# DES PETITS CORPS DU SYSTÈME SOLAIRE

Jean-Marc C. Petit

# THESE

*Présentée à*

l'Université de Nice-Sophia Antipolis

*pour obtenir le diplôme d'Habilitation à Diriger des Recherches*

## DES PETITS CORPS DU SYSTEME SOLAIRE

**Jean-Marc PETIT**

*Soutenue le 19 Septembre 1996 devant la commission d'examen composée de*

| | | |
|---|---|---|
| M. | P. Farinella | |
| M. | R. Greenberg | Rapporteur |
| M. | M. Hénon | |
| M. | G. Iooss | Président |
| M. | J. Laskar | Rapporteur |
| M. | G. Valsecchi | Rapporteur |

**DES PETITS CORPS DU SYSTEME SOLAIRE**

# DOSSIER DE CANDIDATURE A L'HABILITATION A DIRIGER DES RECHERCHES

Présenté à

l'Université de Nice-Sophia Antipolis
par

**Jean-Marc PETIT**
Observatoire de la Côte d'Azur
B.P 229 06304 Nice Cedex 04

Je déclare sur l'honneur que ce dossier constitue ma première et unique candidature au diplôme d'habilitation à diriger des recherches.

Jean-Marc Petit, le 26 Mai 1996.

# *Avant propos*

Ce dossier d'habilitation rassemble une partie des travaux que j'ai effectués postérieurement à ma thèse de Doctorat.

Il contient un curriculum vitae, une liste des publications et une synthèse des travaux. Cette synthèse est présentée dans l'introduction et peut servir de guide de lecture à la sélection des articles contenus dans le document.

Je tiens ici à remercier toutes les personnes grâce auxquelles, ou à cause desquelles, ce travail a été commis.

Je remercie par dessus tout Michel Hénon, mon directeur de recherche qui m'a initié au métier de la recherche et qui, par sa grande compréhension de la physique et de la nature humaine, m'a permis de surmonter certains moments difficiles depuis mes débuts en recherche.

Je remercie aussi toutes les personnes avec qui j'ai pu collaborer et qui m'ont permis d'aborder de nouveaux sujets et d'élargir le champ de mes contacts humains. Je citerai notamment J. Binney, P. Farinella et R. Greenberg qui a de plus accepté d'être un des rapporteurs de cette thèse.

Je présente mes vifs remerciements à J. Laskar et G. Valsecchi qui ont bien voulu rapporter sur ce travail ainsi qu'à G. Iooss qui a accepté de présider la commission d'examen.

Je terminerai ces remerciements en exprimant ma gratitude à tous les amis que je côtoie régulièrement sur mes lieux de travail, que ce soit à Nice ou à Tucson, et qui font que ces moments sont si agréables.

A special note for those in Tucson who may have troubles understanding the above acknowledgements: thanks to all of you.

# CURRICULUM VITAE

**Jean-Marc PETIT**

**Renseignements personnels :**
Né le 26 Mai 1961 à Nice
Nationalité: Française
Situation de famille: divorcé, deux enfants
Adresse: 98 Boulevard de Cessole, 06100 Nice

**Situation professionnelle :**
Chargé de Recherches CNRS 1$^{\text{ère}}$classe
Secteur SDU, section 14
Observatoire de la Côte d'Azur, URA 1362

**Adresse professionnelle :**
Laboratoire G.D. Cassini URA 1362
Observatoire de la Côte d'Azur, B.P. 229, 06304 Nice Cedex 04
Téléphone : 92 00 30 89

**Diplômes :**
**Juillet 1981 :** Admis à l'Ecole Normale Supérieure (Paris)
**Septembre 1982 :** Maîtrise de Physique fondamentale (Paris)
**Juillet 1983 :** Diplôme d'Etudes Approfondies, Mécanique des Fluides (Turbulence et Systèmes Dynamiques, Nice)
**15 Mars 1985 :** Thèse de 3$^{\text{ème}}$cycle (Nice), Mécanique des Fluides. Titre : "Simulation numérique des anneaux de Saturne."
**15 Juillet 1985 :** Agrégation de Physique (Paris)

**Activités professionnelles :**
**Septembre 1985 - 1987 :** Assistant Normalien à l'Université de Nice
**Octobre 1987 - Septembre 1988 :** Chercheur post-doctoral au département de physique théorique, Université d'Oxford, Angleterre
1$^{\text{er}}$**Mars 1988 :** Chargé de Recherches CNRS 2$^{\text{ème}}$classe
1$^{\text{er}}$**Mars 1992 :** Chargé de Recherches CNRS 1$^{\text{ère}}$classe

**Septembre 1993 - 1994 :** Scientifique invité au Lunar and Planetary Laboratory, Université d'Arizona, Tucson, USA

**Encadrements :**
**1990 :** Encadrement du stage de DEA de F. Pereira Gama (Analyse statistique de l'effet de rencontres proches dans les anneaux planétaires)
**1991 :** Co-encadrement de la thèse de 3$^{\text{ème}}$cycle de F. Pereira Gama (Simulation numérique des anneaux planétaires)
**1994 :** Co-encadrement du stage de "graduate student" de T. Hurford (Dynamique du satellite de Ida)

**Enseignements :**
**Septembre 1985 - 1987 :** Cours de Méthodes Numériques et Informatique en DEUG SPI 2$^{\text{ème}}$année
**Octobre 1988 - aujourd'hui :** Cours de Physique et d'Informatique en classes préparatoires au lycée Masséna (M') et au lycée des Eucalyptus (Technologie) à Nice

**Administration :**
**Juin 88 - Avril 90 :** membre élu du conseil d'administration de l'observatoire de la Côte d'Azur, membre élu du conseil scientifique de l'Observatoire de la Côte d'Azur
**Septembre 94 - aujourd'hui :** membre nommé du Comité Informatique et Télématique de l'Observatoire de la Côte d'Azur
**Janvier 95 - aujourd'hui :** membre élu du conseil d'administration de l'Observatoire de la Côte d'Azur, membre élu du conseil scientifique de l'Observatoire de la Côte d'Azur

# LISTE des PUBLICATIONS

**Jean-Marc PETIT**

**Revues à comité de lecture :**

[1] Petit, J-M., 1985. *Simulation numérique des anneaux de Saturne.*. Thèse de Troisième cycle, Université de Nice,

[2] Hénon, M. *&* Petit, J-M., 1986. *Series expansions for encounter-type solutions of Hill's Problem.* Celestial Mechanics **38**, pp. 67–100.

[3] Petit, J-M. *&* Hénon, M., 1986. *Satellite Encounters.*. Icarus **66**, pp. 536–555.

[4] Petit, J-M. *&* Hénon, M., 1987. *A numerical simulation of planetary rings. I. Binary Encounters.*. Astron. Astrophys. **173**, pp. 389–404.

[5] Petit, J-M. *&* Hénon, M., 1987. *A numerical simulation of planetary rings. II. Monte Carlo model.*. Astron. Astrophys. **188**, pp. 198–205.

[6] Petit, J-M. *&* Hénon, M., 1988. *A numerical simulation of planetary rings. III. Mass segregation, ring confinement, and gap formation.*. Astron. Astrophys. **199**, pp. 343–356.

[7] Froeschlé, C. *&* Petit, J-M., 1990. *Polynomial approximations of Poincaré maps for Hamiltonian systems.* Astron. Astrophys. **238**, pp. 413–423.

[8] Spahn, F., Petit, J-M. *&* Bendjoya, Ph., 1993. *The gravitational influence of satellite Pan on the radial distribution of ring–particles in the region of the Encke–division in Saturn's A ring.* Celestial Mechanics and Dynamical Astronomy **57**, pp. 391–402.

[9] Petit, J-M. *&* Farinella, P., 1993. *Modelling the outcomes of high–velocity impacts between small solar system bodies.* Celestial Mechanics and Dynamical Astronomy **57**, pp. 1–28.

[10] Froeschlé, C. & Petit, J-M., 1993. *On a temporary confinement of chaotic orbits of four dimensional symplectic mapping: A test on the validity of the synthetic approach .* Celestial Mechanics and Dynamical Astronomy **57**, pp. 123–130.

[11] Bendjoya, Ph., Petit, J-M. & Spahn, F., 1993. *Wavelets analysis of the Voyager data on planetary rings. I. Description of the method.* Icarus **105**, pp. 385–399.

[12] Petit, J-M. & Froeschlé, C., 1994. *Polynomial approximations of Poincaré maps for Hamiltonian systems II..* Astron. Astrophys. **282**, pp. 291–303.

[13] Campo Bagatin, A., Farinella, P. & Petit, J-M., 1994. *Fragment ejection velocities and the collisional evolution of asteroids.* Planetary and Space Science **42**, 12, pp. 1099–1107.

[14] Belton, M., Chapman, C., Thomas, P., Davies, M., Greenberg, R., Klaasen, K., Byrnes, D., D'Amario, L., Synnott, S., Merline, W., Petit, J-M., Storrs, A., Zellner, B. & the Galileo Imaging Team, 1995. *The bulk density of asteroid 243 Ida from Dactyl's orbit.* Nature **374**, pp. 785–788.

[15] Belton, M., Mueller, B., D'Amario, L., Byrnes, D., Klaasen, K., Synnott, S., Breneman, H., Johnson, T., Thomas, P., Veverka, J., Harch, A., Davies, M., Merline, W., Chapman, C., Davis, D., Denk, T., Petit, J-M., Greenberg, R., McEwen, A., Storrs, A. & Zellner, B., 1996. *The discovery and Orbit of 1993 (243)1 Dactyl.* Icarus **120**, pp. 185–199.

[16] Geissler, P., Petit, J.M., Durda, D., Greenberg, R., Bottke, W. & Nolan, M., 1996. *Erosion and Ejecta Redistribution on 243 Ida and its Moon.* Icarus **120**, pp. 140–157.

[17] Greenberg, R., Bottke, W., Nolan, M., Geissler, P., Petit, J-M., Durda, D., Asphaug, E. & Head, J., 1996. *Collisional and Dynamical History of Ida.* Icarus **120**, pp. 106–118.

[18] Petit, J-M. & Greenberg, R., 1996. *Viscosity in Keplerian Disks: Steady-State Velocity Distribution and Non-Local Collision Effects.* Icarus **123**, pp. 524–535.

[19] Bury, P., Ennode, N., Petit, J-M., Bendjoya, Ph., Martinez, J-P., Pinna, H., Jaud, J. & Balladore, J-L., *Wavelet analysis of X-ray spectroscopic data I : the method.* 1996. Nuclear Instruments and Methods in Physics Research, Section A **383**, pp. 572–588.

**Actes de congrès :**

[20] Hénon, M. *&* Petit, J-M., 1986. *Satellite Encounters.* Bull. of the Amer. Astron. Soc. **18**, p. 755.

[21] Petit, J-M. *&* Hénon, M., 1986. *A numerical simulation of planetary rings.*. Bull. of the Amer. Astron. Soc. **18**, p. 763.

[22] Petit, J-M. *&* Hénon, M., 1988. *Numerical simulation of Ring Dynamics.* in *The Few Body Problem,* ed. M.J. Valtonen, Kluwer Academic Publishers,

[23] Binney, J. *&* Petit, J-M., 1988. in *Dense stellar systems,* ed. D. Merritt, Cambridge Univ. Press,

[24] Petit, J-M. *&* Hénon, M., 1989. *On a Cantor Structure in a Satellite Scattering Problem.* in *Dynamics and Stochastic Processes,* eds. R. Lima, L. Streit, *&* R. Vilela Mendes, Springer Verlag,

[25] Pereira Gama, F. *&* Petit, J.-M., 1991. *Statistical analysis of the effects of close encounters of particles in planetary rings.* in *Predictability, stability and chaos in N-body dynamical systems (NATO ASI SERIES),* ed. A. E. Roy, Plenum Press, **272**, pp. 345–353.

[26] Petit, J-M., 1992. *Monte Carlo methods.* in *Compte-rendu de l'école de Goutelas Physique et Dynamique des petits corps du Système Solaire Mai 1991,* pp. 599–629.

[27] Froeschlé, C. *&* Petit, J-M., 1992. *Polynomial approximations of Poincaré maps for Hamiltonian systems.* in *Asteroids, Comets, Meteors 1991,* eds. Alan W. Harris *&* Edward Bowell, pp. 201–204.

[28] Pereira Gama, F., Petit, J.-M. *&* Scholl, H., 1992. *A collisional and self-gravitational model to simulate numerically the dynamics of planetary disks.* in *Proceedings of IAU Symposium n$^o$ 152 Chaos, Resonance and collective Dynamical Phenomena in the Solar System,* pp. 109–114.

[29] Spahn, Fr., Petit, J-M. *&* Bendjoya, Ph., 1992. *Wavelet analysis of the Encke gap ringlet related to the gravitational action of the satellite Pan.* Bull. Am. Astron. Soc. **24**, p. 1036.

[30] Geissler, P., Petit, J-M. *&* Greenberg, R., 1994. *Ida Block Distribution and Origin: Interim Report.*

[31] Geissler, P., Petit, J-M. & Greenberg, R., 1994. *Ida: Distribution and Origin of Surface Blocks.* L.P.S.C. abstract

[32] Froeschlé, C., Lega, E. & Petit, J-M., 1994. *Polynomial approximation of Poincaré maps for Hamiltonian systems.* in *Evolution dynamique des systèmes gravitationels. Modèles et méthodes. Aussois 1-4 novembre 1994,*

[33] Greenberg, R. & Petit, J-M., *Viscosity in Keplerian Disks: Steady-State Velocity Distribution and Non-Local Collisional Effects.* 1994. Bull. Am. Astron. Soc. **26**, p. 1141.

[34] Geissler, P., Petit, J-M. & Greenberg, R., *Erosion and Regolith Redistribution on 243 Ida and it's Moon.* 1994. Bull. Am. Astron. Soc. **26**, p. 1157.

[35] Petit, J-M., Greenberg, R. & Geissler, P., *Orbits around a small, highly elongated asteroid: Contraints on Ida's Moon.* 1994. Bull. Am. Astron. Soc. **26**, pp. 1157–1158.

[36] Hurford. T., Durda, D., Greenberg, R., Petit, J-M., Geissler, P. & Bottke, W., 1995. *Stable periodic orbits around Ida: Several fit Dactyl's motion.* Bull. Am. Astron. Soc. **27**, pp. 1070–1071.

[37] Petit, J-M. & Bendjoya, Ph., 1996. *A new insight in Uranus rings: a wavelet analysis of the Voyager 2 data.* in *Proceedings of the Pacific Astronomical Society. June 1994,*

[38] Geissler, P., Petit, J-M. & Greenberg, R., 1996. *Ejecta Reaccretion on Rapidly Rotating Asteroids: Implications for 243 Ida and 433 Eros.* in *Proceedings of the Pacific Astronomical Society. June 1994,*

**Revues de vulgarisation :**

[39] Petit, J-M., 1990. *La courte vie des anneaux planétaires.* Sciences et Vie **170**, pp. 12–19.

# TABLE DES MATIERES

# Chapitre 1

# INTRODUCTION GENERALE

## 1.1 Présentation

Au début de mon travail de recherche, je me suis intéressé à la dynamique de différents corps célestes: particules des anneaux planétaires, galaxies. Tout d'abord, j'ai étudié le problème de Hill plan circulaire [2], [3], en collaboration avec M. Hénon. Cette approximation est bien adaptée à l'étude du mouvement de deux petits corps proches l'un de l'autre en orbites planes, quasi-circulaires, autour d'un corp massif. Nous avons ainsi déterminé des formules analytiques donnant les effets de l'interaction gravitationnelle entre deux satellites pour les petits paramètres d'impact (orbites en fer à cheval) et pour les grands paramètres d'impact. Ceci me permit d'aborder la simulation numérique des anneaux planétaires [66], [67], [6]. L'étude détaillée du problème de Hill me permit de mettre en évidence un phénomène ignoré jusque là en mécanique céleste: la diffusion chaotique (chapitre 2). Ce phénomène est lié à la présence d'orbites périodiques instables et il fut possible de montrer les liens avec certains résultats théoriques sur les points homoclines et hétéroclines. Puis, lors de mon séjour à Oxford, j'ai abordé le problème de la dynamique galactique sous deux aspects: modèles analytiques à base de fonctions de distribution dans l'espace des phases et simulations numériques du problème à N corps [23].

Les diverses études que j'avais entreprises m'ont fréquemment amené à étudier des sections de Poincaré. Cet outil puissant peut s'avérer parfois très coûteux en temps de calcul du fait des intégrations numériques d'équations différentielles. Je me suis donc intéréssé au développement d'une méthode "d'application synthétique" qui permet d'accélérer fortement le calcul d'une section de Poincaré (chapitre 3). J'ai montré que cette méthode est efficace pour des systèmes à deux et trois degrés de liberté, aussi bien en ce qui concerne la topologie de l'espace des phases que du point de vue d'indicateurs quantitatifs comme l'exposant de Lyapunov.

Depuis cette époque, j'ai élargi mon approche de l'étude des anneaux planétaires (chapitre 4). Tout d'abord, et afin de pouvoir mieux comparer les résultats des simulations aux observations, j'ai développé une méthode automatique d'analyse de données monodimensionnelles telles que les données PPS des sondes Voyager (section 4.1). Cette méthode, qui repose sur les transformées en ondelettes, donne une mesure objective de la validité des structures détectées. Ce nouvel outil permet la détermination détaillée du profil des structures des anneaux planétaires observés par les sondes Voyager I et II et une comparaison fructueuse avec les modèles numériques et semi-analytiques. Ainsi, l'étude de l'anneau de la division de Encke dans les anneaux de Saturne a permis de mieux déterminer l'excentricité du satellite Pan, et celle des données concernant Uranus de mettre en évidence trois nouveaux anneaux en plus des profils des anneaux $\epsilon$ et $\delta$.

