion model for the giant planet formation

and they discussed the origin of the planet-silicon correlation in planet-host stars. The observations hint that the planets may be easily formed in the silicon-rich cloud.

We investigate the dynamical properties of 81 extrasolar planet systems with the corresponding silicon abundances and discuss the time scale constraints in planet formation. According to the investigation, we propose a hypothesis that the initial cloud with high primordial metallicity may have high primordial silicon abundance and more silicates which make the cloud more sticky to form the planets efficiently, but to the disk with [Fe/H] > 0.5, nature will tune the effect of silicates to the contrary. However, this study still stays at the qualitative "scenario" stage. There exists a number of obscure problems and further quantitative studies are desirable.

Acknowledgements

This work is supported by the National Natural Science Foundation of China under grant No.10433010 and 10521001.

References

Bodaghee, A., Santos, N. C., Israelian, G., Mayor, M. 2003, *A&A*, **404**, 715.
Chakraborty, A., Ge, J., Mahadevan, S. 2004, *ApJ*, **606**, L69.
Fischer, D. A., Valenti, J. 2005, *ApJ*, **622**, 1102.
Gonzalez, G. 1997, *MNRAS*, **285**, 403.
Gonzalez, G. 1998, *A&A*, **334**, 221.
Gonzalez, G., Laws, C., Tyagi, S., Reddy, B. E. 2001, *AJ*, **121**, 432.
Grossman, J. N. 1988, Formation of chondrules (Tuscon: Ariz Univ. Press).
Hartmann, L. 1998, Accretion Processes in Star Formation (Cambridge: Cambridge Univ. Press).
Hollenbach, D., Adams, F. C. 2004, *ASP Conference Series*, Vol. **324**, p. 168.
Huang, C., Zhao, G., Zhang, H. W., Chen, Y. Q. 2005a, *MNRAS*, **363**, 71.
Huang, C., Zhao, G., Zhang, H. W., Chen, Y. Q. 2005b, *Chinese J. Astron. Astrophys.*, **6**, 619.
Israelian, G. 2004, *IAUS*, **219**, 343.
Johnson, K. L., Kendall, K., Roberts, A. D. 1971, *Proc. Roy. Soc. London A*, **324**, 301.
Kornet, K., Stepinski, T. F., Rozyczka, M. 2004, *A&A*, **417**, 151.
Kornet, K., Bodenheimer, P., Rozyczka, M., Stepinski, T. F. 2005, *A&A*, **430**, 1133.
Lissauer, J. J. 1987, *Icarus*, **69**, 249.
Lissauer, J. J. 1993, *ARA&A*, **31**, 129.
Mayor, M., Queloz, D. 1995, *Nat*, **378**, 355.
Pollack, J. B., Hubickyj, O., Bodenheimer, P. *et al.* 1996, *Icarus*, **124**, 62.
Poppe, T., Blum, J., Henning, T. 2000, *ApJ*, **533**, 454.
Robinson, S. E., Laughlin, G., Bodenheimer, P., Fisher, D. 2006, astro-ph/0601656.
Sadakane, K., Ohkubo, M., Takeda, Y. *et al.* 2002, *PASJ*, **54**, 911.
Sadakane, K., Ohkubo, M., Honda, S. 2003, *PASJ*, **55**, 1005.
Shen, Z. X., Jones, B., Lin, D. N. C., Liu, X. W., Li, S. L. 2005, *ApJ*, **635**, 608.
Smith, V., Cunha, K., Lazzaro, D. 2001, *AJ*, **121**, 3207.
Strom, K., Strom, S. E., Edwards, S. *et al.* 1989, *AJ*, **97**, 1451.
Takeda, Y., Sato, B., Kambe, E. *et al.* 2001, *PASJ*, **53**, 1211.
Udry, S., Mayor, M., Santos, N. C. 2003, *A&A*, **407**, 369.
Walter, F. M. 1986, *ApJ*, **306**, 573.
Weidenschilling, S. J. 1977, *MNRAS*, **180**, 57.
Wetherill, G. W. 1990, *Annu. Rev. Earth Planet. Sci.*, **18**, 205.

J. Astrophys. Astr. (2007) **28,** 147–156

Activity Cycle of Solar Filaments

K. J. Li[1,*], Q. X. Li[1,2], P. X. Gao[1,2], J. Mu[1,2], H. D. Chen[1,2], T. W. Su[1]
[1]*National Astronomical Observatories/Yunnan Observatory, CAS, Kunming 650 011, China.*
[2]*School of Graduate, CAS, Beijing 100 863, China.*
[*]*e-mail: lkj@ynao.ac.cn*

Received 2007 February 28; accepted 2007 September 2

Abstract. Long-term variation in the distribution of the solar filaments observed at the Observatorie de Paris, Section de Meudon from March 1919 to December 1989 is presented to compare with sunspot cycle and to study the periodicity in the filament activity, namely the periods of the coronal activity with the Morlet wavelet used. It is inferred that the activity cycle of solar filaments should have the same cycle length as sunspot cycle, but the cycle behavior of solar filaments is globally similar in profile with, but different in detail from, that of sunspot cycles. The amplitude of solar magnetic activity should not keep in phase with the complexity of solar magnetic activity. The possible periods in the filament activity are about 10.44 and 19.20 years. The wavelet local power spectrum of the period 10.44 years is statistically significant during the whole consideration time. The wavelet local power spectrum of the period 19.20 years is under the 95% confidence spectrum during the whole consideration time, but over the mean red-noise spectrum of $\alpha = 0.72$ before approximate Carrington rotation number 1500, and after that the filament activity does not statistically show the period. Wavelet reconstruction indicates that the early data of the filament archive (in and before cycle 16) are more noiseful than the later (in and after cycle 17).

Key words. Sun: activity—filaments—active cycle.

1. Introduction

Solar filaments are prominences observed against the solar surface, and they are cool and dense clouds in the hot and diluted solar corona. The shape, dynamics and physical properties of filaments vary in the wide range of values (Tandberg-hanssen 1995). Solar filaments are distributed around the whole solar disk, from the solar equator to the poles, and during the whole period of the cycle. Such a distribution is not random due to filaments that are closely connected with sites of magnetic fields on the solar surface (Martin 1990). Filaments can provide the possibility for investigation of global properties of large-scale magnetic fields, especially when magnetographic observations are not available. Their property to appear in all heliospheric latitudes and to outline the border between magnetic fields with different polarities makes them suitable tracers for the large-scale pattern of the weak background magnetic field

(McIntosh 1972; Minarovjech *et al.* 1998a, 1998b). On the other hand, study of the occurrence of filaments can help us to better understand distribution of these fields on the solar surface, their development with a cycle activity, and especially, provide useful insights into the nature of the sun's magnetic field (Rusin *et al.* 1998, 2000; Mouradian & Soru-Escaut 1994). Study on filaments through both individual events and statistical analyses are of importance. In the present study, the long-term cyclic variation in the distribution of the solar filaments observed at the Observatorie de Paris, Section de Meudon from March 1919 to December 1989 is presented to compare with the sunspot cycle.

The so-called periods of solar activity, such as the well-known 11-year Schwabe period and the 80–100-year Gleissberg period, are generally referred mainly to those of solar activity in the photosphere and have been widely studied (Ogurtsov *et al.* 2002; Polygiannakis *et al.* 2003; Prabhakaran *et al.* 2002). However, up to now we have known much less about the periods of the solar chromosphere and corona activities. The wavelet transform is a very powerful tool to analyze non-stationary signals. It permits identification of the main periodicities in a time series and the evolution in time of each frequency (De Moortel *et al.* 2004; Torrence & Compo 1998; Faria *et al.* 2004). In the present study, the complex Morlet wavelet transform (Torrence & Compo 1998) is utilized to study the periodicity in the solar filaments, showing the periods of the coronal activity.

2. Activity cycle of solar filaments

2.1 *Data*

Lucien d'Azambuja investigated solar filament and prominence behaviour over many years, and maintained a synoptic program of this phenomena similar to the Zurich sunspot program (d'Azambuja 1923; Coffey & Hanchett 1998b). He published the first "Cartes Synoptiques de la Chromosphere Solaire et Catalogue des Filaments de la Couche Superieure", a compendium of reduced solar observations covering the time period March 1919–January 1920 (d'Azambuja 1928). The compiled database gave both visual and quantitative measures of solar activity beginning with Carrington rotation 876. Since then, data through 1989 have been published in succeeding Carte Synoptiques issues (Coffey & Hanchett 1998b). The World Data Center A (WDC-A) for Solar-Terrestrial Physics has digitized the Carte Synoptiques (Coffey and Hanchett 1998b). The data can be accessed via the World Wide Web at *ftp://ftp.ngdc.noaa.gov/STP/SOLAR_DATA/SOLAR_FILAMENTS*. Utilized here is the Carte Synoptique solar filaments archive (Mouradian 1998), namely the catalogue of solar filaments from March 1919 to December 1989, corresponding Carrington solar rotation numbers 876 to 1823, including 41044 filament regions in total (Coffey & Hanchett 1998b). In fact, in the database, the total 41042 filament regions are recorded with the filaments no 16479 and 17299 were repeatedly recorded, and the filaments No 14678, 16283, 16382, 16383, 18536, 19612, 19622, 30161-30188, and 20655 should be the filaments no 14687, 16282, 16381, 16382, 18546, 19622, 19621, 20161–20188, and 20665, respectively. Each data point represents the central meridian passage data of a filament region that was observed crossing the solar disk (Coffey & Hanchett 1998a).

2.2 *Cycle variation of solar filaments*

The normal solar activity is usually applied to solar active events whose latitudes are less than 50° (Sakurai 1998), and which are found to be anti-correlated with high-latitude solar activity (Li *et al.* 2002). We count the number of the low-latitude solar filaments whose latitudes are less than 50° in each of Carrington solar rotation numbers 876 to 1823, which is shown in Fig. 1. Also given in the figure is the 13-point smoothing values of the filament numbers. The figure shows that the number of the solar filaments waxes and wanes with an approximate 11-year Schwabe cycle, as the sunspot number does, showing a very nice cycle behaviour, called here as activity cycle of solar filaments. But the cycle maxima vary only in a small range (the ratio of the maximum value to the minimum one among the 6 maximum cycle amplitudes is about 2.0 for filaments, but about 2.6 for sunspot relative numbers), this is unlike the behaviour of sunspot numbers, areas or 10.7 cm flux. As we know, there is a good correlation among sunspot numbers, areas or 10.7 cm flux. However, there is a weak positive linear correlation between the maximum smoothed filament number and the maximum smoothed sunspot number, which is given in Fig. 2. The correlation coefficient is 0.578, which is even less than the critical correlation coefficient (0.729) at the 90% confident level. The minimum and maximum times of sunspot cycles are indicated in Fig. 1. The figure shows that the minimum times of activity cycle of solar filaments are very close to those of sunspot cycles, but the maximum times of activity cycle of solar filaments are later than those of sunspot cycles for 4 cycles of the total 6 cycles. And further, filament cycles show some delay on the downward curve of the filament data after sunspot maximum, as noted by d'Azambuja and d'Azambuja (1948). Here, in summary, we infer that activity cycle of solar filaments should have the same cycle length as sunspot cycle, but the cycle behaviour of solar filaments is globally similar in profile with sunspot cycles, but different in detail from sunspot cycles, especially, their corresponding maximum values are not linearly correlated with a statistical significance, and their corresponding maximum times are different from each other (the maximum times are generally later for filaments).

As we know, sunspots are related to magnetic flux, thus, sunspot numbers somewhat represent the "amplitude" of solar magnetic activity. Filaments are usually located at the neutral line of a solar magnetic field with different polarities and related to magnetic non-potentiality or electric currents, therefore, filament numbers are inferred to represent the "complexity" of solar magnetic activity. The magnetic field is observed to be very weak around the minimum of a normal cycle, appearing as simple sunspots, and even almost as single-polarity sunspots, so, it is plausibly expected that both filament activity and sunspot activity should have the same minimum time. Generally, the magnetic field of sunspots is observed to appear more and more complicated with the increase of their magnetic flux, thus, filament numbers seem positively correlated with sunspot relative numbers in general, however, increase of the number of sunspots or sunspot groups does not completely mean increase of pairs of magnetic polarity. The maximum times that filament activity cycles are found to be later than those of sunspot cycles, means that the complexity of sunspots peaks later than the magnetic flux of sunspots do during a normal active cycle. The amplitude of solar magnetic activity should not keep in phase with the complexity of solar magnetic activity.

Time-latitude distribution of the solar filaments was once shown by Coffey & Hanchett (1998b) in their Fig. 2, it is the butterfly plot of the entire Cartes Synoptiques

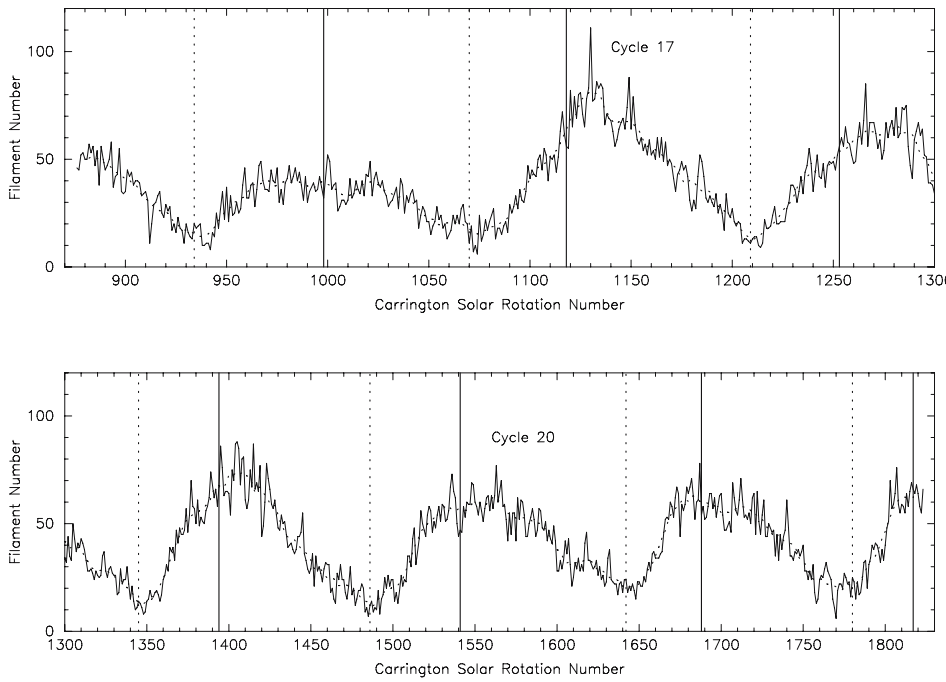

Figure 1. The numbers of the low-latitude solar filaments in Carrington solar rotation numbers 876 to 1823 (the solid thin curve) and their 13-point smoothed values (the dotted curve). The vertical solid and dotted lines are drawn to mark the minimum and maximum times of sunspot cycles, respectively.

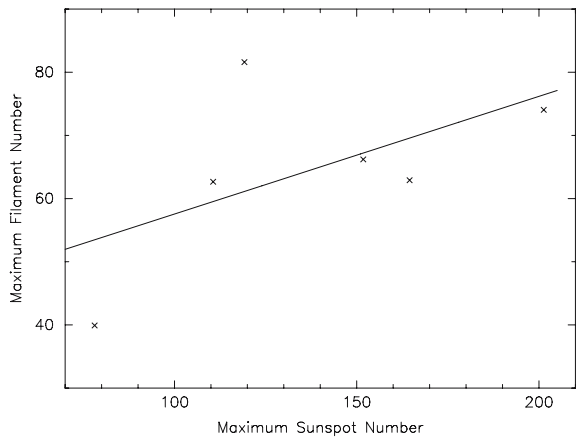

Figure 2. A weak linear correlation between the maximum smoothed filament number and the maximum smoothed sunspot number.

filament data archive. At the mid- and low-heliographic latitudes, filament activity shifts equatorward starting from the beginning of the cycle. But at high latitudes, it migrates polarward. Note the "rush to the pole" close to solar maximum when the

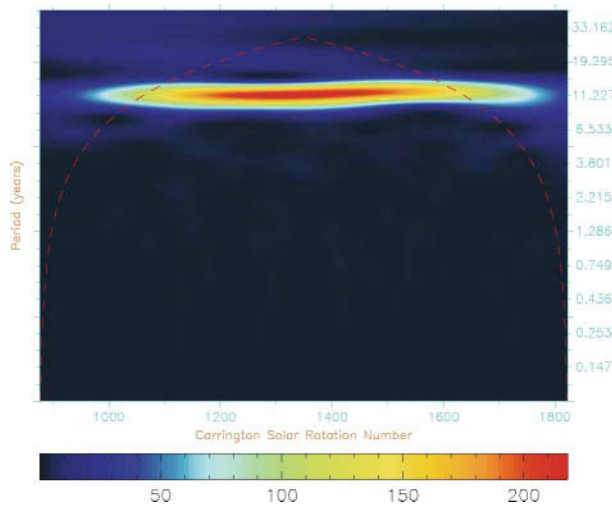

Figure 3. The local wavelet power spectral map of the filament number. The region upwards of the red dashed line indicates the COI.

solar polar magnetic field reversal occurs, and after then, very few filaments appear at high latitudes.

2.3 *Periodicity of activity cycle of solar filaments*

The wavelet analysis is a very powerful tool to analyze non-stationary signals, and it is rapidly gaining popularity as a means of extracting periodicities from observed signals (De Moortel *et al.* 2004a, 2004b; Faria *et al.* 2004). Among the wavelet analyses of solar activity, the Morlet wavelet is usually used, which has a reasonably large number of oscillation and may ensure a good frequency resolution. The complex Morlet wavelet analysis (Torrence & Compo 1998; De Moortel *et al.* 2004a, b) is utilized to study the periodicity of the filament number. Its local wavelet power spectrum is shown in Fig. 3. Through the figure, it can be found that the period belt of the highest power is located around the Schwabe period, as the photospheric sunspot activity does (Ogurtsov *et al.* 2002; Faria *et al.* 2004). It is the Schwabe period of the solar corona activity.

Great care is needed when using a wavelet: the Fourier transform is usually used to speed up the computation in wavelet analysis programme (Torrence & Compo 1998). However, as the Fourier transform assumes that data are cyclic, and most time series are of finite length, this introduces errors at the edges of the transform (De Moortel *et al.* 2004a). The region in which the transform suffers from these edge effects is known as the cone of influence (COI). As in Torrence & Compo (1998), the COI is defined so that the wavelet power for a discontinuity at the edge decreased by a factor e^{-2}. Portions of the transform that are outside the area formed by the 'time' axis and the COI are subject to these edge effects and are, therefore, unreliable (Moortel *et al.* 2004a; Torrence & Compo 1998). In Fig. 3, such an area is marked as the area upwards of the red dashed line. Shown in Fig. 4 is the global wavelet power spectrum of the number of the solar filaments. The period values of the power peaks in the global power spectrum are about 0.36, 1.04, 5.44, 10.44, 19.20, and 35.12 years, which can

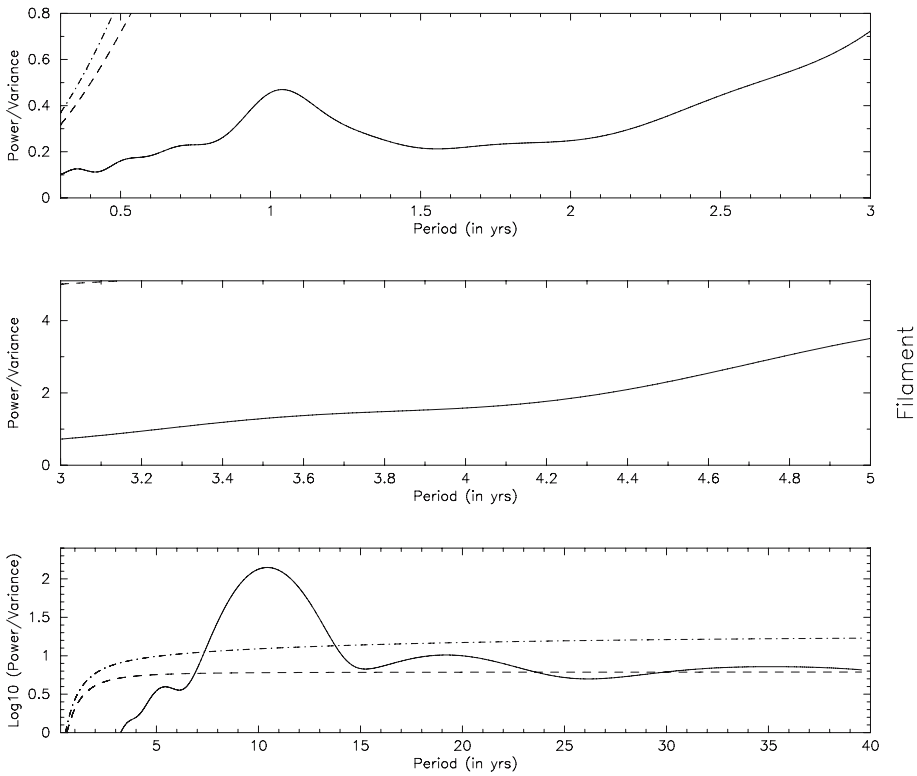

Figure 4. The global Morlet wavelet power spectrum of the filament number (the solid line). The dashed line is the mean red-noise spectrum of $\alpha = 0.72$, while the dotted and dashed line is the 95% confidence level for the global wavelet spectrum.

be all found in the periods of the photospheric sunspot activity (Prabhakaran *et al.* 2002; Polygiannakis *et al.* 2003). The most eminent period is shown to be 10.44 years, whose spectral power is over the 95% confidence level line, it is the Schwabe period of the solar corona activity. The power spectra of the periods of 19.20 and 35.12 years are under the 95% confidence level line but over the mean red-noise spectral line (Torrence & Compo 1998), being possible periods of the filament number, and the other rest periods are even under the mean red-noise spectral line, which are statistically insignificant. However, the period of 35.12 years locates in the COI are thus unreliable. Thus, reliable and possible periods in the filament number are 10.44 and 19.20 years.

The periodicity of 2.98 years is the so-called QBO (quasi-biennial 2–3 year oscillation) (Kane 2005), and a QBO was reported earlier in many solar indices and the solar magnetic field (Obridko & Schelting 2001; Kononovich & Shefov 2002; Knaack & Stenflo 2004; Kane 2005, and references therein). The QBO is considered the main feature of the 11-year cycle and it is suggested that the source of the QBO of the solar magnetic field should be situated rather low at the base of the solar convection zone (Bumba 2003; Kane 2005). However, strangely, such a period has not been found in filament activity.

Recent helioseismic probing of the solar interior has shown that the rotation rate of the Sun near the base of its convective zone changes with a period of roughly 1.3 years

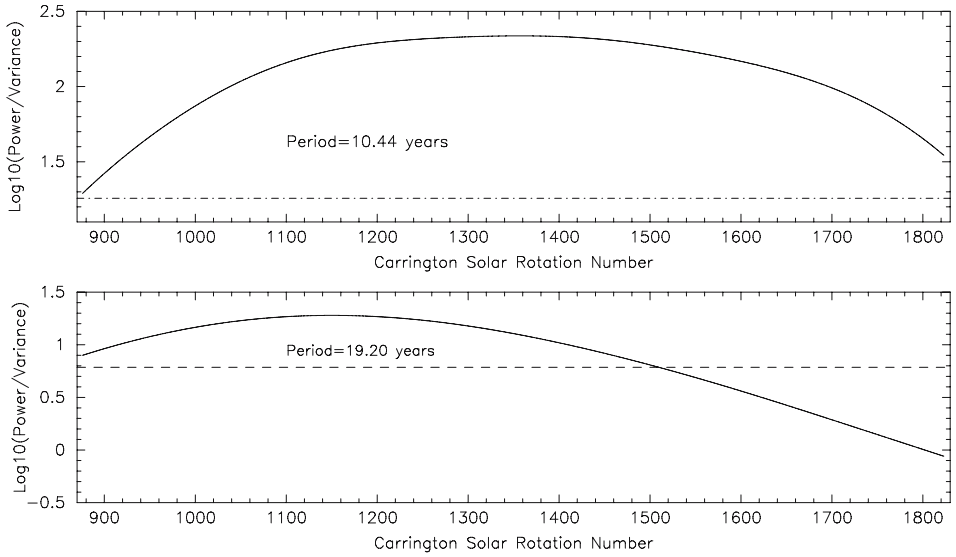

Figure 5. The local power spectra (the solid lines) of the reliable and possible periods in the filament number. The dashed line is the mean red-noise spectrum, while the dotted and dashed line is the 95% confidence level for the local wavelet spectrum.

(Howe *et al.* 2000), and significant power at this period (1.28 years) is indeed found in sunspots and is observed to vary strongly with time (Krivova & Solanki 2002). As well, the periodicity of the 1.3 year and 1.7 year are reported to be present in interplanetary plasma parameters and cosmic ray intensity as a modulation effect of solar activity (Kudala *et al.* 2002; Kato *et al.* 2003; Kane 2005, and references therein). Kane (2003) reported, however, that whereas such periodicities did exist, these were not all similar for solar indices, interplanetary plasma parameters, cosmic ray neutron monitor intensities during the time of the years 1991 to 2001. Recently, Kane (2005) identified the short-term (few tens of months) periodicities of several solar indices and found that, the periodicities of 1.3 year and 1.7 year often mentioned in the literature were seen neither often nor prominently. Here, the two periodicities are not found in the low-latitude filaments. One needs to check whether these periodicities are consistently in all solar indices and all the time. And, the magnetic field around the low-latitude filaments cannot be seemingly inferred to be related to the deep solar convection zone as well.

Shown in Fig. 5 are the local power spectra of the reliable and possible periods in the filament number. The wavelet local power spectrum of the Schwabe period is over the 95% confidence spectrum at the whole consideration time and highly significant. The wavelet local power spectrum of the period 19.20 years is under the 95% confidence spectrum at the whole consideration time, but over the mean red-noise spectrum of $\alpha = 0.72$ before Carrington rotation number 1501. The period is more eminent before Carrington rotation number 1501, and after that it is statistically insignificant.

Since the wavelet transform is a band-pass filter with a known response function (the wavelet function), it is possible to reconstruct the original time series. Such a reconstruction will denoise any low-amplitude regions of the wavelet transform, which are

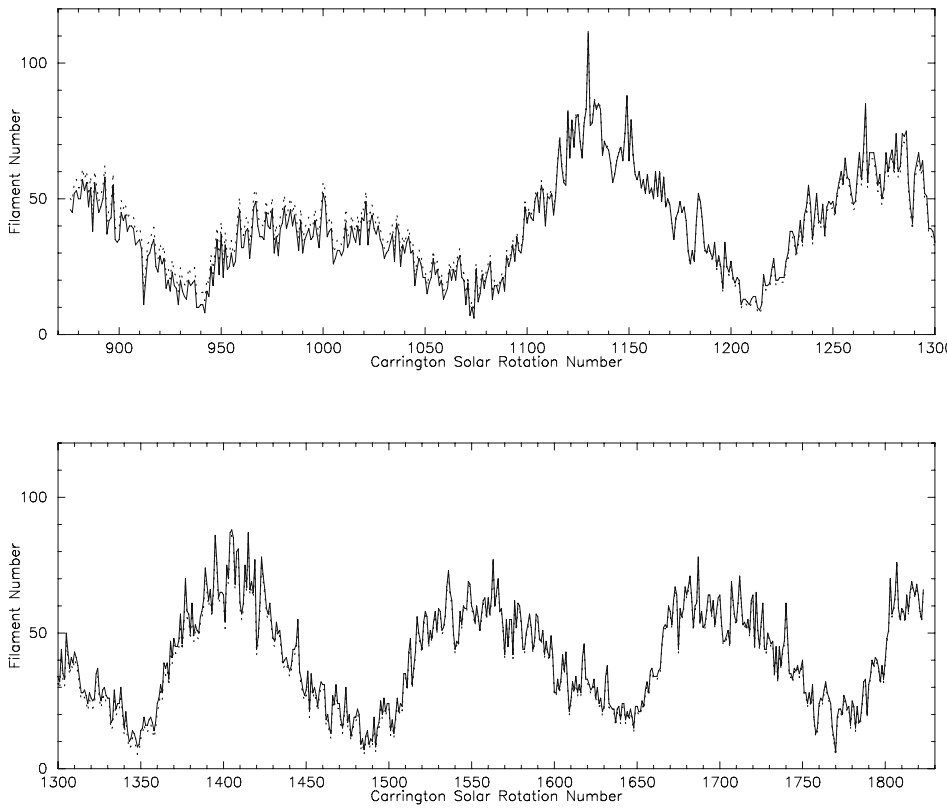

Figure 6. Comparison of the filament number (the solid line) and its reconstruction time series (the dotted line).

presumably due to noise (Torrence & Compo 1998). Shown in Fig. 6 is the comparison of the filament number and its reconstruction time series. The figure shows that the reconstruction time series fits better in and after cycle 17 than before, indicating that the early data of the filament archive are more noiseful.

3. Conclusions and discussions

The Carte Synoptique solar filament archive, namely, the catalogue of solar filaments observed at the Observatorie de Paris, Section de Meudon from March 1919 to December 1989, corresponding Carrington solar rotation numbers 876 to 1823 is used to show the long-term cyclic variation in the space and time distributions of the solar filaments. The number of the low-latitude solar filaments whose latitudes are less than 50° is counted in every Carrington solar rotation number of the consideration time. The time-latitude distribution of the filaments shows the butterfly plot of the coronal activity and the "rush to the pole" close to solar maximum when the solar polar magnetic field reversal occurs. Strangely, very few filaments appear at high latitudes after the solar polar magnetic field reversal.

The temporal distribution of the occurrence of the filaments shows that the filament number waxes and wanes with an approximate 11-year Schwabe cycle, as the

sunspot number does. It is the cyclic activity of solar filaments. Comparing with sunspot cycle,

- the minimum times of activity cycle of solar filaments are very close to those of sunspot cycles;
- the amplitudes of the cycle maximum varies only in a small range, obviously less than that for sunspot cycle;
- there is not a statistical positive linear correlation between the maximum smoothed filament number and the maximum smoothed sunspot number;
- the maximum times of activity cycle of solar filaments are later than those of sunspot cycles for 4 cycles of the total 6 cycles, with some delay on the downward curve of the filament data after sunspot maximum;
- the filament number appears to be more sensitive to solar minima than sunspot number (Coffey & Hanchett 1998b). Thus, it is inferred here that activity cycle of solar filaments should have the same cycle length as sunspot cycle, but the cycle behaviour of solar filaments is globally similar in profile with, but different in detail from, that of sunspot cycles. The amplitude of solar magnetic activity should not keep in phase with the complexity of solar magnetic activity.