J'ai aussi entrepris une approche par un formalisme hydrodynamique de la modélisation des anneaux planétaires (section 4.2). Il s'agit de préparer des simulations numériques dans

lesquelles les effets à grande distance sont calculés de façon déterministe et les effets à courte distance (collisions inélastiques, gravitation locale) sont pris en compte sous la forme d'une "viscosité". Le modèle réalisé évalue la viscosité dans un disque képlérien et permet de déterminer la viscosité locale (transport statistique classique) et non–locale (due à la taille finie non nulle des particules) pour des particules soumises à des collisions inélastiques sans hypothèse préalable sur la distribution des vitesses. Cette approche est particulièrement robuste pour les faibles profondeurs optiques où les simulations classiques atteignent leurs limites. Le formalisme que j'ai développé permet l'étude d'un mélange de deux populations de tailles différentes et de la gravitation entre les particules pour un coût très faible.

Les anneaux planétaires ne représentent qu'un cas parmi d'autres de petits corps dans le système solaire. J'ai aussi porté mon attention sur une autre classe d'objets: les astéroïdes (chapitre 5). Plus encore peut-être que dans les anneaux planétaires, les collisions jouent un rôle important dans l'évolution des astéroïdes et le lien qui existe entre les conditions à l'origine du système et ce que nous pouvons observer maintenant (section 5.1). Les effets qui m'intéressent sont l'érosion (collisions à faible vitesse, anneaux planétaires), la cratérisation (collisions à grande vitesse avec de petits projectiles) et la fragmentation (collisions à grande vitesse de gros projectiles). J'ai donc développé un algorithme évaluant le résultat d'une collision entre deux corps qui tend à harmoniser les différentes théories existantes, à les rendre auto-cohérentes et à produire des données utilisables dans des simulations d'évolution. En particulier, nous considérons des fonctions de distribution de masse sur des intervalles finis et prenons en compte une éventuelle ré-accrétion des fragments. Nous avons ensuite, avec A. Campo-Bagatin et P. Farinella, testé la sensibilité des modèles d'évolution aux différents paramètres décrivant les effets des collisions.

Les images transmises par la sonde Galileo ont permis, à partir de la fin de 1993, de mieux connaître l'astéroïde 243 Ida (section 5.2). Les images à haute résolution nous donnent de très nombreux renseignements sur la surface de Ida en particulier en ce qui concerne les cratères et la présence de blocs rocheux. Il s'agit de la première observation de tels blocs à la surface d'un corps de cette taille. Leur distribution est de plus irrégulière, avec une forte concentration sur la face avant par rapport à la rotation de Ida. En collaboration avec l'équipe d'imagerie de la sonde Galileo, j'ai modélisé la dynamique des fragments éjectés lors des collisions pour étudier la localisation des points de ré-impact. La distribution observée a des implications sur les modèles de collisions et de cratérisation. L'étude de la distribution en taille des cratères nous a aussi apporté des précieux renseignements sur la physique des collisions et ses effets sur l'état de surface d'un astéroïde, ainsi que sur la distribution des projectiles (petits fragments issus d'autres collisions entre les astéroïdes).

Les images de Galileo ont aussi révélé la présence d'un satellite: Dactyl. Cette découverte relance les études sur les satellites d'astéroïdes. Nous avons abordé différents aspects du problème:

quelles informations sur la masse de Ida peut-on déduire de la connaisance partielle de l'orbite de Dactyl; les études dynamiques précédentes indiquant une origine commune pour Ida et Dactyl, comment peut-on créer une telle configuration; quels sont les processus susceptibles de stabiliser ou déstabiliser une orbite dans l'environnement très particulier de Ida (rotation rapide d'un corps irrégulier très allongé) pour des durées allant de quelques dizaines de millions à un milliard d'années; quels phénomènes dynamiques ou physiques doit-on prendre en compte pour expliquer une éventuelle survie de Dactyl pendant un milliard d'années là où les modèles classiques donnent une durée de vie collisionnelle de l'ordre de cent à deux cents millions d'années ?

## 1.2 Diffusion chaotique [24]

Depuis plusieurs années, de nombreuses études ont été menées sur le chaos dans les systèmes hamiltoniens classiques bornés et des outils puissants ont été développés et appliqués. Par opposition, peu de choses étaient connues sur le problème du chaos dans la diffusion classique lorsque j'ai rencontré ce phénomène dans le problème de Hill. Cependant, depuis près de vingt ans, il y avait eu des observations numériques de comportements complexes - chaotiques: modèles classiques de la diffusion moléculaire inélastique [69], [52], [48], [72], [62]; rencontres de satellites [3]; dynamique des vortex [46]; diffusion par un potentiel [47], [57]. Mais le phénomène n'avait pas été étudié pour lui-même.

Le problème physique que nous considérons est fort simple: il s'agit de la rencontre de deux satellites en orbites circulaires proches autour d'une planète. Les équations du mouvement sont non intégrables et sans réelle singularité. Malgré la continuité des équations, le comportement asymptotique présente un grand nombre de discontinuités par rapport aux paramètres initiaux. Tout porte à croire qu'il y a un nombre infini de discontinuités. Ce phénomène est dû à la présence, dans la section de Poincaré, des points homoclines et hétéroclines correspondant à des orbites périodiques instables. Dans le cas précis du problème de Hill, la variété instable d'une orbite périodique coupe la variété stable d'une deuxième orbite périodique instable. Ainsi, en partant près de la variété stable de la première orbite, on s'en éloignera le long de la variété instable. On passera alors proche de la deuxième variété stable, et on repartira à l'infini en suivant l'une ou l'autre des branches de la deuxième variété instable. Le choix de la branche dépend du côté duquel on arrive pour la première fois au voisinage de la deuxième variété stable. Une variation continue des paramètres initiaux fait varier continuement les points le long de la variété instable et donc le premier point au voisinage de la deuxième variété instable passe alternativement d'un côté à l'autre, donnant lieu à un comportement asymptotique complètement différent.

Du fait des difficultés numériques, M. Hénon [55] a développé un modèle simple où une particule ponctuelle soumise à une accélération constante rebondie sur deux disques. Cela conduit à une famille d'orbites à un paramètre avec diffusion chaotique. Dans ce modèle, il est possible de caractériser entièrement l'ensemble des discontinuités qui forment un ensemble de Cantor dont on peut calculer la dimension fractale. On peut ainsi définir une dynamique symbolique analogue au décalage de Bernoulli ou "transformation du boulanger" [61].

## 1.3 Applications de Poincaré des systèmes hamiltoniens [7], [12]

Les applications de Poincaré sont largement utilisées dans tous les domaines de la dynamique non linéaire, et en particulier en mécanique céleste [54]. Mais le calcul de ces applications est très gourmand en temps de calcul car il est nécessaire d'intégrer des équations différentielles pour les obtenir. Plusieurs auteurs ont essayé de remplacer le système hamiltonien qui les intéresse par une application symplectique [81]. Mais ces applications ne sont valables que localement et un important travail analytique doit être accompli pour chaque nouveau problème que l'on souhaite traiter.

Nous avons développé, en collaboration avec Claude Froeschlé, une méthode efficace en terme de temps de calcul pour étudier des applications de Poincaré pour des systèmes hamiltoniens quelconques. Nous découpons la surface de section en parallélogrammes (cellules) et calculons les valeurs de l'application de Poincaré aux sommets. Nous utilisons ensuite cette connaissance parcellaire de l'application pour définir une approximation polynômiale dans chaque cellule. Pour ce faire, nous utilisons soit uniquement les valeurs de l'application sur plusieurs nœuds du réseau, soit les valeurs de l'application et de ses dérivées aux nœuds les plus proches. Différentes techniques ont été mises en œuvre pour minimiser les détériorations dans les cellules situées au bord du domaine de définition. Les polynômes d'ordre élevé qui utilisent plus que les nœuds plus proches voisins nécessitent dans ce cas des formules non-symétriques, ce qui augmente l'erreur. Il est préférable d'utiliser des formules incluant les dérivées et n'utilisant que les nœuds plus proches voisins. Il est aussi possible de diminuer localement la taille des cellules.

La méthode a été testée à deux dimensions sur l'application "standard" et dans le cas du problème de Hill. Dans les deux cas, nous avons obtenu un excellent accord entre l'application interpolée et le cas réel aussi bien sur le plan de la géométrie de l'espace des phases que sur l'exposant caractéristique de Lyapunov. D'une manière générale, la correspondance entre les surfaces de section s'améliore lorsqu'on augmente le nombre de points de la grille et l'ordre de

l'approximation polynômiale.

Nous avons aussi comparé notre application synthétique avec l'application "standard" à 4 dimensions. Nous avons ainsi pu montrer que tous les exposants caractéristiques de Lyapunov sont bien reproduits de même que le temps d'explosion dans un phénomène de "collage-explosion". Pour une application de dimension quatre, les zones stables ne forment pas une barrière hermétique. Une orbite initialement confinée dans une zone chaotique finira après un certain temps par diffuser vers une autre région chaotique. Cette technique diminue le temps de calcul par un facteur d'au moins 10 et peut être utilisée pour l'étude des zones stables et chaotiques dans de nombreux problèmes de dynamique (par exemple, la ceinture des astéroïdes).

## 1.4 Sur les anneaux planétaires

### 1.4.1 Analyse de données de Voyager [8], [11], [36]

Les images transmises par les sondes Voyager ont montré que les anneaux planétaires sont le siège de phénomènes dynamiques d'une grande complexité encore mal compris, en particulier le confinement et les bords raides des anneaux, les lacunes et la distribution en taille des particules dans les anneaux. L'étude approfondie du problème de Hill a servi de base à plusieurs simulations de l'évolution des anneaux. Le phénomène de répulsion entre particules pour les "grands" paramètres d'impact et les orbites en fer à cheval expliquent la création de lacunes, le confinement des anneaux et l'apparition d'annelets au centre des grandes divisions (Encke [73]). Les simulations auto-cohérentes par méthode de Monte Carlo développées dans ma thèse [66], [67], [68], ainsi que les simulations déterministes avec particules tests [75], [76], [77] permettent de relier les paramètres physiques utilisés au profil précis des lacunes ou dépressions et des annelets. Afin de pouvoir départager les différentes possibilités, il convient de comparer les résultats obtenus numériquement avec les observations des sondes Voyager I et II. Les structures fines que nous cherchons à détecter sont de faible amplitude et le rapport signal sur bruit (rapport de l'amplitude de la structure à l'écart type du bruit) est très faible.

Avec Philippe Bendjoya et Frank Spahn, nous avons développé une méthode de détection de signal dans un ensemble de données à une dimension fondée sur une technique de reconnaissance de forme dans l'espace des coefficients de la transformée en ondelettes des données initiales. La comparaison de ces coefficients à ceux obtenus pour un signal aléatoire, présentant le même histogramme de valeurs que le signal étudié, permet de déterminer un ensemble de coefficients correspondant à des structures. Il nous a ainsi été possible de détecter des signaux ayant un rapport signal sur bruit de l'ordre de 1. Cette méthode a d'abord été employée pour analyser les données sur la division de Encke dans les anneaux de Saturne. Nous avons montré que

l'annelet central a une largeur d'environ 14 km et un profil à deux pics, avec une ébauche de troisième pic. En admettant une taille de 20 km pour le satellite Pan [73], cela implique une excentricité de l'orbite de Pan de l'ordre de $10^{-4}$. Nous avons aussi établi un catalogue des structures des anneaux externes de Uranus à partir des données d'occultation de $\sigma$ Sagitarii. L'anneau $\epsilon$, le plus large et le plus externe est constitué d'un plateau central et de pics latéraux. Nous avons également mis en évidence trois nouveaux anneaux; deux situés entre l'anneau $\epsilon$ et l'anneau $\delta$, le troisième à l'intérieur de l'anneau $\delta$. Le plus externe de ces anneaux pourrait correspondre à l'anneau I détecté par Lane *et al.* [58].

### 1.4.2 Viscosité des anneaux planétaires [14]

Le transport de moment angulaire est un mécanisme clef qui gouverne les structures à grande échelle des anneaux planétaires. Il détermine le temps d'étalement radial et le comportement des ondes. Les études analytiques suivent généralement le formalisme conventionnel de la mécanique statistique utilisant une version modifiée de l'équation de Boltzmann pour représenter la dynamique dans un champ de particules en mouvement képlérien [51], [74], [40]. De telles approches nécessitent un formalisme mathématique complexe qui rend difficile la modélisation des effets critiques des collisions. Récemment, une approche heuristique plus physique conduisit à des résultats semblables à ceux obtenus par les modèles analytiques et permit d'étudier la viscosité (coefficient caractérisant le transport de moment angulaire) dans le cas de disques ou anneaux planétaires présentant un gradient radial de densité [64]. Tous ces modèles considèrent des particules ponctuelles et des interactions locales. Or les particules peuvent avoir des tailles importantes et l'interaction peut être à grande distance, comme dans le cas de la gravitation, par exemple.

Richard Greenberg et moi-même avons développé un modèle de collisions entre particules sphériques en orbites képlériennes autour d'un corps central. Les collisions sont caractérisées par un coefficient de rebond radial qui détermine la variation de la composante radiale (par rapport au centre des particules) de la vitesse relative au cours de la collision. L'intégration d'un système intégro-différentiel nous permet d'obtenir la distribution des vitesses à l'équilibre.

A partir de là, nous déterminons le transport "local" et le transport "non-local". Le transport local est évalué en sommant sur la distribution des vitesses le moment angulaire transporté par chaque particule qui passe au travers d'un cercle de référence (on s'intéresse ainsi aux particules qui sont localisées sur ce cercle de référence). Dans notre modèle, nous tenons compte de la taille des particules. De ce fait, une particule située d'un côté du cercle de référence peut subir une collision avec une particule située de l'autre côté, faisant ainsi passer du moment angulaire d'un côté à l'autre, sans pour autant qu'aucune des particules soit jamais située sur le cercle de référence (transport non-local). Cet effet est d'une grande importance lorsque

la taille des particules est du même ordre de grandeur que l'excursion radiale. Nous avons montré que la viscosité locale varie avec la taille des particules contrairement au résultat de la théorie classique. La viscosité non-locale quant à elle est négligeable pour les petites particules, mais devient dominante pour les grosses particules, surtout lorsque la profondeur optique est importante.

Notre modèle va nous permettre d'étudier la viscosité pour un mélange de particules de tailles différentes. Dans les anneaux de Saturne, il semble que la plus grande partie de la masse soit due aux grosses particules, alors que la profondeur optique est dominée par les petites particules. Il est donc intéressant de mesurer la viscosité de petites particules dont la dynamique est contrôlée par de grosses particules. De même, les effets non-locaux doivent être considérés lorsqu'on inclut les effets de la gravitation lors des rencontres proches. Notre formalisme va nous permettre d'estimer la viscosité en présence de la gravitation.

## 1.5 Sur les astéroïdes

### 1.5.1 Evolution physique des petits corps [9], [13]

Les astéroïdes sont souvent considérés comme étant des corps ayant moins évolué que les planètes depuis la formation du système solaire. Ils pourraient ainsi représenter de bons traceurs des conditions initiales lors de la génèse du système solaire. Cependant, même s'ils ne sont pas différenciés et ne présentent pas d'évolution physique propre, ils évoluent dynamiquement. Leurs orbites présentent des inclinaisons relatives et des excentricités non négligeables, ce qui donne lieu à des collisions à grande vitesse, de l'ordre de quelques kilomètres par seconde. Cela a pour effet de modifier la distribution de la taille des corps ainsi que leur état de surface et leur structure interne.

Pour étudier l'évolution des astéroïdes, il est nécessaire de savoir modéliser l'effet de ces collisions à haute vitesse. Des expériences de laboratoire ont été réalisées afin de décrire les résultats des collisions en terme de quelques paramètres d'entrée simples tels que la masse et la vitesse du projectile, la masse de la cible et la réponse du matériau de la cible à un impact [49]. Malheureusement, ces expériences utilisent des cibles de petite taille alors que les corps célestes qui nous intéressent peuvent avoir plusieurs centaines de kilomètres de diamètre. Dans ce cas, l'autogravité peut jouer un rôle important en augmentant la solidité intérieure de la cible et en permettant une ré-accrétion partielle des fragments. Des expériences de fragmentation menées sous forte pression ont été réalisées [56] pour quantifier l'effet sur la solidité et ont permis d'étalonner les lois analytiques pour les grandes cibles.

Du point de vue théorique, notre compréhension des processus physiques complexes qui ont

lieu lors des collisions est encore assez limitée. Les travaux récents portent sur des hypothèses sur les critères de fracture et sur la géométrie du champ de vitesse des fragments [65], ou sur l'intégration numérique d'équations hydrodynamiques adaptées aux phénomènes à deux ou trois dimensions [70], [71], [63], [41].

Les résultats des expériences montrent une large dispersion des tailles et des vitesses des fragments. Cependant, les distributions obtenues ainsi que celles issues des modèles numériques et semi-analytiques sont bien approximées par des lois de puissance. Ma première contribution à l'étude de ce problème fut de développer un modèle numérique consistant qui permet de prédire le résultat d'impacts entre des astéroïdes (ou autre petits corps du système solaire). Il est basé sur un ensemble de paramètres qui peuvent être estimés à partir des évidences expérimentales disponibles. Il inclut aussi la possibilité de ré-accumulation des fragments éjectés possédant une vitesse inférieure à une vitesse d'évasion judicieusement choisie. Toutes les fonctions de distribution de la masse des fragments sont modélisées à partir de lois de puissance tronquées. Dans un modèle cohérent, il faut tenir compte de discontinuités aux deux extrémités des distributions. Les modèles de fracture nous permettent de déterminer la taille du plus gros fragment en fonction des conditions de la collision. La distribution passe donc de 0 à 1 de façon discrète pour cette taille là. De même la taille des particules a une limite inférieure non nulle ce qui évite d'ailleurs une divergence de la masse totale pour des exposant élevés. La vitesse des particules atteint aussi une limite supérieure, liée au mécanisme de transfert de moment cinétique. La limite utilisée dans nos simulations est liée à la vitesse de propagation des ondes sonores. Finalement, une possible corrélation entre la taille des fragments et leur vitesse est prise en compte, soit de façon déterministe, soit de façon probabiliste.

En collaboration avec Paolo Farinella, nous avons testé la dépendance entre les paramètres d'entrée et les distributions obtenues. En particulier, nous avons étudié l'efficacité du processus de ré-accrétion gravitationnelle en fonction des paramètres de collision et de la relation entre taille et vitesse des fragments. Aussi bien la taille au-delà de laquelle la plus grande part de la masse des fragments va se ré-accumuler que l'énergie collisionnelle nécessaire pour disperser les fragments "à l'infini" s'avèrent très sensibles au choix de cette relation.

Nous avons utilisé ce modèle de fragmentation dans un code simulant l'évolution collisionnelle des astéroïdes dans le but de déterminer l'effet de cette sensibilité sur la distribution finale de la taille des astéroïdes. Partant d'une distribution initiale en loi de puissance, nous obtenons à la fin de l'évolution une distribution ondulée qui dépend fortement de la relation entre vitesse moyenne et masse des fragments ainsi que de la dispersion des vitesses. La dépendance est plus prononcée pour les tailles (de quelques dizaines à quelques centaines de kilomètres) auxquelles la ré-accumulation gravitationnelle est la plus effective.

### 1.5.2 Evolution du système Ida/Dactyl [16], [17], [18], [20]

Les images de l'astéroïde Ida 243 prises par la sonde Galileo au mois d'août 1993 nous révélèrent une surface modelée principalement par des cratères dus à des collisions avec d'autres corps de la ceinture d'astéroïdes. En particulier, elles mirent en évidence l'existence de blocs de rocher à la surface de Ida, avec une distribution concentrée à l'un des bouts de l'astéroïde [60]. Il s'agit d'une première évidence observationnelle que les effets dynamiques peuvent se refléter sur la morphologie de la surface, comme cela avait été suggéré dans le cas des deux satellites de Mars: Phobos et Deimos [44], [42]. Du fait de la défaillance de l'antenne principale de communication de la sonde, il fallut attendre jusqu'au 17 février 1994 pour recevoir les images qui nous firent découvrir Dactyl, satellite de Ida. Avant la découverte de Dactyl, de nombreuses observations furent rapportées qui indiquaient que des satellites d'astéroïdes pouvaient exister [79], [80]. Cependant aucune de ces observations ne fut considérée comme convaincante et on commença à douter de l'existence des satellites [50]. Au moins quatre images à haute résolution de Dactyl permettent une étude de sa surface et là aussi, il est clair que son état actuel est le résultat de collisions.

Je me suis donc intéréssé, en compagnie de Paul Geissler, à la dynamique dans le système Ida/Dactyl. Cette dynamique est fortement influencée par la faible densité (entre 2,3 g/cm$^3$ et 3 g/cm$^3$), la forme très allongée (axes principaux de 58,9 km, 25,4 km et 18,6 km pour l'ellipsoïde triaxial qui approxime le mieux Ida) et la rapide rotation de l'astéroïde (période de rotation de 4,63 heures). Nous avons exploré le processus d'évasion et de ré-accrétion de fragments sur Ida et Dactyl à l'aide de simulations numériques en trois dimensions. Nous avons tout d'abord étudié les effets de la rotation, le point de lancement et la vitesse d'éjection des fragments dans le cas d'un ellipsoïde triaxial ayant la forme approximative et la densité de Ida. Les fragments éjectés à faible vitesse ré-impactent près du cratère et forment une couverture d'éjecta bien définie avec une asymétrie entre face avant et face arrière par rapport à la rotation. L'effet global de ce genre de cratérisation est de produire une couche épaisse de régolites uniformément distribuée. Par opposition, aucune couverture n'est créée quand les fragments sont éjectés à grande vitesse, de l'ordre de la vitesse d'évasion. En fait, la plupart des fragments s'échappent du système et ceux qui ré-impactent le font après un temps passé en orbite significatif comparé à la période de rotation de l'astéroïde. L'effet global est de produire une couche de régolites plus mince, moins uniforme, avec des concentrations sur les faces avant.

A l'aide d'un modèle de forme de Ida réaliste [78], nous avons montré qu'une des zones étendues de couleur/albédo uniforme qui domine les parties nord et ouest de Ida peut s'expliquer par la ré-accrétion de fragments issus du très grand et récent cratère "Azzura". La vitesse d'éjection requise pour reproduire les observations des couleurs est de l'ordre de quelques mètres par seconde, en accord avec les modèles qui prévoient que les cratères de plusieurs kilomètres

sur Ida se forment dans un régime dominé par la gravité et produisent une couche locale de régolites. Les fragments éjectés à une vitesse d'environ 10 m/sec dans la direction de rotation s'élèvent au dessus de l'astéroïde et sont ensuite balayés par la face avant de l'autre côté de l'astéroïde. Les lieux d'atterissage de ces fragments correspondent bien à la distribution de blocs observée dans les images à haute résolution.

Le modèle précédent nous permet d'expliquer certaines particularités de la topographie de Ida, liées à son histoire récente. Pour mieux connaître son histoire globale, il faut prendre en compte son appartenance à la famille de Koronis et aussi les données statistiques déduites de l'observation de sa surface. L'étude de l'astéroïde Gaspra [53] a déjà permis de montrer que la distribution en taille des projectiles est particulièrement pentue et que les secousses dues aux grandes collisions affectent la couverture de cratère en effaçant fréquemment tous les petits cratères sur toute la surface. La population de cratères de Ida est très différente de celle de Gaspra. Il nous a donc fallu élaborer un scénario qui tienne compte des différences des propriétés physiques et d'environnement tout en utilisant un même modèle physique pour les collisions. Les contraintes d'environnement nous permettent déjà d'estimer une limite supérieure de l'âge de Ida: 1,5 milliard d'années, âge maximum de la famille de Koronis [45]. La dynamique dans le système de Ida rend très improbable que Dactyl puisse être un morceau de Ida, arraché puis mis en orbite. Il est plus probable que Dactyl soit un compagnon primordial, créé en même temps que Ida lors de la fragmentation de l'astéroïde parent de la famille Koronis. Or l'espérance de vie collisionnelle de Dactyl est de l'ordre de la centaine de millions d'années. Par ailleurs, les orbites de Dactyl déduites des observations sont instables à très court terme (moins d'un an) pour une densité de Ida supérieure à 3 g/cm$^3$. Ida étant situé plus près du centre de la ceinture d'astéroïdes, les projectiles sont environ 40% plus nombreux que dans le cas de Gaspra, mais leur vitesse moyenne est inférieure (3,55 km/sec contre 5,0 km/sec [43]) du fait de la faible excentricité de l'orbite de Ida. Nous avons montré que la distribution des cratères autorise deux âges contradictoires pour Ida: soit de l'ordre de 50 millions d'années, soit plus d'un milliard d'années. L'espérance de vie de Dactyl milite en faveur de la première solution. L'état dégradé de certains cratères favoriserait plus la deuxième solution, à moins d'invoquer une population de projectiles dominée par les petites particules. Un scénario où Dactyl serait le reste d'un satellite érodé ou fragmenté puis ré-accrété autoriserait aussi l'âge le plus vieux.

Du fait des conditions d'observation (ligne de visée proche du plan de l'orbite), il n'est pas possible de définir exactement l'orbite de Dactyl. On ne peut qu'obtenir une famille d'orbites possibles, paramétrée par la masse de Ida. Une grande masse de Ida est à rejeter à cause de l'instabilité à très court terme (densité supérieure à 3,1 g/cm$^3$). Pour une faible masse de Ida, l'orbite serait hyperbolique, ce que nous devons rejeter aussi pour des raisons statistiques: la probabilité de voir un astéroïde étranger au système capturé temporairement sur une orbite hyperbolique est inférieure à $2 \times 10^{-17}$. Une étude à moyen terme de la stabilité nous a permis

de montrer que les orbites correspondant à une masse de Ida légèrement plus grande (densité 2,1 g/cm$^3$ à 2,3 g/cm$^3$) sont chaotiques et finissent par s'échapper à l'infini. On peut donc exclure ces orbites. Pour une densité comprise entre 2,3 g/cm$^3$ et 2,5 g/cm$^3$, les orbites sont aussi chaotiques, mais elles semblent être piégées dans une région non connectée à l'infini. On obtient donc un chaos "stable" similaire à celui mis en évidence par Laskar [59] pour la Terre.

Les intégrations numériques réalisées sur plus de 3000 ans (soit plus de 1 million d'orbites) pour les densités intermédiaires montrent une grande stabilité. Cependant pour le long terme, il faut très certainement tenir compte d'effets dissipatifs. Plusieurs observations laissent penser que le système Ida/Dactyl a évolué par dissipation: l'inclinaison de l'orbite de Dactyl sur le plan équatorial de Ida est faible, de l'ordre de 8°, la période de rotation de Dactyl est longue, supérieure à 8 heures (malgré la faible taille de Dactyl), et il est possible que la rotation soit synchrone avec le mouvement orbital (le grand axe de Dactyl est orienté en direction de Ida). Il est donc important de déterminer les résonances présentes dans le système. Nous avons étudié les commensurabilités entre la période orbitale de Dactyl et la période de rotation de Ida. Pour les résonances d'ordre faible, seule la 5:1 est en accord avec la position de Dactyl au moment de la rencontre de Galileo avec Ida et semble assez large pour capturer Dactyl. Pour les ordres élevés, l'imprécision sur les paramètres osculateurs ne nous permet pas de conclure. Pour des résonances d'ordres supérieurs à 50, les commensurabilités se recouvrent, expliquant le chaos et la diffusion vers des orbites d'évasion. Une première série de simulations d'effets dissipatifs modélisés par des forces de traînée a mis en évidence l'importance de résonances secondaires. Lorsque l'orbite de Dactyl est capturée dans une résonance principale, son excentricité augmente, jusqu'à croiser une commensurabilité entre la fréquence d'oscillation de la direction du péricentre et la période de rotation de Ida. L'orbite devient alors chaotique et en fonction de l'importance de la dissipation, soit sort de la résonance principale, soit part sur une orbite d'évasion. L'étude plus approfondie des forces de traînée (dues aux collisions incessantes avec de petites particules, comme en témoigne l'état de surface de Dactyl) et des forces de marées, que j'ai entreprise, devrait apporter de précieux renseignements sur l'histoire du système Ida/Dactyl et contraindre plus précisement la densité de Ida.

## 1.6 Conclusion et perspectives

Les travaux présentés dans ce dossier traitent de différents aspects des petits corps du système solaire et de méthodes numériques développées pour ces études. Nous nous sommes intéréssés à la dynamique de petits corps tels que les astéroïdes et les anneaux planétaires, ainsi qu'à l'effet des collisions à grande vitesse entre des corps de tailles allant de la poussière à plusieurs centaines de kilomètres. Une nouvelle technique de traitement de signal, utilisant la

transformation en ondelettes, a été développée pour dépouiller des données unidimensionnelles fortement bruitées. L'étude de la dynamique nous a amenés à développer une nouvelle technique d'intégration de systèmes hamiltoniens. Nous avons aussi mis en évidence et expliqué une classe particulière de chaos: la diffusion chaotique.

L'algorithme que nous avons mis au point pour évaluer la viscosité est bien adapté à l'étude de systèmes à deux dimensions et avec une profondeur optique faible ou moyenne. Nous avons ainsi pu mettre en évidence la variation de la viscosité locale et non-locale en fonction de la taille des particules et de la profondeur optique. Cette approche va donner lieu à divers développements qui sont au cœur d'une thèse de Ph.D. à l'Université d'Arizona. Il s'agit tout d'abord de faire une description tri-dimensionnelle des collisions afin de mieux modéliser les anneaux ou disques épais, ou les petites particules dans les anneaux ou disques fins. De par sa structure, ce modèle est particulièrement bien adapté à l'étude de systèmes diffus comme le disque du proto-système solaire. Nous avons aussi la possibilité de remplacer la fonction de diffusion par collisions seules par une fonction de diffusion par gravitation et collisions. Cela permettra une meilleure représentation des anneaux ou disques froids (faible excentricité relative des particules). Afin de pouvoir étudier aussi des systèmes denses, nous allons modifier l'évaluation de la fréquence de collision pour tenir compte de la diminution de l'espace accessible lorsque la profondeur optique augmente ainsi que du phénomène de collage des particules.

L'étude à long terme du système Ida/Dactyl impose la prise en compte de l'évolution physique des astéroïdes. Pour cela il est nécessaire de raffiner la modélisation de la fragmentation et de la cratérisation des corps lors d'impacts à grande vitesse. En particulier, il nous faut mieux comprendre l'effet de la taille de la cible (gravitation) et de son état physique (roc solide ou agrégat). Des expériences avec des corps plus gros (de l'ordre du mètre ou plus) et sous forte pression pour simuler la gravitation devront être menées en conjonction avec des simulations numériques utilisant des codes de type SPH. Il nous faudra ensuite intégrer ces résultats dans nos modèles d'évolution physique et dynamique du système Ida-Dactyl pour obtenir un scénario complet et des contraintes sur la physique des collisions. Cela nous permettra en particulier de mieux dater ces deux corps. Et que dire de l'effet des forces de marée sur un corps très allongé ? Comment varie cet effet en fonction de la structure interne des corps, notamment en ce qui concerne les temps caractéristiques de dissipation ? J'ai entrepris la modélisation des effets des forces de marée sur un corps allongé. La dissipation ainsi obtenue, ajoutée à celle due aux frottements visqueux (collisions avec les poussières de la ceinture d'astéroïdes et les éjecta des cratères de Ida) determine l'évolution de l'orbite de Dactyl, en particulier lors des captures dans les résonances de moyen mouvement. Il est spécialement important de savoir si l'excentricité augmente ou diminue pour prévoir une éventuelle sortie de la résonance.

Comme le montrent les travaux exposés dans cette thèse, les outils numériques représentent un complément indispensable aux observations et expériences physiques de laboratoire. L'astro-

nome moderne doit savoir non seulement utiliser les outils numériques existants, mais aussi en créer d'autre pour répondre aux exigences croissantes de la recherche actuelle. Au vu des gigantesques bases de données emmagasinées par les grands téléscopes et les sondes interplanétaires, il paraît nécessaire de développer de nouveaux outils de traitement du signal capable d'extraire automatiquement de l'information avec un rapport signal sur bruit faible, de l'ordre de 1. Pour être pleinement utilisables, ces méthodes devront donner un niveau de signification des informations extraites et être indépendantes de tout modèle. L'algorithme que j'ai développé permet déjà de traiter quasi automatiquement un signal unidimensionnel, d'éliminer le bruit et de séparer le fond du signal. La prochaine étape de ce travail est de généraliser cette méthode à deux dimensions et d'inclure une détection et un comptage automatique des structures ou objets extraits du signal original.

Parmi les objets susceptibles de nous éclairer sur les origines du système solaire, les astéroïdes forment un groupe important car relativement facilement observables et n'ayant pas eu d'évolution physique interne propre. Cependant ils évoluent physiquement par interaction collisionnelle avec leurs semblables. Les expériences de laboratoires envisageables de nos jours étant limitées à de petits corps, il est important de définir de nouveaux algorithmes de calcul pour la modélisation de la structure interne des corps.

De plus en plus, nous cherchons à déterminer la dynamique à long terme des petits corps du système solaire. Le long terme s'entend comme un temps long par rapport à la plus petite échelle de temps caractéristique du système étudié. Les techniques d'intégrations symplectiques classiques, obtenues par perturbation d'un hamiltonien intégrable, ne sont pas applicables dans un cas tel que le système Ida/Dactyl où l'hamiltonien est très éloigné d'un système intégrable connu. Là encore, il nous faut définir de nouveaux outils purement numériques exempts de toute modélisation physique pour une plus grande versatilité.

La simulation numérique représente dans de nombreux cas nos grands instruments et nos appareils de laboratoire. De même qu'il faut sans cesse améliorer les instruments d'observation, inventer de nouveaux détecteurs, de nouvelles procédures d'acquisitions de données, il faut développer de nouveaux moyens de calculs, tant au niveau matériel que logiciel, en poussant les performances aux limites de la physique des semi-conducteurs et en faisant coopérer de plus en plus de machines. Il est nécessaire de définir de nouveaux algorithmes de calculs pouvant tirer parti du parallélisme hétérogène offert de nos jours par les réseaux. De nombreuses simulations en mécanique céleste et en traitement du signal possèdent l'indépendance voulue pour être parallélisées. A partir de l'expérience que j'ai acquise en parallélisation, aussi bien sur des machines fortement parallèles telles que la Connection Machine sur des problèmes d'intégration de systèmes à N corps, que sur des ensembles de machines hétérogènes pour des problèmes à parallélisme moins massif comme l'intégration d'équations intégro-différentielles, je vais collaborer à la parallélisation de divers gros codes numériques. Le premier exemple

sera probablement le code de formation de planètes dans le disque du proto-système solaire, développé au Planetary Science Institute, à Tucson.

# Références

[40] Araki S. & Tremaine S., 1986. *The dynamics of dense particle disks.* Icarus **65**, pp. 83–109.

[41] Asphaug, E., Benz, W. & Melosh, H.J., 1992. *Asteroid impact studies with SPH 3D (abstract).* Bull. Amer. Astr. Soc. **24**, p. 960.

[42] Banaszkiewicz, M. & Ip, W.-H., 1991. *The Stickney impact of Phobos: a dynamical model.* Icarus **101**, pp. 144–164.

[43] Bottke, W., Nolan, M., Greenberg, R. & Kolvoord, R., 1994. *Velocity distributions among colliding asteroids.* Icarus **107**, pp. 255–268.

[44] Davis, D., Housen, K. & Greenberg, R., 1981. *The unusual dynamical environment of Phobos and Deimos.* Icarus **47**, pp. 220–233.

[45] Durda, D., 1993. *The collisional evolution of teh asteroidal belt and its contribution to the zodiacal cloud.* Ph. D. Thesis Univ. of Florida,

[46] Eckhardt, B. & Aref, H., 1989. Phil. Trans. R. Soc. Lond. A **326**, p. 655.

[47] Eckhardt, B. & Jung, C., 1986. J. Phys. A **19**, p. *L*829.

[48] Fitz, D. E. & Brumer, P., 1979. J. Chem. Phys. **70**, p. 5527.

[49] Fujiwara, A., Cerroni, P., Davis, D.R., Ryan, E., Di Martino, M., Holsapple, K. & Housen, K., 1989. *Experiments and scaling laws on catastrophic collisions.* in *Asteroids II,* eds. R.P. Binzel, T. Gehrels & M.S. Matthews, Univ. of Arizona Press, pp. 240–−265.