The complex Morlet wavelet analysis is utilized to study the periodicity in both the filament number and its smoothed number. The possible periods in the filament activity are about 10.44 (the Schwabe period of coronal activity) and 19.20 years. The wavelet local power spectrum of the period 10.44 years is statistically significant at the whole consideration time. The wavelet local power spectrum of the period 19.20 years is under the 95% confidence spectrum at the whole consideration time, but over the mean red-noise spectrum of $\alpha = 0.72$ before approximate Carrington rotation number 1500, and after then the filament activity does not statistically show the period. The periods of 154 days and 1.3 years, appearing in sunspot activity are not detected in the filament activity, but the period of about 5.4 years is detected. The Gleissberg period is not detected yet due to the short length of the filament data. Wavelet reconstruction indicates that the filament data are more noiseful in and before cycle 16 than after.

One thing should be pointed out that one Carrington solar rotation cycle is here regarded approximately as 27 days, or 0.9 months. The data used here show that 948 Carrington solar rotation cycles from Carrington solar rotation numbers 876 to 1823 correspond to 850 months from March 1919 to December 1989, therefore, one Carrington solar rotation cycle is 0.897 months. The two values are very close to each other.

Acknowledgements

The authors would like to thank the referee for useful comments. The wavelet transform software was provided by C. Torrence and G. Campo and is available at URL: http://paos.colorado.edu/research/wavelets/. This work is supported by the 973 project (2006CB806300), the National Natural Science Foundations of China (40636031 and 10573034), and the Chinese Academy of Sciences.

References

Bumba, V. 2003, Cyclic changes of the solar global and local magnetic field patterns, In: *Proceedings of ISCS 2003 Symposium, Solar Variability as an Input to the Earth's Environment*, Tatranska Lomnica, Slovakia, ESA SP-535, pp. 23–28.

Coffey, H. E., Hanchett, C. D. 1998a, *ASP Conference Series*, **140**, 447.
Coffey, H. E., Hanchett, C. D. 1998b, *ASP Conference Series*, **150**, 488.
d'Azambuja, L. 1928, Annales de l'Observatoire de Paris, Section de Meudon, Tom VI, Fascicule I.
d'Azambuja, L. 1923, *Comptes Rendus*, **176**, 950.
d'Azambuja, L., d'Azambuja, M. 1948, Annales de l'Observatoire de Paris, Section de Meudon, Tom VII, Fascicule VII.
De Moortel, I., McAteer, R. T. J. 2004a, *Solar Phys.*, **222**, 203.
De Moortel, I., Munday, S. A., Hood, A. W. 2004b, *Solar Phys.*, **223**, 1.
Faria, H. H., *et al.* 2004, *Solar Phys.*, **223**, 305.
Howe, R., Christensen-Dalgaard, J., Hill, F. *et al.* 2000, *Science*, **287**, 2456.
Kane, P. R. 2003, *JGR*, **108(A1)**, 1046.
Kane, P. R. 2005, *Solar Phys.*, **227**, 155.
Knaack, R., Stenflo, J. O. 2004, In *Large-Scale Solar Magnetic Fields: Temporal Variations, Syars and Suns: Activity, Evolution and Planets*; Dupree, A. K., Benz, A. O., (eds.) *IAU Symposium*, **219**, pp. 552–556.
Kononovich, E. V., Shefov, N. N. 2002, Quasi-biennial variations of the solar activity during the 11-year cycle and their display in the variations of the temperature of the middle atmosphere, "Physics of Auroral Phenomena". In: *Proceedings of XXV Annual Seminar, Apatity,* Kola Science Center, 143, Polar Geophysical Institute, Russian Academy of Science, pp. 143–146.
Krivova, N. A., Solanki, S. K. 2002, *AA*, **394**, 701.
Li, K. J., Liang, H. F., Yun, H. S., Gu, X. M. 2002, *Solar Phys.*, **205**, 361.
Martin, S. F. 1990, *IAUS*, **138**, 129.
McIntosh, P. S. 1972, *Rev. Geophys. Space Sci.*, **10**, 837.
Minarovjech, M., Rybansky, M., Rusin, V. 1998a, *Solar Phys.*, **177**, 357.
Minarovjech, M., Rybansky, M., Rusin, V. 1998b, *ASP Conference Series*, **150**, 484.
Mouradian, Z. 1998, *ASP Conference Series*, **140**, 179.
Mouradian, Z., Soru-Escaut, I. 1994, *A & A*, **290**, 279.
Obridko, V. N., Shelting, B. D. 2001, *Astron. Rep.*, **45**, 1012.
Ogurtsov, M. G., Nagovitsyn, Kocharov, Jungner 2002, *Solar Phys.*, **211**, 371.
Polygiannakis, J., Preka-Papadema, P., Moussas, X. 2003, *Mon. Not. R. Astron. Soc.*, **343**, 725.
Prabhakaran, S. R., Radhika, Revathy, Ramadas 2002, *Solar Phys.*, **208**, 359.
Sakurai, T., 1998, *ASP Conf. Ser.*, **140**, 483.
Rusin, V., Rybansky, M., Minarovjech, M. 1998, *ASP Conference Series*, **140**, 353.
Rusin, V., Minarovjech, M., Rybansky, M. 2000, *J. Astrophys. Astr.*, **21**, 201.
Tandberg-hanssen, E. 1995, The nature of solar prominences, Kluwer Acad. Publ., Dordrecht, Holland.
Torrence, C., Compo, G. P. 1998, *Bull. Am. Meteorol. Soc.*, **79**, 61.

J. Astrophys. Astr. (2007) **28,** 157–166

Forbidden Transition Probabilities of Astrophysical Interest among Low-lying States of V III

Andrei Irimia
Department of Physics and Astronomy, Vanderbilt University, Nashville, Tennessee, 37235, USA.
e-mail: andrei.irimia@vanderbilt.edu

Received 2005 May 15; accepted 2007 August 7

Abstract. Electric and magnetic multipole transitions among low-lying states of doubly ionized vanadium were computed using the multi-configuration Hartree–Fock (MCHF) method with Breit–Pauli (BP) corrections to a non-relativistic Hamiltonian. Energy levels were determined up to and including $3d^2(^1G)4s$ b $^2G_{7/2}$ and computed energies were found to be in good agreement with experiment and other theories. In addition to Einstein A_{ki} coefficients for some E2 and M1 transitions, lifetime data and selected weighted oscillator strengths are also reported.

Key words. Atomic structure calculations—atomic data—oscillator strength—lifetimes.

1. Introduction

Vanadium transition data are useful in a wide range of scientific applications. Ionized vanadium, for example, has been used in plasma diagnosis (Wagatsuma & Danzaki 1999), where physical properties such as electron temperature and concentration can be determined by analyzing the intensity and width of spectral lines (Griem 1964). Lines of forbidden ionized vanadium were observed in the spectra of B[e] stars by Jaschek & Andrillat (1997) while Lodders (2002) studied the abundance of this element in low-mass dwarf stars. A chemical abundance study of brown dwarfs made use of vanadium spectroscopic data to determine its depletion in these astrophysical objects (Burrows *et al.* 2000; Gounelle *et al.* 2001) used spectral data of vanadium for a study of cosmic-ray irradiation in early solar-system rocks. Yields of this element in massive stars were modeled by Samland (1998), who made the observation that an accurate interpretation of the observational database for vanadium is restricted at the present time by the fact that abundance determinations for metal-poor and solar-metalicity stars are missing. Transition data in vanadium were also used by Whaling *et al.* (1985) and Biemont *et al.* (1989) to determine the solar abundance of this element.

Experimental studies of the Sc-like isoelectronic sequence are few in spite of dating as far back as the 1920s (Gibbs & White 1927; White 1929a, 1929b; Goly 1978;

Bromage 1978). Gibbs & White (1927) determined the wavelengths and relative positions of multiplets in the V III spectrum while White (1929a, 1929b) focused on investigating term separations for this ion. Ekberg (1976) observed the spectrum of V III in a vacuum sliding-spark discharge experiment and made contributions to our present understanding of this ion. Radiative lifetimes for excited states in V III were measured by Andersen *et al.* (1977) for a few $3d^2\ 4p$ levels using a beam-foil experiment, but their data do not include lifetimes for the metastable levels of the $3d^3$ and $3d^2\ 4s$ configurations. In recent years, Stark broadening parameters for spectral lines of V I, V II and V III were determined by Popovic & Dimitrijevic (2000).

Early theoretical investigations of V III include a spectrum calculation by Many (1946) and a study of d^3 and d^4 configurations of vanadium by Meshkov (1954). In 1963, Mendlowitz computed intermediate-coupling transition strengths for $d^2\ d \rightarrow d^2 p$ and $d^3 \rightarrow d^2 p$ allowed transitions.

A spin-dependent interaction calculation for this spectrum was performed by Pasternak & Godschmidt (1972) and some oscillator strengths were determined by Roberts (1973), who also computed the lifetimes of several $3d^2(a\ ^3F)4p$ levels. An atomic partition function calculation for some Ti and V ions – including V III – was done in 1988 by Halenka (1988), followed by Zilitis (2001), who employed the Dirac-Fock method to study the structure of the energy level system and to determine ionization potentials for 39 Sc-like ions.

In the scandium sequence, some forbidden transitions were computed by Luke (1997) for singly ionized vanadium and atomic data have also been published for V II (Roberts 1973; Halenka 1988) and V V (Berry 1976). Nevertheless, a similar study involving doubly ionized vanadium has not yet been performed. In a 2003 compilation of resonance absorption lines for wavelengths longward of the Lyman limit, Morton (2003) mentioned the lack of sufficient atomic data for V III and indicated the work of Kurucz (1998) as one of the very few recent sources of reliable tabulations for this ion.

Forbidden transitions play an important role in many astrophysical applications, particularly in the interpretation of spectra of diluted astrophysical plasmas, where energy is transfered to the plasma via excitation and ionization processes. Because astrophysical plasma densities are very low, the probability of collisions is small and many states decay by M1 or E2 transition radiation. This is why forbidden transitions are useful indicators of plasma densities, indicating a need for both experimental and theoretical data for such transitions (Hartman *et al.* 2003).

Forbidden lines of the Sc-like ions have recently come under much attention as a result of the FERRUM project (Hartman *et al.* 2003), in which the laser probing technique was extended at the CRYRING storage ring to measure extremely long lifetimes, such as that of the metastable $3d^2(^3P)4s$ b $^4P_{5/3}$ level in Ti II. With the recent lifetime measurement of a very long-lived (520^{+310}_{-140} s) metastable state in strontium (Yasuda & Katori 2004), the growing importance of atomic data for forbidden lines – including for the Sc-like sequence – cannot be overstated. In this paper, we present the results of a V III spectrum calculation for states up to and including $3d^2(^1G)4s$ b $^2G_{7/2}$ using the multi-configuration Hartree Fock (MCHF) method with Breit–Pauli corrections. In addition to energy levels and non-relativistic oscillator strengths, we report lifetimes for the excited states computed as well as A_{ki} coefficients for E2 and M1 transitions in this atom. This is the first theoretical investigation of forbidden lines and metastable state lifetimes for the lower portion of the V III spectrum.

2. Computational approach

In the non-relativistic MCHF approach (Froese Fischer 1991), the wave function Ψ of state γLS is written as

$$\Psi(\gamma LS) = \sum_j c_j \Phi\left(\gamma_j LS\right), \tag{1}$$

where γ denotes the dominant configuration as well as any other quantum numbers required to specify the state uniquely. The wave function Ψ is expanded in terms of configuration state functions (CSFs) that have an identical LS symmetry but different electronic configurations γ_j. A basis consisting of one-electron spin-orbital functions

$$\phi_{nlm_l m_s} = \frac{1}{r} P_{nl}(r) Y_{lm_l}(\theta, \phi)\, \chi_{m_s} \tag{2}$$

is used to build the CSFs. Radial functions are determined by the CSFs included in the expansion and they are subjected to the requirement of orthonormality within each l symmetry:

$$\int_0^\infty P_{n'l}(r) P_{nl}(r) dr = \delta_{n'n}. \tag{3}$$

Since the orbitals obtained using the MCHF procedure are eventually used in a Breit–Pauli configuration interaction (CI) calculation involving many LS terms, our method optimizes on a weighted linear combination of energy expressions thereby optimizing simultaneously for a group of terms. In the multiconfiguration self-consistent field (MC–SCF) approach, both radial functions and expansion coefficients are determined so that the energy functional

$$\langle \gamma LS | \mathcal{H} | \gamma LS \rangle \tag{4}$$

is left stationary. The iterative Davidson algorithm (Davidson 1975) is used to determine the desired eigenvalues and eigenvectors.

Because the Breit–Pauli interaction matrix must be computed from one orthonormal orbital basis, LS terms are grouped according to the term interaction. Simultaneous optimization of the radial functions is performed on all LS states that were grouped together. After obtaining the set of radial orbitals, relativistic corrections are taken into account within the BP approximation by diagonalizing the BP Hamiltonian (Froese Fischer 1997) to obtain the intermediate coupling functions

$$\Psi(\gamma J) = \sum_{LS} \sum_j c_j(LSJ) \Phi(\gamma_j LSJ). \tag{5}$$

The expansion coefficients $c_j(LSJ)$ and the corresponding energy $E(LSJ)$, are an eigenvector and eigenvalue, respectively, of the interaction matrix. All contributors to the Breit–Pauli Hamiltonian were included in the present calculation with the exception of the orbit–orbit interaction term, which does not contribute to interactions between CSFs from different LS terms. It is convenient to think of the Breit–Pauli interaction matrix as having a block structure, in that diagonal blocks are the interactions within an LS term, while off-diagonal blocks represent the interaction between the

terms. Two LS terms are referred to as interacting if, for some value of J, there are non-zero interactions in the off-diagonal block for the pair of LS terms.

Weighted oscillator strengths gf are computed here using the length and velocity gauges given by

$$g_i f_l(ik) = \frac{2}{3}\Delta E_{ik}\left|\left\langle \Psi_i \left|\left| \sum_j \mathbf{r}_j \right|\right| \Psi_k \right\rangle\right|^2 \tag{6}$$

and

$$g_i f_v(ik) = \frac{2}{3}\frac{1}{\Delta E_{ij}}\left|\left\langle \Psi_i \left|\left| \sum_j \nabla_j \right|\right| \Psi_k \right\rangle\right|^2, \tag{7}$$

respectively. In the equations above, g_i is the degeneracy factor, i.e., $g_i = (2L_i + 1)(2S_i + 1)$ for LS-coupled wave functions and $g_i = (2J_i + 1)$ for BP wave functions.

In the computational approach employed here, an active set (AS) of orbitals was used to determine wave function expansions. The latter were obtained for increasing sizes of the principal quantum number n and the convergence of the calculation was monitored by observing the agreement of the non-relativistic length and velocity gauges of the oscillator strength. Since the length gauge is correct to $\mathcal{O}(\alpha^2)$ while the velocity gauge requires a relativistic correction to the gradient operator (Drake 1972), both gauges are reported in this work. Throughout this paper, we refer to the $\{1s, 2s, 2p, \ldots, 3s, 3p, 3d\}$ set of orbitals as the $n = 3$ orbital set, $\{1s, 2s, 2p, \ldots, 4s, 4p, 4d, 4f\}$ as $n = 4$, etc. In the case of the present calculation, expansions were obtained for orbital sets with $n = 4, \ldots, 6$ and $s \leq l \leq g$ by means of generating single (S) and double (D) excitations from a multi-reference set, where $1s^2 2s^2 2p^6$ was considered to form a closed common core and the multi-reference set contained the configurations $3d^3$ and $3d^2 4s$. Configuration states from SD excitations which did not interact with at least one member of the multi-reference set were discarded.

The number of CSFs included in the calculation increases rapidly both with n and with the number of electrons outside the common core. Since large orbital sets can result in a considerable increase in computational time required for the problem, appropriate restrictions are necessary. For this reason, for each of $n = 7, 8, 9$, only expansions over even $3d\ nl\ n'l'$ configuration states (where nl and $n'l'$ are orbitals from the orbital set) were added to the existing set of CSFs.

Once radial functions were determined for each group, the Breit–Pauli CI calculations were performed and transition data determined for these wave functions. All E2, E3, M1, M2 and M3 transitions between the states targeted by the calculation were computed. To improve the reliability of the mixing of LS terms in the Breit–Pauli interaction matrix, fine-tuning (Hibbert 1993) was applied to its elements. In this process, adjustments were made to the matrix elements that shift all diagonal components of a given LS block by a fixed amount. This shift was the difference of an *ab initio* energy level and the observed level. If more than one eigenstate in a group had the same LS value as, for example, $3d^3$ a ^4F and $3d^2$b $4s^4F$, energy adjustments were made separately for each of the eigenstates with different shifts for ^4F.

Table 1. Breit–Pauli energies, energy differences (computed – observed (NIST), in cm^{-1}), splittings (in cm^{-1}), and lifetimes (in s) for excited states up to and including $3d^2(^1G)4s$ b $^2D_{3/2}$ after fine-tuning.

Config.	Term	J	Energy	Diff.	Splitting	Lifetime
$3d^3$	a ^4F	3/2	0.00			
		5/2	176.72	31.22	176.72	4.1984E+03
		7/2	400.29	58.79	400.29	2.0642E+03
		9/2	646.92	63.12	646.92	2.4716E+03
	a ^4P	1/2	11482.92	−30.88		2.3032E+01
		3/2	11618.03	26.23	135.12	2.3673E+01
		5/2	11774.37	4.67	291.46	2.5194E+01
	a ^2G	7/2	11948.37	−17.93		7.6451E+02
		9/2	12204.93	17.93	256.57	4.9393E+03
	a ^2P	3/2	15459.73	−90.57		9.8242E+02
		1/2	15670.38	90.58	210.65	1.1899E+04
	a ^2D	3/2	16205.37	−125.13		4.4811E+04
		5/2	16502.87	128.17	297.51	3.2689E+03
	a ^2H	9/2	16793.77	−17.13		1.4307E+04
		11/2	16994.74	17.14	200.97	5.9964E+03
	a ^2F	7/2	27724.56	−3.24		8.2104E+00
		5/2	27850.05	3.25	125.50	7.2835E+00
	b ^2D	5/2	42261.58	−5.82		1.6962E−01
		3/2	42371.21	0.01	109.64	1.6224E−01
$3d^2(^3F)4s$	b ^4F	3/2	43893.05	−49.44		1.1551E−02
		5/2	44095.10	−14.94	202.06	1.1435E−02
		7/2	44354.29	8.47	461.25	1.1297E−02
		9/2	44646.95	−0.01	753.90	1.1160E−02
	b ^2F	5/2	49314.12	−13.62		1.7603E−02
		7/2	49818.90	13.61	504.78	1.6899E−02
$3d^2(^1D)4s$	c ^2D	5/2	56204.84	44.42		1.1082E−02
		3/2	56212.33	−44.42	7.49	1.1234E−02
$3d^2(^3P)4s$	b ^4P	1/2	56567.40	38.10		4.5039E−03
		3/2	56703.31	34.26	135.91	4.4934E−03
		5/2	56850.12	−72.38	282.72	4.4967E−03
	b ^2P	1/2	61570.04	−8.70		1.2980E−02
		3/2	61785.86	8.71	215.82	1.2495E−02
$3d^2(^1G)4s$	b ^2G	9/2	63307.35	4.23		7.2629E−03
		7/2	63310.81	−4.24	3.45	7.1578E−03

Table 2. Convergence trends for some non-relativistic weighted oscillator strengths (gf values) in V III. Per cent differences were computed according to the formula $[gf(l) - gf(v)] \times 100/\max\{gf(l),\ gf(v)\}$.

$3d^3$ a $^4F - 3d^2(^1G)4s$ b ^4P				$3d^3$ a $^2H - 3d^2(^1G)4s$ b ^2G			
n	$gf(l)$	$gf(v)$	Diff. (%)	n	$gf(l)$	$gf(v)$	Diff. (%)
4	9.82E−7	7.74E−7	21.2	4	1.21E−6	8.99E−7	25.6
5	8.80E−7	7.32E−7	16.8	5	1.06E−6	8.78E−7	17.2
6	8.70E−7	7.55E−7	13.3	6	1.03E−6	9.00E−7	12.5
7	8.60E−7	7.90E−7	8.2	7	1.01E−6	9.46E−7	6.3
8	8.60E−7	7.91E−7	8.1	8	1.01E−6	9.49E−7	6.0
9	8.59E−7	7.92E−7	7.8	9	1.01E−6	9.50E−7	5.8

Table 3. Non-relativistic weighted oscillator strengths (gf values) in V III. Per cent differences were computed according to the formula $[gf(l) - gf(v)] \times 100/\max\{gf(l),\ gf(v)\}$.

Transition			$gf(l)$	$gf(v)$	Diff. (%)
$3d$ a ^2G	–	$3d$ a ^2D	1.518E−11	1.683E−11	9.8
$3d$ a ^2G	–	$3d$ a ^2G	5.290E−07	5.546E−07	4.6
$3d$ a ^2G	–	$3d^2(^3F)4s$ c ^2D	2.416E−07	2.545E−07	5.1
$3d$ a ^2G	–	$3d^2(^3P)4s$ b ^2P	9.975E−08	1.050E−07	5.0
$3d$ a ^2D	–	$3d^2(^1G)4s$ b ^2G	3.468E−08	3.548E−08	2.3
$3d^2(^3F)4s$ b ^2F	–	$3d^2(^3P)4s$ b ^2P	1.862E−09	1.775E−09	4.7

3. Results and discussion

In Table 1, we present our computed energies for the V III spectrum up to and including $3d^2(^1G)4s$ b $^2G_{7/2}$. In addition to LS term splittings, we also report energy differences (computed – observed (NIST Online Database)) and lifetimes for the excited states. The differences from observed are very good for these term energies. The separation between the levels of each term determines its spread and the computed separation would be the same as observed if term separations were exact.

One feature that makes the Sc-like isoelectronic sequence somewhat unusual is the fact that the ground state configuration changes along each of the first three members of the sequence. Thus, the ground state configuration changes from $3d\ 4s^2$ in Sc I to $3d^2\ 4s$ in Ti II to $3d^3$ in V III. The lowest odd-parity configuration in Sc I is $3d\ 4s\ 4p$, replaced by $3d^2\ 4p$ in both Ti II and V III. Although the separation between $3d^2\ 4s$ even-parity levels and the ground state increases from $Z = 21$ to $Z = 23$, this phenomenon occurs slower than it does for odd-parity $3d^2\ 4p$ levels. The net result of this is that $3d^2\ 4s$ remains the lowest excited configuration in V III, excluding the ground state configuration $3d^3$. Thus, the lowest portion of the V III spectrum is populated by 34 even-parity levels that belong to the $3d^3$ and $3d^2\ 4s$ configurations. As Table 1 shows, these levels are metastable because the excited electron can only decay via forbidden electric (E2, E4) and magnetic (M1, M3) multipole transitions. For

Table 4. Einstein A_{ki}-coefficients in s^{-1} for selected E2 transitions after fine-tuning[a]. Per cent differences were computed according to the formula $[A_{ki}(L) - A_{ki}(V)] \times 100/\max\{A_{ki}(L),\ A_{ki}(V)\}$.

Transition			$A_{ki}(L)$	$A_{ki}(V)$	Diff. (%)
$3d^3$ a $^2D_{5/2}$	–	$3d^2(^1D)4s$ c $^2D_{5/2}$	1.79E+01	1.86E+01	−3.85
	–	$3d^2(^1D)4s$ c $^2D_{3/2}$	6.69E+00	6.97E+00	−4.03
	–	$3d^2(^1D)4s$ b $^2P_{1/2}$	8.16E+00	8.74E+00	−6.55
$3d^3$ a $^2D_{3/2}$	–	$3d^2(^1D)4s$ c $^2D_{5/2}$	4.62E+00	4.77E+00	−3.21
	–	$3d^2(^1D)4s$ c $^2D_{3/2}$	1.61E+01	1.67E+01	−3.39
	–	$3d^2(^1D)4s$ b $^2P_{3/2}$	6.19E+00	6.58E+00	−5.94
	–	$3d^2(^1D)4s$ b $^2P_{1/2}$	1.21E+01	1.30E+01	−6.84
$3d^3$ a $^2D_{3/2}$	–	$3d^2(^1G)4s$ b $^2G_{7/2}$	2.27E+00	2.32E+00	−2.14
$3d^3$ a $^2D_{5/2}$	–	$3d^2(^1G)4s$ b $^2G_{9/2}$	2.61E+00	2.64E+00	−0.97
	–	$3d^2(^1G)4s$ b $^2G_{7/2}$	2.62E+00	2.65E+00	−1.31
	–	$3d^2(^1G)4s$ b $^2D_{5/2}$	1.79E+01	1.86E+01	−3.80
$3d^3$ a $^2D_{5/2}$	–	$3d^2(^1G)4s$ b $^2G_{9/2}$	2.61E+00	2.64E+00	−0.97
	–	$3d^2(^1D)4s$ b $^2G_{7/2}$	2.62E−01	2.65E−01	−1.31
$3d^3$ a $^2F_{7/2}$	–	$3d^2(^3F)4s$ b $^2F_{7/2}$	3.33E−01	3.19E−01	4.20
	–	$3d^2(^3F)4s$ b $^2F_{5/2}$	4.76E−02	4.77E−02	−0.21
$3d^3$ a $^2F_{5/2}$	–	$3d^2(^3F)4s$ b $^2F_{7/2}$	3.88E−02	3.74E−02	3.64
	–	$3d^2(^3F)4s$ b $^2F_{5/2}$	2.77E−01	2.79E−01	−0.83
$3d^3$ a $^2H_{11/2}$	–	$3d^2(^1G)4s$ b $^2G_{9/2}$	6.70E+01	6.70E+01	0.00
	–	$3d^2(^1G)4s$ b $^2G_{7/2}$	3.22E+00	3.22E+00	0.05
$3d^3$ a $^2H_{9/2}$	–	$3d^2(^1G)4s$ b $^2G_{9/2}$	3.93E+00	3.94E+00	−0.19
	–	$3d^2(^1G)4s$ b $^2G_{7/2}$	6.89E+01	6.90E+01	−0.15
$3d^2(^1D)4s$ c $^2D_{3/2}$	–	$3d^2(^1D)4s$ c $^2D_{5/2}$	7.98E−20	8.50E−20	−6.15
$3d^2(^3F)4s$ b $^2F_{7/2}$	–	$3d^2(^1D)4s$ c $^2D_{5/2}$	1.34E−05	1.29E−05	4.09
$3d^2(^3F)4s$ b $^2F_{5/2}$	–	$3d^2(^1D)4s$ c $^2D_{3/2}$	2.00E−05	1.95E−05	2.60

[a]The velocity form of the transition operator has neglected some relativistic corrections and hence, unlike the length form, is not correct in that, some terms of order α^2 have been omitted. The values of the length and velocity gauges have significance when term mixing is small and the transition is spin-allowed.

example, the lowest excited level $3d^3$ a $^4F_{5/2}$ has a very long lifetime of approximately 69 minutes. In the case of several other excited levels, the nature of the spectrum causes their primary decay channels to involve magnetic octupole (M3) transitions, which dramatically lengthens their associated lifetimes. As the table suggests, $3d^3$ levels are all very long-lived, while $3d^2$ $4s$ levels have shorter lifetimes, *albeit* still of the order of hundredths of a second.

Because the lifetimes of excited levels in V III involve highly forbidden transitions that have very small Einstein A_{ki} coefficients, the computation of transition proba-

Table 5. Einstein A_{ki}-coefficients in s^{-1} for selected M1 transitions after fine-tuning.

Transition			A_{ki}
$3d^3$ a $^4P_{3/2}$	–	$3d^3$ b $^4P_{5/2}$	6.18456E−05
$3d^3$ a $^2P_{1/2}$	–	$3d^3$ b $^2P_{3/2}$	8.40396E−05
$3d^3$ a $^2D_{3/2}$	–	$3d^3$ a $^2D_{5/2}$	2.84062E−04
$3d^3$ a $^2F_{7/2}$	–	$3d^3$ a $^2F_{5/2}$	3.04621E−05
$3d^3$ a $^2G_{7/2}$	–	$3d^3$ a $^2G_{9/2}$	2.02458E−04
$3d^3$ a $^2D_{5/2}$	–	$3d^3$ b $^2D_{5/2}$	8.46173E−11
$3d^3$ a $^2D_{5/2}$	–	$3d^3$ b $^2D_{3/2}$	4.33143E−02
$3d^3$ a $^2D_{3/2}$	–	$3d^3$ b $^2D_{5/2}$	2.95112E−02
$3d^3$ a $^2D_{3/2}$	–	$3d^3$ b $^2D_{3/2}$	1.40610E−11
$3d^3$ a $^2F_{7/2}$	–	$3d^2(^3F)4s$ b $^2F_{7/2}$	1.94999E−12
$3d^3$ a $^2F_{7/2}$	–	$3d^2(^3F)4s$ b $^2F_{5/2}$	3.96734E−04
$3d^3$ a $^2F_{5/2}$	–	$3d^2(^3F)4s$ b $^2F_{7/2}$	3.13588E−04
$3d^3$ a $^2F_{5/2}$	–	$3d^2(^3F)4s$ b $^2F_{5/2}$	2.37724E−13
$3d^3$ a $^4P_{5/2}$	–	$3d^2(^3F)4s$ b $^4P_{3/2}$	2.39644E−05
$3d^3$ a $^2D_{5/2}$	–	$3d^2(^1D)4s$ c $^2D_{5/2}$	3.32393E−13
$3d^3$ a $^2D_{5/2}$	–	$3d^2(^1D)4s$ c $^2D_{3/2}$	6.43609E−06
$3d^3$ a $^2D_{3/2}$	–	$3d^2(^1D)4s$ c $^2D_{5/2}$	6.21414E−06
$3d^3$ a $^2D_{3/2}$	–	$3d^2(^1D)4s$ c $^2D_{3/2}$	6.78299E−14
$3d^3(^3F)4s$ b $^2F_{5/2}$	–	$3d^2(^3F)4s$ b $^2F_{7/2}$	1.48676E−03
$3d^3(^3F)4s$ b $^4F_{7/2}$	–	$3d^2(^3F)4s$ b $^4F_{9/2}$	6.76096E−04
$3d^3(^3P)4s$ b $^4P_{3/2}$	–	$3d^2(^3P)4s$ b $^4P_{5/2}$	5.12022E−05

bilities in this ion requires a very high level of accuracy. To ensure the convergence of our calculation, the length and velocity gauges of the non-relativistic weighted oscillator strength (gf values) were monitored. If convergence is achieved as the active set of orbitals is expanded, the agreement between the two gauges should improve. In Table 2, we present convergence trends for two forbidden transitions between excited levels. In both cases, the label change is $3d \rightarrow 4s$, which makes such transitions suitable for testing how well correlation was captured by our calculation.