[50] Gehrels, T., Drummond, J. & Levenson, N., 1987. *The absence of satellites of asteroids.* Icarus **70**, pp. 257–263.

[51] Goldreich P. & Tremaine S., 1978. *The velocity dispersion in Saturn's rings.* Icarus **34**, pp. 227–239.

[52] Gottdiener, L., 1975. Molecular Physics **29**, p. 1585.

[53] Greenberg, R., Nolan, M., Bottke, W. & Kolvoord, R., 1994. *Collisional history of Gaspra.* Icarus **107**, pp. 84–97.

[54] Hénon, M., 1981. in *Cours des Houches,* **XXXVI**, North Holland, Amsterdam, p. 57.

[55] Hénon, M., 1988. Physica D **33**, p. 132.

[56] Housen, K.R., Schmidt, R.M. & Holsappple, K.A., 1991. *Laboratory simulations of large scale fragmentation events.* Icarus **94**, pp. 180–−190.

[57] Jung, C. & Scholz, H-J., 1987. J. Phys. A **20**, p. 3607.

[58] Lane, A. L. *et al.*, 1986. Science **233**, p. 65.

[59] Laskar, J., 1993. *??.* ?? pp. ??–??.

[60] Lee, P., Veverka, J., Thomas, P., Helfstein, P., Belton, M., Chapman, C., Greeley, R., Pappalardo, R., Sullivan, R. & Head, J., 1996. *Ejecta Blocks on 243 Ida and on Other Asteroids.* Icarus **120**, pp. 87–105.

[61] Moser, J., 1973. *Stable and Random motions in Dynamical Systems.* in *Annals of Mathematics Studies,* **77**, Princeton University Press,

[62] Noid, D. W., Gray, S. K. & Rice,, S. A., 1986. J. Chem. Phys. **84**, p. 2649.

[63] Nolan, M.C., Asphaug, E. & Greenberg, R., 1992. *Numerical simulation of impacts on small asteroids (abstract).* Bull. Amer. Astr. Soc. **24**, pp. 959–960.

[64] Ojakangas G.W. & Greenberg R., 1990. *Viscosity and mass transport in nonuniform keplerian disks.* Icarus **88**, pp. 146–171.

[65] Paolicchi, P., Verlicchi, A. & Cellino, A., 1992. *A new semiempirical model for catastrophic impact processes (abstract).* Bull. Amer. Astr. Soc. **24**, p. 960.

[66] Petit, J.M. & Hénon, M., 1987*a*. *A numerical simulation of planetary rings I. Binary encounters.* Astron. Astrophys. **173**, pp. 389–404.

[67] Petit, J.M. & Hénon, M., 1987*b*. *A numerical simulation of planetary rings II. Monte Carlo model.* Astron. Astrophys. **188**, pp. 198–205.

[68] Petit, J.M. & Hénon, M., 1988. *A numerical simulation of planetary rings III. Mass segregation, ring confinement, and gap formation.* Astron. Astrophys. **199**, pp. 343–356.

[69] Rankin, C. C. & Miller, W. H., 1971. J. Chem. Phys. **55**, p. 3150.

[70] Ryan, E.V., Asphaug, E. & Melosh, H.J., 1990. *Hydrocode predictions of collisional outcomes: Effects of target size.* Lunar Planet. Sci. **XXI**, pp. 1155–−1156.

[71] Ryan, E.V. & Melosh, H.J., 1992. *Understanding asteroid collisions: The effect of target size on collisional outcome (abstract).* Bull. Amer. Astr. Soc. **24**, p. 961.

[72] Schlier, C. G., 1983. Chemical Physics **77**, p. 267.
**scme86**, Showalter, M., Cuzzi, J.N., Marouf, E.A. & Esposito, L.W., 1986. *Satellite 'wakes' and the orbit of the Encke gap moonlet.* Icarus **66**, pp. 297–323.

[73] Showalter, M., 1991. *Visual detection of 1981 S13, Saturn's eighteenth satellite, and its role in the Encke gap.* Nature **351**, pp. 709–713.

[74] Shu, F.H. & G.R. Stewart, 1985. *The collisional dynamics of particulate disks.* Icarus **62**, pp. 360–383.

[75] Spahn, F., 1987. *Scattering Properties of a Moonlet (Satellite) Embedded in a Particle Ring: Application to the Rings of Saturn.* Icarus **71**, pp. 69–77.

[76] Spahn, F. & Wiebicke, H.-J., 1989. *Long-Term Gravitational Influence of Moonlets in Planetary Rings.* Icarus **77**, pp. 124–134.

[77] Spahn, F., Saar, A., Schmidt, S. & Schwarz, U., 1992. *The influence of various moonlets on the optical depth profile in planetary rings.* Icarus **100**, pp. 143–153.

[78] 1996. Thomas, P., Belton, M., Carcich, B., Chapman, C., Davies, M., Sullivan, R. & Veverka, J., *The shape of Ida.* Icarus **120**, pp. 20–32.

[79] Van Flandern, T., Tedesco, E. & Binzel, R., 1979. *Satellites of Asteroids.* in *Asteroids,* ed. Gehrels, T., Univ. of Arizona Press, Tucson, pp. 443–465.

[80] Weidenschilling, S., Paolicchi, P. & Zappala, V., 1989. *Do asteroids have satellites ?.* in *Asteroids II,* eds. Binzel, R., Gehrels, T. & Matthews M., Univ. of Arizona Press, Tucson, pp. 643–660.

[81] Wisdom, J., 1982. *The origin of the Kirkwood gaps: a mapping for asteroidal motion near the 3/1 commensurability.* Astron. J. **87**, pp. 577–593.

# Chapitre 2

# DIFFUSION CHAOTIQUE

**[24]**

"On a Cantor structure in a satellite scattering problem."

PETIT, J-M. *&* HENON, M.

in *Dynamics and Stochastic Processes*, eds. R. Lima, L. Streit *&* R. Vilela Mendes (1989)

# On a Cantor structure in a satellite scattering problem.

Jean-Marc Petit
CNRS, Observatoire de Nice, France

Michel Hénon
CNRS, Observatoire de Nice, France

**Abstract**

The phenomenon of chaotic scattering is described in the context of satellite encounters. We consider a one-parameter family of orbits obtained by starting with two satellites on circular, coplanar and close orbits. We numerically find that this family exhibits a large number of discontinuities, probably an infinite number. This phenomenon seems to be due to the existence of homoclinic and heteroclinic points of unstable periodic orbits. We model the chaotic scattering by a simple billiard: a point particle bounces on two disks and in addition is subjected to a constant acceleration. This leads to a one-parameter family with chaotic scattering. With the help of symbolic dynamics, the structure of the family can be completely elucidated.

## 1 Introduction

In the last few years, many studies have been carried out on the chaos in bound classical hamiltonian systems and powerful methods have been developed and applied. In contrast, less work has been done on chaos in classical scattering systems. However, for nearly twenty years, there have been numerical observations of complicated - chaotic - behavior in continuous scattering problems: classical models for inelastic molecular scattering (Rankin and Miller 1971, Gottdiener 1975, Fitz and Brumer 1979, Schlier 1983, Noid et al. 1986), satellite encounters (Petit and Hénon 1986), vortex dynamics (Eckhardt and Aref 1989), potential scattering (Eckhardt and Jung 1986, Jung and Scholz 1987). But until recently, this phenomenon had not been studied for itself.

We found this kind of behavior in a simple physical problem: the encounter of two satellites on close circular orbits around a planet (namely, Saturn). In section 2, we describe in detail the physical problem and derive the equations of motion. One would notice that the equations

of motion are non integrable and contain no true singularities. In section 3, we present more precisely the chaos that appears in our problem: the asymptotic behavior of the system is discontinuous with respect to the initial parameters. Then, in section 4, we give some hints on what creates this phenomenon. In view of the numerical difficulties which were encountered in exploring this problem, a simple "model" problem was developped (Hénon, 1988) which is just complex enough to exhibit all the features we are interested in. This model is described in section 5.

# 2 The physical problem

The physical problem we consider is a particular case of the three body problem. Two light bodies $M_2$ and $M_3$ describe initially coplanar and circular orbits, with slightly different radii, around a heavy central body $M_1$. Bodies $M_2$ and $M_3$ are initially far apart, so that their mutual attraction is negligible. However, the inner body has a slightly larger angular velocity and eventually catches up with the outer body; the distance from $M_2$ to $M_3$ becomes small and their mutual attraction is no longer negligible. We shall call this an *encounter*. This study can be applied to different problems in astronomy: planetary rings dynamics, motion of coorbital satellites (Janus and Epimetheus of Saturn) or planetary formation. For convenience, $M_1$ will be called the *planet* and $M_2$, $M_3$ will be called the *satellites*. The difference between the radii of the initial circular orbits will be called the *impact parameter*.

Analytic approximations of the solution are available in two cases:

(i) When the impact parameter is sufficiently large, the result of the encounter is only a slight deflection of $M_2$ and $M_3$ from their previous circular orbits. These deflections can then be obtained by a perturbation theory (Goldreich and Tremaine 1979, 1980).

(ii) When the impact parameter is very small, the interaction of $M_2$ and $M_3$ produces a "horseshoe" motion: $M_2$ and $M_3$ "repel" each other azimutally and never come in close proximity. This case can also be treated by a perturbation theory (Dermott and Murray 1981, Yoder et al. 1983).

Between these two asymptotic cases, however, no theory exists, and apparently only a numerical integration of the equations of motion can give the answer. In order to have an accurate numerical study, we first reduce the equations to a simpler form: the classical set of Hill's equations. Only a brief review of this reduction will be given here; details can be found in Hénon and Petit 1986. We assume that the mass of either satellite is small compared to the mass of the planet:

$$m_2 \ll m_1, \qquad m_3 \ll m_1, \tag{1}$$

where $m_i$ is the mass of body $M_i$. We assume also that the distance between the two satellites

is small compared to their distance to the planet. In a zero-order approximation, the two satellites can then be considered as a single body in orbit around the planet. This orbit will be called the *mean orbit*, and will be assumed to be circular. We call $a_0$ the radius of the mean orbit. (The precise definition of $a_0$ does not matter, as long as it is nearly equal to the radii of the satellite orbits). The angular velocity on the mean orbit is

$$\omega_0 = \sqrt{Gma_0^{-3}}, \tag{2}$$

where $m$ is the total mass of the system:

$$m = m_1 + m_2 + m_3. \tag{3}$$

We define

$$\mu = \frac{m_2 + m_3}{m}. \tag{4}$$

Let $X_i$, $Y_i$ be the coordinates of body $i$ in an inertial system. We introduce dimensionless coordinates by

$$X_i' = \frac{X_i}{a_0}, \qquad Y_i' = \frac{Y_i}{a_0}, \qquad m_i' = \frac{m_i}{m}, \qquad t' = \omega_0 t, \tag{5}$$

and for simplicity we drop the primes in what follows. In the new variables, the radius of the orbit, the angular velocity, the mass of the system, and the gravitational constant are all equal to 1. We choose the origin of time so that the two satellites are in the vicinity of $X = 1$, $Y = 0$ at $t = 0$. We introduce new coordinates $\xi$, $\eta$, which will be called *Hill's coordinates:*

$$\begin{aligned} X_i - X_1 &= (1 + \mu^{1/3}\xi_i)\cos t - \mu^{1/3}\eta_i \sin t, \\ Y_i - Y_1 &= (1 + \mu^{1/3}\xi_i)\sin t - \mu^{1/3}\eta_i \cos t, \qquad (i = 2, 3) \end{aligned} \tag{6}$$

We go over to new coordinates $\xi^*$, $\eta^*$, $\xi$, $\eta$, describing, respectively, the position of the center of mass and the relative position of the two satellites:

$$\xi^* = \frac{m_2\xi_2 + m_3\xi_3}{m_2 + m_3}, \qquad \eta^* = \frac{m_2\eta_2 + m_3\eta_3}{m_2 + m_3}, \qquad \xi = \xi_3 - \xi_2, \qquad \eta = \eta_3 - \eta_2. \tag{7}$$

The equations for the motion of the center of mass are linear and easily integrated (Hénon and Petit 1986). The equations of relative motion are approximately

$$\ddot{\xi} = 2\dot{\eta} + 3\xi - \frac{\xi}{\rho^3}, \qquad \ddot{\eta} = -2\dot{\xi} - \frac{\eta}{\rho^3}, \qquad \rho = \sqrt{\xi^2 + \eta^2}, \tag{8}$$

which are *Hill's equations* (Hill 1978). The error in these equations is of order of $\mu^{1/3}$. They become exact in the limit of vanishing satellite masses. Taking this limit is equivalent to zoom on the two satellites and the main effect is to repel the planet to infinity and transform circular orbits in straight lines.

The most important points to notice on these equations are:

(i) There is no parameter left in the equations (the same equations are valid in every physical case).

(ii) As is easily shown, the initial conditions for relative motion of circular orbits are given by only one parameter: the impact parameter $h$. Therefore, the set of solutions is a one-parameter family, and it seems reasonable to try to study it. We can even reduce the study to positive values of $h$ because of the symmetries of the equations of motion.

Hill's equations admit the integral

$$\Gamma = 3\eta^2 + \frac{2}{\rho} - \dot{\xi}^2 - \dot{\eta}^2 \tag{9}$$

which can be called the *Jacobi integral* by analogy with the restricted problem. We can write the Jacobi integral in terms of the initial conditions:

$$\Gamma = \frac{3}{4}h^2. \tag{10}$$

We have performed a detailed study of the entire familly $0 < h < \infty$. Figure 1, taken from a collection of several hundred pictures, represents the relative motion of the two satellites. For convenience, we shall think of the special case $m_2 \gg m_3$, and identify the origin of the $(\xi, \eta)$ system with the satellite $M_2$; the curves then simply represent the motion of $M_3$. An interesting feature is that the third body always escapes either upward or downward, but never stays close for ever. This is in agreement with a general result by Marchal (1977) which shows that the set of "capture orbits" is of zero measure. For a more detailed explanation of the equation of motion and of the orbits, see Hénon and Petit (1986) and Petit and Hénon (1986).

# 3 Chaotic scattering

The familly exhibits an interesting feature that we call "transitions". Roughly speaking, when $h$ varies continuously, one can observe discontinuities. For given values of $h$, the orbit shape changes suddenly and an orbit that used to escape downward starts to escape upward, or the converse. This is the phenomenon that we want to develop now.

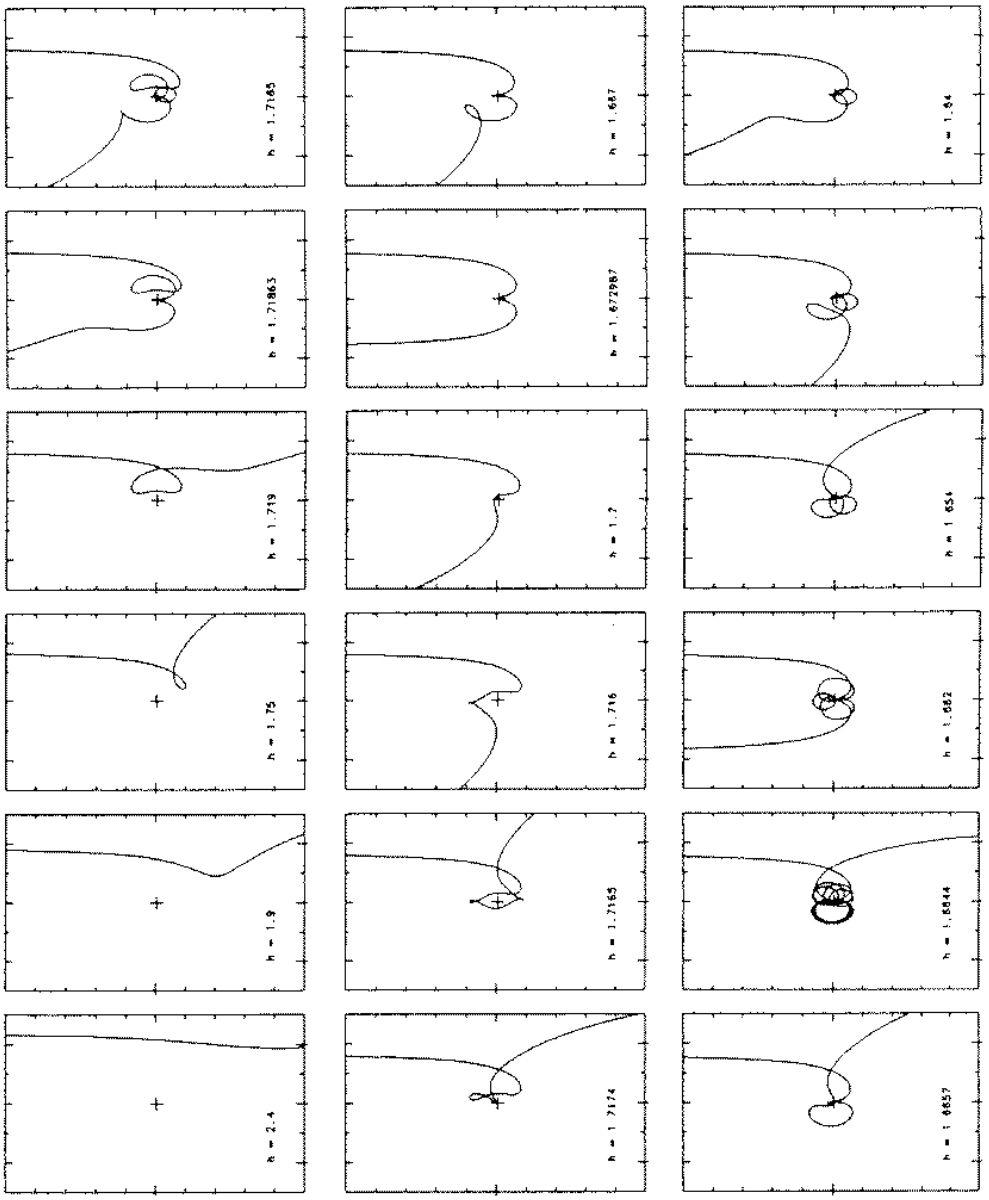


**Fig. 1:** Family of encounter orbits. Each frame corresponds to one particular value of the reduced impact parameter $h$. The curve represents the relative motion of one satellite with respect to the other, in Hill's coordinate ($\xi$ in abscissa, $\eta$ in ordinate). The first approach is downwards from $\eta = +\infty$.

Consider an example. When $h$ decreases from large values, the shape of the orbit changes continuously with $h$ and the third body always escapes downward (first four plots of fig. 1). Suddenly, something happens and it escapes upward. The change (transition) occurs for $h_{max} = 1.718779940$. It can be thought of as a discontinuity of the shape of the orbit. But we need a more quantitative description of this discontinuity. If we look at the equations of motion, we find that there is no true singularity in it (the $1/\rho^2$ singularity can be removed by the Levy-Cevita regularization). So the position of the third body at a given time $t$ is a continuous function of the impact parameter $h$. The orbits are absolutely not chaotic. But if we look at parameters describing the asymptotic motion as functions of $h$, we see very sharp variations at values of $h$ corresponding to the changes of escape side. Especially, consider the final impact parameter $h'$ (defined from the mean motion for $t \to \infty$). Using the Jacobi integral, it can be shown that $|h'| \geq h$. In the general case of asymptotically eccentric orbits, we call $k$ the "reduced" (transformed into Hill's coordinates) eccentricity of the relative motion. One can show that

$$\Gamma = \frac{3}{4}h^2 - k^2. \tag{11}$$

When we start with circular orbits and finish with eccentric orbits, the equality

$$\Gamma = \frac{3}{4}h^2 = \frac{3}{4}h'^2 - k'^2 \tag{12}$$

holds. Therefore, $h'^2 \geq h^2$, which leads to the previous inequality. A downward escape corresponds to $h' > 0$ and an upward escape to $h' < 0$. Therefore a change of escape side leads to a change of sign for $h'$ and a discontinuity of step at least $2h$. This is what we mean by discontinuity. The set of discontinuity values of $h$ being very complex, we shall speak of "chaotic" behavior of the familly.

We will now describe rapidly the set of discontinuities. Consider an orbit defined by an arbitrary value $h_0$. Typically, the following happens: when decreasing $h$ from $h_0$, the orbit changes continuously down to $h_1$ where there is a discontinuity in the sense defined above. We call this a "transition value". Numerically, it is not difficult (even if time consuming) to localize this value with any accuracy. Similarly, if we increase $h$ from $h_0$, we reach a second transition value $h_2$. The interval between $h_1$ and $h_2$ is called a "continuity interval". There are two particular cases: a continuity interval ranges from $h_{max}$ to $\infty$, an other one ranges from 0 to $h_{min} = 1.336117188$ (figure 2). Suppose we have localized an interval of continuity. We do it again, starting from another value $h_0$ out of the range $[h_1, h_2]$. We find another interval of continuity and so on. One could expect to find all the intervals to be contiguous. This would give an exhaustive description of the orbit family.

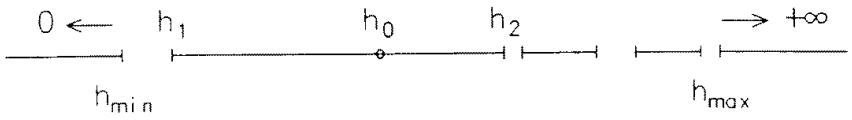



**Fig. 2:** A Schematic representation of the largest continuity intervals.

But life is not that simple. Experiment shows that intervals are never contiguous. If one takes a point in an unexplored interval, one will find a new continuity interval which doesn't touch a previous interval neither on the left nor on the right. This gives birth to two new unexplored intervals. This goes on and on to infinity. One result is that there is no possible exhaustive description of the family. This must remind the reader of the classical definition of the Cantor set. The difference here is that the intervals are not regularly ordered. The actual structure can be seen on the function $h' = f(h)$ (figure 3). We can easily see the two external intervals $[0, h_{min}]$ and $[h_{max}, +\infty]$ and the three largest inner continuity intervals.

# 4 Some hints

We shall now try to explain how the discontinuities occur. In order to reconcile the continuity of the orbits with the discontinuity of the asymptotic behavior ($h'$), the family must go through an orbit with infinite capture time. This is achieved by having an orbit asymptotic to a periodic orbit. For example, figure 4 represents the transition orbit we find when decreasing $h$ from large values: $h = h_{max}$. The orbit tends to a bean shaped periodic orbit. This limiting orbit is easily identified: it belongs to the one-parameter family $a$ of periodic orbits, emanating from the Lagrangian point $L_2$ (Hénon 1969, Fig. 2). It is an unstable periodic orbit, which is necessary in a Hamiltonian system since it admits an asymptotic orbit.

It will be helpful to introduce at this point a *surface of section* defined for instance by $\eta = 0$ and $\dot{\eta} > 0$: for each crossing of an orbit with the $\xi$ axis in the positive direction ($\eta$ increasing), we plot a point with the coordinates $\xi$, $\dot{\xi}$ (figure 5). An orbit is then represented by a sequence of points. For a given value of $\Gamma$, a point in the surface of section defines completely the corresponding orbit: $\xi$, $\dot{\xi}$, $\eta$ are immediately known and $\dot{\eta}$ can be computed from (**??**). In particular the next intersection point can be found. This defines a mapping of the surface of section onto itself, known as a *Poincaré map*.

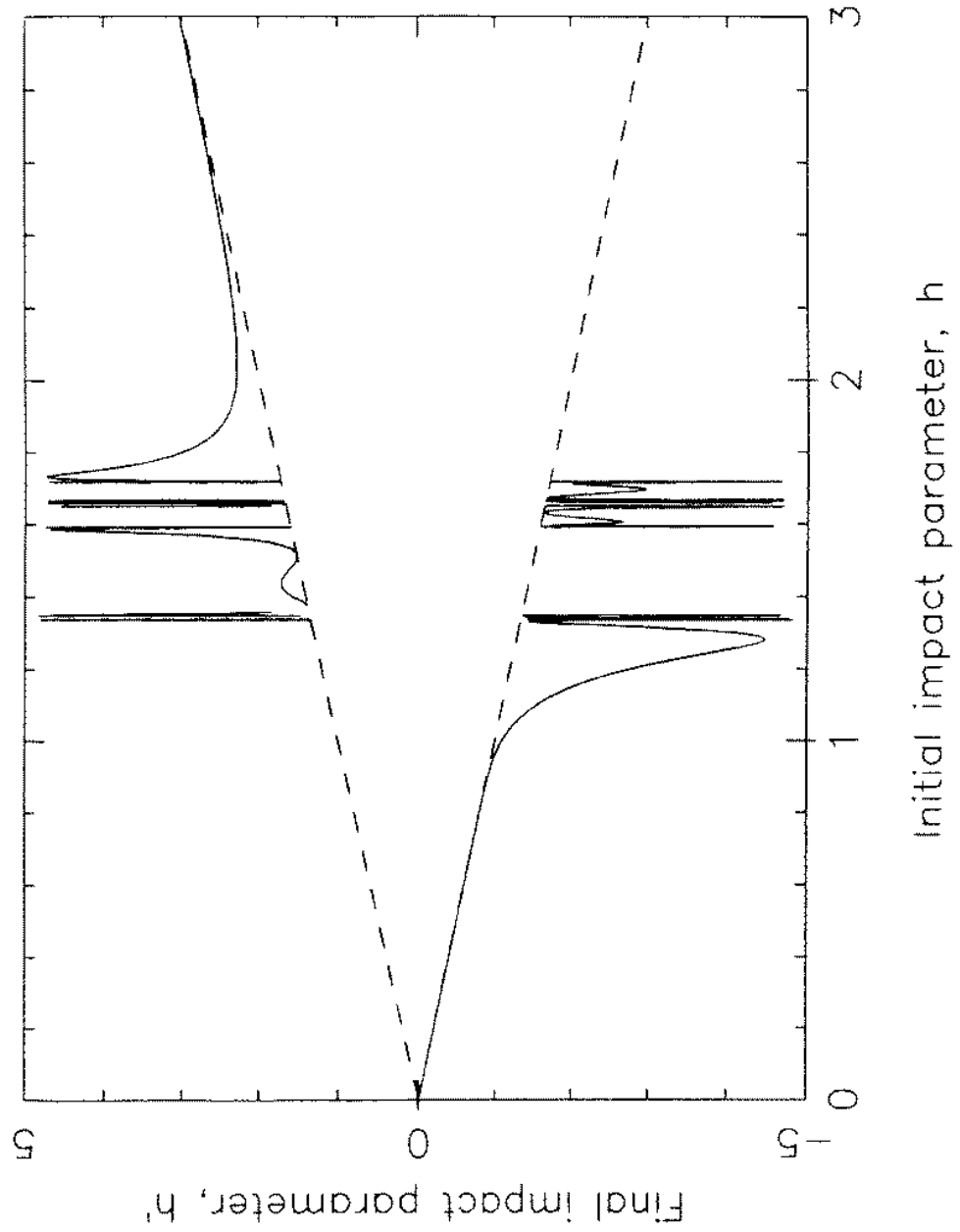



**Fig. 3:** Final impact parameter $h'$ as a function of the initial impact parameter $h$. The region between the two dashed lines is forbidden.

Note that for the orbits we are concerned with, we get generally a finite (small) number of points in the surface of section: three points for the orbit with $h = 1.71863$ for instance. An orbit can also have no point at all in that surface ($h > 2.4$). The periodic orbit is represented by a fixed point $P$ (figure 5). The stability index of that orbit is of order 320, corresponding to two real eigenvalues $\lambda_1 = 1/640$ and $\lambda_2 = 640$. The eigenvalue smaller than 1 in modulus ($\lambda_1$) is associated with a one-parameter family of *incoming orbits* tending towards the periodic orbit. The orbit of figure 4 is a member of this family. An orbit of this family is represented by an infinite sequence of points on the *stable invariant manifold* $W_s$ of $P$ and converges exponentially towards $P$ ($Y_0$, $Y_1$, $Y_2$, ...). Since the periodic orbit is unstable, there are also *outgoing orbits*, tending towards the periodic orbit for $t \to -\infty$. They form a one-parameter family associated with the eigenvalue larger than 1 in modulus ($\lambda_2$). They are represented by points (..., $Z_{-2}$, $Z_{-1}$, $Z_0$, ...) located on the *unstable invariant manifold* $W_u$ of $P$ and which diverge exponentially from $P$.

Consider now an orbit of our family with $h$ slightly different from $h_{max}$, say larger. The points in the surface of section are slightly beside $W_s$ (crosses on the picture). They stay close to $W_s$ until they reach the vicinity of $P$, then they go away along $W_u$. An important point is that $\lambda_2$ is positive. So the points go along only one branch of $W_u$. Here, it is the upper right branch. The corresponding orbits are quite regular. Particularly, they all escape downward and vary continuously when $h$ increases (figure 6a). This accounts for the continuity interval for $h > h_{max}$.

For $h = h_{max}$, the point crosses $W_s$ and for $h < h_{max}$, the points escape along the left branch of $W_u$. The two branches of $W_u$ are in two different parts of phase space. This explains the transitions. The orbits for $h < h_{max}$ are shown on figure 6b. Things are much more complicated than before. Sometimes orbits escape upward, some time downward. So there is no continuity interval on the left of $h_{max}$. This explains the complex structure of the continuity intervals. For $h < h_{max}$, instead of escaping directly, the orbit will first go in the vicinity of an other unstable periodic orbit. This orbit will itself give birth to a transition phenomenon, that we shall call a *second order transition*. In this way, one can construct a hierarchical structure of transitions of higher and higher order. Suppose we have an orbit going close to one periodic orbit then close to a second one. By changing $h$, we can push the points in the surface of section closer to the first fixed point. Particularly, one can manage to have the same pattern along $W_u$ and one or more additional points in the vicinity of $P$. This corresponds to orbits with the same escape but with one or more additional turns around the first periodic orbit (figure 7). In the first plot, the orbit follows the periodic orbit during half a turn, in the second during one and a half and in the third during two and a half (even if this is not visible one the figure). This gives rise to a geometrical progression of ratio $\lambda_1$ in the values of $h$.

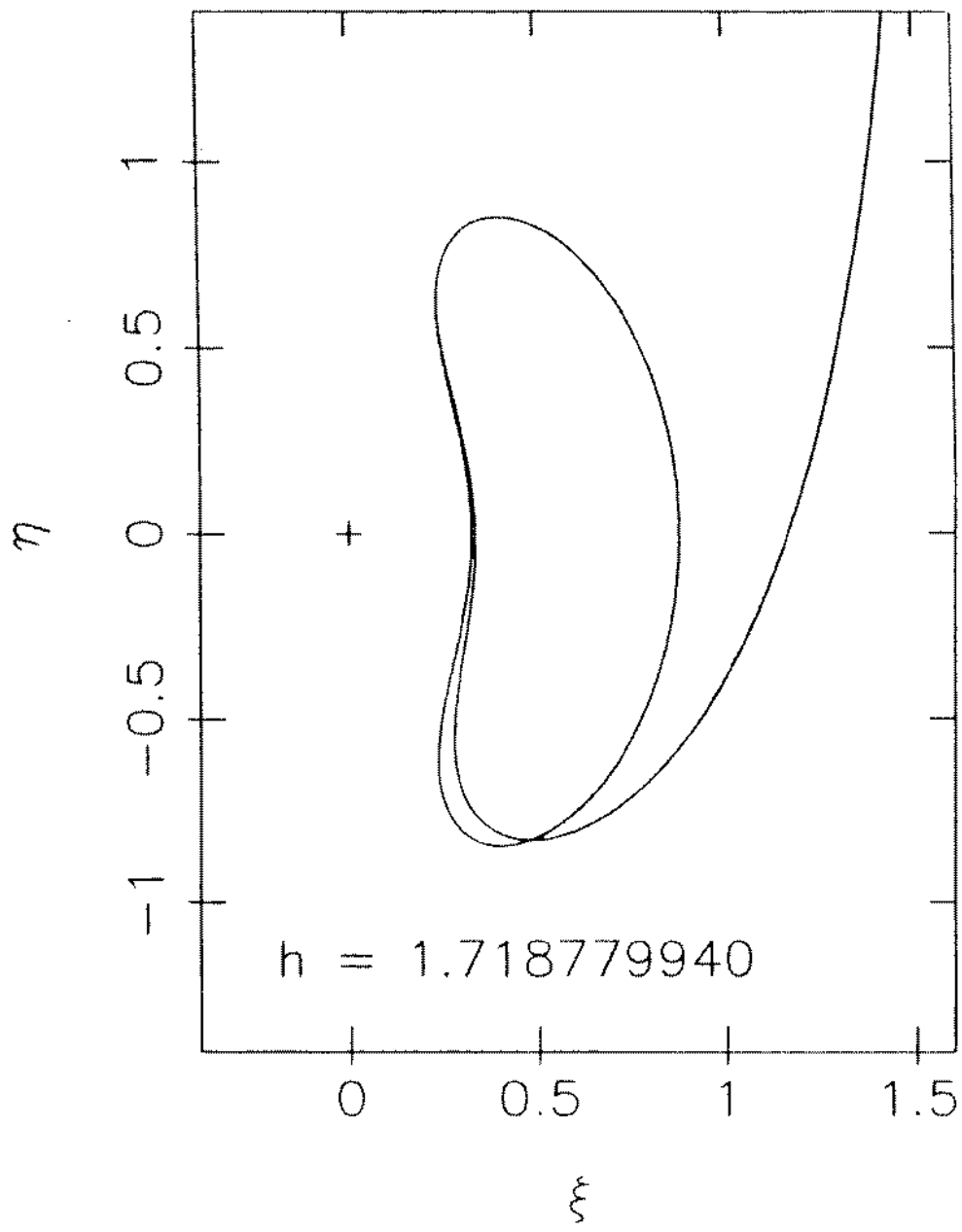



**Fig. 4:** An orbit of the Satellite Encounter family which is asymptotic to an unstable periodic orbit.

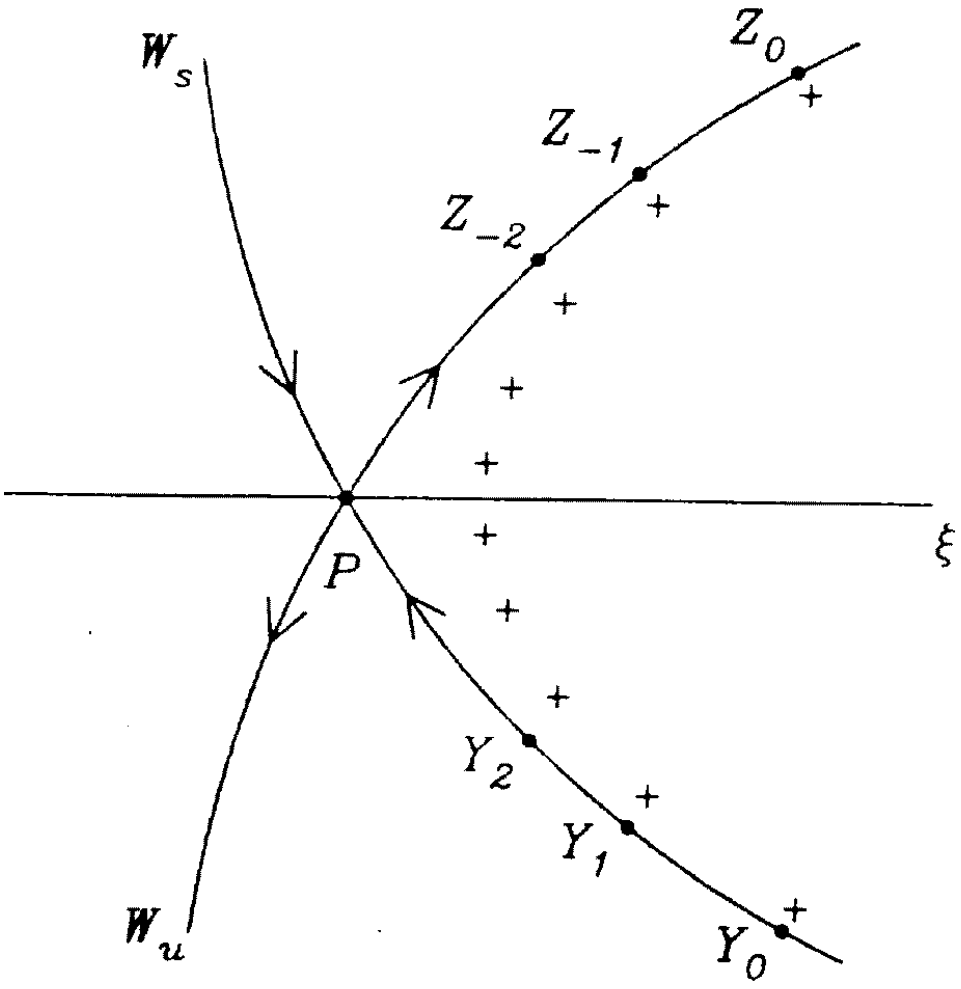



**Fig. 5:** Sketch of the surface of section. The value of $\lambda_1$ has been artificially increased to show the structure more clearly.