In Table 3, we present several other computed oscillator strengths for transitions between excited levels where the change $3d \rightarrow 4s$ occurs. We also include one case of a parent-changing ($^3F \rightarrow ^3P$) transition between two excited levels of the $3d^2\ 4s$ configuration. For all these transitions, the agreement of the two gauges is excellent in spite of their small magnitudes.

Some Einstein A_{ki} coefficients for electric quadrupole and magnetic dipole transitions in V III are made available in Tables 4 and 5, respectively. The selection of these A_{ki} values was made on the basis of their magnitudes and contributions to the associated decay channels. For many of the computed levels in V III, the accuracy of A_{ki}

values for E2 and M1 transitions is essential in determining the accuracy of lifetime data. This is because, for the most part, A_{ki} coefficients for E4 and M3 transitions were found to have much lower values than their E2 and M1 counterparts. This indicates that the corresponding E4 and M3 decay channels have negligible contributions to the reported lifetimes whenever E2 and/or M1 decay channels are also available for a certain level.

4. Conclusion

In summary, we have computed forbidden transitions of astrophysical interest among low-lying states of V III using the MCHF approach with BP corrections to a non-relativistic Hamiltonian. Excellent agreement was obtained between computed and observed energy levels and convergence trends for the weighted oscillator strengths were also satisfactory. Good agreement between the length and velocity gauges of these quantities was obtained, although typically this is notoriously difficult to achieve for forbidden transitions since the g_f values associated with them are usually small. Some Einstein A_{ki} coefficients were also presented.

References

Andersen, T., Petersen, P., Biemont, E. 1977, *J. Quant. Spectrosc. Radiat. Transfer*, **17**, 389.
Berry, H. G. 1976, *Phys. Scripta*, **13**, 36.
Biemont, E., Grevesse, N., Faires, L. M., Marsden, G., Lawler, J. E., Whaling, W. 1989, *Astron. Astrophys.*, **209**, 391.
Bromage, G. E. 1980, *Astron. Astrophys. Suppl.*, **41**, 79.
Burrows, A., Marley, M. S., Sharp, C. M. 2000, *Astrophys. J.*, **531**, 438.
Davidson, E. R. 1975, *J. Comput. Phys.*, **17**, 87.
Drake, G. W. R. 1972, *Phys. Rev. A*, **5**, 1979.
Ekberg, J. O. 1976, *Phys. Scripta*, **13**, 111.
Froese Fischer C. 1991, *Comput. Phys. Commun.*, **74**, 415.
Froese Fischer, C., Brage, T., Jönsson, P. 1997, *Computational Atomic Structure – An MCHF Approach* Bristol: IOP Publishing.
Gibbs, R. C., White, H. E. 1927, *Phys. Rev.*, **29**, 655.
Goly, A. 1978, *Astron. Astrophys.*, **68**, 79.
Griem, H. R. 1964, *Plasma Spectroscopy*, McGraw-Hill, NY.
Gounelle, M., Shu, F. H., Shang, H., Glassgold, A. E., Rehm, K. E., Lee, T. 2001, *Astrophys. J.*, **548**, 1051.
Halenka, J. 1988, *Astron. Astrophys. Suppl.*, **75**, 47.
Hartman, H. *et al.* 2003, *J. Phys. B: At. Mol. Opt. Phys.*, **36**, L197.
Hibbert, A. 1993, *Phys. Scripta*, **T65**, 104.
Jaschek, C., Andrillat, Y. 1997, *Astron. Astrophys. Suppl.*, **128**, 475.
Kurucz, R. L. 1998, http://kurucz.harvard.edu/atoms.html.
Lodders, K. 2002, *Astrophys. J.*, **577**, 974.
Luke, T. M. 1997, *J. Phys. B: At. Mol. Opt. Phys.*, **30**, 4223.
Many, A. 1946, *Phys. Rev.*, **70**, 511.
Mendlowitz, H. 1963, *Astrophys. J.*, **138**, 1277.
Meshkov, S. 1954, *Phys. Rev.*, **93**, 270.
Morton, D. C. 2003, *Astrophys. J. Suppl.*, **149**, 205.
NIST atomic spectra database (http://physics.nist.gov/cgi-bin/AtData/main_asd)
Pasternak, A., Goldschmidt, Z. B. 1972, *Phys. Rev. A*, **6**, 55.
Popovic, L. C., Dimitrijevic, M. S. 2000, *Phys. Scripta*, **61**, 192.
Roberts, J. R. 1973, *Astrophys. J.*, **181**, 578.
Samland, M. 1998, *Astrophys. J.*, **496**, 155.

Wagatsuma, K., Danzaki, Y. 1999, *J. Anal. At. Spect.*, **14**, 1727.
Whaling, W., Hannaford, P., Lowe, R. M., Biemont, E., Grevesse, N. 1985, *Astron. Astrophys.*, **153**, 109.
White, H. E. 1929a, *Phys. Rev.*, **33**, 672.
White, H. E. 1929b, *Phys. Rev.*, **33**, 914.
Yasuda, M., Katori, H. 2004, *Phys. Rev. Lett.*, **92**, 153004-1.
Zilitis, V. A. 2001, *Opt. Spectrosc.* **92**, 353.

J. Astrophys. Astr. (2007) **28,** 167–174

Causal Temperature Profiles in Horizon-free Collapse

N. F. Naidu* & M. Govender**
Astrophysics and Cosmology Research Unit, School of Mathematics, University of KwaZulu Natal, Durban 4041, South Africa.
**e-mail: 203507365@ukzn.ac.za*
***e-mail: govenderm43@ukzn.ac.za*

Received 2006 March 24; accepted 2008 February 12

Abstract. We investigate the causal temperature profiles in a recent model of a radiating star undergoing dissipative gravitational collapse without the formation of a horizon. It is shown that this simple exact model provides physically reasonable behaviour for the temperature profile within the framework of extended irreversible thermodynamics.

Key words. Horizon-free collapse—thermodynamics.

1. Introduction

The Cosmic Censorship Conjecture occupies center stage within the realms of relativistic astrophysics. The final outcome of the gravitational collapse of a star is still very much open to debate with the discovery of models admitting naked singularities (Harada *et al.* 1998; Kudoh *et al.* 2000). Various scenarios of gravitational collapse have been considered in which the energy momentum tensor is taken to be either a perfect fluid or an imperfect fluid with heat flux and anisotropic pressure (Bonnor *et al.* 1989; Herrera & Santos 1997a; Naidu *et al.* 2006). It has been shown that shearing effects delay the formation of the apparent horizon by making the final stages of collapse incoherent thus leading to the generation of naked singularities (Joshi *et al.* 2002). In this paper we revisit a radiating stellar model proposed by Banerjee *et al.* (Banerjee *et al.* 2002), (hereafter referred to as the *BCD* model) in which the horizon is never encountered. The interior matter distribution is that of an imperfect fluid with heat flux and the exterior spacetime is described by the radiating Vaidya metric (Vaidya 1951). The junction conditions required for the smooth matching of the interior and exterior spacetimes across a four-dimensional time-like hypersurface are solved exactly.

In this paper we investigate the physical viability of the *BCD* model. In particular, we analyse the relaxational effects on the temperature profiles within the framework of extended irreversible thermodynamics. We are in a position to obtain exact solutions to the causal heat transport equation for both the special case of constant collision time as well as variable collision time. Our results are in agreement with earlier thermodynamical investigations of radiating stellar models. We find that relaxational effects enhance the temperature at each interior point of the stellar configuration. Our investigations show that the *BCD* model displays physically reasonable temperature profiles throughout the evolution of the star.

2. The *BCD* radiating model revisited

In the *BCD* model the following form of the metric for the interior spacetime is assumed

$$ds^2 = -A^2(r,t)dt^2 + B^2(r,t)[dr^2 + r^2 d\theta^2 + r^2 \sin^2\theta d\phi^2], \tag{1}$$

in which the metric functions A and B are yet to be determined. The energy momentum tensor for the interior matter distribution is given by

$$T^{\mu\nu} = (\rho + p)\, v^\mu v^\nu + p g^{\mu\nu} + q^\mu v^\nu + q^\nu v^\mu. \tag{2}$$

The heat flow vector q^μ is orthogonal to the velocity vector so that $q^\mu v_\mu = 0$. In order to generate an exact model of radiative gravitational collapse, the following ansatz was adopted for the metric functions in (equation 1)

$$A = a(r), \tag{3}$$

$$B = b(r)R(t), \tag{4}$$

which reduces the Einstein field equations for the interior matter distribution to

$$\rho = \frac{1}{R^2}\left[\frac{3}{a^2}\dot{R}^2 - \frac{1}{b^2}\left(\frac{2b''}{b} - \frac{b'^2}{b^2} + \frac{4b'}{rb}\right)\right], \tag{5}$$

$$p = \frac{1}{R^2}\left[-\frac{1}{a^2}(2R\ddot{R} + \dot{R}^2) + \frac{1}{b^2}\left(\frac{b'^2}{b^2} + \frac{2a'b'}{ab} + \frac{2}{r}\left(\frac{a'}{a} + \frac{b'}{b}\right)\right)\right], \tag{6}$$

$$q^1 = -\frac{2a'\dot{R}}{R^3 a^2 b^2}, \tag{7}$$

where '$\cdot$' and '$'$' indicate derivatives with respect to time and the radial coordinate respectively. The condition of pressure isotropy yields

$$\frac{a''}{a} + \frac{b''}{b} - 2\frac{b'^2}{b^2} - 2\frac{a'b'}{ab} - \frac{a'}{ra} - \frac{b'}{rb} = 0. \tag{8}$$

Since the star is radiating energy the exterior spacetime is described by the Vaidya metric given explicitly in the form

$$ds^2 = -\left(1 - \frac{2M(v)}{\bar{r}}\right)dv^2 - 2d\bar{r}dv + \bar{r}^2(d\theta^2 + \sin^2\theta d\phi^2), \tag{9}$$

where v is the retarded time and $M(v)$ is the exterior Vaidya mass. The junction conditions required for the smooth matching of the interior metric (equation 1) and the exterior Vaidya metric (equation 9) across a time-like hypersurface Σ are given by

$$(rB)_\Sigma = \bar{r}_\Sigma, \tag{10}$$

$$p_\Sigma = (q^1 B)_\Sigma, \tag{11}$$

$$m_\Sigma = \left[\frac{r^3 B\dot{B}^2}{2A^2} - r^2 B' - \frac{r^3 B'^2}{2B}\right]_\Sigma, \tag{12}$$

where m_Σ represents the total mass of the stellar configuration of radius r inside Σ. Utilising equations (6) and (7) in the boundary condition (equation 11) yields

$$2R\ddot{R} + \dot{R}^2 + m\dot{R} = n, \tag{13}$$

where m and n are constants. A simple particular solution of equation (13) is

$$R(t) = -Ct, \tag{14}$$

where $C > 0$ is a constant of integration. As pointed in (Banerjee *et al.* 2002) the mass-to-radius ratio, $m_\Sigma/\bar{r}_\Sigma$, is independent of time. A simple calculation yields

$$\frac{2m_\Sigma}{\bar{r}_\Sigma} = \frac{2m_\Sigma}{(rB)_\Sigma} = 2\left[\frac{C^2 r_0^2 b_0^2}{2a_0^2} - \frac{r_0 b_0'}{b_0} - \frac{r_0^2 b_0'^2}{2b_0^2}\right], \tag{15}$$

where $b(r_0) = b_0$ and r_0 defines the boundary of the stellar configuration. It is interesting to note that the parameters in equation (15) may be chosen so that $2m_\Sigma/\bar{r}_\Sigma < 1$ in order to avoid the appearance of horizon at the boundary.

3. Causal temperature profiles

In this section we consider the physical viability of the *BCD* model. In order to satisfy the condition of pressure isotropy (equation 8), the *BCD* model assumes $b(r) = 1$ and

$$A = a(r) = 1 + \xi_0 r^2. \tag{16}$$

The fluid volume collapse rate is

$$\Theta = \frac{3}{A}\frac{\dot{B}}{B} = \frac{3}{(1+\xi_0 r^2)t}, \tag{17}$$

and in the absence of shear is the same in both the radial and tangential directions. The proper stellar radius is given by

$$r_p(t) = \int_0^{b_0} B dr = -Ctb_0. \tag{18}$$

Since the star is collapsing we require that C be positive which corresponds to $-\infty < t < 0$. We further have

$$C^2 < 4\xi_0(1 + \xi_0 r_0^2). \tag{19}$$

The Einstein field equations (5)–(7) reduce to

$$\rho = \frac{3}{t^2(1+\xi_0 r^2)^2}, \tag{20}$$

$$p = \frac{1}{t^2(1+\xi_0 r^2)^2}\left[\frac{4\xi_0}{C^2}(1+\xi_0 r^2) - 1\right], \tag{21}$$

$$q^1 = -\frac{4\xi_0 r}{(1+\xi_0 r^2)^2}\frac{1}{C^2 t^3}. \tag{22}$$

We note that all the above thermodynamical quantities diverge as $t \to 0$. The regularity conditions $\rho > 0$, $p > 0$ and $\rho' < 0$, and $p' < 0$ together with the dominant energy condition, $(\rho - p) > 0$ and the more stringent requirement $(\rho + p) > 2|q|$ are all satisfied when

$$\left[1 - \frac{2\xi_0 r}{C}\right]^2 > -\frac{2\xi_0}{C^2}(1-\xi_0 r^2). \tag{23}$$

We can now write

$$1 - \frac{2m_\Sigma}{\bar{r}_\Sigma} = \left[1 - \frac{C^2 r_0^2}{(1+\xi_0 r_0^2)^2}\right]. \tag{24}$$

We note that when

$$C^2 < \frac{1}{r_0^2} + \xi_0^2 r_0^2 + 2\xi_0, \tag{25}$$

the boundary surface can never reach the horizon (Banerjee *et al.* 2002). Furthermore, the surface redshift is given by

$$1 + z_\Sigma = \left(1 + r_0\frac{b_0'}{b_0} + r_0\dot{b}_0\right)^{-1}, \tag{26}$$

which diverges for an observer at infinity at the time of the appearance of the horizon. For the *BCD* model, equation (26) reduces to

$$1 + z_\Sigma = (1 - Cr_0)^{-1}, \tag{27}$$

which diverges when $C = 1/r_0$. In order to avoid the divergence of the surface redshift we must have

$$\frac{1}{r_0^2} < C^2 < \frac{1}{r_0^2} + \xi_0^2 r_0^2 + 2\xi_0, \tag{28}$$

where we have taken equation (25) into account. The luminosity of the star as perceived by an observer at infinity is given by

$$L = -\frac{dm}{dv} = \frac{c^3 r^3(1+\xi_0 r^2 - rC)}{(1+\xi_0 r^2)^4}, \tag{29}$$

which is independent of time. We now turn our attention to the evolution of the temperature profiles of the *BCD* model. To this end we employ the causal transport equation for the heat flux, which in the absence of rotation and viscous stress is given by

$$\tau h_a^b \dot{q}_b + q_a = -\kappa (D_a T + T\dot{u}^a)\,, \tag{30}$$

where τ is the relaxation time for the thermal signals. Setting $\tau = 0$ in the above, we regain the so-called Eckart transport equations which predict infinite propagation velocities for the dissipative fluxes. For the line element (equation 1) the causal transport equation (equation 30) reduces to

$$\tau (qB)_{,t} + A(qB) = -\kappa \frac{(AT)_{,r}}{B}, \tag{31}$$

which governs the behaviour of the temperature. Setting $\tau = 0$ in equation (31) we obtain the familiar Fourier heat transport equation

$$A(qB) = -\kappa \frac{(AT)_{,r}}{B}, \tag{32}$$

which predicts reasonable temperatures when the fluid is close to quasi-stationary equilibrium. In order to study the evolution of the temperature in the *BCD* model, we employ the thermodynamic coefficients for radiative transfer as outlined in Maharaj & Govender (2005). The thermal conductivity takes the form

$$\kappa = \gamma T^3 \tau_{\rm c}, \tag{33}$$

where γ (≥ 0) is a constant and $\tau_{\rm c}$ is the mean collision time between the massless and massive particles. We further adopt the generalised power-law behaviour for $\tau_{\rm c}$

$$\tau_{\rm c} = \left(\frac{\alpha}{\gamma}\right) T^{-\omega}, \tag{34}$$

where α (≥ 0) and ω (≥ 0) are constants. The velocity of thermal dissipative signals is assumed to be comparable to the adiabatic sound speed, which is satisfied if the relaxation time is proportional to the collision time:

$$\tau = \left(\frac{\beta\gamma}{\alpha}\right) \tau_{\rm c}, \tag{35}$$

where τ (≥ 0) is a constant. The constant β is a measure of the strength of relaxational effects, with $\beta = 0$ giving the noncausal case. Using the above definitions for τ and κ, equation (31) takes the form

$$\beta (qB)_{,t} T^{-\omega} + A(qB) = -\alpha \frac{T^{3-\omega}(AT)_{,r}}{B}. \tag{36}$$

The Eckart temperature is readily obtained by setting $\beta = 0$ in equation (36). We are in a position to integrate equation (36) for the special case $\omega = 0$ which corresponds to constant collision time and more interestingly, the case $\omega = 4$ which gives a variable collision time. For constant collision time, the causal temperature profile is given by

$$T^4(r,t) = \left(\frac{L}{4\pi\delta}\right)\frac{1}{r_0^2c^2t^2}\left(\frac{1+\xi_0 r_0^2}{1+\xi_0 r^2}\right)^4 + \frac{8\beta\xi_0\left[2(r_0^2-r^2)+\xi_0(r_0^4-r^4)\right]}{\alpha t^2(1+\xi_0 r^2)^4} + \frac{8\xi_0\left[3(r^2-r_0^2)+3\xi_0(r^4-r_0^4)+\xi_0^2(r^6-r_0^6)\right]}{3\alpha t(1+\xi_0 r^2)^4}. \tag{37}$$

For $\omega = 4$, the causal temperature is given by

$$T^4(r,t) = \frac{8\beta\xi_0}{\alpha t^2(1+\xi_0 r^2)^3}\left[\left(\frac{1+\xi_0 r_0^2}{1+\xi_0 r^2}\right)[8+\alpha t(1+\xi_0 r_0^2)]\mathrm{e}^{\frac{8\xi_0}{\alpha t}\left(\frac{r^2-r_0^2}{(1+\xi_0 r^2)(1+\xi_0 r_0^2)}\right)}\right] - \frac{8\beta\xi_0}{\alpha t^2(1+\xi_0 r^2)^3}[8+\alpha t(1+\xi_0 r^2)] + \frac{512\beta\xi_0\mathrm{e}^{-\left(\frac{8}{\alpha t(1+\xi_0 r^2)}\right)}}{\alpha^2 t^3(1+\xi_0 r^2)^4}\left[\text{ExpIntegral } Ei\left(\frac{8}{\alpha t(1+\xi_0 r^2)}\right)\right] - \frac{512\beta\xi_0\mathrm{e}^{-\left(\frac{8}{\alpha t(1+\xi_0 r^2)}\right)}}{\alpha^2 t^3(1+\xi_0 r^2)^4}\text{ ExpIntegral } Ei\left(\frac{8}{\alpha t(1+\xi_0 r_0^2)}\right) + \left[\frac{1+\xi_0 r_0^2}{1+\xi_0 r^2}\right]^4\frac{L}{(4\pi\delta)r_0^2c^2t^2}\mathrm{e}^{\frac{8\xi_0}{\alpha t}\left[\frac{r^2-r_0^2}{(1+\xi_0 r^2)(1+\xi_0 r_0^2)}\right]}, \tag{38}$$

where L is given by equation (29) and δ is a constant. We note that the noncausal temperature ($\beta = 0$) and causal temperature are equal at the boundary ($r = r_0$). Figure 1 shows that the relaxational effects are dominant when the stellar fluid is far from equilibrium (large values of β). In the case of variable collision time, Fig. 2, we see that the causal temperature is everywhere greater than the corresponding noncausal temperature within the stellar interior. Furthermore, Figs. 1 and 2 indicate that the causal temperatures at late times (large values of β) decrease more rapidly than the causal temperatures when the star is close to quasi-static equilibrium. This is in agreement with the perturbative results of Herrera & Santos (1997b) as well as the acceleration-free model studied in Govender *et al.* (1998).

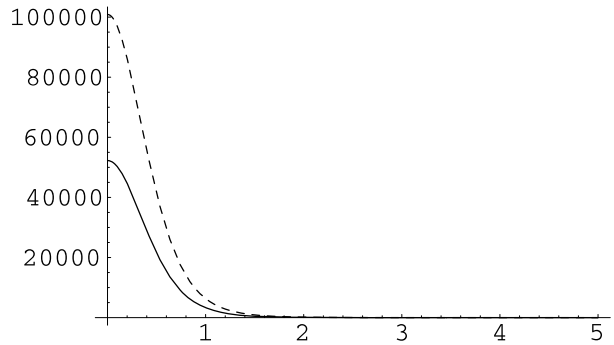

Figure 1. Temperature profiles for constant collision time (close to equilibrium – solid line), (far from equilibrium – dashed line) *versus* r.

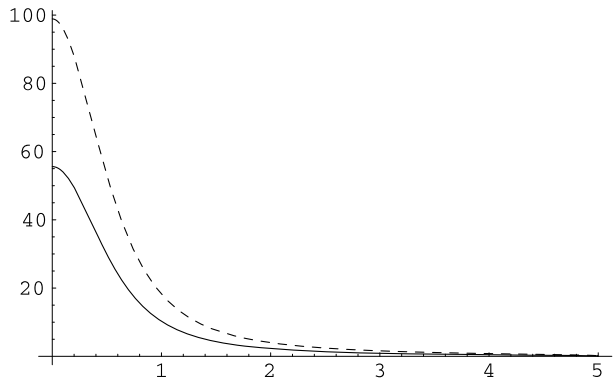

Figure 2. Temperature profiles for variable collision time (close to equilibrium – solid line), (far from equilibrium – dashed line) *versus* r.

4. Concluding remarks

We have investigated the physical viability of the *BCD* model within the framework of extended irreversible thermodynamics. We have shown that this simple model allows us greater insight into the evolution of the temperature for different collision times. More importantly, we were able to confirm earlier findings that the causal temperature dominates the Eckart temperature within the stellar core, even for variable collision time. As pointed out in earlier treatments, the constant collision time approximation is only valid for a limited period of the stellar evolution (Naidu *et al.* 2006). One expects that the collision time between the particles making up the stellar fluid will change with temperature. Such effects on the evolution of the temperature profiles were clearly demonstrated with the variable collision time solution. It must be pointed out that the truncation of the transport equations leads naturally to an implicitly defined temperature law (Govender & Govinder 2001). Such a temperature law may only be valid for a limited period of collapse. What remains is to investigate the behaviour of the temperature by employing the full transport equation for the heat flux as well as to include the effects of shear. The general framework for such an investigation has recently been provided in Herrera & Santos (2004).

References

Banerjee, A., Chatterjee, S., Dadhich, N. 2002, *Mod. Phys. Lett. A*, **35**, 2335.
Bonnor, W. B., de Oliveira, A. K., Santos, N. O. 1989, *Phys. Rep.*, **181**, 269.
Govender, M., Maharaj, S. D., Maartens, R. 1998, *Class. Quantum Grav.*, **15**, 323.
Govender, M., Govinder, K. S. 2001, *Gen. Rel. and Grav.*, **33**, 2015.
Harada, T., Iguchi, H., Nakao, K. 1998, *Phys. Rev.*, **D58**, 041502-1.
Herrera, L., Santos, N. O. 1997a, *Phys. Rep.*, **286**, 53.
Herrera, L., Santos, N. O. 1997b, *MNRAS*, **287**, 161.
Herrera, L., Santos, N. O. 2004, *Phys. Rev.*, **D70**, 084004.
Joshi, P. S., Dadhich, N., Maartens, R. 2002, *Phys. Rev.*, **D65**, 101501.
Kudoh, H., Harada, T., Iguchi, H. 2000, *Phys. Rev.*, **D62**, 104016-1.
Maharaj, S. D., Govender, M. 2005, *IJMP*, **D14**, 667.
Naidu, N. F., Govender, M., Govinder, K. S. 2006, *IJMP*, **D**, 1053.
Vaidya, P. C. 1951, *Proc. Indian Acad. Sci.*, **A33**, 264.

J. Astrophys. Astr. (2007) **28,** 175–184

Search for Orbital Motion of the Pulsar 4U 1626-67: Candidate for a Neutron Star with a Supernova Fall-back Accretion Disk

Chetana Jain[1,*], Biswajit Paul[2], Kaustubh Joshi[3], Anjan Dutta[1] & Harsha Raichur[2,4]

[1]*Department of Physics and Astrophysics, University of Delhi, Delhi 110 007, India.*
[2]*Raman Research Institute, Sadashivnagar, C. V. Raman Avenue, Bangalore 560 080, India.*
[3]*Department of Electrical Engineering, Indian Institute of Technology – Bombay, Powai, Mumbai 400 076, India.*
[4]*Joint Astronomy Programme, Indian Institute of Science, Bangalore 560 012, India.*
[*]*e-mail: chetanajain11@yahoo.co.in*

Received 2007 June 1; revised 2008 January 31; accepted 2008 February 1

Abstract. We report here results from a new search for orbital motion of the accretion powered X-ray pulsar 4U 1626-67 using two different analysis techniques. X-ray light curve obtained with the Proportional Counter Array of the Rossi X-ray Timing Explorer during a long observation carried out in February 1996, was used in this work. The spin period and the local period derivative were first determined from the broad 2–60 keV energy band light curve and these were used for all subsequent timing analysis. In the first technique, the orbital phase dependent pulse arrival times were determined for different trial orbital periods in the range of 500 to 10,000 s. We have determined a 3σ upper limit of 13 lt-ms on the projected semi-major axis of the orbit of the neutron star for most of the orbital period range, while in some narrow orbital period ranges, covering about 10% of the total orbital period range, it is 20 lt-ms. In the second method, we have measured the pulse arrival times at intervals of 100 s over the entire duration of the observation. The pulse arrival time data were used to put an upper limit on any periodic arrival time delay using the Lomb–Scargle periodogram. We have obtained a similar upper limit of 10 lt-ms using the second method over the orbital period range of 500–10,000 s. This puts very stringent upper limits for the mass of the compact object except for the unlikely case of a complete face-on orientation of the binary system with respect to our line-of-sight. In the light of this measurement and the earlier reports, we discuss the possibility of this system being a neutron star with a supernovae fall-back accretion disk.

Key words. Accretion, accretion disks—binaries: general—stars: individual (4U 1626-67)—stars: neutron—X-rays: stars.

1. Introduction

4U 1626-67 is a very well studied accretion powered X-ray pulsar. The X-ray pulsations were discovered by Rappaport *et al.* (1977) during SAS3 observations. The presence of an accretion disk around the high magnetic field neutron star in this system is very well established from X-ray spectroscopic measurements and also from the detection of quasi-periodic oscillations in X-rays (Shinoda *et al.* 1990; Kaur *et al.* 2008). The X-ray spectrum of 4U 1626-67 was studied in detail first with the ASCA spectrometers (Angelini *et al.* 1995) and later with the High Energy Transmission Grating instrument of Chandra (Schulz *et al.* 2001; Krauss *et al.* 2007). The spectrum shows clear signatures of irradiation of the accretion disk by the central X-ray source. The spectrum is rich in emission lines and also shows Doppler and thermal broadening of the line features. The neutron star has a magnetic field strength of a few times 10^{12} Gauss, determined from the cyclotron resonance absorption features in the hard X-ray band (Orlandini *et al.* 1998). In addition to the spectral signatures, the optical emission also shows reprocessing of the X-ray pulses. The Fourier power spectrum of the optical light curve of 4U 1626-67 clearly shows the reprocessed X-ray pulses at 130.38 mHz (Middleditch *et al.* 1981; Chakrabarty *et al.* 2001). Additionally, the pulse peak in the power spectrum has multiple sidebands, the main ones are separated by about 0.4 mHz. If the sidebands in the power spectrum are taken to be signatures of the binary companion, this indicates an orbital period of about 2485 ± 60 s. Based on these studies, it is generally felt that the neutron star most probably, has a low mass binary companion and a small orbital period of about 42 minutes. However, orbital motion of the neutron star (usually observed in binary pulsars in the form of Doppler shift in the pulse period, or delay in the pulse arrival time) has never been detected in 4U 1626-67. Using a long X-ray observation with the Medium Energy (ME) detectors of the EXOSAT observatory, Levine *et al.* (1988) determined an upper limit of 10 lt-ms for $a_x \sin(i)$, which is the most stringent upper limit reported so far for this object. In the present work, we have used the Lomb–Scargle periodogram technique for detection of a periodic signal in an unevenly sampled time series with RXTE–PCA. Compared to the Fourier transform, Lomb–Scargle method allows us to search for a periodic signal with the desired resolution in orbital period.

We have analyzed data of this pulsar from a very long observation carried out in February 1996, with the RXTE–PCA. We have used two different pulse arrival time analysis techniques to search for any possible binary motion of the neutron star. Results obtained from this long RXTE–PCA observation are presented here.