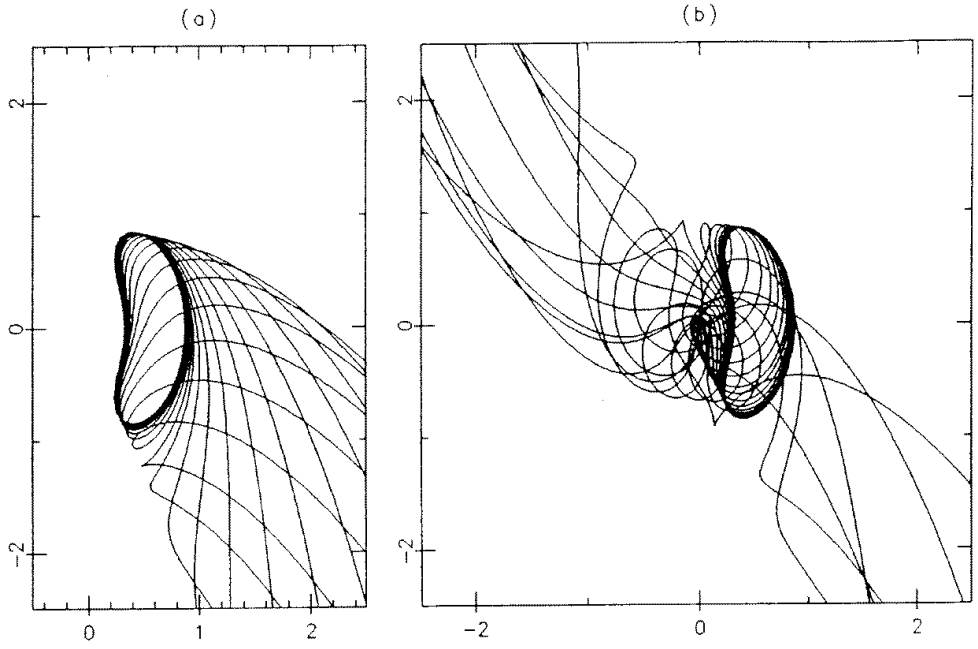



**Fig. 6:** (a) Outgoing orbits for $h$ above the critical value $h_{max}$. (b) Outgoing orbits for $h$ below $h_{max}$.

From all our numerical integrations, it seems that only two family of periodic orbits are involved: family $a$ mentioned above and the symmetrical family $b$ also described in Hénon 1969.

The necessary ingredients for this kind of behavior is the existence of periodic orbits and heteroclinic or homoclinic points (intersection points of invariant manifolds of two different or one single periodic orbit). But it is very difficult to go any further with this problem due to the large value of the eigenvalue ($\sim 640$).

Other authors have observed similar behavior in scattering problems. Recently, Jung and Scholz (1988) have studied the scattering of a charged particle by a magnetic dipole. Due to the smaller value of the eigenvalue, they have been able to compute the stable and unstable manifolds of the periodic orbit with great details (solid line in Fig. 6 of their paper). It happens in their case that they need only one periodic orbit due to the presence of a homoclinic point. The dots in that figure represent the first intersection of the surface of section when varying the initial parameter. They also transported the invariant manifold into the space of initial parameters (Fig. 11 of the same paper). If now we draw a line across the figure, corresponding to the variation of one of the initial parameters, we shall see immediately the existence of a complex structure. For comparison, we did the same thing in the satellite problem (figure 8) but the structure is far too thin to be seen.

# 5 The inclined billiard

According to Moser (1973), in the vicinity of a homoclinic point, it is theorically possible to define a symbolic dynamics which is Bernoulli. This gives a better description of the dynamics of the system. But in our problem, we haven't been able to define it so far. So a model problem was designed which is complex enough to exhibit all the features we are interested in, and simple enough so that all the calculations can be done analytically. This model is the *inclined billiard* (Hénon, 1988). It is defined as follows: a particle moves in the $(X, Y)$ plane and bounces elastically on two fixed disks with radius $r$ and with their centers in $(-1, -r)$ and $(1, -r)$ respectively. In addition, it is subjected to a constant acceleration $g$ which pulls it in the negative $Y$ direction. To make the computation affordable, one considers the limit where $r$ is large and approximates the circles (disks) by parabolas. The "disks" extend then from $-\infty$ to $\infty$ in the $X$ direction and the number of rebounds of the particle on them is now always infinite. We suppose that initially the particle is started at $(h, Y_0)$ where $Y_0$ is large. So we can neglect the thickness of the profile of the disks. Only the slopes are of consequence.

We consider the surface of section defined by the collisions. The coordinates in this surface

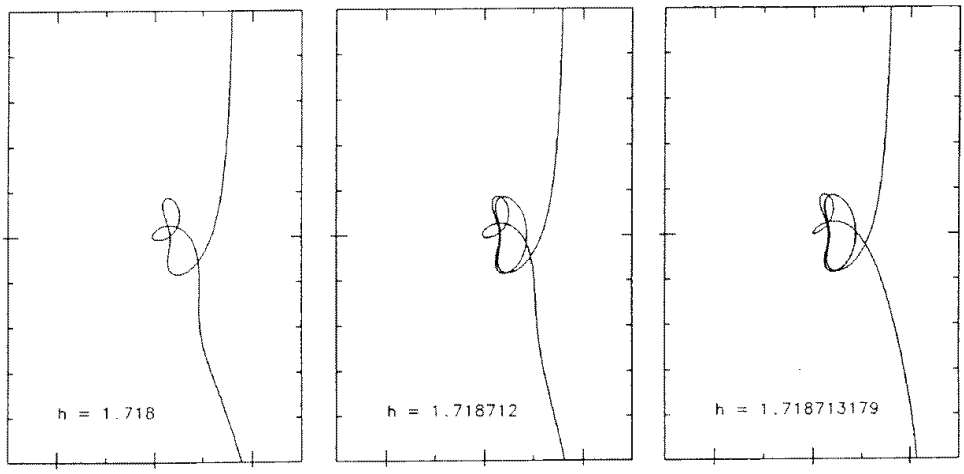



**Fig. 7:** Three orbits with essentially the same outgoing but different behavior during the close encounter. The orbit goes along the unstable periodic orbit for half a turn on the left, for one and a half in the middle and for two and a half on the right plot.

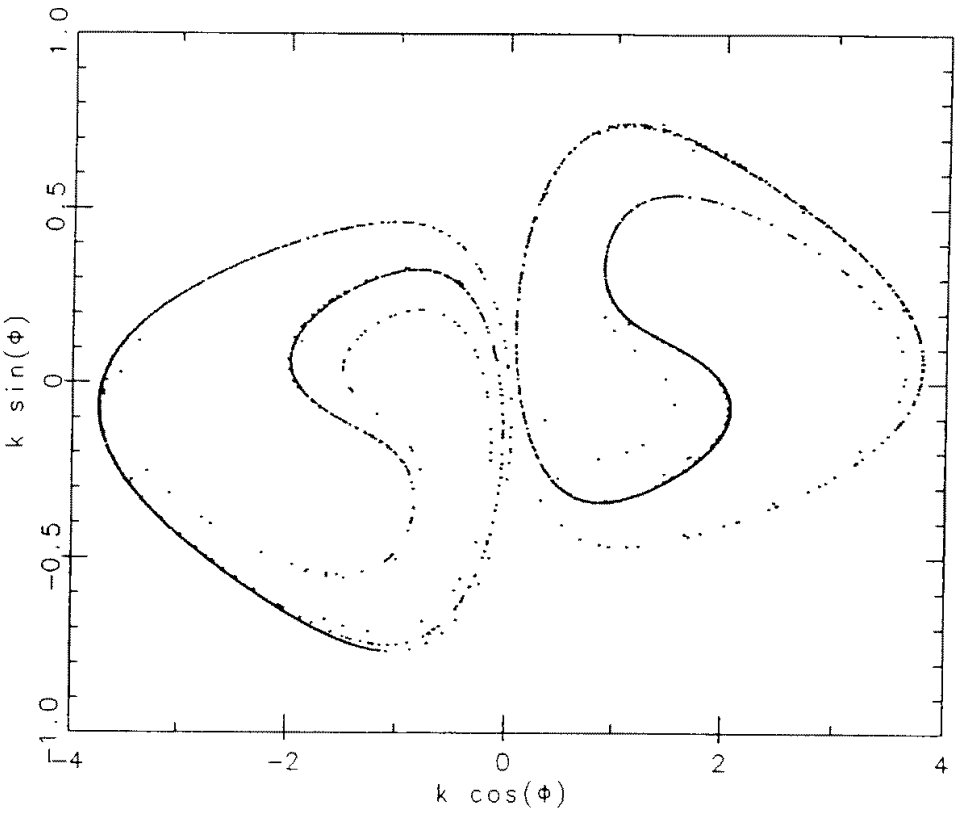



**Fig. 8:** Image of the invariant manifolds in the space of initial parameters.

are $X$, the horizontal abscissa of the collision and $W$, the tangential velocity of the particle. After a tedious calculation, one can finally obtain the mapping:

$$\begin{aligned} u_{j+1/2} &= W_j + \frac{\sqrt{2E}}{r} f(X_j), \\ X_{j+1} &= X_j + \frac{2\sqrt{2E}}{g} U_{j+1/2}, \\ W_{j+1} &= U_{j+1/2} + \frac{\sqrt{2E}}{r} f(X_{j+1}), \end{aligned} \tag{13}$$

with

$$f(X) = \begin{cases} X+1 & \text{for } X < 0, \\ X-1 & \text{for } X \geq 0, \end{cases}$$

where $U_{j+1/2}$ is the horizontal velocity between the collisions and $E$ the energy of the particle. Arbitrarily, we have decided that the intersection of the two disks belongs to the right disk. We introduce a new parameter $\Phi$, related to the energy by

$$\cosh \Phi = 1 + \frac{4E}{gr}, \tag{14}$$

and the new variables $u$ and $w$

$$U = u\sqrt{\frac{g}{2r}\left(2 + \frac{4E}{gr}\right)}, \qquad W = w\sqrt{\frac{g}{2r}\left(2 + \frac{4E}{gr}\right)}, \tag{15}$$

in order to rewrite the mapping in a dimensionless form:

$$\begin{aligned} X_{j+1} &= X_j \cosh\Phi + w_j \sinh\Phi - s_j(\cosh\Phi - 1), \\ w_{j+1} &= X_j \sinh\Phi + w_j \cosh\Phi - (s_j \cosh\Phi + s_{j+1}) \tanh\frac{\Phi}{2}, \\ s_j &= \operatorname{sign} X_j. \end{aligned} \tag{16}$$

The parameter $\Phi$ cannot be eliminated since it is related to the eigenvalue of the fixed point. It is easy to show that there are five kinds of asymptotic regimes:

1. right-escaping orbit: $X_j \to +\infty$, $w_j \to +\infty$.
2. right-asymptotic orbit: $X_j \to +1$, $w_j \to 0$.
3. left-escaping orbit: $X_j \to -\infty$, $w_j \to -\infty$.
4. left-asymptotic orbit: $X_j \to -1$, $w_j \to 0$.
5. oscillating orbit: $X_j$ and $w_j$ are bounded but have no limit.

We will now define a symbolic dynamics to represent the essence of the dynamics of the billiard. To each orbit we associate a semi-infinite sequence $d_1$, $d_2$, $\cdots$ of 0 and 1. The orbit is described by the sequence of points in the surface of section and each point is represented by 0 if it is a collision on the left disk and 1 if it is a collision on the right disk, rather than by its coordinates. Then we define a number $A$ by its binary representation:

$$A = 0.d_1 d_2 d_3 \cdots = \sum_{j=1}^{\infty} 2^{-j} d_j.$$

Clearly, $0 \leq A \leq 1$. A given sequence defines one value of $A$, but there might be *two* sequences with the same value of $A$:

- If $A$ is of the form $k.2^{-n}$, where $k$ and $n$ are natural numbers, $A$ is called a round number and has two representations: $0.d_1 \cdots d_{n-1}0111\cdots$ and $0.d_1 \cdots d_{n-1}1000\cdots$.
- In the other case, $A$ is a non-round number and the sequence is oscillating.

There is a simple correspondence between the types of orbits, the $D$ sequence and $A$.

| orbit | $D$ sequence | $A$ |
|---|---|---|
| right-escaping | 1-ending | round |
| right-asymptotic | 1-ending | round |
| left-escaping | 0-ending | round |
| left-asymptotic | 0-ending | round |
| oscillating | oscillating | non-round |

In a continuity interval, the orbit changes continuously, so $A$ is constant. This suggests to look at the function $A(h)$. Figure 9 shows the numerical result for $\lambda = e^{\Phi} = 3$. The reader will have recognized a *Devil's staircase*. It is possible to explain completely this figure, provided that $e^{\Phi} \geq 3$. One can show the following:

- $A$ is a non-decreasing function of $h$.
- $A$ is a continuous function of $h$.
- If $A$ is non-round, it corresponds to a unique value of $h$ defined by:

$$h = (e^{\Phi} - 1) \sum_{j=1}^{+\infty} e^{-j\Phi} s_j. \tag{17}$$

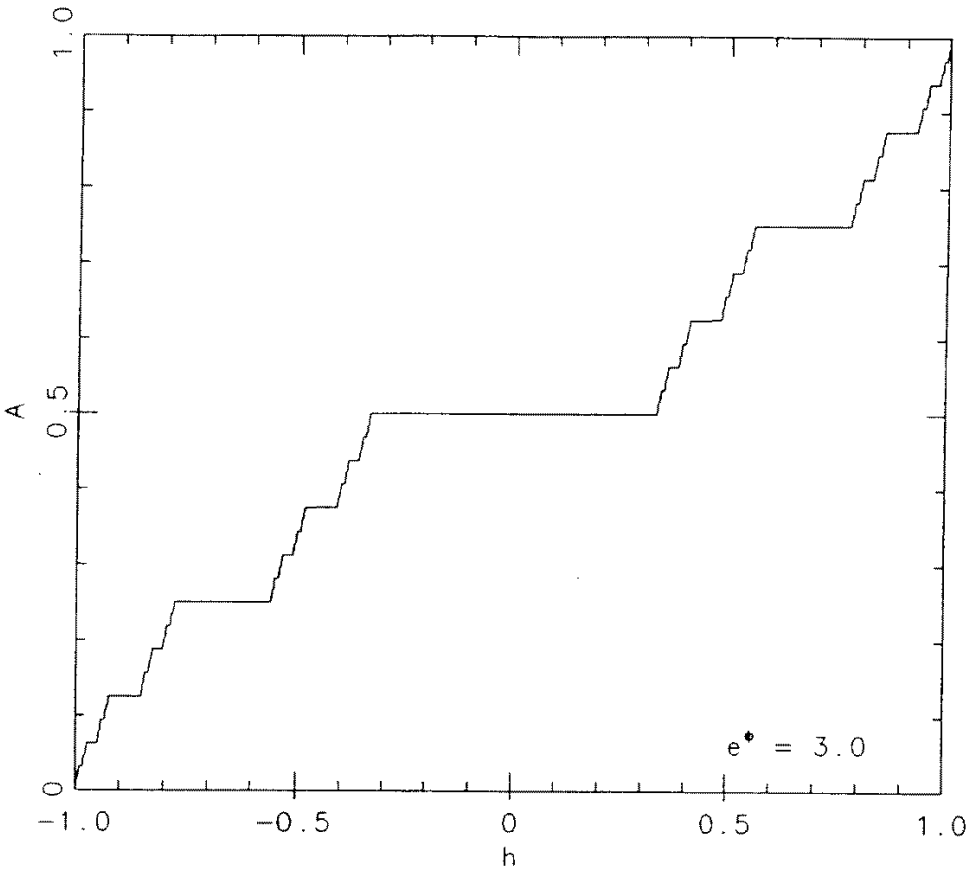



**Fig. 9:** The function $A(h)$ for $e^{\Phi} = 3$.

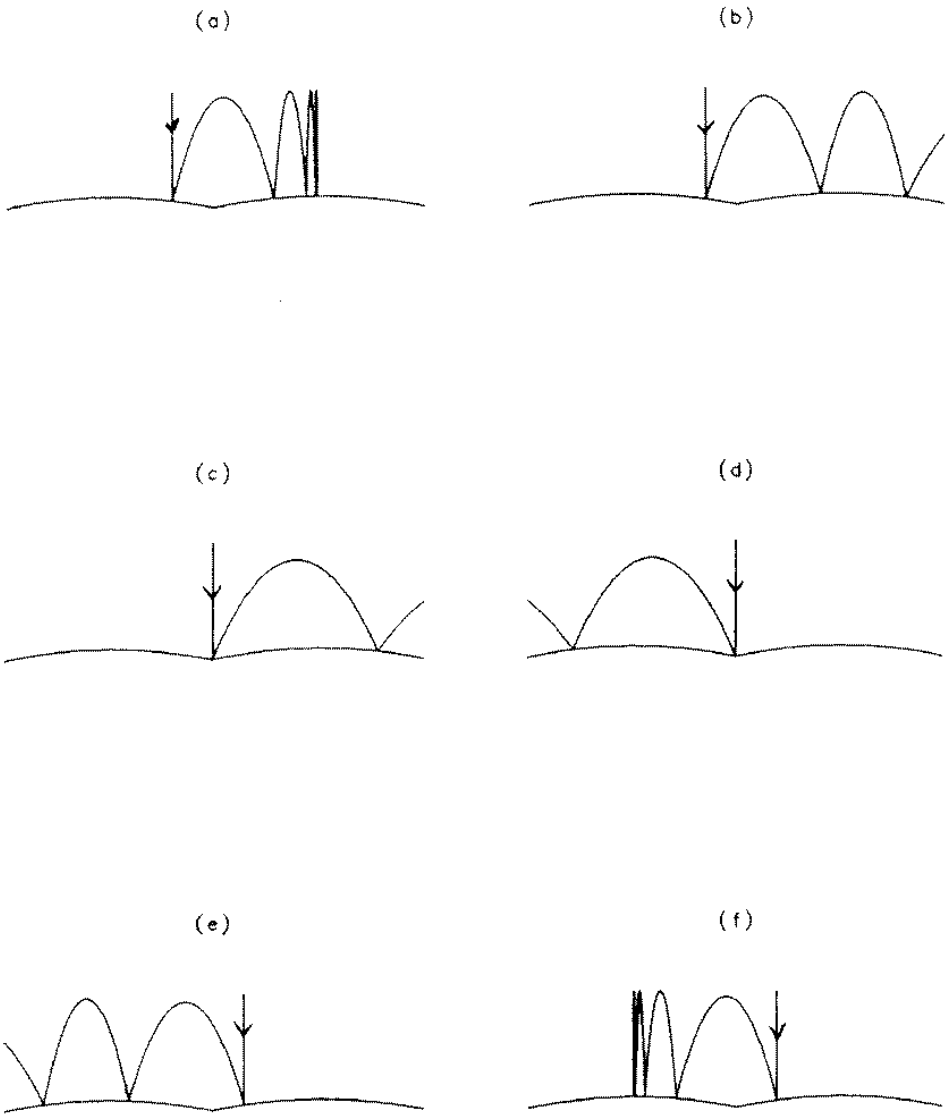



**Fig. 10** Structure of the $h$-orbits corresponding to a round value of $A$ (here $A = \frac{1}{2}$). (a) right-asymptotic orbit; (b), (c) right-escaping orbits; (d), (e) left-escaping orbits; (f) left-asymptotic orbit.

- If $A$ is round, things are more interesting. $A$ corresponds to a horizontal step, that is to a continuity interval $h_- \leq h \leq h_+$. All the steps have the same structure. Consider the central step on figure 9 (the largest). The leftmost point corresponds to an asymptotic orbit ($a$) (figure 10) on the right: $A = 0.0111\cdots = 1/2$. When $h$ increases, the orbit escapes rightward and the sequence doesn't change. This is true for interval ($b$). After orbit ($c$), there is a discontinuity: the slope for the first collision changes suddenly and we go to orbit ($d$). Now, $A = 0.1000\cdots$, that is still $1/2$. Here, there is a small miracle: the orbit changes completely through ($c$) but not $A$. The discontinuity at $X = 0$ disappears completely in $A$. Then there is the interval ($e$) of left escaping orbits and finally the left asymptotic orbit ($f$). One can easily compute the values of $h_-$ and $h_+$ from $A$:

$$\begin{aligned} h_- &= (e^\Phi - 1)\sum_{j=1}^{n-1} e^{-j\Phi} s_j - (e^\Phi - 2)e^{-n\Phi}, \\ h_+ &= (e^\Phi - 1)\sum_{j=1}^{n-1} e^{-j\Phi} s_j + (e^\Phi - 2)e^{-n\Phi}. \end{aligned} \tag{18}$$

We have thus fully solved the "inverse problem" and obtained a complete classification of $h$-orbits. For non-round $A$, there is a single $h$ given by (**??**), corresponding to an oscillating orbit. For round $A$, there is a finite closed interval of $h$ values $h_- \leq h \leq h_+$ (or an infinite interval in the special cases $A = 0$ and $A = 1$). This explains the *Devil's staircase* appearance of the figures.

There are two additional results that can be proved:

- The curve $A(h)$ has exact self-similarity. The curve as a whole extends from $h = -1$ to $h = +1$ and from $A = 0$ to $A = 1$. In the lower left corner is an exact replica of the whole picture, reduced by a factor $e^\Phi$ horizontally and 2 vertically, extending from $h = -1$ to $h = -1 + 2e^{-\Phi}$ and from $A = 0$ to $A = 1/2$. There is an identical replica in the upper right corner.

- The set of values of $h$ corresponding to bounded orbits forms a Cantor set, with measure 0 and with fractal dimension

$$\ln(2)/\Phi. \tag{19}$$

  In the borderline case $e^\Phi = 3$, we obtain exactly the classical Cantor set (repeated exclusion of the middle third). The asymptotic orbits form an enumerable subset of the bounded orbits; this subset also has the dimension (**??**).

# References


[1] Dermott S. F. & Murray C. D., 1981. Icarus **48**, p. 1.

[2] Eckhardt B. & Aref H., 1989. Phil. Trans. R. Soc. Lond. A **326**, p. 655.

[3] Eckhardt B. & Jung C., 1986. J. Phys. A **19**, p. *L*829.

[4] Fitz D. E. & Brumer P., 1979. J. Chem. Phys. **70**, p. 5527.

[5] Goldreich P. & Tremaine S., 1979. Nature **277**, p. 97.

[6] Goldreich P. & Tremaine S., 1980. Astrophys. J. **241**, p. 425.

[7] Gottdiener L., 1975. Molecular Physics **29**, p. 1585.

[8] Hénon M., 1969. Astron. Astrophys. **1**, p. 223.

[9] Hénon M., 1988. Physica D **33**, p. 132.

[10] Hénon M. & Petit J-M., 1986. Celestial Mechanics **38**, p. 67.

[11] Jung C. & Scholz H-J., 1987. J. Phys. A **20**, p. 3607.

[12] Jung C. & Scholz H-J., 1988. J. Phys. A **21**, p. 2301.

[13] Marchal C., 1977. J. Differ. Equations **23**, p. 387.

[14] Noid D. W., Gray S. K. & Rice S. A., 1986. J. Chem. Phys. **84**, p. 2649.

[15] Petit J-M. & Hénon M., 1986. Icarus **66**, p. 536.

[16] Rankin C. C. & Miller W. H., 1971. J. Chem. Phys. **55**, p. 3150.

[17] Schlier C. G., 1983. Chemical Physics **77**, p. 267.

[18] Yoder C. F., Colombo G., Synnott S. P. & Yoder K. A., 1983. Icarus **53**, p. 431.

# Chapitre 3

# APPLICATIONS DE POINCARE DES SYSTEMES HAMILTONIENS

**[7]**

# "Polynomial approximations of Poincaré maps for Hamiltonian systems."

FROESCHLE, C. & PETIT, J-M.

*Astron. Astrophys.* **238** (1990)

# Polynomial approximations of Poincaré maps for Hamiltonian systems.

Froeschlé C.
Observatoire de Nice, France
claude@obs-nice.fr

Petit J-M.
C.N.R.S., Observatoire de Nice, France
petit@obs-nice.fr

**Abstract**

Different methods are proposed and tested for transforming a non linear differential system, and more particulary a hamiltonian one, into a map without having to integrate the whole orbit as in the well known Poincaré return map technique. We construct piecewise polynomial maps by coarse-graining the phase surface of section into parallelograms using values of the Poincaré maps at the vertices to define a polynomial approximation within each cell. The numerical experiments are in good agreement with both the real symplectic and Poincaré maps. The agreement is better when the number of vertices and the order of the polynomial fit increase. Computations of Lyapunov Characteristic Exponents give a measure of how well the fit approximate the different maps.

## 1 Introduction

Poincare return maps are widely used in all fields of non linear dynamics including celestial mechanics (see Hénon 1981). However return maps often require a lot of computer time since it is necessary to integrate the differential equations to plot them. In order to investigate the long-term stability of test particles in the solar system many authors have therefore sought simpler and more tractable equations which approximate at least qualitatively the original newtonian equations. This is the aim of classical perturbation theory which uses canonical transformations of the variables in order to successively push the non integrable part of the

hamiltonian to higher and higher orders. Such an approach continued until the implications of Poincaré's theoretical work in this field were fully appreciated in the sixties with the computer revolution (see Hénon 1981). Poincaré showed that most hamiltonian systems do not possess the integrals required to reduce the solution to quadratures and that trajectories near unstable periodic orbits possess extraordinary complexity. Then in order to keep this complexity one has to resort to numerical experiments. Therefore it is important to increase the computing speed. A first attempt was made using numerical averaging which led to an order of magnitude increase in speed. Let us emphasize that this kind of approach,used by Schubart (1968) in the framework of asteroidal resonant motions, differs radically from the so called averaging principle (see Arnold 1976): Schubart does not drop short periodic terms and does not use series expansions. Therefore with Schubart's averaging there is no restriction for eccentricities or resonance type, and chaotic behavior is still present (see Froeschle and Scholl 1981). The next step which drastically increases the speed of the computation is to replace the hamiltonian dynamical system by a symplectic mapping (an area preserving mapping for systems with two degrees of freedom). This programme has been carried out quite successfully to analyze the long-term behavior of asteroids near resonances with Jupiter (see Wisdom 1982, Murray 1986, Sidlichovsky 1986). However the so-called Chirikov- Wisdom method for building such mappings near resonances is not reliable with the same accuracy all over the phase space. The key point is: the truncated secular part of the hamiltonian should be solvable analytically as well as not changing the topology of the phase space too much. Actually the higher the eccentricity becomes the less reliable is the mapping. Therefore even for the 3/1 resonance, (whose shape was brilliantly explained by Wisdom using his mapping), after the first increase of eccentricity the features of the real orbits may be significantly different from those of the mapping orbits. In the same spirit Sagdeev and Zaslavsky (1987) and Petrosky and Broucke (1988) have derived a two dimensional mapping which describes the dynamics of long period cometary motion within the framework of the restricted three body problem. The basic idea is again to separate out an integrable part (the two body motion) valid for large period of time where the comet is far from the "collision zone" (i.e. the zone where Jupiter interacts strongly with the comet) and a collisional part. This collision induces a change in the energy and angular momentum of the comet and this is modelled by taking into account, using Fourier analysis, the fact that the time interval between "collisions" is very large compared with the "collision" duration. Hadjidemetriou (1986) also solved the unperturbed equations and found the corresponding mapping on a surface of section. Then he perturbed the mapping in such a way that some structures of the phase space known through numerical experiments (stability or unstability of some families of periodic orbits for the given example: 3/1 commensurability) was included in the perturbed mapping. Finally, and again in the framework of the restricted three-body problem (Sun-Jupiter-test particle), a mapping has been derived by Duncan, Quinn

and Tremaine (1989) in order to study the long term evolution of planetesimal like orbits i.e. orbits with small eccentricities in the solar system. They used the approximation based on the fact that planets have small eccentricities and the fractional difference in semi major axes between the orbit of a test particle which is nearly circular and the nearest planet is small. Then the perturbations to the motion of the test particle are localised near conjunction with the planet and the first-order perturbations to the orbital elements can be computed analytically using Hill's equations (Hénon and Petit 1986) or equivalently Gauss' or Lagrange's equations. Away from conjunction, as for the nearly parabolic case, the perturbing effect of the planet is assumed to be negligible and hence the orbital elements after conjunction remain unchanged until the next conjunction. All these mappings have been devised using the same kind of approach: use of special features of the motion i.e. resonant orbits, near parabolic or near circular orbits. These mappings are reliable as long as one remains within the domain of validity of the approximations, made in order to isolate an integrable part and some instantaneous perturbations. In any case, these mappings are not reliable for close approaches, which are at the origin of the stochastic motion of short-period comets and of Appolo-Amor asteroids (Milani et al 1989). This underlying stochasticity is at the basis of the mappings worked out by Rickman and Vaghi (1976) and Froeschle and Rickman (1980). Within the framework of the elliptic restricted three body problem their mapping consists in regarding the perturbations of the orbital elements for one orbital revolution as a stochastic variable. More precisely, for different cells in phase space, separate distributions of perturbations are calculated. The orbital evolution is then simulated as a sequence of perturbations randomly chosen according to the calculated distribution. The validity of such procedures has been recently studied by Froeschle and Baille (1989) who have shown some drawbacks of the method which may be considered as being valid only from a statistical point of view and again only within some regions of phase space. The point to be emphasized is that even inside regular regions the method introduces a diffusion which has no physical meaning.

All these mappings are ad hoc and are valid only in some regions of phase space and for specific purposes. We wish to build a mapping valid everywhere in phase space following an idea already used by Varosi et al (1987) but in the framework of non hamiltonian systems, i.e. systems where attractors do exist. The method consists of coarse-graining the phase space surface of section and then interpolating the values of the image of a point. In section 2 we present and discuss the different interpolations we have used in addition of the linear one already used by Varosi et al (1987). In section 3 we give the numerical results concerning the different synthetic mappings compared with both the standard map and the Poincare map of the restricted three body problem.

# 2 The synthetic mappings

The basic idea of the synthetic mapping is to interpolate the image of a point, when given the images of a set of points located at the vertices of a grid. The simplest way to achieve this is to use linear interpolation (Varosi et al, 1987). Unfortunately, this requires a rather fine graining of the phase space, i.e. we have to compute a lot of points on the grid. We shall try to find more accurate methods. The more commonly used methods of higher order are the Taylor and Spline interpolations.

## 2.1 Taylor interpolation

We know the value of the functions $x^{(n+1)} = f(x^{(n)}, y^{(n)})$ and $y^{(n+1)} = g(x^{(n)}, y^{(n)})$ at discrete points $x_i$ and $y_j$. The points are equally spaced in $x$ and $y$. We want to interpolate the values of $f$ and $g$ at any point. We first derive the interpolation formula in one dimension. We write the Taylor expansion for $f$ about a point in the middle of a given interval, $x' = (x_i + x_{i+1})/2$:

$$\begin{aligned} f(x) &= f(x') + (x-x')f'(x') + \frac{(x-x')^2}{2}f''(x') + \frac{(x-x')^3}{6}f'''(x') \\ &\quad + \frac{(x-x')^4}{24}f^{(4)}(x') + \frac{(x-x')^5}{120}f^{(5)}(x') + \mathrm{O}\left((x-x')^6\right). \end{aligned} \tag{1}$$

We use a centered point to insure the symmetry of the formula. We need to know the value of $f$ and its derivatives at $x'$. For this we write (**??**) for $x_{i-2}$, $x_{i-1}$ and $x_i$:

$$\begin{aligned} f(x_{i-2}) &= f(x') - \frac{5h}{2}f'(x') + \frac{25h^2}{8}f''(x') - \frac{125h^3}{48}f'''(x') \\ &\quad + \frac{625h^4}{384}f^{(4)}(x') - \frac{3125h^5}{3840}f^{(5)}(x') + \mathrm{O}(h^6), \\ f(x_{i-1}) &= f(x') - \frac{3h}{2}f'(x') + \frac{9h^2}{8}f''(x') - \frac{27h^3}{48}f'''(x') \\ &\quad + \frac{81h^4}{384}f^{(4)}(x') - \frac{243h^5}{3840}f^{(5)}(x') + \mathrm{O}(h^6), \\ f(x_i) &= f(x') - \frac{h}{2}f'(x') + \frac{h^2}{8}f''(x') - \frac{h^3}{48}f'''(x') \\ &\quad + \frac{h^4}{384}f^{(4)}(x') - \frac{h^5}{3840}f^{(5)}(x') + \mathrm{O}(h^6), \end{aligned} \tag{2}$$

where $h = x_{i+1} - x_i$. The corresponding formulae for $f(x_{i+1})$, $f(x_{i+2})$ and $f(x_{i+3})$ are obtained from those for $f(x_i)$, $f(x_{i-1})$ and $f(x_{i-2})$ respectively by changing the minus signs into plus

signs. Neglecting terms of order 6 or higher, one can inverts (**??**) and write $f(x')$, $f'(x')$, $f''(x')$, $f'''(x')$, $f^{(4)}(x')$and $f^{(5)}(x')$ as functions of $f(x_{i-2})$, $f(x_{i-1})$, $f(x_i)$, $f(x_{i+1})$, $f(x_{i+2})$ and $f(x_{i+3})$.

$$\begin{aligned}
f(x') &= \frac{1}{256}[3f(x_{i-2}) - 25f(x_{i-1}) + 150f(x_i) \\
&\quad +150f(x_{i+1}) - 25f(x_{i+2}) + 3f(x_{i+3})], \qquad (3) \\
hf'(x') &= \frac{1}{1920}[-9f(x_{i-2}) + 125f(x_{i-1}) - 2250f(x_i) \\
&\quad +2250f(x_{i+1}) - 125f(x_{i+2}) + 9f(x_{i+3})], \\
h^2f''(x') &= \frac{1}{48}[-5f(x_{i-2}) + 39f(x_{i-1}) - 34f(x_i) \\
&\quad -34f(x_{i+1}) + 39f(x_{i+2}) - 5f(x_{i+3})], \qquad (4) \\
h^3f'''(x') &= \frac{1}{8}[f(x_{i-2}) - 13f(x_{i-1}) + 34f(x_i) \\
&\quad -34f(x_{i+1}) + 13f(x_{i+2}) - f(x_{i+3})], \\
h^4f^{(4)}(x') &= \frac{1}{2}[f(x_{i-2}) - 3f(x_{i-1}) + 2f(x_i) \\
&\quad +2f(x_{i+1}) - 3f(x_{i+2}) + f(x_{i+3})], \\
h^5f^{(5)}(x') &= [-f(x_{i-2}) + 5f(x_{i-1}) - 10f(x_i) \\
&\quad +10f(x_{i+1}) - 5f(x_{i+2}) + f(x_{i+3})].
\end{aligned}$$

By substituting into (**??**), we obtain:

$$\begin{aligned}
f(x) &= \Phi_{-2}(\frac{x-x'}{h})f(x_{i-2}) + \Phi_{-1}(\frac{x-x'}{h})f(x_{i-1}) \\
&\quad +\Phi_0(\frac{x-x'}{h})f(x_i) + \Phi_1(\frac{x-x'}{h})f(x_{i+1}) \qquad (5) \\
&\quad +\Phi_2(\frac{x-x'}{h})f(x_{i+2}) + \Phi_3(\frac{x-x'}{h})f(x_{i+3}),
\end{aligned}$$

where the functions $\Phi_n(X)$, $n = -2, -1, 0, 1, 2, 3$ read:

$$\begin{aligned}
\Phi_{-2}(X) &= \frac{1}{3840}[45 - 18X - 200X^2 \\
&\quad +80X^3 + 80X^4 - 32X^5], \\
\Phi_{-1}(X) &= \frac{1}{3840}[-375 + 250X + 1560X^2 \\
&\quad -1040X^3 - 240X^4 + 160X^5],
\end{aligned}$$

$$
\begin{aligned}
\Phi_0(X) &= \frac{1}{3840}[2250 - 4500X - 1360X^2 \\
&\quad +2720X^3 + 160X^4 - 320X^5], \\
\Phi_1(X) &= \frac{1}{3840}[2250 + 4500X - 1360X^2 \\
&\quad -2720X^3 + 160X^4 + 320X^5], \\
\Phi_2(X) &= \frac{1}{3840}[-375 - 250X + 1560X^2 \\
&\quad +1040X^3 - 240X^4 - 160X^5], \\
\Phi_3(X) &= \frac{1}{3840}[45 + 18X - 200X^2 \\
&\quad -80X^3 + 80X^4 + 32X^5].
\end{aligned} \tag{6}
$$

To interpolate a two dimensional function, we need an array of 36 values of $f$. For the 6 different values of $y$ (viz. $y_{i-2}$, $y_{i-1}$, $y_i$, $y_{i+1}$, $y_{i+2}$ and $y_{i+3}$), one evaluates the value of $f$ at the given $x$. Then one uses these values to interpolate along the $y$-axis.