2. Observations and analysis

The source 4U 1626-67 was observed with RXTE from February 10, 1996 to February 14, 1996 in three segments with useful exposures of 95561 s, 63393 s and 50255 s respectively. We have used light curves obtained with the Proportional Counter Array (PCA) of the RXTE, which has a total photon collection area of 6500 cm^2 (Jahoda *et al.* 1996). This dataset is well suited for the search of orbital motion using the pulse arrival time measurements around the proposed orbital period of about 42 minutes. Light curves were extracted from data stored in the event mode, which has the highest energy and timing information. For data screening, the filtering criteria applied were: all the five Proportional Counter Units (PCU) were ON, time since South Atlantic Anomaly

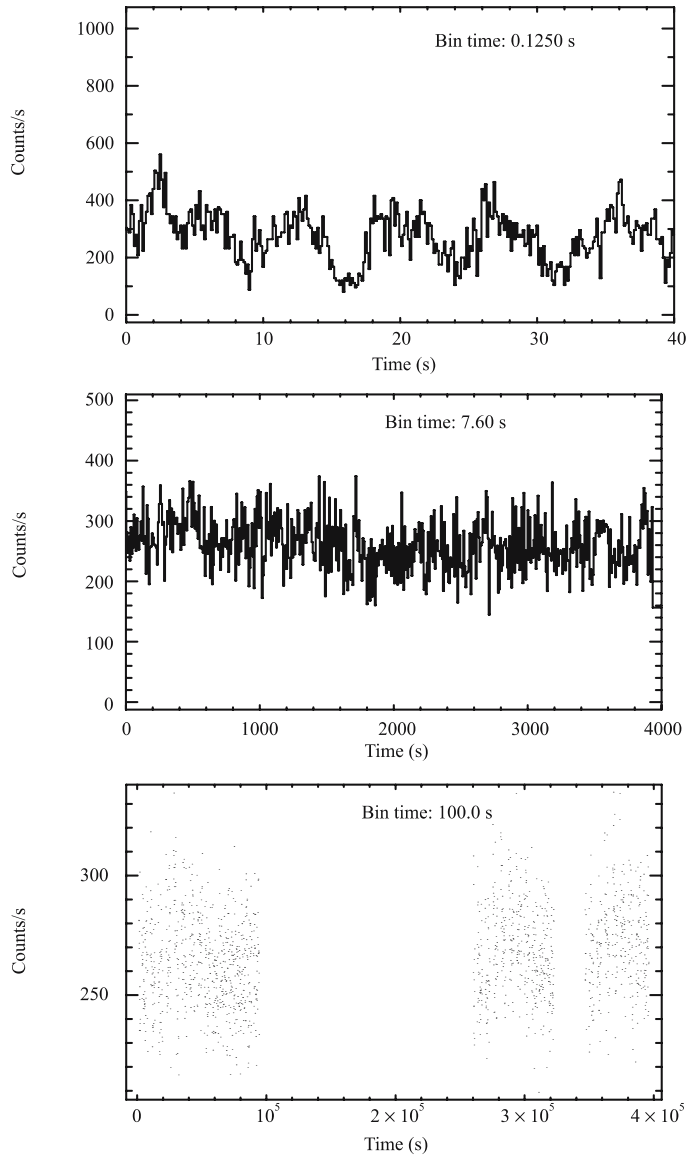

Figure 1. 2–60 keV RXTE/PCA light curves with different bin sizes for 4U 1626-67. The bottom panel shows the light curve obtained from the entire light curve. The upper and middle panels show a section of the light curve.

(SAA) was greater than 10 minutes, pointing offset was less than 0.02 degrees, and Earth elevation angle was greater than 10 degrees.

The 2–60 keV light curve obtained from the entire observation is shown in the bottom panel of Fig. 1 with a time resolution of 100 s. In the two other panels of the same figure, sections of the light curve are shown with a time resolution of 0.125 s and 7.6 s respectively. For all analysis described here, the photon arrival times were first converted to the solar system barycenter. The pulsar 4U 1626-67 shows monotonic

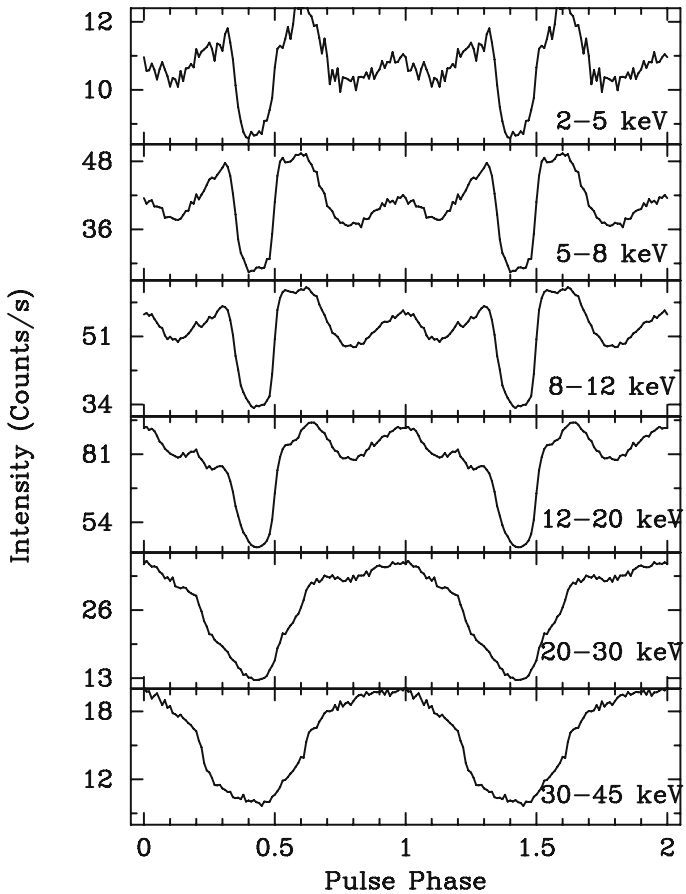

Figure 2. Energy-dependent pulse profiles of 4U 1626-67 covering an energy range of 2–45 keV. The highest signal-to-noise ratio is observed for 5–45 keV energy band.

spin-up or spin-down for long periods (Chakrabarty *et al.* 1997). To establish the pulse period and the period derivative during the present observation, we have carried out the pulse folding and χ^2 maximization analysis assuming different trial period derivatives (Paul *et al.* 2001). From this analysis we have determined a spin period of 7.66735 s (MJD = 50123.00) and a period derivative of 3.83×10^{-11} s s^{-1}. The period derivative derived here was used in all further analyses described below. These values are consistent with the same reported in Chakrabarty *et al.* (1997), derived as a long term solution of pulse period evolution of 4U 1626-67. The pulse profiles in different energy bands are shown in Fig. 2. The pulse profile of 4U 1626-67 is known to be energy dependent (Levine *et al.* 1988); bi-horned profile in 2–15 keV range and sinusoidal at higher energies. As shown in Fig. 2, at energies below 15 keV, we also found a sharp dip in the pulse profile in addition to bi-horned nature. From the Chandra-HETGS and XMM-RGS observations, Krauss *et al.* (2007), have also reported a dip in the pulse profile, after the 1990 torque reversal. The light curve has poor statistics below 5 keV and there is a phase shift in the pulse profile above 10 keV. This source has a very hard spectrum, therefore, although the quantum efficiency of the PCA detectors drops

rapidly after about 25 keV, we have observed the highest signal-to-noise ratio in the pulse profile for 5–45 keV energy band. After examining the χ^2 obtained from fitting a constant to the pulse profiles of various energy bands, we chose this wide energy band for pulse arrival time analysis.

2.1 *Pulse arrival time analysis with trial orbital periods*

To determine the orbital parameters of the system 4U 1626-67, we first searched for pulse arrival time delays as a function of the orbital period. To reduce the error of individual pulse arrival time measurements, we have added the pulse profiles coherently from identical orbital phases with trial orbital periods in the range of 500–10,000 s. The consecutive trial periods were chosen in a way to ensure that when data from this long RXTE observation are folded, the orbital phase error at the beginning and at the end is not more than 5%. The entire observation duration was divided into segments of length equal to a trial orbital period. Each segment was further broken into 30 equal orbital phases. The ensemble of the corresponding phases from each orbital segment was then folded with the spin period and period derivative mentioned above and pulse profiles were created with 512 phase bins. A template profile was also created using the entire dataset. The advantage of using this template profile instead of any one of the phase specific profiles is better statistical accuracy. Finally we cross-correlated these orbital phase resolved pulse profiles with the template profile to determine the pulse arrival time delay. This process was repeated for the large number of trial orbital periods from 500–10,000 s in a fashion to avoid more than 5% error in orbital phase.

If any of the trial orbital period is close to the actual orbital period of the system, the plot of arrival times against the phase will show the pulse arrival time delay for a general elliptical orbit. The amplitude of the pulse arrival time delay should give us the projection of semi-major axis onto the line of sight $a_x \sin(i)$. We have also determined the orbital phase resolved pulse arrival times for 10 arbitrary trial orbital periods spread over the entire period range covered in the analysis. The average standard deviation of the 30 pulse arrival times for these arbitrary trial periods is found to be 0.002 s, which was used as an error in the pulse arrival time measurement for all the other trial orbital periods. Subsequently, the relative pulse arrival times determined at 30 different orbital phases were fitted to a constant value. The χ^2 derived from the constant fit for different trial orbital periods is shown in Fig. 3. It was observed that for most of the cases, the χ^2 was less than 40. But for some trial orbital periods, the χ^2 was in the range of 40–70 for 29 degrees of freedom. We examined the pulse arrival time curve for these periods and also energy resolved data. From the nature of the pulse arrival times as a function of the trial orbital phase, it is evident that the high χ^2 is due to random variation of pulse arrival times. The 1σ and 3σ upper limits of $a_x \sin(i)$, determined from the pulse arrival time analysis from this RXTE-PCA light curve of 4U 1626-67 are shown in Fig. 4 for a trial orbital period range of 500–10,000 s. As can be seen from the figure, we determined a 3σ upper limit of 13 lt-ms on the projected semi-major axis of the orbit of the neutron star for most of the orbital period range, while in some narrow orbital period ranges, covering about 10% of the total orbital period range, it is 20 lt-ms.

2.2 *Pulse arrival time analysis*

In this method, we searched for the presence of an orbital motion using the periodogram technique. We created continuous pulse profiles for about 100 s (13 pulses) in the

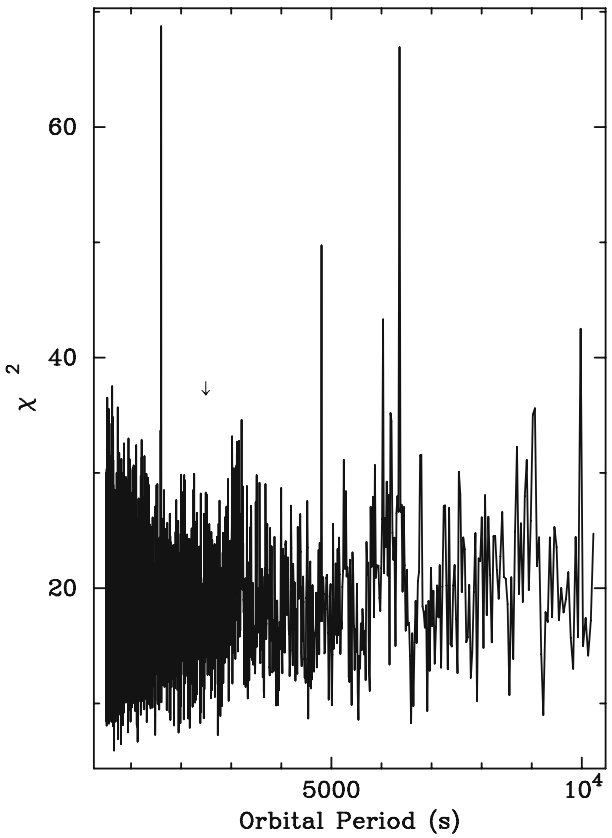

Figure 3. The χ^2 values obtained after fitting a constant to the pulse arrival times at 30 different phases of the orbital period is shown here for trial orbital periods in the range of 500–10,000 s. The arrow marks the orbital period (2485 s) as inferred from the optical observations.

energy range of 5–45 keV. These pulse profiles were then cross-correlated with a master template pulse profile created using the entire stretch of observation. The peaks of the cross-correlation curves were used to determine the pulse arrival times in each interval of 100 s with respect to the phase coherent solution described by a spin period of 7.66735 s and a period derivative of 3.83×10^{-11} s s^{-1}. We searched for a periodic delay in the pulse arrival times using the Lomb–Scargle technique. The normalized Lomb–Scargle periodogram is shown in Fig. 5. There is an absence of a peak at any frequency and hence it rules out the possibility of any orbital motion. The frequency resolution chosen for the analysis was such that the total orbital phase offset is less than 1% of the trial orbital period, which is much less as compared to the earlier work. Figure 6 is a histogram of the power distribution made from the normalized periodogram. It follows an exponential distribution and a power of 13 or above in the simulated dataset would therefore indicate a signal detection. To set an upper limit on the projected semi-major axis of the orbit of the neutron star, we added sinusoidal signals of different amplitudes for periods of 500 s, 1250 s, 2500 s, 5000 s and 10,000 s to the entire data. As shown in Fig. 5, peaks were observed at the signal frequency for a signal amplitude of 10 ms. There was no peak for an added sinusoidal signal of

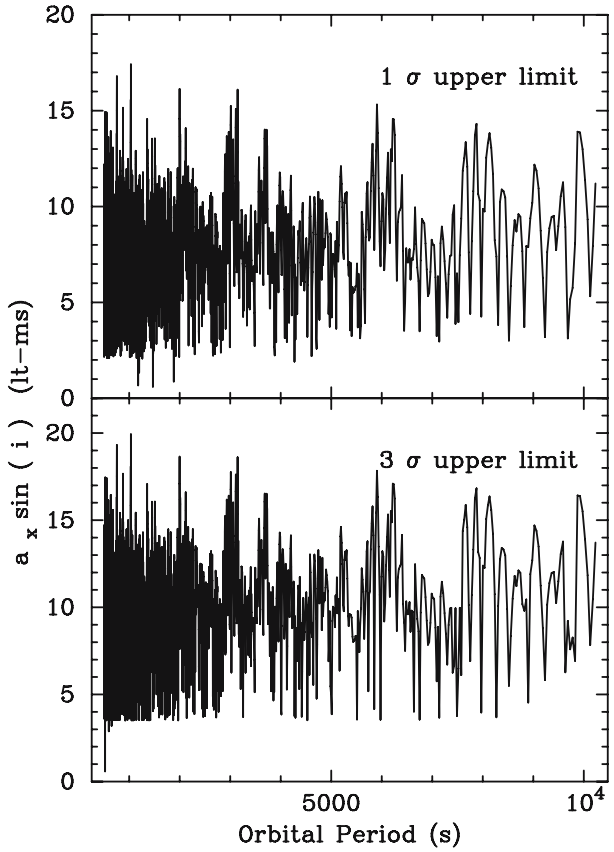

Figure 4. Upper limit of $a_x \sin(i)$ (in lt-ms) for different trial orbital periods. The top panel shows the 1σ upper limit of $a_x \sin(i)$ while the bottom panel shows the 3σ upper limits of $a_x \sin(i)$ for an orbital period range of 500–10,000 s.

amplitude less than 10 ms. This sets an upper limit of 10 lt-ms on the projected semi-major axis of the orbit of the neutron star.

3. Discussions

The new upper limit of 10 lt-ms for $a_x \sin(i)$ of the persistent X-ray pulsar 4U 1626-67 gives a very small upper limit of $2.85 \times 10^{-6}\,M_\odot$ for the mass function. This puts a strong constraint on the companion mass.

The blue and red Doppler shifted emission lines of Ne and O detected in the X-ray spectrum of 4U 1626-67 indicate that the lines originate in the accretion disk and there has been no significant variation in the widths of the emission lines (Krauss *et al.* 2007). The disk velocity component along our line-of-sight has a magnitude of a few hundred km s^{-1}. Therefore, the disk axis cannot be very close to our line-of-sight. One can assume that the orbital plane and the accretion disk plane are co-aligned. Otherwise, the forced precession of the accretion should result into a modulation of the X-ray emission as a super-orbital period, which is not seen in the long term X-ray

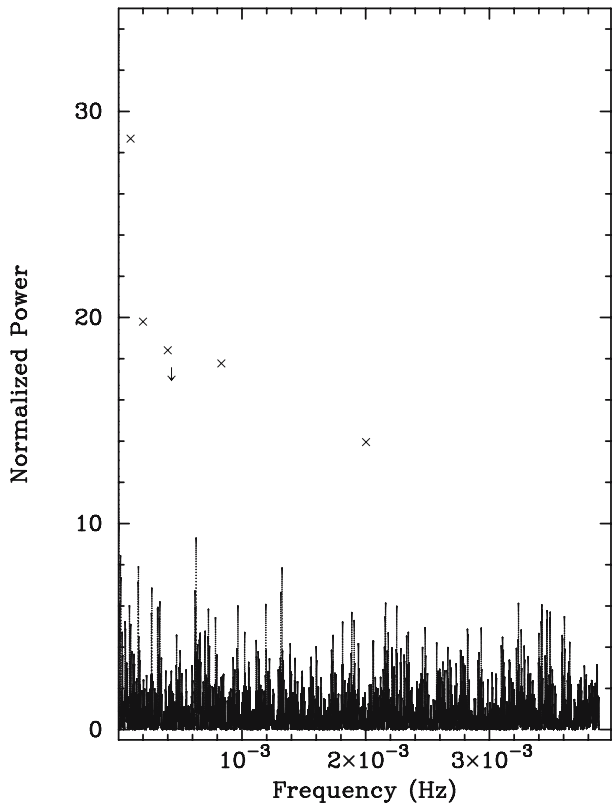

Figure 5. The normalized Lomb–Scargle periodogram for the pulse arrival time series of 4U 1626-67. Absence of a clear peak rules out the possibility of the presence of an orbital motion of the neutron star. With the addition of sinusoidal signals (of time period 500 s, 1200 s, 2500 s, 5000 s and 10,000 s) of amplitude 10 ms, peaks were observed at the corresponding frequencies. The crosses in the figure indicate the peak power at those frequencies. The arrow marks the orbital period (2485 s) as inferred from the optical observations.

light-curve of 4U 1626-67. The non-detection of orbital motion from pulse arrival time delay therefore leaves scope for only a very small mass companion star.

In view of the results presented here, we point out that 4U 1626-67 is a strong candidate for a neutron star with a supernova fall-back accretion disk. Possibility of a supernova fall back accretion disk or fossil disk has been considered in many different contexts before. The possibility of supernova fall back accretion disk around neutron stars and black holes has also been found from modeling of SN explosion (Perna *et al.* 2000; Nomoto *et al.* 2006). Based on this model, Lin *et al.* (1991) explained the formation of planets around the radio pulsars. A unified picture for dim isolated thermal neutron stars, anomalous X-ray pulsars and soft gamma-ray repeaters can also be obtained from this model (Alpar *et al.* 2001). The spin down of anomalous X-ray pulsars were also considered in the context of fall-back disk model (Chatterjee *et al.* 2000). All neutron stars can acquire debris through fall-back but the subsequent evolution of neutron star depends on the interaction between the stellar magnetic field and the disk flow (Chatterjee *et al.* 2000). Presence of a fall-back disk has been

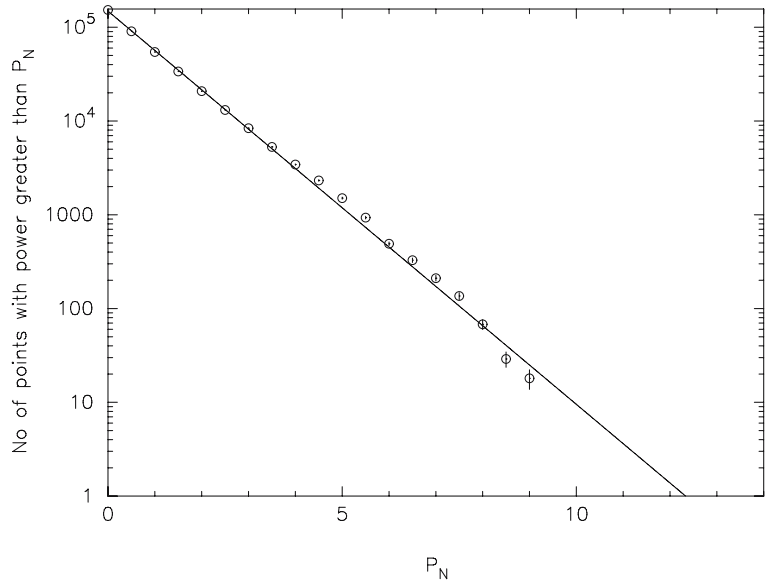

Figure 6. A plot of the normalized power distribution over all the frequencies in the Lomb–Scargle periodogram. The distribution is found to be exponential with highest power not exceeding 10.

proposed by Wang *et al.* (2006), around the anomalous X-ray pulsar 4U 0142 + 61. However, Ertan *et al.* (2007) have suggested that the optical and near IR observations are consistent with the gaseous disk model. For a spin period of 7.66 s and magnetic field strength of 10^{12} G, a fossil accretion disk can feed the neutron star for a long period of time (Wang *et al.* 2006). If a significant fraction of the compact objects in core collapse supernovae have fall-back accretion disks, it influences the chemical evolution history of the Universe. However, since 4U 1626-67 seems to be only the second candidate for such a system, this does not give us any clue about what fraction of the neutron stars may have such a disk, how long the disks last, what fraction of the heavy metal falls back, etc. In this context, we also point out here that the accretion disk of the pulsar 4U 1626-67 is indeed metal-rich.

Acknowledgement

We thank the RXTE team for making the data available and an anonymous referee for valuable comments on the manuscript.

References

Alpar, M. A. 2001, *ApJ*, **554**, 1245.
Angelini, L., White, N. E., Nagase, F., Kallman, T. R., Yoshida, A., Takeshima, T., Becker, C., Paerels, F. 1995, *ApJ*, **449**, L41.
Chakrabarty, D. *et al.* 1997, *ApJ*, **474**, 414.
Chakrabarty, D., Homer, L., Charles, P. A., O'Donoghue, D. 2001, *ApJ*, **562**, 985C.
Chatterjee, P., Hernquist, L., Narayan, R. 2000, *ApJ*, **534**, 373.

Ertan, U., Erkut, M. H., Eksi, K. Y., Alpar, M. A. 2007, *ApJ*, **657**, 441.
Jahoda, K., Swank, J. H., Giles, A. B., Stark, M. J., Strohmayor, T., Zhang, W., Morgan, E. H. 1996, *Proc. SPIE*, **2808**, 59.
Kaur, R., Paul, B., Kumar, B., Sagar, R. 2008, *ApJ* (in press) astro-ph/0801.1761.
Krauss, M. I., Schulz, N. S., Chakrabarty, D., Juett, A. M., Cottam, J. 2007, *ApJ*, **660**, 605.
Levine, A., Ma, C. P., McClintock, J., Rappaport, S., Van Der Klis, M., Verbunt, F. 1988, *ApJ*, **327**, 732.
Lin, D. N. C., Woosley, S. E., Bodenheimer, P. H. 1991, *Nature*, **353**, 827.
Middleditch, J., Mason, K., Nelson, J., White, N. 1981, *ApJ*, **244**, 1001.
Nomoto, K., Tominaga, N., Umeda, H., Kobayashi, C., Maeda, K. 2006, *Nucl. Phys. A.*, **777**, 424.
Orlandini, M., Fiume, D., Frontera, F., del Sordo, S., Piraino, S., Santangelo, A., Segreto, A., Oosterbroek, T., Parmar, A. N. 1998, *ApJ*, **500**, L163.
Paul, B., Agrawal, P. C., Mukerjee, K., Rao, A. R., Seetha, S., Kasturirangan, K. 2001, *A& A*, **370**, 529.
Rappaport, S., Markert, T., Li, F., Clark, G., Jernigan, J., McClintock, J. 1977, *ApJ*, **217**, 29.
Perna, R., Hernquist, L., Narayan, R. 2000, *ApJ*, **541**, 344P.
Schulz, N. S., Chakrabarty, D., Marshall, H. L., Canizares, C. R., Lee, J. C., Houck, J. 2001, *ApJ*, **563**, 941S.
Shinoda, K., Kii, T., Mitsuda, K., Nagase, F., Tanaka, Y., Makishima, K., Shibazaki, N. 1990, *PASJ*, **42**, L27.
Wang, Z., Chakrabarty, D., Kaplan, D. L. 2006, *Nature*, **440**, 772.

J. Astrophys. Astr. (2007) **28,** 185–196

Pulse Phase Dependence of the Magnetar Bursts

Chetana Jain[1,*], Anjan Dutta[1] & Biswajit Paul[2]
[1]*Department of Physics and Astrophysics, University of Delhi, Delhi 110 007, India.*
[2]*Raman Research Institute, Sadashivnagar, C. V. Raman Avenue, Bangalore 560 080, India.*
[*]*e-mail: chetanajain11@yahoo.co.in*

Received 2007 October 3; revised 2008 February 6; accepted 2008 February 8

Abstract. We report here results from a study of X-ray bursts from 3 magnetar candidates (SGR 1806–20, SGR 1900+14 and AXP 1E 2259+586). We have searched for a pulse phase dependence of the X-ray burst rate from these sources. X-ray light curves were obtained with the Proportional Counter Array on-board the Rossi X-ray Timing Explorer during the periods of intense burst activity in these sources. On detailed analysis of the three sources, we found a very significant burst rate for all pulsar phases. However, some locations appear to produce bursts slightly more often, rendering the non-isotropic distribution. Only in the case of SGR 1900+14, there is a clear pulse phase dependence of burst rate.

Key words. Gamma rays: bursts—neutron stars, magnetars: individual (SGR 1806–20, SGR 1900+14 and AXP 1E 2259+586)—X-rays: bursts.

1. Introduction

Soft Gamma-ray Repeaters (SGRs) are a small group of neutron stars with repetitive bursts of soft gamma rays at random intervals. These bursts are of a short duration ~ 0.1 s (Kouveliotou 1995) and have a peak luminosity of $\sim 10^{41}$ erg s^{-1} (Hurley 2000) – well above the standard Eddington limit of $\sim 2 \times 10^{38}$ erg s^{-1} for a $1.4\,M_{\odot}$ neutron star. Other than bursts, these sources show quite stable pulsations though the pulse profiles and pulsed and unpulsed X-ray flux are known to vary with time. Enhancement in pulsed and persistent flux during the burst active phase of the SGRs is believed to result from the back heating of the outer crust by the burst emission in the magnetosphere. Burst repetition timescales have been observed to vary from seconds to years. Three giant flares were observed from SGRs 0526–66 (Mazets *et al.* 1999), 1900+14 (Hurley *et al.* 1999) and 1806–20 (Palmer *et al.* 2005). In SGR 1900+14, the quiescent pulse profile changed from a complex multi-peaked morphology before the August 27, 1998 giant flare to a sinusoidal profile after the burst active episode (Woods *et al.* 2001; Gogus *et al.* 2002). The pulse profile of the SGR 1806–20 became more complex after the December 27, 2004 giant flare (Palmer *et al.* 2005). These X-ray pulsars exhibit a rapid spin down which can be attributed to the magnetic braking of a strongly magnetized neutron star (with magnetic fields $\sim 10^{14}$–10^{15} G), also referred to as 'Magnetars' (Kouveliotou *et al.* 1998). The magnetar theory was proposed (Duncan & Thompson 1992; Thompson & Duncan 1995) to explain the SGR

burst activity and the quiescent X-ray emission. These sources have properties similar to another group of neutron stars known as Anomalous X-ray Pulsars (AXPs) which have also been suggested to be magnetars (Thompson & Duncan 1996). The statistical properties of the bursts in SGR 1806–20 (Gogus *et al.* 2000) and SGR 1900+14 (Gogus *et al.* 1999) imply that the SGR bursts resemble earthquakes and solar flares. SGRs have X-ray spectra comparable to that of AXPs when in quiescence. Both AXPs and SGRs have pulse periods ranging from 5–12 s. The pulsed X-ray luminosities of AXPs lie in the range $\sim 10^{33}$–10^{35} erg s^{-1}. It was also proposed (Chatterjee *et al.* 2000; Corbet *et al.* 1995; van Paradijs *et al.* 1995) that they are powered by accretion of fall back material from a supernova. The absence of Doppler shift (or pulse arrival time delay) of the X-ray pulses down to a few millisecond (Mereghetti *et al.* 1998) and the very faint optical counterparts (Hulleman *et al.* 2000a, 2000b; Israel *et al.* 2002) rule out the presence of a binary companion. Most AXPs and SGRs have X-ray spectra that can be modeled by thermal emission of kT $\sim$0.5 keV along with a power law tail (Woods *et al.* 1999; Paul *et al.* 2000; Kouveliotou *et al.* 2001). At energies above 10 keV, Kuiper *et al.* (2006), have reported a very hard component in the X-ray spectra of the AXPs. Similar hard spectra have also been detected with INTEGRAL for SGRs (Gotz *et al.* 2006; Mereghetti *et al.* 2005; Molkov *et al.* 2005). However, this very hard X-ray component will not be addressed in this paper.

SGR 1900+14 became extremely active in May, 1998 after a long period of quiescence. A giant flare was recorded on 1998, August 27 (Hurley *et al.* 1999; Feroci *et al.* 2001; Mazets *et al.* 1999). The pulsed flux increased by a factor of ~ 2 above its pre-burst level and the total energy exceeded 10^{44} erg. A total of 200 events were detected with the Burst and Transient Source Experiment (BATSE) on-board the Compton Gamma Ray Observatory (CGRO) during May 1998–January 1999 (Woods *et al.* 1999). PCA on-board RXTE recorded $\sim$800 bursts between 1998, May 31 and December 21. After a long quiescence of almost 2 years, a burst was detected on 2001, April 18. The energy released during this intermediate flare was much less than the 1998 flare (Kouveliotou *et al.* 2001). The recurrence time distribution of these bursts is characterized by a log–normal function which peaks at $\sim$49 s (Gogus *et al.* 1999). A positive correlation exists between the burst intensity and the waiting times till the next burst. The burst duration and the energy also show a definite relation. Palmer *et al.* (2000) had reported an isotropic distribution of the bursts in SGR 1900+14 using a small fraction of the RXTE observations.

The burst activity of SGR 1806–20 increased during November, 1996 and bursts continued to occur occasionally, unlike the burst activity in SGR 1900+14. SGR 1806–20 entered a phase of increased burst activity in May, 2004 that persisted for ~ 1 year. A giant flare was observed during this period on 2004, December 27 (Hurley *et al.* 2005; Merghetti *et al.* 2005; Palmer *et al.* 2005). This burst had a peak luminosity of $\sim 2 \times 10^{47}$ erg s^{-1}, a total energy of $\sim 5 \times 10^{46}$ erg and a duration of $\sim$5 minutes. More than 300 bursts were recorded from the all-sky instruments within the Interplanetary Network (IPN). The SGR 1900+14 bursts are consistent with a power law index of 1.66 while the RXTE bursts of SGR 1806–20 have an index of 1.43. These bursts resemble a self-organized critical system (Bak *et al.* 1998) which if perturbed from the critical state, can cause a chain reaction in the system. SGRs are believed to be strained by the evolving magnetic stresses (Thompson & Duncan 1995). Palmer (1999) reported an isotropic distribution as a function of pulsar phase of bursts in SGR 1806–20 based on 33 bursts detected with the International Cometary Explorer.

AXP 1E 2259+586 is the second AXP after 1E 1048.1–5937 which experienced a major outburst on June 18, 2002. However, very few X-ray bursts were detected from the AXP 1E 1048.1–5937 (Gavriil *et al.* 2002) and therefore the data available from this source are not sufficient for a pulse phase dependence analysis. More than 80 SGR-like bursts were detected from 1E 2259+586 with the RXTE-PCA during a brief observation spanning over 14 ks (Kaspi *et al.* 2003; Woods *et al.* 2004). A significant change in the pulse morphology was observed during and following the outburst. The temporal and statistical properties of these bursts (Gavriil *et al.* 2004) were quite similar to those observed in SGRs. Gavriil *et al.* (2004) claimed detection of a pulse phase dependence of the bursts in 1E 2259+586 with a low statistical significance.

In the present work, we have done a detailed analysis to search for a pulse phase dependence of the bursts from the three magnetar candidates (SGR 1806–20, SGR 1900+14 and AXP 1E 2259+586). The surface of a neutron star may be weaker at some areas which can lead to volcanic sites at these places. In such a scenario, the detection of the bursts is expected to be dependent on the pulse phase.

2. Observations and analysis

We have used archival data from observations of SGR 1806–20, SGR 1900+14 and the magnetar candidate 1E 2259+586 using the Proportional Counter Array (PCA) on-board the Rossi X-ray Timing Explorer (RXTE). The PCA consists of an array of five collimated xenon/methane multi-anode proportional counter units (PCUs) and with a total photon collection area of 6500 cm^2 (Jahoda *et al.* 1996). Data used in the analysis were taken from the event mode and the Good-Xenon-with-Propane mode which records the photon arrival times with 1 μs resolution.

A large number of X-ray bursts were detected from SGR 1806–20 at different times (Gogus *et al.* 2002; Hurley *et al.* 2005; Israel *et al.* 2005; Strohmayer *et al.* 2006). For the present work, we have analyzed data taken from the observations made during 1996, November 5–18 and 2004, May 24 – November 22. During this period, there was a long exposure with RXTE-PCA and the pulse profile was detected with a good signal-to-noise ratio. The 2–10 keV light curve was obtained with a time resolution of 0.03125 s. The total exposure for the 1996 observation was 146.5 ks and for the 2004 observation, the total exposure was 486.6 ks. The source has a variable spin down history. Therefore, from the series of observations in 2004 spread over a period of $\sim$6 months, it was not possible to have a consistent pulse phase connected solution for period evolution for the entire stretch of data. Woods *et al.* (2007) have carried out extensive analysis to obtain the correct spin parameters of SGR 1806–20 during the 2004 observations. We divided the data into small segments of $\sim$10 ks and measured the local spin periods and period derivatives which were found to be consistent with the reported values (Woods *et al.* 2007). For all subsequent analysis of the 2004 dataset, we have used the published pulse ephemerides by Woods *et al.* (2007). To establish the pulse period and the period derivative for the 1996 observation, we carried out pulse folding and χ^2 maximization analysis assuming different trial period derivatives (see Paul *et al.* 2001 for details of the method). The pulse period and the period derivative thus obtained (table 1) was used in further analysis from the 1996 observations.

The source SGR 1900+14 was observed with the RXTE-PCA from August 28 – October 08, 1998 and from April 19 – 30, 2001 during the two active periods of this SGR. The 2–10 keV light curve was created from the PCA data with a time resolution

Table 1. Pulse frequency ephemerides for SGR 1806–20 for the 1996 and 2004 observations, SGR 1900+14 for observations between August 28 and October 08, 1998 and between April 04 and 30, 2001 and for the June 18, 2002 outburst of AXP 1E 2259+586.

Source	Epoch (MJD TDB)	Time range (MJD TDB)	P (s)	$\dot{P}$ (10^{-11} s s^{-1})	Total exposure (ks)
SGR 1806-20	50398	50392–50405	7.4765534	6.8	146.5
SGR 1806-20*	53153	53149–53156	7.5499048	55.8	
SGR 1806-20*	53160	53156–53163	7.5502525	58.7	
SGR 1806-20*	53167	53163–53170	7.5506146	61.0	
SGR 1806-20*	53174	53170–53177	7.5510279	75.2	
SGR 1806-20*	53181	53177–53184	7.5514493	63.8	
SGR 1806-20*	53230	53226–53233	7.5545127	81.0	
SGR 1806-20*	53237	53233–53240	7.5549819	74.2	
SGR 1806-20*	53244	53240–53247	7.5553889	60.5	
SGR 1806-20*	53251	53247–53254	7.5557577	61.0	
SGR 1806-20*	53258	53254–53261	7.5561191	58.2	
SGR 1900+14	51070	50879–51094	5.16026572	5.93	142.2
SGR 1900+14	52021	52019–52030	5.1728219	15.8	106
1E 2259+586	52443	52443–52443	6.9785984	0.047855	10.8

* P and P: As given in Woods *et al.* (2007).

of 0.03125 s. The total exposure for the 1998 observation was 142.2 ks and for the 2001 observation, it was 106 ks. The pulse ephemeris obtained after the period search and χ^2 maximization is shown in Table 1.

The anomalous X-ray pulsar 1E 2259+586 was observed on 2002, June 18, during which SGR-like bursts were detected for the first time in this source (Kaspi *et al.* 2003). We extracted 2–20 keV band light curve from the PCA data with a time resolution of 0.03125 s. The exposure time for the observation was 10.8 ks. The pulse period and the period derivative was determined from this light curve after the removal of the bursts and are shown in Table 1.

To study the pulse phase dependence of the X-ray bursts in these three sources, estimated background count rates in the same energy band were subtracted from the light curves. The photon arrival times of the background subtracted light curves were converted to the solar system barycenter. To identify the bursts, the counts in the ith time bin were compared with a windowed (one pulse cycle) local running mean spanning about four pulses around the ith time bin (Gavriil *et al.* 2004). A typical burst lasts ~ 0.1 s and a bin size of 0.03125 s was taken for the analysis. Therefore, a window of one pulse period (~ 200 bins) around the time bin under consideration is enough to account for the rise and decay of a burst. Hence, the time bins immediately around this point were excluded from contributing to the local running mean. Events for which the count rate was significantly greater than the local running mean were labeled as bursts. The light curves of the three sources were screened and all the bursts (3σ and above) detected using this method were first removed from the light curves. The light curves obtained after the removal of 125 ms of data around these 3σ bursts were folded

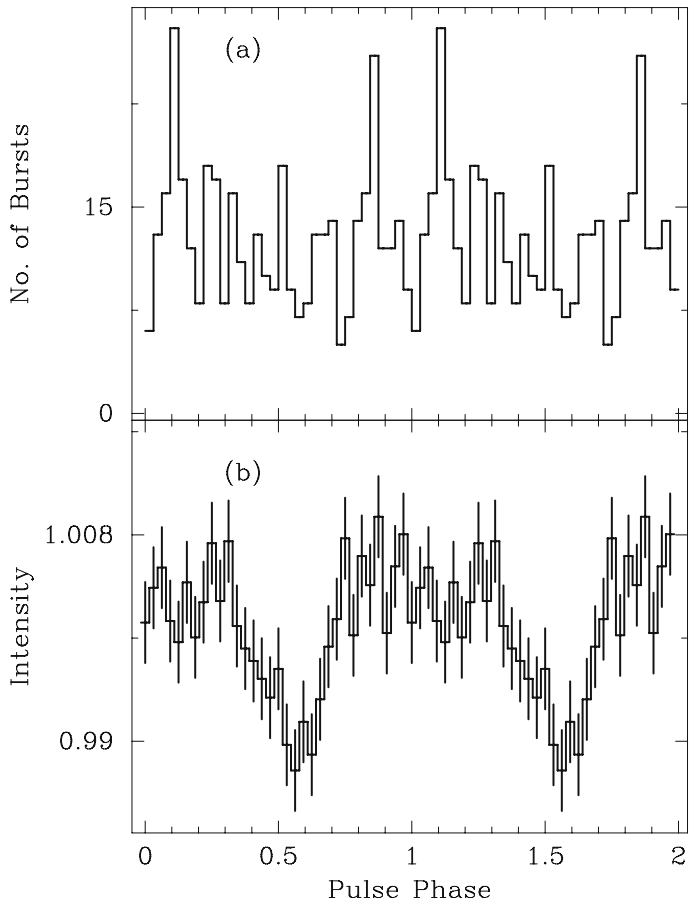

Figure 1. Result of the burst search analysis for the source SGR 1806–20 observed between November 5 and November 18, 1996. The upper panel shows the number of bursts observed for $> 3\sigma$ level of significance. The light curve obtained after the removal of 3σ bursts was folded with 32 phase bins. The bottom panel shows the normalized pulse profile of the source.

with 32 phase bins, along with the parameters listed in Table 1. The folded pulse profiles of all the three sources (SGR 1806–20, SGR 1900+14 and AXP 1E 2259+586) after the removal of all the bursts (3σ and above) are shown in Figs. 1(b), 2(b), 3(b), 4(b) and 5(b). The X-ray profiles of the magnetar candidates varied from a complex multi-peaked morphology to a simple sinusoidal variation. The 2–10 keV pulse profile of the 1996 observation of the SGR 1806–20 is dominated by a broad maximum, while the pulse profile became sinusoidal during the 2004 outburst over the same energy range. The X-ray pulse profile of SGR 1900+14 during the 1998 observation is a single-peaked sinusoidal. The pulse broadened and became more complex during the 2001 outburst. The pulse profile consists of two prominent peaks during the 2002 outburst of the anomalous X-ray pulsar 1E 2259+586. The burst times were converted to the corresponding pulse phase and histograms were made for bursts with detection significance above the 3σ level. This is shown in Figs. 1(a), 2(a), 3(a), 4(a) and 5(a) above the pulse profiles of each observation.

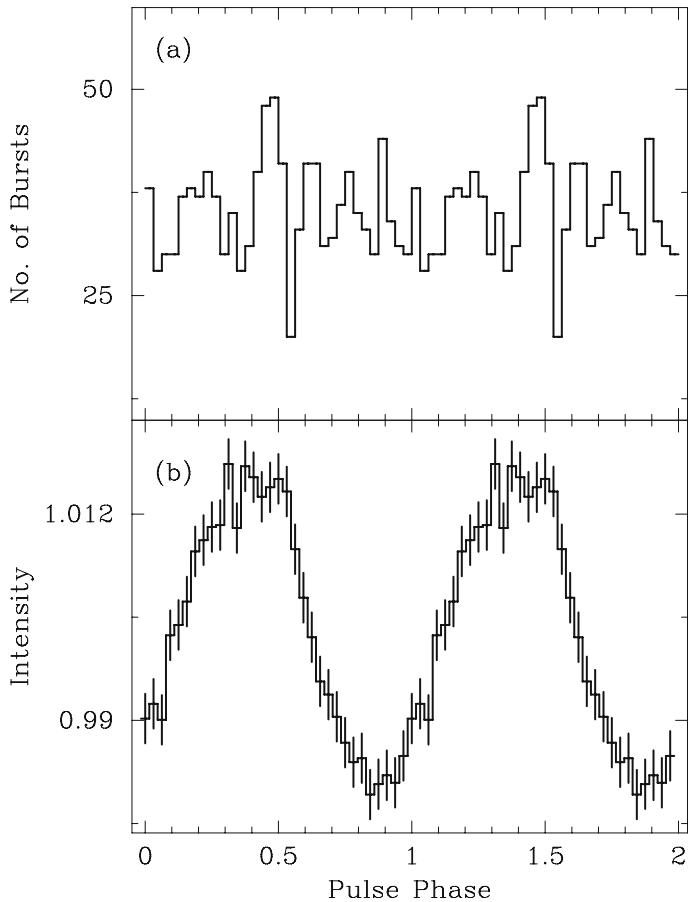

Figure 2. As in Fig. 1, this figure shows the result of the burst search analysis for the source SGR 1806–20 observed between May 24 and November 22, 2004.

As can be seen from Fig. 1, during the 1996 observation of SGR 1806–20 the burst frequency shows two spikes both falling within the peak of the pulse profile. But, fewer bursts were observed during the minimum of the pulse profile. Similarly for the 2004 observation of the same source, a spike is seen in the burst distribution that coincides with the peak of the pulse profile (Fig. 2). For both the observations, the total number of $> 3\sigma$ bursts (N) observed were quite high. During the 1996 observation, a total of $407 > 3\sigma$ bursts were observed, whereas it was 1128 during the 2004 observations (Table 2). The maximum number of $> 3\sigma$ bursts observed in an interval ($N_{\rm max}$) were 28 and 49 against an average number (N) of 12.72 and 35.25, respectively for the 1996 and 2004 observations.

From the 1998 observations of SGR 1900+14, shown in Fig. 3, it is evident that the bursts frequency shows a pulse phase dependence with an offset with respect to the pulse profile. The burst frequency at the peak of the distribution is 41 ± 1.3 per bin (calculated from the 14 points near the peak) compared to a value of 29 ± 1.5 (calculated from the 14 points near the minimum), showing a clear pulse phase dependence of the burst rate. There is a clustering of higher and lower burst rates in two phase intervals. In

Table 2. The burst statistics for the magnetar candidates SGR 1806–20 and SGR 1900+14; and the AXP 1E 2259+586.

Source	SGR 1806–20		SGR 1900+14		AXP 1E 2259+586
Year of observation	1996	2004	1998	2001	2002
Total no. of bursts (N)	407	1128	1119	86	87
Max. no. of bursts ($N_{\rm max}$) in a phase bin	28	49	50	7	8
Av. no. of bursts (N)	12.72	35.25	34.97	2.69	2.72
Standard deviation ($\sigma_{\rm real}$)	5.24	6.15	7.40	1.96	2.16
Percentage probability $P(> N_{\rm max})$ of occurrence of more than $N_{\rm max}$ bursts in 32 bins*	0.46	52.19	31.05	64.67	22.09
No. of simulated sets (out of 100,000) with $\sigma_{\rm simulated} > \sigma^{**}_{\rm real}$	33	35938	2400	7239	1458

*Assuming Poisson distribution, $P(> N_{\rm max})$ is the probability of occurrence of more than $N_{\rm max}$ events (in 32 phase bins) about an average number ($\bar{N}$) of random events.
**$\sigma_{\rm real}$ represents the standard deviation observed in the real dataset, while $\sigma_{\rm simulated}$ represents the standard deviation observed in the randomly generated datasets.

SGR 1900+14, the bursts are localized but are out of phase with the pulse profile. The detection of a significant burst rate over all pulse phases for both SGR 1806–20 and SGR 1900+14, could imply that the bursts spread over the neutron star surface after the initial outburst (Hurley *et al.* 2002). In the 1998 observations of SGR 1900+14, we observed 1119 bursts (N) spread over the 32 phase bins with an average (N) of 34.97 per interval. A maximum of 50 bursts ($N_{\rm max}$) were observed in an interval. The 2001 observation of SGR 1900+14 (Fig. 4) suggests a correlation between the burst frequency and the pulse phase. But this is inconclusive due to poor burst statistics.

The burst statistics for the AXP 1E 2259+586 during the 2002 observation (Fig. 5 and Table 2) is also poor. But, it appears that there are two phase intervals with high burst rates (Fig. 5(a)), one coinciding in phase with the first pulse, while the other appears to be leading the second pulse of the profile (Fig. 5b). In this case, a maximum of 8 bursts were observed in an interval.

An analysis was carried out to estimate the deviation of the burst frequency distribution from a constant and detection significance of any peak in the burst distribution. This was carried out independently for the 3σ bursts from each of the observations and the result is shown in Table 2.

For the 1996 and 2004 observations of SGR 1806–20, assuming a Poisson distribution of bursts during the pulse phase, the probability of occurrence of more than $N_{\rm max}$ bursts (in 32 bins) is about 0.46% and 52.19% respectively. The small probability in the case of the 1996 observation indicates a detection significance of the narrow peak in the burst frequency distribution at the 3σ level. In the case of 1998 observations of SGR 1900+14, the probability of occurrence of 50 or more number of bursts, assuming Poisson statistics, is 31.05%. Further, for the 2001 observation, the probability of occurrence of 7 bursts against an average of 2.69 bursts is 64.67%. A large

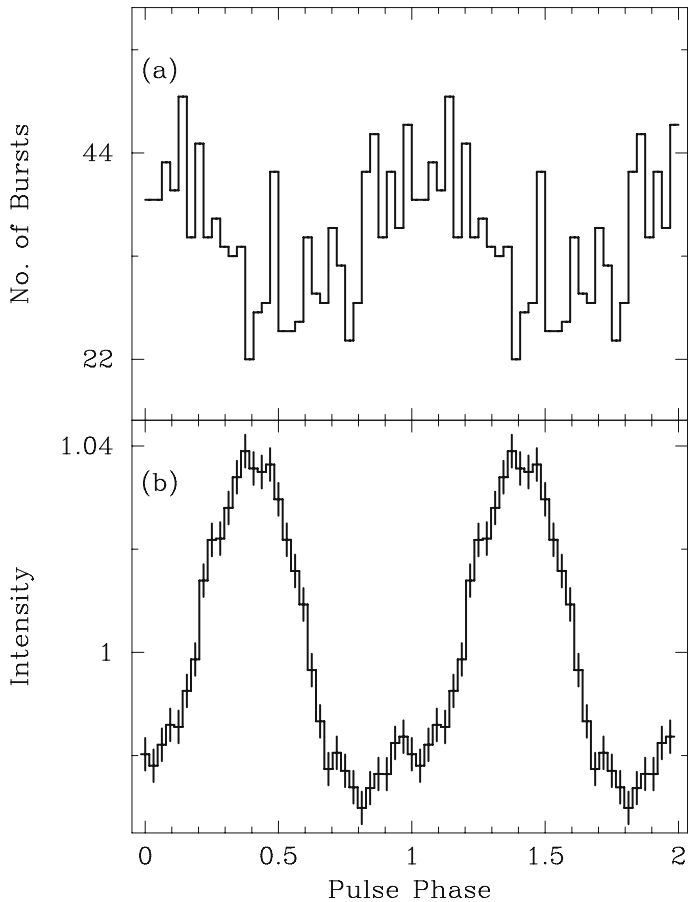

Figure 3. As in Fig. 1, this figure shows the result of the burst search analysis for the source SGR 1900+14 observed between August 28 and October 08, 1998.

probability of occurrence of more than $N_{\max}$ bursts, rules out any significant peak in the burst frequency. Hence, nothing definite can be inferred from these observations regarding the presence of narrow peaks in the burst distribution of SGR 1900+14. For AXP 1E 2259+586, the probability of occurrence of more than 8 random events in the 32 bins is 22.09%. Since, this observation has poor statistics, the presence of any peak in the burst frequency cannot be inferred.

Even in the absence of any clear peaks in the burst distribution, the burst frequency can differ from a constant. We have tested the constant frequency hypothesis by generating 100,000 sets of burst distributions for each observation. These sets were simulated using a random number generator by assuming a Poisson distribution about the average number (N) of events. We compared the standard deviation ($\sigma_{\text{simulated}}$) of each simulated dataset with the standard deviation (σ_{real}) of the real dataset. In the case of SGR 1806–20, we observed that for the 1996 observations, out of 100,000 simulations, only 33 datasets had a standard deviation greater than the real standard deviation. While it was 35,938 in the case of 2004 observations (Table 2). The results of the two analyses confirm a non-uniform spin phase dependence of the burst distribution during the 1996

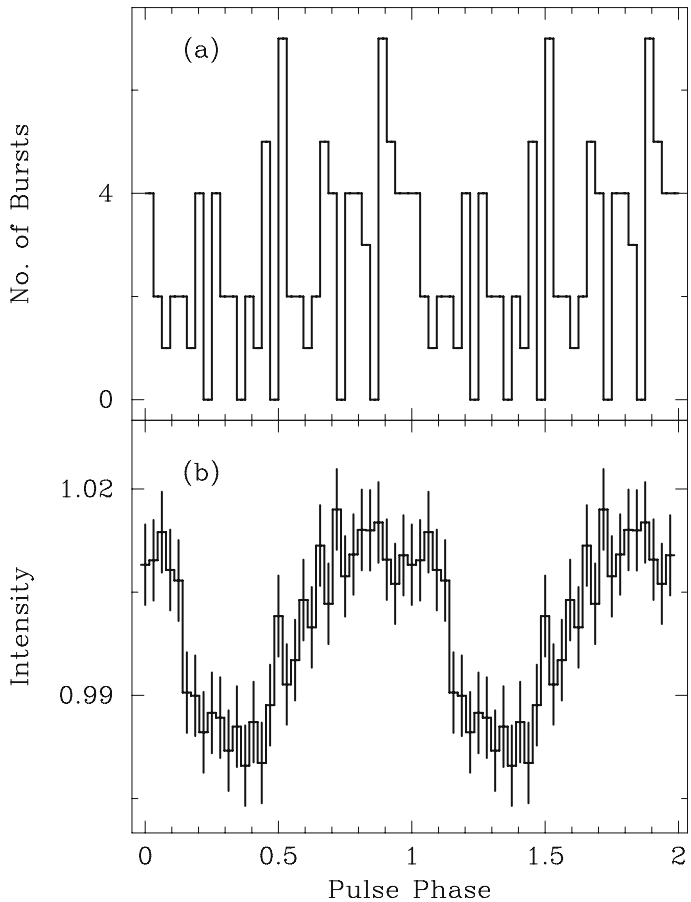

Figure 4. As in Fig. 1, this figure shows the result of the burst search analysis for the source SGR 1900+14 observed between April 19 and April 30, 2001.

observations, however there is no apparent strong clustering in phase. The results from the 2004 observations are inconclusive. Similarly, from the 1998 observations of SGR 1900+14, only 2400 had a standard deviation greater than $\sigma_{\rm real}$, which is an indication at the 2.4% level of a non-constant distribution. This is in agreement with the clustering of high and low burst rates reported above. In the case of AXP 1E 2259+586, out of the 100,000 sets, only 1458 had a standard deviation greater than the standard deviation of the real dataset. This also suggests a deviation of the burst frequency from a constant distribution in agreement with indication for the non-uniformity presented above.

3. Discussion

AXP and SGR bursts have many qualitative similarities as has been predicted by the magnetar model. The magnetar model postulates that the short duration SGR bursts are triggered by starquakes induced by magnetic stresses in the neutron star crust (Thompson & Duncan 1996, 2001). The burst active phases vary in both intensity and duration. Burst activity in SGRs is known to be concentrated in time. There is

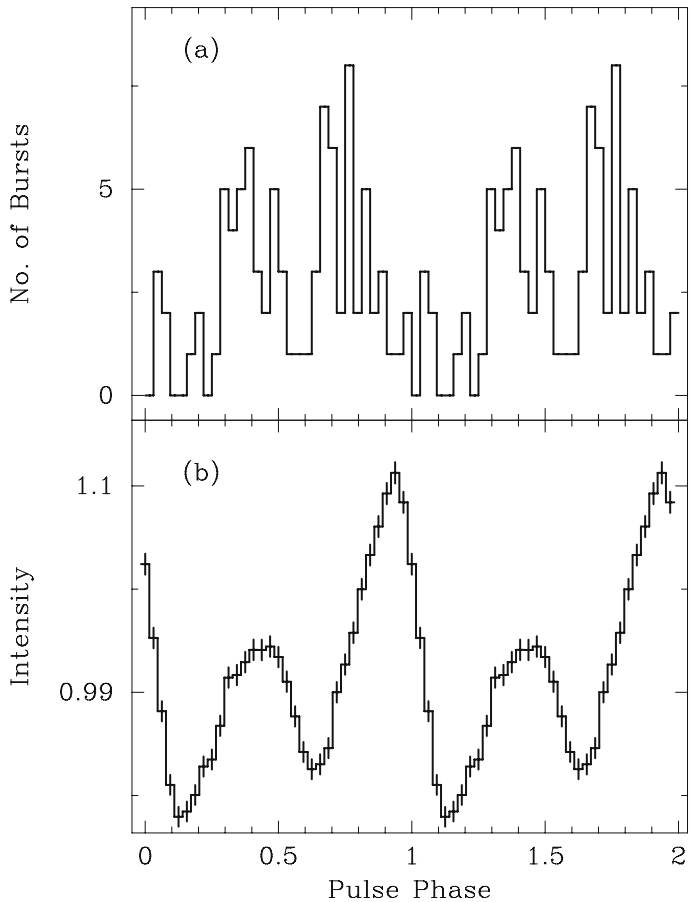

Figure 5. As in Fig. 1, this figure shows the result of the burst search analysis for the magnetar candiate 1E 2259+586 observed on June 18, 2002.

no correlation between the burst energy and the time before the next burst in either the SGRs (Laros *et al.* 1987; Gogus *et al.* 1999), or in 1E 2259+586 (Gavriil *et al.* 2004). But, there is an evidence for a correlation between the burst activity and the time elapsed since the previous burst for SGR 1900+14 (Gogus *et al.* 1999). This suggests that these bursts are not accretion-powered but rather they are triggered by an instability of the stellar magnetic field.

Bursts have been observed to occur for all the pulse phases. In the case of SGR 1806–20, a spike is observed above 3σ significance level when the pulsed intensity is high. Since the spike in the burst distribution is narrow, the phase coincidence with the broad X-ray pulse profile may also happen randomly without any physical significance. Furthermore, in all the three sources, a deviation from a constant burst distribution was observed, ranging in significance from 0.03% in SGR 1806–20, to 2.4% for SGR 1900+14, and 1.5% for AXP 1E 2259+586.

Short and long outbursts of magnetars are believed to arise from injection of energy into the magnetosphere, through a rearrangement of the magnetic field and the formation of dissipation of strong localized currents. The crust provides a site for

the initial loss of equilibrium that triggers an outburst (Palmer 1999). The thermal blackbody component of the X-ray emission observed from the magnetar candidates comes from $\sim 1\%$ of the stellar surface (SGR 1900+14: Ibrahim *et al.* 2001; 1E 2259+586: Woods *et al.* 2004). This suggests that the crust develops a network of small dislocations. In view of the results presented here, we propose that the crust of the neutron star has volcanic sites all over the viewed surface. However, few hot spots appear to burst slightly more often. There is a significant detection of a pulse phase dependence, only in the case of SGR 1900+14. For the AXP 1E 2259+586, there is an indication of a possible pulse phase dependence, but the results are inconclusive due to poor statistics. More detailed observations of the magnetars during their active phases, perhaps with the upcoming mission ASTROSAT, will enable us to better understand the bursting behaviour of magnetars.

Acknowledgements

This research has made use of data obtained from the High Energy Astrophysics Science Archive Research Center (HEASARC), provided by NASA's Goddard Space Flight Center. We thank Darshana Ramachandran for help with some preliminary analysis. We also thank the anonymous referee for his valuable advice about the manuscript.

References

Bak, P., Tang, C., Wiessenfeld, K. 1998, *Phys. Rev. A.*, **38**, 364.
Chatterjee, P., Hernquist, L., Narayan, R. 2000, *ApJ*, **534**, 373.
Corbet, R. H. D., Smale, A. P., Qzaki, M., Koyama, K., Iwasawa, K. 1995, *ApJ*, **443**, 786.
Duncan, R. C., Thompson, C. 1992, *ApJ*, **392**, L9.
Feroci, M., Hurley, K., Duncan, R. C., Thompson, C. 2001, *ApJ*, **549**, 1021.
Gavriil, F. P., Kaspi, V. M., Woods, P. M. 2002, *Nature*, **419**, 142.
Gavriil, F. P., Kaspi, V. M., Woods, P. M. 2004, *ApJ*, **607**, 959.
Gogus, E., Woods, P. M., Kouveliotou, C., van Paradijs, J., Briggs, M. S., Duncan, R. C., Thompson, C. 1999, *ApJ*, **526**, L93.
Gogus, E., Woods, P. M., Kouveliotou, C., van Paradijs, J., Briggs, M. S., Duncan, R. C., Thompson, C. 2000, *ApJ*, **532**, L121.
Gogus, E., Kouveliotou, C., Woods, P. M., Finger, M. H., van der Klis, M. 2002, *ApJ*, **577**, 929.
Gotz, D., Mereghetti, S., Tiengo, A., Esposito, P. 2006, *A&A*, **449**, L31.
Hulleman, F., van Kerkwijk, M. H., Verbunt, F. W. M., Kulkarni, S. R. 2000a, *A&A*, **358**, 605.
Hulleman, F., van Kerkwijk, M.H., Kulkarni, S. R. 2000b, *Nature*, **408**, 689.
Hurley, K. *et al.* 1999, *Nature*, **397**, 41.
Hurley, K. 2000, In: AIP Conf. Proc. 526, 5th Hunstville Symp. on Gamma-ray Bursts (eds) Kippen, R. M., Mallozzi, R. S., Fishman, G. F. (New York: AIP), 763.
Hurley, K., Sari, R., Djorgovski, S. G. 2002, astro-ph/11620H.
Hurley, K. *et al.* 2005, *Nature*, **434**, 1098.
Ibrahim, A. I., Strohmayer, T. E., Woods, P. M., Kouveliotou, C., Thompson, C., Duncan, R. C., Dieters, S., Swank, J. H., van Paradijs, J., Finger, M. A. 2001, *ApJ*, **558**, 237.
Israel, G., Mereghetti, S., Stella, L. 2002, *Mem. della Soc. Ast. It.*, **73**, 465.
Israel, G. L., Belloni, T., Stella, L., Rephaeli, Y., Gruber, D. E., Casella, P., Dall'Osso, S., Rea, N., Persic, M., Rothschild, R. E. 2005, *ApJ*, **628**, L53.
Jahoda, K., Swank, J. H., Giles, A. B., Stark, M. J., Strohmayor, T., Zhang, W., Morgan, E. H. 1996, *Proc. SPIE*, **2808**, 59.
Kaspi, V. M., Gavriil, F. P. 2003, *ApJ*, **596**, L71.

Kaspi, V. M., Gavriil, F. P., Woods, P. M., Jensen, J. B., Roberts, M. S. E., Chakrabarty, D. 2003, *ApJ*, **588**, L93.
Kouveliotou, C. 1995, *Ap&SS*, **231**, 49.
Kouveliotou, C. *et al.* 1998, *Nature*, **393**, 235.
Kouveliotou, C., Tennant, A., Woods, P. M., Weisskopf, M. C., Hurley, K., Fender, R. P., Garrington, S. T., Patel, S. K. 2001, *ApJ*, **558**, L47.
Kuiper, L., Hermsen, W., den Hartog, P. R., Collmar, W. 2006, *ApJ*, **645**, 556.
Laros, J. G. *et al.* 1987, *ApJ*, **320**, L111.
Mazets, E. P., Cline, T. L., Aptekar, R. L., Butterworth, P. S., Frederiks, D. D., Golenetskii, S. V., Il'Inskii, V. N., Pal'Shin, V. D. 1999, *Astron. Lett.*, **25**, 635.
Mereghetti, S., Israel, G. L., Stella, L. 1998, *MNRAS*, **296**, 689.
Mereghetti, S., Gotz, D., von Kielin, A., Rau, A., Lichtii, G., Weidensponter, G., Jean, P. 2005, *ApJ*, **624**, L105.
Molkov, S., Hurley, K., Sunyaev, R., Shtykovsky, P., Revnivtsev, M., Kouveliotou, C. 2005, *A&A*, **433**, 13.
Palmer, D. M. 1999, *ApJ*, **512**, L113.
Palmer, D. M. 2000, In: *AIP Conf. Proc. 526, 5th Hunstville Symp. on Gamma-ray Bursts* (eds) Kippen, R. M., Mallozzi, R. S., Fishman, G. F. (New York: AIP), 791.
Palmer, D. M., *et al.* 2005, *Nature*, **434**, 1107.
Paul, B., Kawasaki, M., Dotani, T., Nagase, F. 2000, *ApJ*, **537**, 319.
Paul, B., Agrawal, P. C., Mukerjee, K., Rao, A. R., Seetha, S., Kasturirangan, K. 2001, *A&A*, **370**, 529.
Strohmayer, T. E., Watts, A. L. 2006, *ApJ*, **653**, 593.
Thompson, C., Duncan, R. C. 1995, *MNRAS*, **275**, 255.
Thompson, C., Duncan, R. C. 1996, *ApJ*, **473**, 322.
Thompson, C., Duncan, R. C. 2001, *ApJ*, **561**, 980.
van Paradijs, J., Taam, R. E., van den Heuvel, E. P. J. 1995, *A&A*, **299**, L41.
Woods, P. M., Kouveliotou, C., van Paradijs, J., Finger, M. H., Thompson, C., Duncan, R. C., Hurley, K., Strohmayer, T., Swank, J., Murakami, T. 1999, *ApJ*, **524**, L55.
Woods, P. M., Kouveliotou, C., Gogus, E., Finger, M. H., Swank, J., Smith, D. A., Hurley, K., Thompson, C. 2001, *ApJ*, **552**, 748.
Woods, P. M., Kaspi, V. M., Thompson, C., Gavriil, F. P., Marshall, H. L., Chakrabarty, D., Flanagan, K., Heyl, J., Hernquist, L. 2004, *ApJ*, **605**, 378.
Woods, P. M., Kouveliotou, C., Finger, M. H., Gogus, E., Wilson, C. A., Patel, S. K., Hurley, K., Swank J. H. 2007, *ApJ*, **654**, 470.

J. Astrophys. Astr. (2007) **28,** 197–206

Orbital Evolution Measurement of the Accreting Millisecond X-ray Pulsar SAX J1808.4–3658

Chetana Jain[1,*], Anjan Dutta[1] & Biswajit Paul[2]
[1]*Department of Physics and Astrophysics, University of Delhi, Delhi 110 007, India.*
[2]*Raman Research Institute, Sadashivnagar, C.V. Raman Avenue, Bangalore 560 080, India.*
[*]*e-mail: chetanajain11@yahoo.co.in*

Received 2007 October 1; revised 2007 December 5; accepted 2007 December 6

Abstract. We present results from a pulse timing analysis of the accretion-powered millisecond X-ray pulsar SAX J1808.4–3658 using X-ray data obtained during four outbursts of this source. Extensive observations were made with the proportional counter array of the Rossi X-ray Timing Explorer (RXTE) during the four outbursts that occurred in 1998, 2000, 2002 and 2005. Instead of measuring the arrival times of individual pulses or the pulse arrival time delay measurement that is commonly used to determine the orbital parameters of binary pulsars, we have determined the orbital ephemeris during each observation by optimizing the pulse detection against a range of trial ephemeris values. The source exhibits a significant pulse shape variability during the outbursts. The technique used by us does not depend on the pulse profile evolution, and is therefore, different from the standard pulse timing analysis. Using 27 measurements of orbital ephemerides during the four outbursts spread over more than 7 years and more than 31,000 binary orbits, we have derived an accurate value of the orbital period of 7249.156862(5) s (MJD = 50915) and detected an orbital period derivative of $(3.14 \pm 0.21) \times 10^{-12}$ s s^{-1}. We have included a table of the 27 mid-eclipse time measurements of this source that will be valuable for further studies of the orbital evolution of the source, especially with ASTROSAT. We point out that the measured rate of orbital period evolution is considerably faster than the most commonly discussed mechanisms of orbital period evolution like mass transfer, mass loss from the companion star and gravitational wave radiation. The present time scale of orbital period change, 73 Myr is therefore likely to be a transient high value of period evolution and similar measurements during subsequent outbursts of SAX J1808.4–3658 will help us to resolve this.

Key words. Stars: neutron stars—pulsars: individual (SAX J1808.4–3658), magnetic fields—X-rays: binaries—accretion.

1. Introduction

In accretion-powered binary X-ray pulsars, the orbital evolution can often be measured using pulse timing. Though the accretion-powered X-ray pulsars often have variable

and broad pulse profiles, the uncertainties in the measurement of orbital ephemeris is usually much smaller than the rate of advancement or delay of the ephemeris in a few years time scale. In the last three decades, repeated measurements of the orbital ephemeris of many binary X-ray pulsars have led to accurate determination of the orbital evolution time scales (Cen X-3: Nagase *et al.* 1992, Paul *et al.* 2007; LMC X-4: Naik & Paul 2004; SMC X-1: Wojdowski *et al.* 1998, Paul *et al.* 2007; 4U 1538–52: Mukherjee *et al.* 2006; Her X-1: Paul *et al.* 2004). In some of the non-pulsating X-ray binaries, the orbital evolution has been measured using eclipse timing or non-eclipsing orbital modulation of the X-ray light curve (Cyg X-3: Singh *et al.* 2002; EXO 0748-676: Wolff *et al.* 2002; 4U 1822-371: Parmar *et al.* 2000). In the High Mass X-ray Binaries (HMXB), the orbital evolution time-scale is in the range of 10^5–10^6 years and in the low mass X-ray binaries it is more than 10^8 years. In HMXB sources, the orbital evolution is believed to be driven by tidal interaction and mass loss from the companion star while in LMXB sources, it is due to mass transfer or magnetic interaction between the two stellar components. In very compact X-ray binaries, like some of the millisecond X-ray pulsars with orbital period of about 40 minutes, orbital evolution due to gravitational wave radiation may also become dominant.

All the X-ray sources mentioned above are persistent X-ray pulsars while all the millisecond accreting X-ray pulsars are transient X-ray sources. It is therefore difficult to measure the orbital evolution of the later class of sources. Only one of the eight accreting millisecond X-ray pulsars, SAX J1808.4–3658 has shown repeated outbursts that makes it suitable for investigation of orbital evolution. SAX J1808.4–3658 is the first of the eight known accretion-powered millisecond pulsars. It was discovered during an outburst in 1996 by the *Beppo*SAX Wide Field Camera (in 't Zand *et al.* 1998). The X-ray transient was again detected during an outburst in April 1998 (Marshall 1998) with the Proportional Counter Array (PCA) on-board the Rossi X-ray Timing Explorer (RXTE). During this outburst, coherent millisecond X-ray pulsations at 401 Hz were detected, first in any accretion-powered X-ray pulsar (Wijnands & van der Klis 1998). From the same observation a binary period of 2 h was also measured (Chakrabarty & Morgan 1998). Outbursts were again detected from SAX J1808.4–3658 in 2000 (van der Klis *et al.* 2000; Wijnands *et al.* 2001), 2002 (Wijnands 2004 for a review) and 2005 (Wijnands 2005).

Chakrabarty & Morgan (1998) had measured the orbital parameters using the observations made by the RXTE during the 1998 outbursts. The orbital parameters were subsequently revised by Papitto *et al.* (2005), who reduced the error in the earlier reported orbital period by an order of magnitude. Papitto *et al.* (2005) performed the timing analysis for observations spanning 5 years, covering the outbursts that occurred in 1998, 2000 and 2002. However, the data from these three outbursts were insufficient for the determination of orbital period derivative and Papitto *et al.* (2005) reported an orbital period derivative in the range of -6.6×10^{-12} s s^{-1} to 0.8×10^{-12} s s^{-1}.

In the present work described in the following chapters, we have analyzed all the long RXTE–PCA observations of this source during the four outbursts. We have made an accurate measurement of the orbital period evolution of this compact X-ray binary system during the period 1998–2005.

2. Observations and analysis

We have analyzed all the sufficiently long archival data from observations of the millisecond pulsar SAX J1808.4–3658 using the RXTE–PCA. The PCA consists of an array of five collimated xenon/methane multi-anode proportional counter units (PCU) with a total photon collection area of 6500 cm^2 (Jahoda *et al.* 1996). Data used in the present analysis were taken from the event mode of PCA with a time resolution of 122 μs and 64 channel energy information in the entire PCA energy band of 2–60 keV. Table 1 lists all the observation IDs used in the present work along with the average number of PCUs available during each observation and the good exposure time for each of the observations.

Table 1. Log of RXTE observations analyzed in the present work.

Outburst	Data range (MJD)	Observation IDs	Average number of PCU ON	Exposure time (ks)
1998	50914–50932	30411-01-05-00	4–5	15.8
		30411-01-06-00	3–4	24.9
		30411-01-07-00	3–4	5.5
		30411-01-08-00	3–4	17.9
		30411-01-09-00	3–4	17.2
		30411-01-09-02	3–4	14.4
		30411-01-09-04	3–4	13.5
		30411-01-10-00	3–4	8.9
		30411-01-10-01	3–4	7.3
2000	51585–51593	40035-01-02-08	3–4	6.5
		40035-01-03-00	4–5	8.2
2002	52564–52573	70080-01-01-020	2–3	20.1
		70080-01-01-02	3.0	
		70080-01-02-04	3–4	17.3
		70080-01-02-06	3–4	17.2
		70080-01-02-08	3–4	17.2
		70080-01-02-14	2–3	13.8
		70080-01-03-02	3–4	17.4
2005	53524–53537	91056-01-02-01	2–3	17.5
		91056-01-02-02	3–4	13.9
		91056-01-02-03	3–4	20.4
		91056-01-02-04	2–3	13.8
		91056-01-02-05	2–3	17.2
		91056-01-02-06	3–4	10.2
		91056-01-02-08	3–4	13.8
		91056-01-03-00	3–4	13.2
		91056-01-03-06	2–3	13.8
		91056-01-03-10	3–4	13.5

The source count rate during most of the observations varied from 500–900 counts/s. But during the 2000 outburst, the count rate was lower, about 270 counts/s. Observations shorter than a duration of 5 ks were not used in this analysis, as they would not provide a good measurement of the orbital ephemeris.

Orbital parameters of binary X-ray pulsars are usually determined by measuring the delay in arrival times of individual or a group of pulses due to orbital motion. The maximum group length is kept small so as to avoid smearing of the pulses due to orbital motion. In the case of millisecond accreting pulsars, especially when the source is in low intensity state, folding very short data segments may lead to non-detection of pulses or significant error in determination of the pulse phase (or arrival time). Therefore, to detect the pulsations even in the low state and also to simultaneously measure the local orbital ephemeris from data stretches comparable to the orbital period, we adopted a technique in which the photon arrival times are corrected for the binary motion. For this correction, a range of trial ephemerides were taken around the value extrapolated from earlier measurements (Chakrabarty & Morgan 1998; Papitto *et al.* 2005) and the folded pulse profiles were analyzed to determine the most appropriate orbital ephemeris. The orbital period (7249.1569 s) which was initially taken to be a constant and the orbital semi-amplitude values (62.809 lt–ms) were also taken from the same earlier reports. Though the orbital period is evolving, the rate of evolution is too small to affect the local orbital motion correction. We developed this technique for orbital evolution studies of X-ray pulsars and a similar technique has also been applied to *Beppo*-SAX, ASCA and IXAE observations of several sources like LMC X-4, SMC X-1 and XTE J1946 + 27 (Paul *et al.* 2001, 2002; Naik & Paul 2004).

From each observation listed in Table 1, we generated energy resolved pulse folded light curves using the FASEBIN tool in the Ftools version 5.2. This tool corrects the photon arrival times for the orbital motion of the pulsar as described above and also converts the photon arrival times to the solar system barycenter. The output is a two-dimensional histogram of photon count rate against the pulse phase and energy. It applies all the known RXTE clock corrections and converts the photon arrival times to the solar system barycenter using JPL DE-200 ephemeris. The position coordinates used were R.A. $= 18^h\ 08^m\ 27^s.6$ and dec $= -36^\circ\ 58'\ 43''.3$. This is the best estimate for the source position and is compatible with the optical counterpart (Rupen *et al.* 2002; Giles *et al.* 1999).

We have then used another Ftool – FBSSUM, which enables us to use the output of FASEBIN and integrate the pulse profiles over different energy ranges. Figure 1 shows a sample of the pulse profiles in the 2–10 keV energy band during each of the four outbursts obtained with the ephemerides determined from the present analysis. The technique used in this analysis is independent of the pulse profile changes that are seen in this pulsar. The pulse profiles thus obtained were fitted with a constant and the χ^2 was measured as a function of the orbital ephemeris (or the time of passage of the neutron star at the ascending node T^*). This variation of χ^2 with the orbital ephemeris is shown in Fig. 2 for one of the observations. We made a Gaussian fit to this χ^2 distribution and the best value of the orbital ephemeris (T^*) and its error were determined for each set of observations. We took a long dataset from one of the 1998 observations and measured T^* and its error using the pulse profiles in different energy ranges. The energy range of 2–10 keV was chosen for all further analysis because this energy gave us the minimum error in measurement of T^*. We also independently measured the spin period of the pulsar during each of the outbursts. The T^* obtained for each of the dataset and

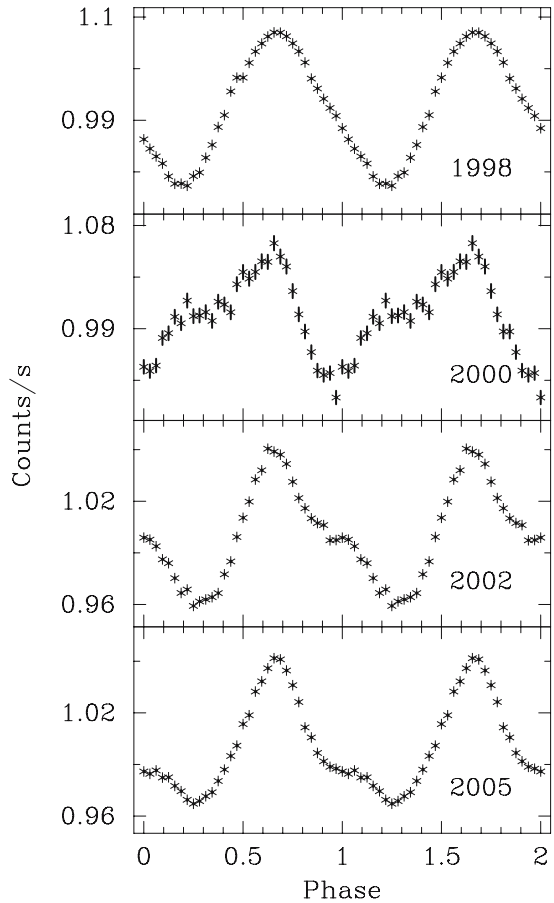

Figure 1. X-ray pulse profile of the millisecond pulsar SAX J1808.4–3658 obtained from observations with RXTE during the four outbursts are shown here in the energy band of 2–10 keV.

the corresponding 1σ measurement uncertainty are shown in Table 2 against the corresponding binary orbit number $N (N = \text{int}[(T^*–T_0^*)/P_{\text{orb}}]$, where T_0^* is the time of passage of the neutron star at the ascending node at the reference time (start of the 1998 observation). The orbital ephemerides history clearly shows a quadratic nature. To estimate the orbital period and the period derivative valid through the entire outburst history of the millisecond pulsar, we fitted a quadratic curve to the ephemerides history. The parameters obtained for the best fitted curve are given in Table 3. We have obtained an orbital period of 7249.156862(5) s on MJD = 50914.878425(1) and an orbital period derivative of $(3.14 \pm 0.21) \times 10^{-12}$ s s^{-1}. We subtracted the best fitted linear component from the ephemerides history and the residual is plotted in Fig. 3. It clearly shows the quadratic component.

3. Discussion

We have performed a pulse timing analysis of the accretion-powered millisecond pulsar SAX J1808.4–3658 during the four X-ray outbursts. The RXTE–PCA observations

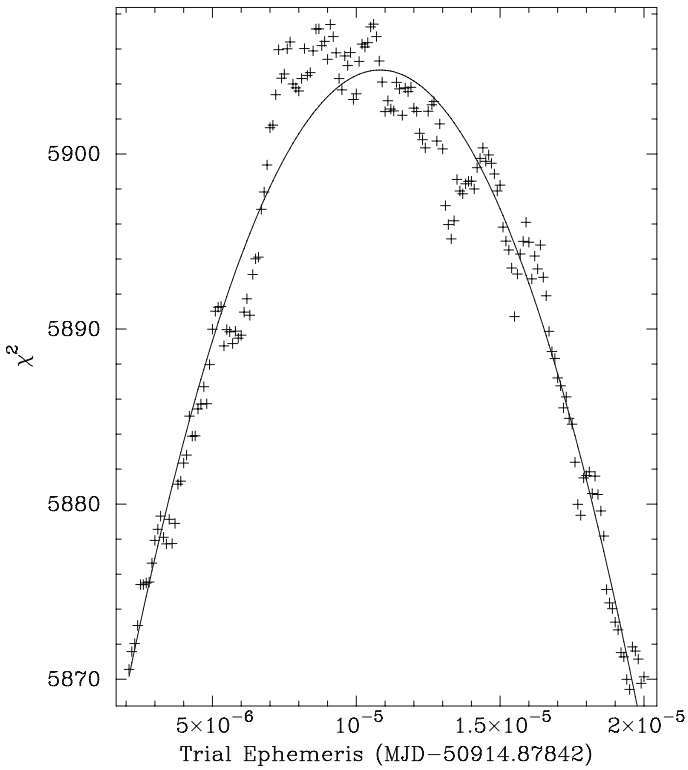

Figure 2. The variation of χ^2 as a function of the trial ephemeris for one of the RXTE observations of the 1998 outburst from SAX J1808.4–3658.

used for the present work covered a span of more than 7 years and more than 31,000 binary orbits. The orbital parameters were measured by correcting the light curves for the binary motion of the pulsar and then optimizing the pulse detection. We have obtained an orbital period of 7249.156862(5) s and a significant orbital period derivative of $(3.14 \pm 0.21) \times 10^{-12}$ s s^{-1}. We have been able to reduce the error in the measurement of the orbital period by 2 orders of magnitude as compared to the published reports (Papitto *et al.* 2005). The observed orbital period derivative implies an orbital evolution time-scale of $(P_{\rm orb}/\dot{P}_{\rm orb})$ 73 Myr. Analysing the same set of RXTE-PCA observations with different/independent analysis techniques, Di Salvo *et al.* (2007) and Hartman *et al.* (2007) have also obtained a large orbital period derivative. Hartman *et al.* (2007) have used the technique of measuring the arrival times of individual pulses to determine the orbital evolution. But, the standard pulse timing analysis has inherent limitations when there is a pulse profile evolution, as in the case with SAX J1808.4–3658. The present work is independent of the pulse profile evolution observed in this pulsar.

The evolution of close binaries crucially depends on the exchange of mass and the exchange of angular momentum between the components of the interacting binaries. A change in the orbital angular momentum results in the orbital evolution of the binary system. A significant mass loss may occur from the binary system through stellar winds or a sudden catastrophic event like a supernova explosion on one of the components of

Table 2. Variation of T^* as a function of the orbit cycle.

Orbit no.	T^* (MJD)	1σ uncertainty in measurement of T^*
0	50914.8784309	0.0000026
0	50914.8784342	0.0000047
107	50923.8559762	0.0000041
140	50926.6247502	0.0000033
153	50927.7154825	0.0000043
164	50928.6384020	0.0000045
176	50929.6452306	0.0000045
188	50930.6520566	0.0000053
211	50932.5818133	0.0000054
7997	51585.8449364	0.0000068
8089	51593.5639393	0.0000036
19657	52564.1454912	0.0000056
19680	52566.0752480	0.0000039
19692	52567.0820713	0.0000044
19704	52568.0888978	0.0000037
19739	52571.0254824	0.0000046
19774	52573.9620552	0.0000055
31105	53524.6588101	0.0000042
31117	53525.6656400	0.0000038
31128	53526.5885608	0.0000047
31141	53527.6792960	0.0000062
31152	53528.6022180	0.0000055
31164	53529.6090365	0.0000033
31176	53530.6158673	0.0000052
31186	53531.4548906	0.0000052
31235	53535.5661091	0.0000068
31257	53537.4119610	0.0000106

the binary. This could result in a change in the energy of the binary system. A stellar wind from the co-rotating companion star carries away the angular momentum which can cause an evolution of the orbit of the binary system on a time-scale of a few Myr (Kelley 1986).

The total energy of the system may also decrease due to tidal friction. The time-scales of orbital evolution due to tidal interaction between the stars' rotation and their mutual orbit, ranges from a few Myr to more than a Hubble time ($\sim 10^{10}$ yr) which is much greater than the observed orbital evolution time-scale of 73 Myr. But tides do not always lead to rapid circularization during the early phase of mass transfer. Thus, phases of episodic mass transfer may occur at the periastron passages and may persist for a longer period. But since the source is observed to be active for a short period of time, therefore, mass transfer cannot be the decisive factor for the observed orbital evolution.

Table 3. Orbital parameters of SAX J1808.4–3658.

Parameter	Chakrabarty & Morgan (1998)	Papitto *et al.* (2005)	Best fit value from the present analysis
$P_{\rm orb}$ (s)	7249.119(1)	7249.1569(1)	7249.156862(5)
$\dot{P}_{\rm orb}(10^{-12}$ s s$^{-1})$	–	$-6.6 \leq \dot{P}_{\rm orb} \leq 0.8$	3.14(0.21)
T_0^* (MJD)	50914.899440(1)[a]	50914.878468(4)	50914.878425(1)

[a] $T_{pi/2}$ value.

Note: The numbers in parentheses are the 1σ uncertainties in the last significant figure.

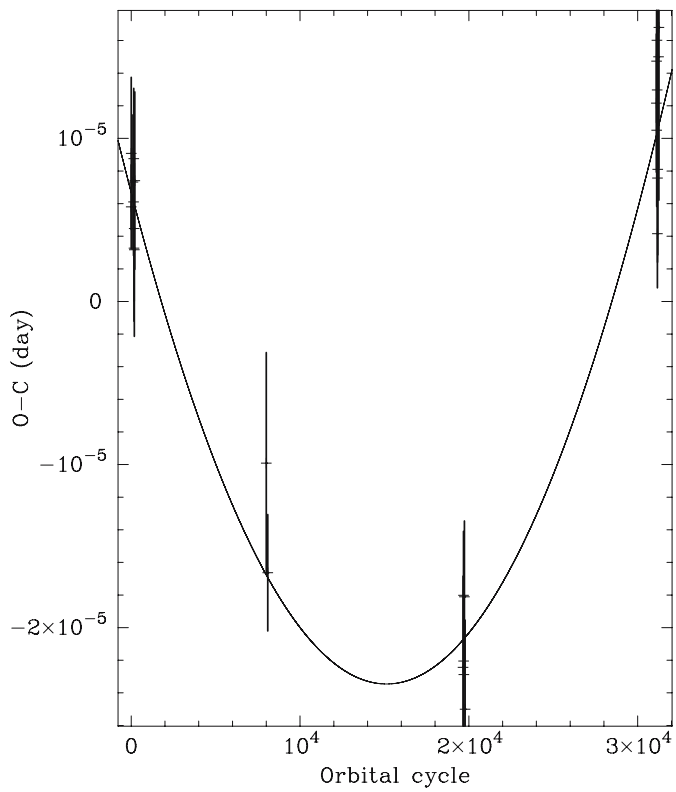

Figure 3. The time of ascending node of SAX J1808.4–3658 has been plotted here as a function of the orbital cycle, relative to an orbital period of 0.083902278 d. The curvature of the O–C curve is a measure of the orbital period derivative of the binary system.

The orbital energy of a binary is also expected to diminish through gravitational wave radiation, thus decreasing the orbital period. The rate of change of the orbital period due to the gravitational wave radiation, in a binary system with a conservative mass transfer, is of the order of 10^{-13} s s^{-1} (Verbunt 1993). But from the present analysis, we have obtained an orbital period derivative of $(3.14 \pm 0.21) \times 10^{-12}$ s s^{-1}, which is higher by more than a factor of 10. It implies that the gravitational wave radiation is not enough to explain the observed parameters.

Further, Chakrabarty & Morgan (1998), have proposed that SAX J1808.4–3658 behaves as a black widow pulsar during the periods of X-ray quiescence and is evaporating its companion star. But the time-scales of the orbital evolution due to conservative mass transfer is of the order of a few Gyr. Therefore, mass loss during quiescence is not sufficient to explain the rapid orbital evolution of the system.

In view of the results presented here, it can be proposed that either the tidal effect in the system is strong or the rapid orbital evolution is a transient high value and an analysis of any further episode of an outburst from SAX J1808.4–3658 can explain the observed evolutionary scenarios.

Acknowledgements

This research has made use of data obtained from the High Energy Astrophysics Science Archive Research Center (HEASARC), provided by NASA's Goddard Space Flight Center. One of the authors, AD, acknowledges the R & D Doctoral Research Scheme of Delhi University.

References

Chakrabarty, D., Morgan, E. H. 1998, *Nature*, **394**, 346.
Di Salvo, T., Burderi, L., Riggio, A., Papitto, A., Menna, M. T. 2007, *MNRAS*, submitted (astro–ph/0708.0498).
Giles, A. B., Hill, K. M., Greenhill, J. G. 1999, *MNRAS*, **304**, 47.
Hartman, J. M., Patruno, A., Chakrabarty, D., Kaplan, D. L., Markwardt, C. B., Morgan, E. H., Ray, P. S., van der Klis, M., Wijnands, R. 2007, *ApJ*, submitted (astro–ph/0708.0211).
in 't Zand, J. J. M., Heise, J., Muller, J. M., Bazzano, A., Cocchi, M., Natalucci, L., Ubertini, P. 1998, *A&A*, **331**, L25.
Jahoda, K., Swank, J. H., Giles, A. B., Stark, M. J., Strohmayor, T., Zhang, W., Morgan, E. H. 1996, *Proc. SPIE*, **2808**, 59.
Kelley, R. L. 1986, In: The Evolution of Galactic X-ray Binaries (eds) Trumper, J., Lewin, W. H. G., Brinkmann, W. (Dordrecht, Reidel) 75.
Marshall, F. E. 1998, *IAUC*, 6876.
Mukherjee, U., Raichur, H., Paul, B., Naik, S., Bhatt, N. 2006, *JApA*, **27**, 411.
Nagase, F., Corbet, R. H. D., Day, C. S. R., Inoue, H., Takeshima, T., Yoshida, K., Mihara, T. 1992, *ApJ*, **396**, 147.
Naik, S., Paul, B. 2004, *ApJ*, **600**, 351.
Papitto, A., Menna, M. T., Burderi, L., Di Salvo, T., D'Antona, F., Robba, N. R. 2005, *ApJ*, **621**, L113.
Parmar, A. N., Oosterbroek, T., DelSordo, S., Segreto, A., Santangelo, A., Dal Fiume, O., Orlandini, M. 2000, *A&A*, **356**, 175.
Paul, B., Agrawal, P. C., Mukerjee, K., Rao, A. R., Seetha, S., Kasturirangan, K. 2001, *A&A*, **370**, 529.
Paul, B., Nagase, F., Endo, T., Dotani, T., Yokogawa, J., Nishiuchi, M. 2002, *ApJ*, **579**, 411.
Paul, B., Naik, S., Bhatt, N. 2004, *Nuclear Physics B – Proceedings Supplements*, **132**, 548.
Paul, B., Raichur, H., Naik, S., Bhatt, N. 2007, The Extreme Universe in the Suzaku Era, 2006, Kyoto (ed.) K. Hayashida.
Rupen, M. P., Dhawan, V., Mioduszewski, A. J., Stappers, B. W., Gaensler, B. M. 2002, *IAUC*, **7997**, 2.
Singh, N. S., Naik, S., Paul, B., Agrawal, P. C., Rao, A. R., Singh, K. Y. 2002, *A&A*, **392**, 161S.
van der Klis, M., Chakrabarty, D., Lee, J. C., Morgan, E. H., Wijnands, R., Markwadt, C. B., Swank, J. H. 2000, *IAUC*, 7358.
Verbunt, F. 1993, *ARA & A*, **31**, 93.
Wijnands R., van der Klis, M. 1998, *Nature*, **394**, 344.

Wijnands, R., Mendez M., Markwadt, C., van der Klis, M., Chakrabarty, D., Morgan, E. 2001, *ApJ*, **560**, 892.
Wijnands, R. 2004, In: *AIP Conf Proc.*, **714**, 209.
Wijnands, R. 2005, In: *Pulsar New Research*, Nova Science Publishers (NY).
Wojdowski, P., Clark, G. W., Levine, A. M., Woo, J. W., Zhang, S. N. 1998, *ApJ*, **502**, 253.
Wolff, M. T., Hertz, P., Wood, K. S., Ray, P. S., Bandhopadhyay, R. M. 2002, *ApJ*, **575**, 384.

J. Astrophys. Astr. (2007) **28,** 207–215

Distribution of Latitudes and Speeds of Coronal Mass Ejections in the Northern and Southern Hemispheres in Cycle 23

P. X. Gao[1,2,*], Q. X. Li[1,2] & S. H. Zhong[1]
[1]*National Astronomical Observatories/Yunnan Observatory, CAS, Yunnan, China.*
[2]*School of Graduate, CAS, Beijing, China.*
[*]*e-mail: gpx@ynao.ac.cn*

Received 2007 August 26; revised 2008 February 12; accepted 2008 February 13

Abstract. Distribution of latitudes and speeds of Coronal Mass Ejections (CMEs) in the northern and southern hemispheres in cycle 23, from September 1996 to December 2006, have been analyzed. By calculating the actual probability of the hemispheric distribution of the activity of the CME, we find that a southern dominance of the activity of the CME is shown to occur in cycle 23 from September 1996 to December 2006. The CME activity occurs at all latitudes and is most common at low latitudes. This should furnish evidence to support that CMEs are associated with source magnetic structures on a large spatial scale, even with transequatorial source magnetic structures on a large spatial scale. The latitudinal distribution of CMEs in the northern and southern hemispheres are no different from a statistical point of view. The speed distribution in the northern and southern hemispheres are nearly identical and to a good approximation they can be fitted with a single lognormal distribution. This finding implies that, statistically, there is no physical distinction between the CME events in the southern and northern hemispheres and the same mechanism of a nonlinear nature acting in both the CME events in the northern and southern hemispheres. Our conclusions seem to suggest that the northern–southern asymmetry of the CME events is related to the northern–southern asymmetry in solar dynamo theory (Jiang *et al.* 2007).

Key words. Sun: activity—Sun: Coronal Mass Ejections (CMEs).

1. Introduction

Coronal Mass Ejections (CMEs), known as the most energetic form of solar magnetic activity, are now believed to be the main sources of the strong interplanetary disturbances that cause moderate-to-intense geomagnetic storms. So, a study on CMEs is an important topic that relates directly to near-Earth space environments. Since CMEs were first discovered in 1971 using the seventh Orbiting Solar Observatory (OSO-7) coronagraph (Tousey 1973), they have been observed by several space-borne coronagraphs and ground-based instruments. The Large Angle and Spectrometric Coronagraph (LASCO) onboard the Solar and Heliospheric Observatory (SOHO) mission

has observed the most CMEs, providing us with a great opportunity to examine the statistical properties of CMEs, such as their speed, acceleration, travel time (Sun to 1 AU), width, latitude, and initial location (Gopalswamy *et al.* 2000, 2001; Wang *et al.* 2002; Cane & Richardson 2003; Yashiro *et al.* 2004; Yurchyshyn *et al.* 2005).

Howard *et al.* (1985) examined the Solwind coronagraph images obtained during the interval period March 28, 1979 to December 31, 1981 and found that the CME activity occurs at all latitudes and is most common at low latitudes. A similar result was reported from the SMM coronagraph/polarimeter obtained images of the solar corona in 1980 and from 1984 to 1989 by Hundhausen (1993). Yashiro *et al.* (2004) found a N–S asymmetry in 1996–2002 from the latitudinal distribution, especially in 2001–2002. By analyzing the data of CMEs observed by SOHO/LASCO in the interval from October 19, 1996 to December 26, 2001, Aoki *et al.* (2003) found that the distribution of the speeds of CMEs is very similar to lognormal distribution. More recently, Yurchyshyn *et al.* (2005) found that the speed distribution for accelerating and decelerating CME events can be modeled by a single lognormal distribution with the use of CMEs in the interval of the years 1998 to 2001 (ascending period of cycle 23). The accelerating CMEs, as a group, are slower than the decelerating ones (Yurchyshyn *et al.* 2005). In this paper, based on the data of CMEs observed by SOHO/LASCO, we will consider the distribution of latitudes and speeds of CMEs in northern and southern hemispheres, and further, the dates used are extended almost to a cycle, in cycle 23.

2. Asymmetry of the activity of CMEs

The data of CME used here come from a catalog of observations of the SOHO/LASCO, which are available at http://cdaw.gsfc.nasa.gov/CME_list/index.html. This CME catalog, generated and maintained at the CDAW Data Center by NASA and The Catholic University of America in co-operation with the Naval Research Laboratory, covers the period from 1996 January to the present. For each CME event, the catalog contains central position angle (CPA), plane-of-sky speeds, and so on. CME latitude is obtained from the CPA of the CME, assuming that CMEs propagate radially away from the solar source region (Gopalswamy *et al.* 2003; Yashiro *et al.* 2004). Halo CMEs are excluded for which the CPAs cannot be determined. The LASCO/SOHO has observed 393 halo CMEs from January 1996 to December 2006, which are 3.4% of all the 11584 CMEs. The minimum between cycles 22 and 23 occurred in September 1996 (Harvey & White 1999), so we study the activity of CMEs from September 1996 to December 2006.

From September 1996 to December 2006, there are 5256 CME events in the northern hemisphere and 5705 CME events in the southern hemisphere. We can find that a southern dominance of the activity of CMEs is shown to occur in solar cycle 23. To be sure that this result cannot be obtained purely by chance we check by calculating the actual probability of obtaining a result. Let us consider a distribution of n objects (CMEs) in 2 classes. The probability that one CME (one object) occurs in one hemisphere (class one) by chance is $p = 1/2$. We use the following binomial formula to derive the probability $P(k)$ of getting k objects in class 1 and $(n - k)$ objects in class 2 (Li *et al.* 1998).

$$P(k) = \binom{n}{k} p^k (1 - p)^{n-k}. \tag{1}$$

In our case for a total number $n = N_N + N_S = 10961$, we have for example:

$$P(k) = \frac{10961!}{(10961-k)!k!}\frac{1}{2^{10961}}.$$

The probability to get more than d objects in one class is:

$$P(\geq d) = \sum_{k=d}^{n} P(k). \tag{2}$$

In general, when $P(\geq d) < 5\%$ we have a statistically significant result, and when $P(\geq d) < 1\%$ the result is highly significant. From September 1996 to December 2006, we have 5705 CME events occurring in the southern hemisphere within the total 10961 CME events, $P(\geq 5705) = 9.4 \times 10^{-6}$, which means that a southern dominance of the activity of the CME is highly significant, that is to say, this result cannot be obtained purely by chance. This conclusion was predicted by Ataç and Özgüc in 1996 and confirms that the hypothesis about a 12-cycle periodicity is valid (Li *et al.* 2002).

It must be pointed out that, in this paper, in order to study the latitudes of CMEs, we convert CPAs to projected heliographic latitudes. It is likely that some CME events in the northern (southern) hemisphere may be misidentified as in the southern (northern) hemisphere because of projection effects. To minimize this, we group CMEs with projected heliographic latitudes $\geq 10^\circ$ as the CME events in the northern hemisphere and those with projected heliographic latitudes $\leq -10^\circ$ as CME events in the southern hemisphere. Based on the above grouping, we calculate the actual probability of the hemispheric distribution of the activity of the CME again. From September 1996 to December 2006, we have 4538 CME events occurring in the southern hemisphere within the total 8677 CME events, $P(\geq 4538) = 9.6 \times 10^{-6}$. We also find that a southern dominance of the activity of the CME is shown to occur in cycle 23 from September 1996 to December 2006.

3. The latitudinal distribution of CMEs

We divide CMEs into 18 unsigned latitudinal bands (one band per 5°, from 0° to 90°) and count the number of CMEs within each band. The result of this binning is shown in Fig. 1. In the figure, the number of CMEs for each band is divided by the total number of CMEs from September 1996 to December 2006, so that the histograms show ratios that define the probability percentage of CMEs for each latitudinal band. We can also find that the CME activity occurs at all latitudes and is most common at low latitudes, augmenting the results obtained by Hundhausen (1993). However, the statistical characteristics of such latitudinal distribution of CMEs are different from those of sunspot groups. The latitudinal distribution of sunspot groups can be represented by a probability density function of the Γ distribution having maximum probability at about 15.5° (Li *et al.* 2003). Sunspots seem to be specified on active regions (AR), but CMEs are a kind of large-scale solar activity. So, our conclusion — the statistical characteristics of such latitudinal distribution of CMEs are different from those of sunspot groups — should furnish evidence to support that CMEs are intrinsically associated with source magnetic structures on a large spatial scale, even with transequatorial source magnetic structures on a large spatial scale (Zhou *et al.* 2006).

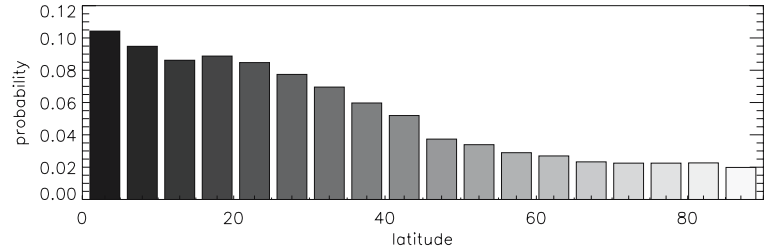

Figure 1. Percentages of CMEs to occur at the 18 latitude bands in 23 cycles.

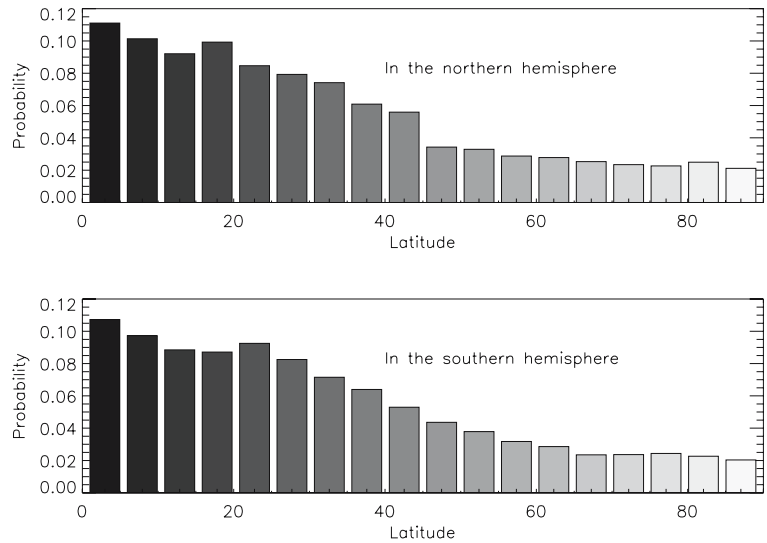

Figure 2. Percentages of CMEs to occur at the 18 latitude bands in the northern and southern hemispheres in 23 cycles.

We divide CMEs into 18 latitudinal bands (one band per 5°, from 0° to 90°) in the northern and southern hemispheres respectively and count the probability percentages of CMEs within each band. The result of this binning is shown in Fig. 2. Then we inspect whether the latitudinal distributions in the northern and southern hemispheres have a statistically significant difference or not. The method used here is the so-called rank sum test (Liu 1996). Firstly, we take the number difference between the corresponding latitudinal bands in the northern and southern hemispheres. The rank sum test has been made by first arranging the 18 points in decreasing order after taking their absolute values to assign a serial number from 1, 2, 3, 4, ... to ≤ 18. If number n and number $n + 1$ points have the same value, then both share the same serial number $n + 0.5$, the average of their serial numbers. If three points have the same value, then the three points share the same serial number, the average of their serial numbers, and so on. It is found that the sum of the serial numbers of the points with a positive sign is $T_+ = 25$ and the number of the points with a positive sign is $n_+ = 2$, while the sum of the serial numbers of the points with a minus sign is $T_- = 146$ and the number of the points with a minus sign is $n_- = 16$. Accordingly, we have

$$U_+ = n_+ n_- + \frac{n_+(n_+ + 1)}{2} - T_+ = 10,$$

$$U_- = n_+ n_- + \frac{n_-(n_- + 1)}{2} - T_- = 22.$$

$U_+ = 10 > U_\alpha = 6$, where U_α is the critical value of the rank sum test at the level 0.05 (Liu 1996). Therefore, it is concluded that there is no difference between the two distributions from the statistical point of view.

Then we grouped CMEs with projected heliographic latitudes $\geq 10^\circ$ as the CME events in the northern hemisphere and those with projected heliographic latitudes $\leq -10^\circ$ as CME events in the southern hemisphere. We can obtain $n_+ = 2, T_+ = 22, n_- = 14, T_- = 114, U_+ = 9, U_- = 19$, and $U_+ = 9 > U_\alpha = 5$. That is to say, it is concluded that there is no difference between the two distributions from the statistical point of view.

4. Speed distribution of CMEs in the northern and southern hemispheres

In that catalog, the CME plane-of-sky speeds are determined from both linear and quadratic fits to the height-time measurements. Yurchyshyn *et al.* (2005), through analyzing the linear (constant speed) fit, found that constant speed is preferable for 90% of all analyzed CMEs. In addition, Zhang & Dere (2006) found that the acceleration of CME in the outer coronal region (larger than $2\,R_\odot$) can be almost neglected in contrast to that in the inner coronal (less than $2\,R_\odot$). LASCO has three telescopes C1, C2, and C3, but the catalog contains all the CMEs detected by C2 and C3, which cover a combined field of view of $2.1\,R_\odot$–$32\,R_\odot$, because C1 was disabled in June 1998. That is, the speed observed is almost constant in the field of view of $2.1\,R_\odot$–$32\,R_\odot$. So, when we analyze the distribution of CME speeds, constant speeds are adopted.

Firstly, we divide all CMEs into those in the northern and southern hemispheres according to their latitudes. Then, we plot the 5-point smoothed distributions of $\ln(v)$ in the northern and southern hemispheres respectively in Fig. 3. We find that the statistical distributions of $\ln(v)$ in the northern and southern hemispheres are all very similar to Gaussian distribution, that is, the statistical distribution of speeds of CMEs should be very similar to a lognormal distribution. When a random variable, in our case the speed of a CME, v, is lognormally distributed, its natural logarithm $\ln(v)$ is normally distributed (Yurchyshyn *et al.* 2005):

$$P = A_0 \exp -\frac{1}{2}\left(\frac{\ln(v) - \mu}{\sigma}\right)^2, \tag{3}$$

where P is the probability density function and σ^2 and μ are the variance and the mean of $\ln(v)$. The distributions of $\ln(v)$ are then approximated with a normal fit by the equation (3). The solid line and the dashed line in Fig. 3 are the normal fit to the distributions of $\ln(v)$ for the CMEs in the northern and southern hemispheres respectively. The fitting parameters for two normal distributions are collected in Table 1, where the last four columns show the goodness of the fit, χ^2, probability, p, reduced χ^2, and degree of freedom, dof.

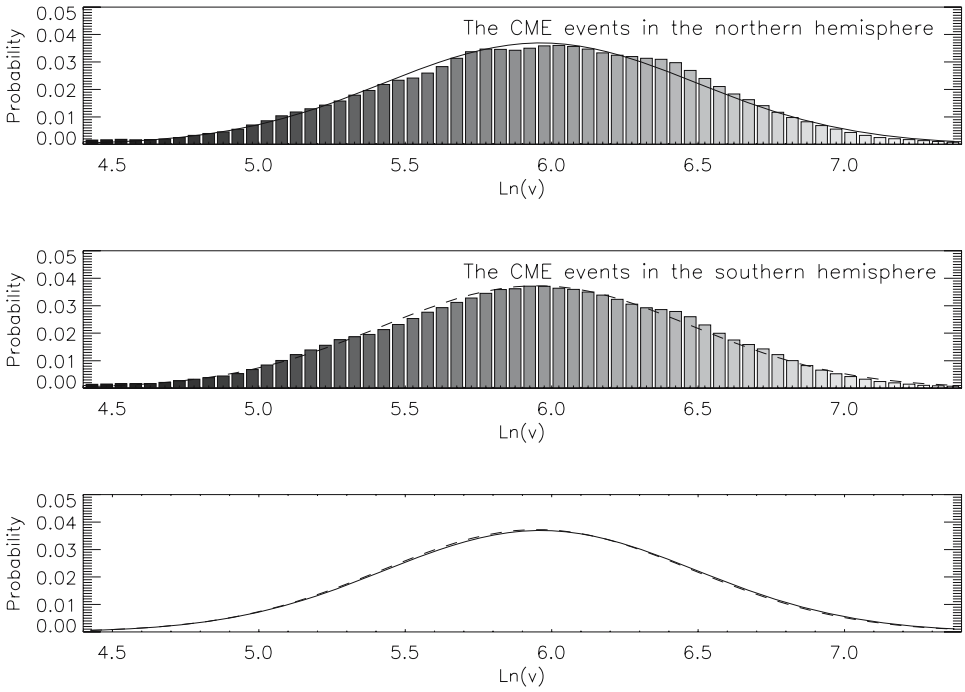

Figure 3. The 5-point smoothed distribution of $\ln(v)$ in the northern (top panel, vertical bars) and southern hemispheres (second panel,vertical bars) observed by SOHO/LASCO in cycle 23 from September 1996 to December 2006. The solid line in the top panel and the dashed line in the second panel show the corresponding Gaussian fits. The bottom panel shows the same Gaussian fits as in the top (solid line) and the second panels (dashed line).

Table 1. Fitting parameters.

	A_0	$\ln(v)$	μ(km s^{-1})	σ	χ^2	P	Reduced χ^2	dof
Northern hemisphere	0.0369	5.9619	388.3	0.5328	36.9	0.98	0.65	57
Southern hemisphere	0.0373	5.9510	384.1	0.5285	31.2	0.99	0.55	57

From Fig. 3 and Table 1, we find that the normal approximations for the $\ln(v)$ distribution of CMEs in the northern and southern hemispheres are nearly identical ($\chi^2 = 34.6400$, with degrees of freedom $= 57$ and probability $p = 0.99$). And the speed distribution in the northern and southern can be fitted with a single lognormal distribution. This finding implies that, statistically, there is no physical distinction between the CME events in the northern and southern hemispheres. The lognormal distribution of the CME speeds suggests that the same driving mechanism of a nonlinear nature is acting in the two groups of CME events (Abramenko *et al.* 2003; Kaufmann *et al.* 2003).

Then we grouped CMEs with projected heliographic latitudes $\geq 10^\circ$ as the CME events in the northern hemisphere and those with projected heliographic latitudes $\leq -10^\circ$ as CME events in the southern hemisphere. We also find that the speed

Table 2. Fitting parameters.

	A_0	$\ln(v)$	$\mu(\mathrm{kms}^{-1})$	σ	χ^2	P	Reduced χ^2	dof
Northern hemisphere	0.0369	5.9713	392.0	0.5317	41.2	0.94	0.72	57
Southern hemisphere	0.0371	5.9577	386.7	0.5299	31.4	0.99	0.55	57

distribution in the northern and southern hemispheres are nearly identical ($\chi^2 = 35.9366$, with degrees of freedom $= 57$ and probability $p = 0.99$) and to a good approximation they can be fitted with a single lognormal distribution. The fitting parameters are collected in Table 2. When the CMEs at the band -10°–10° are excluded, the fitting parameters have a slight change, such as, that the parameter μ that descries the position of the maximum in a distribution is slightly greater than that when all the CMEs are considered.

5. Conclusions and discussions

In the present work, the distribution of latitudes and speeds of Coronal Mass Ejections (CMEs) in the northern and southern hemispheres in cycle 23, from September 1996 to December 2006, has been analyzed. The following conclusions are obtained:

- Through calculating the actual probability of the hemispheric distribution of the activity of the CME, we find that a southern dominance of the activity of the CME is shown to occur in cycle 23 from September 1996 to December 2006.
- In a solar cycle, the CME activity occurs at all latitudes and is most common at low latitudes.
- The latitudinal distribution of CMEs in the northern and southern hemispheres are no different from the statistical point of view.
- The speed distribution in the northern and southern hemispheres are nearly identical and to a good approximation they can be fitted with a single lognormal distribution.

By calculating the actual probability of the hemispheric distribution of the activity of CMEs, we find that a southern dominance of the activity of CMEs is shown to occur in cycle 23 from September 1996 to December 2006. This conclusion was predicted by Ataç and Özgüc in 1996 and confirms that the hypothesis about a 12-cycle periodicity is valid (Li *et al.* 2002). The physical interpretation of the asymmetry in the solar activity has not been clearly given so far. Waldmeier (1971) found that the N–S asymmetries might be caused by phase differences between the magnetic activity in both hemispheres. Haber *et al.* (2002) found that the meridional flows in the northern and southern hemispheres exhibit a striking asymmetry during the years 1998–2001 (ascending period of cycle 23). Zhao & Kosovichev (2004) have recently concluded that the extremely rich and complicated dynamics of the upper convection zone reveals remarkable organization on the large scale, which can be correlated with the magnetic activity zones. They found that the zonal flows in the southern and northern hemispheres are not symmetrical.

The CME activity occurs at all latitudes and is most common at low latitudes. The statistical characteristics of such latitudinal distribution of CMEs are different

from those of sunspot groups. The latitudinal distribution of sunspot groups can be represented by a probability density function of the Γ distribution having maximum probability at about 15.5° (Li *et al.* 2003). We know that sunspots seem to be specified on active regions (AR), but CMEs are a kind of large-scale solar activity. So, our conclusion – the statistical characteristics of the latitudinal distribution of CMEs are different from those of sunspot groups – should furnish evidence to support that CMEs are intrinsically associated with source magnetic structures on a large spatial scale, even with transequatorial source magnetic structures on a large spatial scale (Zhou *et al.* 2006). The latitudinal distribution of CMEs in the northern and southern hemispheres are no different from the statistical point of view.

We find that the speed distributions in the northern and southern hemispheres are nearly identical and to a good approximation they can be fitted with a single lognormal distribution. This finding implies that the presence of nonlinear interactions and multiplicative processes in a system of many flux loops (Yurchyshyn *et al.* 2005), and statistically, there is no physical distinction between the CME events in the northern and southern hemispheres. The lognormal distribution of the CME speeds suggests that the same driving mechanism of a nonlinear nature is acting in the two groups of CME events.

Although we find that a southern dominance of the activity of CMEs is shown to occur in cycle 23 from September 1996 to December 2006, the latitudinal distribution of CMEs in the northern and southern hemispheres are no different from the statistical point of view and the speed distribution in the northern and southern hemispheres are nearly identical. The result should furnish evidence to support that the N–S asymmetry of the CME events is related to the N–S asymmetry of dynamo theory.

It must be pointed out that, in this paper, we adopt the plane-of-sky speeds when we analyze the speed distribution of CMEs in the northern and southern hemispheres. Recently, for non-halo CMEs, Hundhausen *et al.* (1994), Leblanc (2001) and Yeh *et al.* (2005) developed a method to correct the projection effect and obtain the real CME speeds. Further efforts are being undertaken to inspect whether the real speed distributions in the northern and southern hemispheres are nearly identical and to a good approximation whether they can be fitted with a single lognormal distribution or not.

Acknowledgements

The authors thank the referee for useful and helpful comments. This CME catalog is generated and maintained at the CDAW Data Center by NASA and The Catholic University of America in co-operation with the Naval Research Laboratory. SOHO is a project of international cooperation between ESA and NASA. This work is supported by the 973 project (2006CB806300), the Natural Science Funds of China (40636031 and 10573034), and the Chinese Academy of Sciences.

References

Abramenko, V. I., Yurchyshyn, V. B., Wang, H., Spirock, T. J., Goode, P. R. 2003, *ApJ*, **597**, 1135.

Aoki, S. I., Yashiro, S., Shibata, K. 2003, In: *Proc. 28th Int. Cosmic Ray Conf.* (eds) T. Kajita *et al.* (Japan) 2729.

Ataç, T., Özgüç, A. 1996, *Solar Phys.*, **166**, 201.
Ataç, T., Özgüç, A. 1998, *Solar Phys.*, **180**, 397.
Cane, H. V., Richardson, I. G. 2003, *JGR*, **108, A4**, 1156.
Gopalswamy, N., Lara, A., Lepping, R. P., Kaiser, M. L., Berdichevsky, D., St. Cry, O. C. 2000, *GRL*, **27**, 145.
Gopalswamy, N., Lara, A., Yashiro, S., Kaiser, M. L., Howard, R. A. 2001, *JGR*, **106**, 29207.
Gopalswamy, N., Lara, A., Yashiro, S., Howard, R. A. 2003, *ApJ*, **598**, L63.
Harvey, K. L., White, O. R. 1999, *JGR*, **104**, 19759.
Haber, D. A., Hindman, B. H., Toomer, J. 2002, *ApJ*, **570**, 855.
Howard, R. A., Sheeley, Jr. N. R., Koomen, M. J., Michels, D. J. 1985, *JGR*, **90**, 8173.
Hundhausen, A. J. 1993, *JGR*, **98**, 13177.
Hundhausen, A. J., Burkepile, J. T., St. Cry, O. C. 1994, *JGR*, **99**, 6543.
Jiang, J., Chatterjee, P., Choudhuri, A. R. 2007, *MNRAS*, **381**, 1527.
Kaufmann, P., Gimenez de Castro, C. G., Makhmutov, V. S., Raulin, J.-P., Schwenn, R., Levato, H., Rovira, M. 2003, *JGR*, **108**, 1280.
Leblanc, Y., Dulk, G. A., Vourlidas, A., Bougeret, J. L. 2001, *JGR*, **106**, 25301.
Li, K. J., Schmieder, B., Li, Q.-Sh. 1998, *Astron. Astrophys. Suppl. Ser.*, **131**, 99.
Li, K. J., Wang, J. X., Xiong, S. Y., Liang, H. F., Yun, H. S., Gu, X. M. 2000, *A&A*, **383**, 648.
Li, K. J., Irie, M., Wang, J. X., Xiong, S. Y., Yun, H. S., Liang, H. F., Zhan, L. S., Zhao, H. J. 2002, *PASJ*, **54**, 787.
Li, K. J., Wang, J. X., Zhan, L. S., Yun, H. S., Liang, H. F., Zhao, H. J., Gu, X. M. 2003, *Solar Phys.*, **215**, 99.
Liu, H. L. 1996, Tong Ji Xue Jiao Cheng, Shanghai: Shanghai Finance Univ. Press.
Michalek, G., Gopalswamy, N., Lara, A., Manoharan, P. K. 2004, *A&A*, **423**, 729.
Tousey, R. 1973, The solar corona, In: *Space Research XIII* (eds) Rocroft, M. J., Runcorn, S. K. (publishers, Akademie-Verlag, Berlin) 713.
Wang, Y. M., Ye, P. Z., Wang, S., Zhou, G. P., Wang, J. X. 2002, *JGR*, **107, A11**, 1340.
Waldmeier, M. 1971, *Solar Phys.*, **20**, 332.
Yashiro, S., Gopalswamy, N., Michalek, G., St. Cry, O. C., Plunkett, S. P., Rich, N. B., Howard, R. A. 2004, *JGR*, **109**, A07105.
Yeh, Chin-teh, Ding, M. D., Chen, P. F. 2005, *Solar Phys.*, **229**, 313.
Yurchyshyn, V., Yashiro, S., Abramenko, V., Wang, H., Gopalswamy, N. 2005, *ApJ*, **619**, 599.
Zhang, J., Dere, K. P. 2006, *ApJ*, **649**, 1100.
Zhao, J. W., Kosovichev, A. G. 2004, *ApJ*, **603**, 776.
Zhou, G. P., Wang, J. X., Zhang, J. 2006, *A&A*, **445**, 1133.

J. Astrophys. Astr. (2007) **28,** 217–230

Velocity Curve Analysis of the Spectroscopic Binary Stars PV Pup, HD 141929, EE Cet and V921 Her by Nonlinear Regression

K. Karami[1,2,3,*] & R. Mohebi[1,**]
[1]*Department of Physics, University of Kurdistan, Pasdaran St., Sanandaj, Iran.*
[2]*Research Institute for Astronomy & Astrophysics of Maragha (RIAAM), Maragha, Iran.*
[3]*Institute for Advanced Studies in Basic Sciences (IASBS), Gava Zang, Zanjan, Iran.*
[*]*e-mail: karami@iasbs.ac.ir*
[**]*e-mail: rozitamohebi@yahoo.com*

Received 2006 November 15; revised 2007 November 20; accepted 2007 November 30

Abstract. We use the method introduced by Karami & Mohebi (2007), and Karami & Teimoorinia (2007) which enable us to derive the orbital parameters of the spectroscopic binary stars by the nonlinear least squares of observed *vs.* curve fitting (o–c). Using the measured experimental data for radial velocities of the four double-lined spectroscopic binary systems PV Pup, HD 141929, EE Cet and V921 Her, we find both the orbital and the combined spectroscopic elements of these systems. Our numerical results are in good agreement with those obtained using the method of Lehmann-Filhés.

Key words. Stars: binaries: eclipsing—stars: binaries: spectroscopic.

1. Introduction

Determining the orbital elements of binary stars helps us to obtain necessary information such as the mass and radius of stars which play important roles during the evolution of stellar structures. Analyzing both the light and radial velocity curves deduced from the photometric and spectroscopic observations respectively, helps to derive the orbital parameters. One of the usual rules to analyze the velocity curve is the method of Lehmann-Filhés, see Smart (1990). Karami & Mohebi (2007), and Karami & Teimoorinia (2007) introduced a new method to derive the orbital parameters by the nonlinear regression of observed *vs.* curve fitting, hereafter (o–c). Here we use their method to obtain the orbital elements of the four double-lined spectroscopic binary systems PV Pup, HD 141929, EE Cet and V921 Her which have the following properties.

PV Pup is a detached double-lined late A-type eclipsing binary of fairly short period $P = 1.67$ days and small orbital eccentricity ($e = 0.05$). The spectral type is A0V and A2V for the primary and secondary component, respectively. The mean effective temperature for primary is 6920 K and for secondary is 6931 K. The angle of inclination is $83.09 \pm 18^{\circ}$ (Vaz & Andersen 1984). HD 141929 is a double-lined spectroscopic binary with a period of $P = 49.699$ days. The orbit is eccentric ($e = 0.393$). The

effective temperature for both components is estimated to 9500 ± 250 K. Both components have the same spectral type A0/1V, however, the secondary is rotating slower than the primary and the inclination of the orbit is about 11° (Carrier 2002). EE cet is contact double-lined spectroscopic binary with a period of $P = 0.339917$ days (Rucinski *et al.* 2002a,b). V921 Her belongs to the A type of contact binaries, with the more massive component eclipsed during the deeper minimum. The spectral type is A7IV with a period of $P = 0.877$ days (Rucinski *et al.* 2003a,b).

This paper is organized as follows. In section 2, we reduce the problem to solving an equation which is a nonlinear function in terms of the orbital parameters. In section 3, the nonlinear regression technique for estimating the orbital elements is discussed. In section 4, the numerical results implemented for the four different binary systems are reported. Section 5 is devoted to conclusions.

2. Formulation of the problem

The radial velocity of star in a binary system is defined as follows:

$$\mathrm{RV} = V_{cm} + \dot{Z}, \tag{1}$$

where V_{cm} is the radial velocity of the center of mass of system with respect to the Sun and

$$\dot{Z} = K[\cos(\theta + \omega) + e\cos\omega], \tag{2}$$

is the radial velocity of star with reference to the center of mass of the binary, see Smart (1990). In equation (2), the dot denotes the time derivative and θ, ω and e are the angular polar coordinate (true anomaly), the longitude of periastron and the eccentricity, respectively. Note that the quantities θ and ω are measured from the periastron point and the spectroscopic reference line (plane of sky), respectively. Also

$$K = \frac{2\pi}{P}\frac{a\sin i}{\sqrt{1-e^2}}, \tag{3}$$

where P is the period of motion and inclination i is the angle between the line of sight and the normal of the orbital plane.

Observation shows that the photometric phase, ϕ, which is measured from the photometric reference point (line of sight), is a measurable quantity. Hence, one has to try to express θ appearing in equation (2), in terms of ϕ. But the main difficulty here is that this is not so easy in practice, unless in the two following cases:

i) Following Smart (1990) and Rucinski (2002), the photometric (orbital) phase, ϕ, is generally related to the eccentric anomaly, ψ, according to Kepler's equation as

$$\psi - e\sin\psi = 2\pi\left(\frac{t - T_0}{P}\right) = 2\pi\phi, \tag{4}$$

where T_0 is moment of the primary eclipse. Both T_0 and P is usually taken from literature sources and is fixed. Also θ and ψ satisfy the following relation

$$\tan\frac{\theta}{2} = \sqrt{\frac{1+e}{1-e}}\tan\frac{\psi}{2}. \tag{5}$$

To obtain ψ for a given ϕ, one may solve equation (4) numerically. Then θ is derived from equation (5), however e should be fixed before. So we see that deriving θ in terms of ϕ depends on knowing the eccentricity. For a small eccentricity, $e < 1$, expanding equations (4) and (5) reduces to the following relation

$$\theta = 2\pi\phi + \left(2e - \frac{e^3}{4}\right)\sin(2\pi\phi) + \frac{5}{4}e^2\sin(4\pi\phi) + \frac{13}{12}e^3\sin(6\pi\phi) + O(e^4). \tag{6}$$

Equation (6) shows that when $e \ll 1$, then one can use $\theta = 2\pi\phi$ in equation (2).

ii) According to Budding (1993), for most normal eclipsing binaries, the line of sight inclines at a low angle to the orbital plane. This yields to the relationship between θ, ϕ and ω as follows:

$$\theta = 2\pi\phi - \omega \pm \frac{\pi}{2}, \tag{7}$$

where $\pm$ refers to the primary and the secondary components, and hereafter are shown by subscripts p and s, respectively. Therefore equations (1) and (2) for the two components reduce to

$$V_{r,p} = V_{cm} + K_p(e\cos\omega_p - \sin 2\pi\phi),$$
$$V_{r,s} = V_{cm} + K_s(e\cos\omega_s + \sin 2\pi\phi). \tag{8}$$

For a circle like orbit, in which its eccentricity is much less than unity, neglecting the term $e\cos\omega$ in equations (8) reduces to two simple sine curves. See equation (5) in Rucinski (2002).

To avoid the mentioned difficulties in obtaining θ in terms of ϕ, we try to remove it in our equations. To do this, first we take the time derivative of equation (2) as

$$\ddot{Z} = -K\sin(\theta + \omega)\dot{\theta}. \tag{9}$$

Then, using Kepler's second law and the relations obtained from the orbital parameters in the inverse-square field as

$$\dot{\theta} = \frac{h}{r^2}, \tag{10}$$

$$h = \frac{2\pi}{P}a^2\sqrt{1-e^2}, \tag{11}$$

$$r = \frac{a(1-e^2)}{1+e\cos\theta}, \tag{12}$$

one may show that equation (9), yields to

$$\ddot{Z} = \frac{-2\pi K}{P(1-e^2)^{3/2}}\sin(\theta+\omega)(1+e\cos\theta)^2, \tag{13}$$

where r, a and h are the radial polar coordinate, the semi-major axis of the orbit and the angular momentum per unit of mass, respectively.

Using equation (2), one can remove θ from equation (13) as follows:

$$P\ddot{Z} = \frac{-2\pi K}{(1-e^2)^{3/2}} \sin\left(\cos^{-1}\left(\frac{\dot{Z}}{K} - e\cos\omega\right)\right) \times \left\{1 + e\cos\left(-\omega + \cos^{-1}\left(\frac{\dot{Z}}{K} - e\cos\omega\right)\right)\right\}^2. \tag{14}$$

To simplify the notation further, we let $Y = P\ddot{Z}$ and $X = \dot{Z}$. Equation (14) describes a nonlinear relation, $Y = Y(X, K, e, \omega)$, in terms of the orbital elements K, e and ω. Using the nonlinear regression of equation (14), one can estimate the parameters K, e and ω, simultaneously.

One may show that the adopted spectroscopic elements are related to the orbital parameters. First, according to definition of the center of mass, the mass ratio in the system is obtained as

$$\frac{m_p}{m_s} = \frac{a_s \sin i}{a_p \sin i}. \tag{15}$$

According to Kepler's third law and equation (15), the following relation

$$m_p \sin^3 i = a_s \sin i \left(\frac{a_p \sin i + a_s \sin i}{P}\right)^2, \tag{16}$$

is obtained where a, P and m are expressed in AU, years and solar mass, respectively. A similar relation is obtained for the secondary component only by replacing p with s and *vice versa*, in equation (16). Note that in equations (15) and (16) parameter $a \sin i$ is related to the orbital parameters by the aid of equation (3).

3. Nonlinear least squares of (o–c)

To obtain the orbital parameters K, e and ω in equation (14), we use the nonlinear regression method. In this approach, the sum of squares of errors (SSE) for the number of N measured data is calculated as

$$\mathrm{SSE} = \sum_{i=1}^{N} (Y_i - \hat{Y}_i)^2 = \sum_{i=1}^{N} [Y_i - Y(X_i, K, e, \omega)]^2, \tag{17}$$

where Y_i and $\hat{Y}_i$ are the real and the predicted values, respectively. To obtain the model parameters, the SSE should be minimized in terms of K, e and ω as

$$\frac{\partial \mathrm{SSE}}{\partial K} = \frac{\partial \mathrm{SSE}}{\partial e} = \frac{\partial \mathrm{SSE}}{\partial \omega} = 0. \tag{18}$$

To solve equation (18), we use the SAS (Statistical Analysis System) software. Note that the nonlinear models are more difficult to specify and estimate than linear models.

For instance, in contrast to the linear regression, the nonlinear models are very sensitive to the initial guesses for the parameters. Because in practice, SSE might have to be minimized in the several points in the three dimensional parametric space including K, e and ω. However the final goal is finding the absolute minimum. Hence choosing the relevant initial parameters yields to the absolute minimum of SSE which is also stationary. This means that if one changes the initial guesses slightly, then the result reduces to the previous values for the parameters. But if SSE converged at the local minimum, the model would not be stationary. See Sen & Srivastava (1990) and Christensen (1996). Whatever the method used to determine the elements, it will always be necessary to test the elements as derived and make small changes in them before they can be regarded as a satisfactory representation of the observations. To do this a test is quickly computed for ten or twelve points of the RV curve by substituting the derived orbital elements in equations (1), (2), (4) and (5). See Karami & Teimoorinia (2007).

4. Numerical results

Here we use the method introduced by Karami & Mohebi (2007), and Karami & Teimoorinia (2007) to derive both the orbital and combined elements of the four different double-lined spectroscopic systems PV Pup, HD 141929, EE Cet and V921 Her. Using the measured experimental data for radial velocities of the two components of these systems obtained by Vaz & Andersen (1984) for PV Pup, Carrier (2002) for HD 141929, Rucinski *et al.* (2002a,b) for EE Cet and Rucinski *et al.* (2003a,b) for V921 Her, the suitable fitted velocity curves are plotted in terms of the photometric phase in Figs. 1, 2, 3 and 4, respectively.

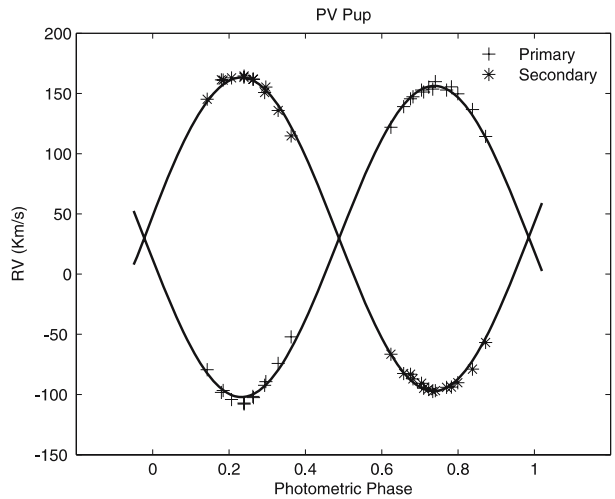

Figure 1. Radial velocities of the primary and secondary components of PV Pup plotted against the photometric phase. The observational data have been derived from Vaz & Andersen (1984).

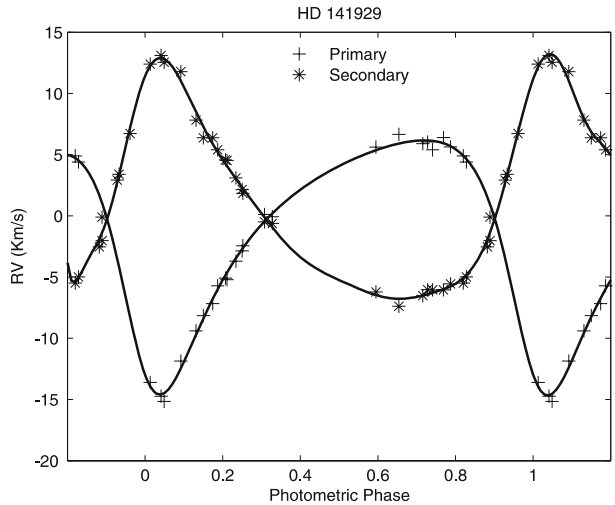

Figure 2. Same as Fig. 1 but for HD 141929. The observational data have been derived from Carrier (2002).

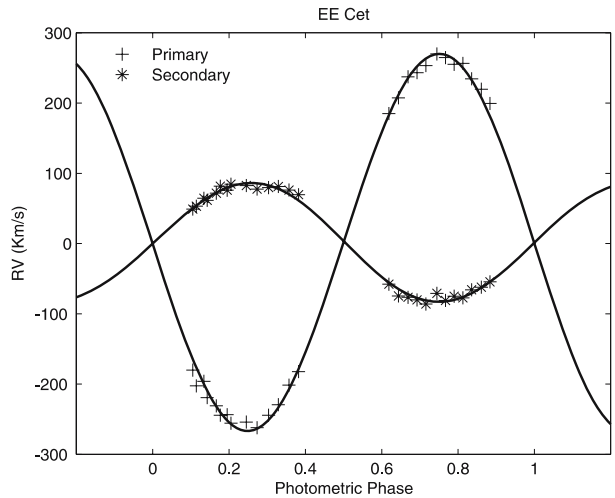

Figure 3. Same as Fig. 1 but for EE Cet. The observational data have been derived from Rucinski *et al.* (2002a,b).

The scaled radial acceleration values, $P\ddot{Z} = P\mathrm{d}\dot{Z}/\mathrm{dt} = \mathrm{dRV}/\mathrm{d}\phi$, corresponding to the radial velocity data in equation (14), are obtained only by taking the phase derivative of the RV without knowing the period P. To do this we implemented two methods:

1) A numerical differentiation method such as *Forward*, *Central* and *Extrapolated* differences, depending on desirable precision, the $P\ddot{Z}$ data corresponding with the RV data are obtained.

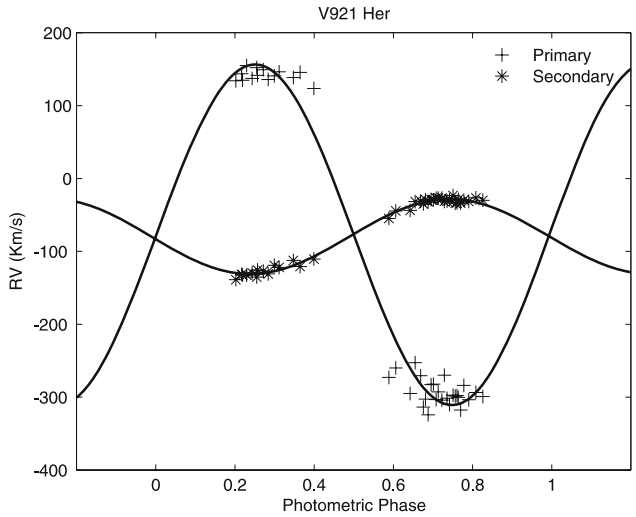

Figure 4. Same as Fig. 1 but for V921 Her. The observational data have been derived from Rucinski *et al.* (2003a,b).

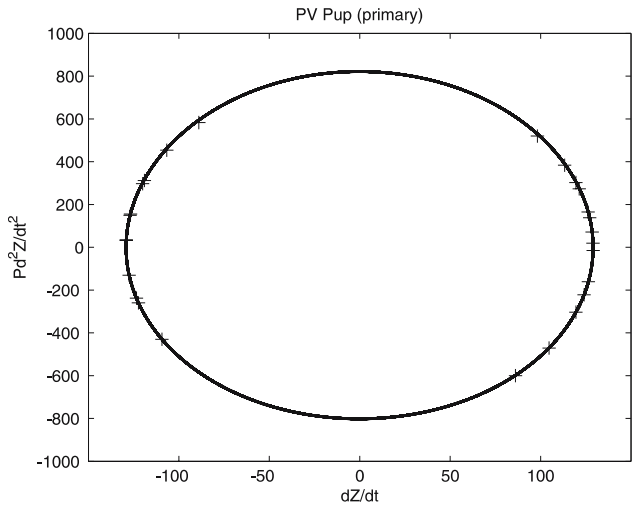

Figure 5. The radial acceleration scaled by the period *versus* the radial velocity of the primary component of PV Pup. The solid curve is obtained from the nonlinear regression of equation (14). The plus points are the experimental data.

2) A *Curve Fitting* program: the suitable velocity curves are fitted on the RV data and then one can obtain the corresponding $P\ddot{Z}$ data. Our experience shows that **the** accuracy of both approaches is very similar but the time consumed in the second approach is considerably shorter than in the first one.

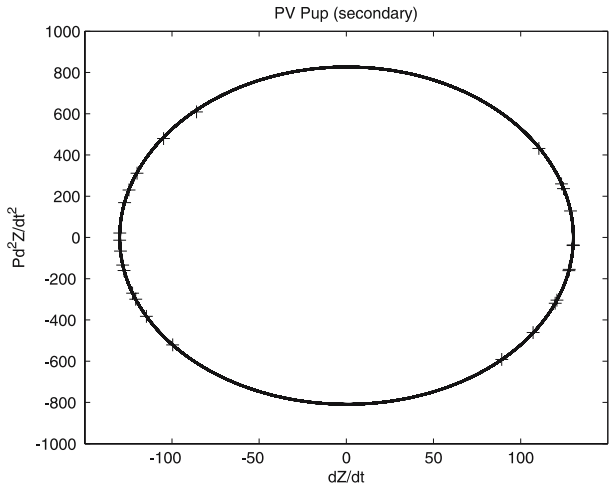

Figure 6. Same as Fig. 5 but for the secondary component of PV Pup.

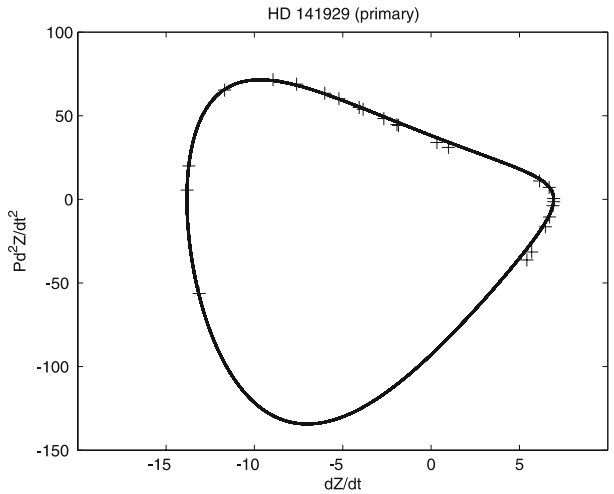

Figure 7. Same as Fig. 5 but for the primary component of HD 141929.

Figures 5–12 show the radial acceleration scaled by the period *versus* the radial velocity for the primary and secondary components of PV Pup, HD 141929, EE Cet and V921 Her, respectively. The solid closed curves are the result of the nonlinear regression of equation (14), which their good coincidence with the measured data yields to derive the optimized parameters K, e and ω. Figures show that also for PV Pup, EE Cet and V921 Her due to having small eccentricities, their radial velocity–acceleration curves display an elliptical shape. But in contrast for HD 141929 which is an eccentric system, its radial velocity–acceleration curve shows a noticeable deviation from an elliptic.

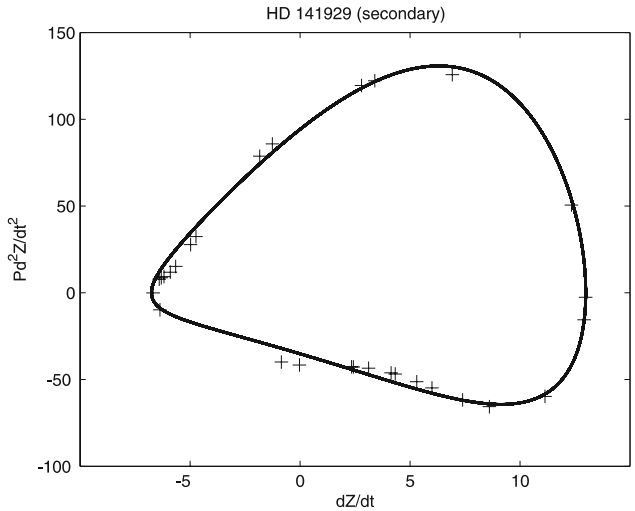

Figure 8. Same as Fig. 5 but for the secondary component of HD 141929.

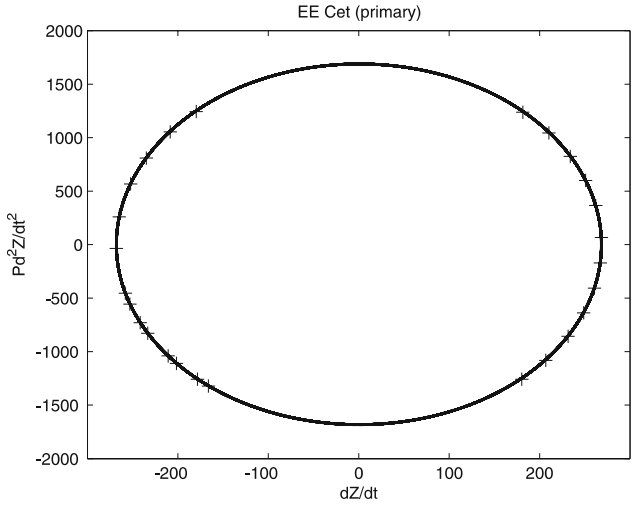

Figure 9. Same as Fig. 5 but for the primary component of EE Cet.

The orbital parameters, K, e and ω, obtained from the nonlinear least squares of equation (14) for PV Pup, HD 141929, EE Cet and V921 Her, are tabulated in Tables 1, 3, 5 and 7, respectively. The velocity of center of mass, V_{cm}, is obtained by calculating the areas above and below the radial velocity curve. Where these areas become equal to each other, the velocity of center of mass is obtained. Tables 1, 3, 5 and 7 show that the results are in good accordance with those obtained by Vaz & Andersen (1984) for PV Pup, Carrier (2002) for HD 141929, Rucinski *et al.* (2002a,b) for EE Cet and Rucinski *et al.* (2003a,b) for V921 Her.

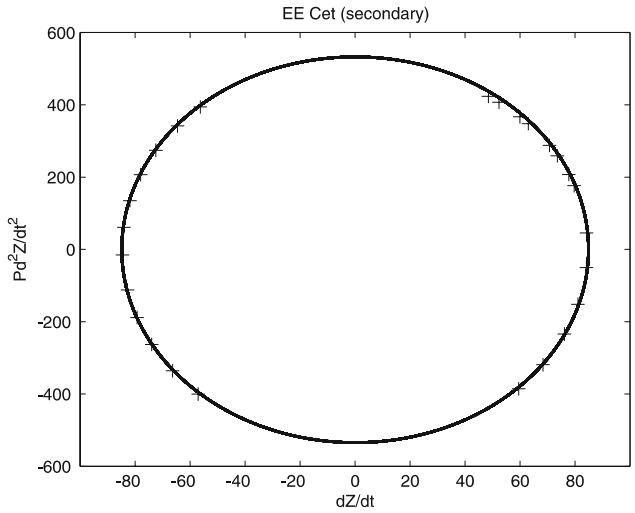

Figure 10. Same as Fig. 5 but for the secondary component of EE Cet.

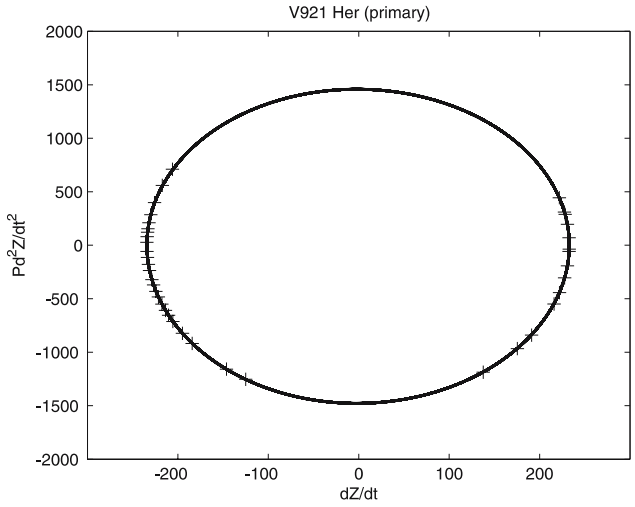

Figure 11. Same as Fig. 5 but for the primary component of V921 Her.

The combined spectroscopic elements including $m_p \sin^3 i$, $m_s \sin^3 i$, $(a_p + a_s) \sin i$ and m_p/m_s are calculated by substituting the estimated parameters K, e and ω in equations (3), (15) and (16). The results obtained for the same four previous systems are tabulated in Tables 2, 4, 6 and 8. Tables show that our results are in good agreement with those obtained by Vaz & Andersen (1984), Carrier (2002), Rucinski *et al.* (2002, 2003) for PV Pup, HD 141929, EE Cet and V921 Her, respectively.

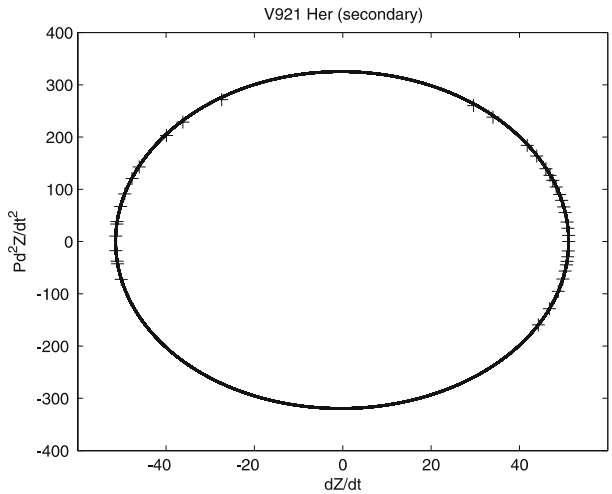

Figure 12. Same as Fig. 5 but for the secondary component of V921 Her.

Table 1. Spectroscopic orbit of PV Pup.

	This paper	Vaz & Andersen (1984)
Primary		
V_{cm}(km s^{-1})	27.16 ± 0.63	31 ± 0.5
K_p(km s^{-1})	129.14 ± 0.92	130 ± 0.5
e	0.01 ± 0.01	0.05 (fixed)
$\omega(^\circ)$	96.33 ± 0.24	149.7 (fixed)
Secondary		
V_{cm}(km s^{-1})	32.93 ± 0.85	29 ± 0.4
K_s(km s^{-1})	130.16 ± 0.39	130.9 ± 0.4
e	0.01 ± 0.01	0.05 (fixed)
$\omega(^\circ)$	278.27 ± 0.29	329.7 (fixed)

Table 2. Combined spectroscopic orbit of PV Pup.

Parameter	This paper	Vaz & Andersen (1984)
$m_p \sin^3 i/\mathrm{M}_\odot$	1.514 ± 0.019	1.531 ± 0.011
$m_s \sin^3 i/\mathrm{M}_\odot$	1.502 ± 0.026	1.52 ± 0.013
$(a_p + a_s) \sin i/\mathrm{R}_\odot$	8.55 ± 0.04	8.55 ± 0.022
m_s/m_p	0.992 ± 0.009	0.993 ± 0.005

Table 3. Same as Table 1 but for HD 141929.

	This paper	Carrier (2002)
Primary		
V_{cm}(km s^{-1})	-0.44 ± 0.12	-0.33 ± 0.08
K_p(km s^{-1})	10.38 ± 0.01	10.58 ± 0.16
e	0.391 ± 0.001	0.393 ± 0.008
$\omega(^\circ)$	148.04 ± 0.34	145.7 ± 1.7
Secondary		
V_{cm}(km s^{-1})	-0.44 ± 0.12	-0.33 ± 0.08
K_s(km s^{-1})	9.87 ± 0.01	9.95 ± 0.17
e	0.389 ± 0.001	0.393 ± 0.008
$\omega(^\circ)$	324.03 ± 0.34	325.7 ± 1.7

Table 4. Same as Table 2 but for HD 141929.

Parameter	This paper	Carrier (2002)
$m_p \sin^3 i/\mathrm{M}_\odot$	0.0163 ± 0.0001	0.01681 ± 0.00064
$m_s \sin^3 i/\mathrm{M}_\odot$	0.0171 ± 0.0001	0.01789 ± 0.00067
$a_p \sin i/10^6$ km	6.53 ± 0.01	6.65 ± 0.11
$a_s \sin i/10^6$ km	6.21 ± 0.02	6.25 ± 0.11
m_p/m_s	0.95 ± 0.004	0.94 ± 0.022

Table 5. Same as Table 1 but for EE Cet.

	This paper	Rucinski *et al.* (2002)
Primary		
V_{cm}(km s^{-1})	1.52 ± 0.85	1.6(0.93)
K_p(km s^{-1})	268.29 ± 0.52	266.92(1.54)
e	0.001 ± 0.001	–
$\omega(^\circ)$	305.59 ± 13.67	–
Secondary		
V_{cm}(km s^{-1})	1.52 ± 0.85	1.6(0.63)
K_s(km s^{-1})	84.91 ± 0.07	84.05(1.24)
e	0.001 ± 0.001	–
$\omega(^\circ)$	75.77 ± 4.1	–

Table 6. Same as Table 2 but for EE Cet.

Parameter	This paper	Rucinski *et al.* (2002)
$m_p \sin^3 i/\mathrm{M}_\odot$	0.417 ± 0.002	–
$m_s \sin^3 i/\mathrm{M}_\odot$	1.317 ± 0.007	–
$(a_p + a_s) \sin i/\mathrm{R}_\odot$	2.651 ± 0.004	–
m_p/m_s	0.316 ± 0.001	0.315(5)
$(m_p + m_s) \sin^3 i/\mathrm{M}_\odot$	1.73 ± 0.01	1.706(41)

Table 7. Same as Table 1 but for V921 Her.

	This paper	Rucinski *et al.* (2003)
Primary		
$V_{cm}(\mathrm{km\,s^{-1}})$	-78.05 ± 0.65	$-79.04(1.05)$
$K_p(\mathrm{km\,s^{-1}})$	233.62 ± 0.49	227.17(1.97)
e	0.005 ± 0.004	–
$\omega(^\circ)$	132.42 ± 1.16	–
Secondary		
$V_{cm}(\mathrm{km\,s^{-1}})$	-78.05 ± 0.65	$-79.04(1.05)$
$K_s(\mathrm{km\,s^{-1}})$	51.3 ± 0.7	51.45(0.89)
e	0.009 ± 0.004	–
$\omega(^\circ)$	237.2 ± 0.8	–

Table 8. Same as Table 2 but for V921 Her.

Parameter	This paper	Rucinski *et al.* (2003)
$m_p \sin^3 i/\mathrm{M}_\odot$	0.379 ± 0.008	–
$m_s \sin^3 i/\mathrm{M}_\odot$	1.724 ± 0.018	–
$(a_p + a_s) \sin i/\mathrm{R}_\odot$	4.94 ± 0.02	–
m_p/m_s	0.2169 ± 0.0035	0.226(5)
$(m_p + m_s) \sin^3 i/\mathrm{M}_\odot$	2.103 ± 0.027	1.971(61)

5. Conclusions

To derive the orbital elements of the spectroscopic binary stars, we used the the method introduced by Karami & Mohebi (2007), and Karami & Teimoorinia (2007). This method is applicable to orbits of all eccentricities and inclination angles. In this method, the time consumed is considerably shorter than in the method of Lehmann-Filhés. It should be more accurate as the orbital elements are deduced from all points of the velocity curve instead of four as in the method of Lehmann-Filhés. The present method enables one to vary all of the unknown parameters K, e and ω simultaneously instead of one or two of them at a time. It is possible to make adjustments in the elements before the final presentation. There are some cases, in which the method of Lehmann-Filhés

is inapplicable, and in these cases the present one may be found useful. One such case would occur when observations are incomplete because certain phases could not have been observed. Another case in which this method is useful is that of a star attended by two dark companions with commensurable periods. In this case the resultant velocity curve may have several unequal maxima and the method of Lehmann-Filhés would fail altogether.

Using the measured experimental data for radial velocities of PV Pup, HD 141929, EE Cet and V921 Her obtained by Vaz & Andersen (1984), Carrier (2002), Rucinski *et al.* (2002) and Rucinski *et al.* (2003), respectively, we find the orbital elements of these systems by the mentioned method. Our numerical calculations show that the results obtained for both the orbital elements and combined spectroscopic parameters are in good agreement with the those obtained via the method of Lehmann-Filhés. In a subsequent paper we intend to test numerically our method for the other different systems.

Acknowledgements

This work was supported by the Department of Physics, University of Kurdistan, Sanandaj, Iran; the Research Institute for Astronomy & Astrophysics of Maragha (RIAAM), Maragha, Iran.

References

Budding, E. 1993, *An Introduction to Astronomical Photometry* (Cambridge Univ. Press), 199.
Carrier, F. 2002, *A&A*, **389**, 475.
Christensen, R. 1996, Analysis of Variance, Design and Regression (Chapman and Hall/CRC Press), 521.
Karami, K., Mohebi, R. 2007, *ChJAA*, **7(4)**, 558.
Karami, K., Teimoorinia, H. 2007, *Ap&SS*, **311**, 435.
Rucinski, M. S. 2002, *AJ*, **124**, 1746.
Rucinski, M. S., Capobianco, C. C., Lu, W., Mochnaki, S. W., Blake, R. M., Thomson, J. R., Ogloza, W., Stachowski, G. 2002a, *AJ*, **124**, 1738.
Rucinski, M. S., Capobianco, C. C., Lu, W., Mochnaki, S. W., Blake, R. M., Thomson, J. R., Ogloza, W., Stachowski, G. 2002b, preprint (astro-ph/0201213).
Rucinski, M. S., Capobianco, C. C., Lu, W., Debond, H., Thomson, J. R., Mochnacki, S. W., Blake, R. M., Ogloza, W., Stachowski, G., Rogoziecki, P. 2003a, *AJ*, **125**, 3258.
Rucinski, M. S., Capobianco, C. C., Lu, W., Debond, H., Thomson, J. R., Mochnacki, S. W., Blake, R. M., Ogloza, W., Stachowski, G., Rogoziecki, P. 2003b preprint (astro-ph/0302399).
Sen, A., Srivastava, M. 1990, *Regression Analysis: Theory, Methods, and Applications* (Springer-Verlag New York Inc.), 298.
Smart, W. M. 1990, Textbook on Spherical Astronomy, 6th edn, revised by Green R. M. (Cambridge Univ. Press), 360.
Vaz, L. P. R., Andersen, J. 1984, *A&A*, **132**, 219.