## 2.2 Accuracy and cost

The main goal of this work is to accelerate the iterations of a mapping from a Hamiltonian flow. Therefore one of the crucial points for long term iterations is the area preservation. The error in the position of the image of a point is of order 6. Let $(x_0, y_0)$ be the real image and $(x, y)$ the interpolated one. We can write:

$$
x = x_0 + \epsilon_x(x_0, y_0)h^6, \qquad y = y_0 + \epsilon_y(x_0, y_0)h^6. \tag{7}
$$

A small displacement from the interpolated point is related to that from the real image by:

$$
\begin{aligned}
dx &= dx_0\left[1 + \frac{\partial \epsilon_x}{\partial x}h^6\right] + \frac{\partial \epsilon_x}{\partial y}h^6 dy_0, \\
dy &= dy_0\left[1 + \frac{\partial \epsilon_y}{\partial y}h^6\right] + \frac{\partial \epsilon_y}{\partial x}h^6 dx_0.
\end{aligned} \tag{8}
$$

Therefore the relative error of the small area is of order 6:

$$
dx.dy = dx_0.dy_0\left[1 + \mathrm{O}(h^6)\right]. \tag{9}
$$

Futhermore, suppose we are interested in details of size $\epsilon$ in the phase portrait. The size of the grid needed to represent such a detail is of order $\epsilon^{1/5}$. For example, to represent a detail of

size 0.001, one needs a grid of size 0.25. So one needs to know the real value of the image at only a few points. This is very interesting for hamiltonian flows that are very long to integrate numerically.

But one never gets something from nothing, and one has to pay for this advantage. The cost is the number of operations involved. For a 2D mapping (two dimensional mapping), we first evaluate the six polynomials (**??**), i.e. 30 multiplications and 30 additions for the $x$ component. Then, we evaluate six times (**??**) for the six different values of $y$. And this is repeated for the 2 components of the image. This second step requires 72 multiplications and 60 additions. Then we evaluate the six polynomials (**??**) for the $y$ component and evaluate (**??**) for the 2 components of the image. The total cost is 144 multiplications and 130 additions. This is far from negligible. For a four dimensional mapping, the cost increases to 6336 multiplications and 5300 additions.

## 2.3 Taylor interpolation of order 3

All the above calculations can be done for a lower order approximation, say for a Taylor expansion of order 3. In this case the error is of order $h^4$. To interpolate we need to know the image of four points for one dimension. The polynomials involved are:

$$\begin{aligned}
\Phi_{-1}(X) &= \frac{1}{48}\left[-3+2X+12X^2-8X^3\right], \\
\Phi_{0}(X) &= \frac{1}{48}\left[27-54X-12X^2+24X^3\right], \\
\Phi_{1}(X) &= \frac{1}{48}\left[27+54X-12X^2-24X^3\right], \\
\Phi_{2}(X) &= \frac{1}{48}\left[-3-2X+12X^2+8X^3\right].
\end{aligned} \tag{10}$$

This amounts to 64 multiplications and 54 additions for a 2D mapping, and 1408 multiplications and 1068 additions for a 4D one.

## 2.4 Spline interpolation

The idea underlying the use of a Spline is to represent a curve by a piece-wise polynomial curve which is as smooth as possible. For a Spline of order 3, we use four points and write the interpolated value as a linear combination of the function at those points:

$$f(x) = \Phi_{-1}(\frac{x-x'}{h})f(x_{i-1}) + \Phi_0(\frac{x-x'}{h})f(x_i)$$

$$+\Phi_1(\frac{x-x'}{h})f(x_{i+1}) + \Phi_2(\frac{x-x'}{h})f(x_{i+2}), \tag{11}$$

where functions $\Phi_i$ are polynomials of degree 3. Futhermore these polynomials are chosen so that: the resulting functions and their respective first derivative are continuous. This is achieved if:

$$\begin{aligned} f'(x_i) &= \frac{f(x_{i+1}) - f(x_{i-1})}{2h}, \\ f'(x_{i+1}) &= \frac{f(x_{i+2}) - f(x_i)}{2h}, \end{aligned} \tag{12}$$

which implies the following values for the $\Phi_i$'s:

$$\begin{array}{llll} \Phi_{-1}(0) = 0, & \Phi_{-1}(1) = 0, & \Phi'_{-1}(0) = -0.5, & \Phi'_{-1}(1) = 0, \\ \Phi_0(0) = 1, & \Phi_0(1) = 0, & \Phi'_0(0) = 0, & \Phi'_0(1) = -0.5, \\ \Phi_1(0) = 0, & \Phi_1(1) = 1, & \Phi'_1(0) = 0.5, & \Phi'_1(1) = 0, \\ \Phi_2(0) = 0, & \Phi_2(1) = 0, & \Phi'_2(0) = 0, & \Phi'_2(1) = 0.5. \end{array} \tag{13}$$

Solving these equations, one finally finds:

$$\begin{aligned} \Phi_{-1}(X) &= -0.5X\left(1 - X^2\right), \\ \Phi_0(X) &= (1 - X)\left(1 + X - 1.5X^2\right), \\ \Phi_1(X) &= X\left(0.5 + 2X - 1.5X^2\right), \\ \Phi_2(X) &= -0.5X^2(1 - X). \end{aligned} \tag{14}$$

A straightforward calculation shows that this is a second order method, i.e, the error is of order $h^3$. Despite the continuity of the first derivative, this is worse than the Taylor expansion of order 3, because we are more interested in the accuracy than in the smoothness of the interpolation.

All these schemes use more points than just the vertices of the cell in question. In cases where there is a limit for one variable, these points may fall in a region where the mapping is not defined. Therefore we need asymmetrical formulae to handle the border cells. These formulae are given in Appendix A.

# 3 Results

Obviously the different synthetic mappings which have been described in section 2, have two key parameters: the number of divisions in each direction $N = (\text{total number of cells})^{(1/D)}$,

where $D$ ($D = 2$ in this paper) is the dimension of the surface of section and the order $M$ of the Taylor expansion. In order to explore the validity of the synthetic approach we applied our method to two cases:

1 An algebraic area preserving mapping for which the computation of orbits is very fast. This allows us to calculate a lot of orbits and to perform enough iterations for a meaningful comparison.

2 A special case of the restricted three body problem studied by Duncan et al (1989).

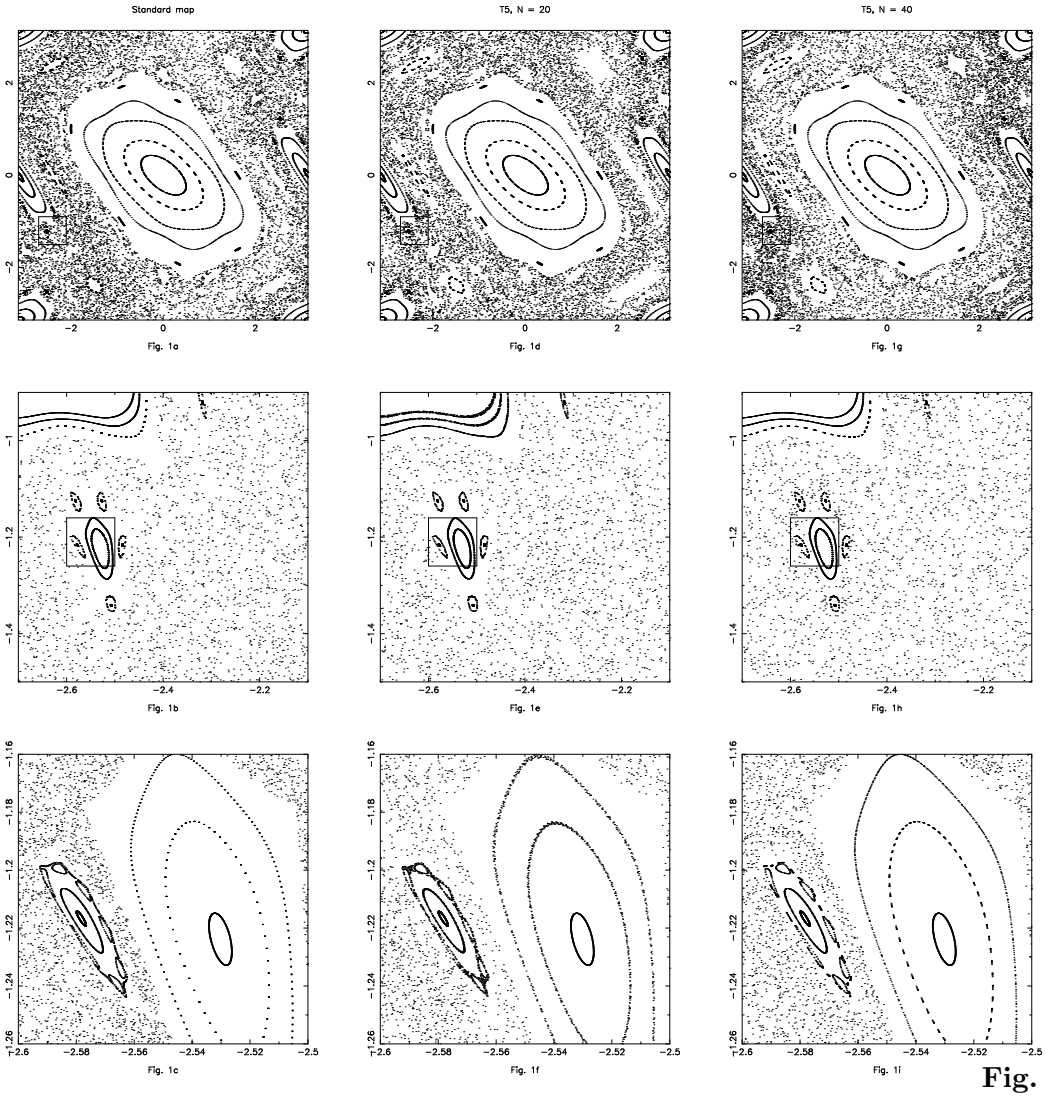



**Fig. 1:** (a)

Standard mapping pictures for $a = -1.3$. (b and c) are enlargements of the small boxes shown in figs. 1a and 1b respectively. (d, e and f) same as figs. 1a,b,c for the synthetic mapping T5 and $N$=20. (g, h and i) same as figs. 1d,e,f for $N$=40.

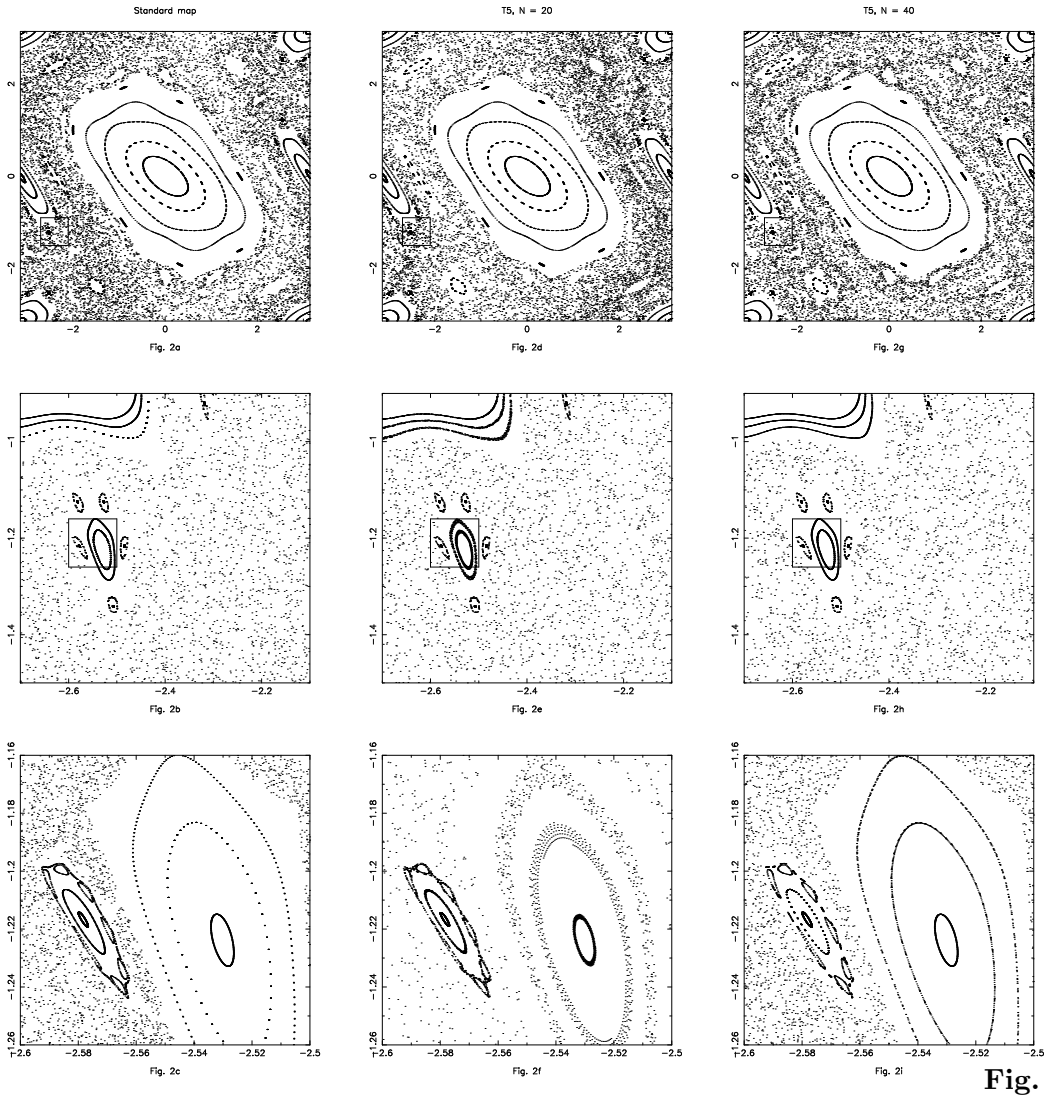



**Fig. 2:** (a-i) same as figs. 1a-i with the asymmetrical formulae for the T5 maps.

First we used the well-known standard mapping (Froeschle, 1970; Chirikov, 1979; Lichten-

berg and Lieberman, 1983):

$$\begin{aligned} x^{(n+1)} &= x^{(n)} + a\sin(x^{(n)} + y^{(n)}) \qquad (\mathrm{mod}\ 2\Pi) \\ y^{(n+1)} &= x^{(n)} + y^{(n)} \qquad (\mathrm{mod}\ 2\Pi). \end{aligned}$$

Fig. 1a shows orbits of the standard mapping for $a = -1.3$. Indeed such a mapping exhibits all the characteristics and well known features of problems with two degrees of freedom such as invariant curves, islands and stochastic zones where the points wander about in a chaotic way. Figs 1b and 1c are magnifications of the small boxes indicated in figs 1a and 1b respectively. At this magnification details like second order islands become evident and the quality of the approximation for the synthetic map is easily visualised. Figs 1g, 1h and 1i correspond respectively to the same orbits and same magnifications as 1a, 1b and 1c for the Taylor interpolation mapping of order $M = 5$ (T5) where the grid is characterized by $N = 40$. We obtain results which look the same as for the original map. For a grid corresponding to $N = 20$ (figs 1d, 1e and 1f), it is interesting to notice the agreement of the details in magnifications e and f. This agreement is remarkable since the dimension of the cell (0.31) is much bigger than the dimension of the smallest islands seen in fig 1f (0.005). Of course the invariant curve of the largest island of 1f shows some thickness which disappears for $N = 40$ but would reappear under further magnification. This phenomenon is due to the fact that synthetic mappings like T5 only conserve areas up to order 6. Let us remark that the computation corresponding to fig. 1 were performed using an extended grid so that we only used the symmetrical interpolation formulae. Since there are cases where one can't cross a given limit, we have tested the asymmetrical interpolation formulae as explained in section 2 (see Appendix A for formulae). Fig. 2 shows the same orbits as fig. 1 for the new formulae. The discrepancy in the sharpness of the invariant curves is particularly obvious in fig. 2f where most of the points are in a region corresponding to the asymmetrical formulae. The problem is less stringent for larger values of $N$ since the portion of phase space needing asymmetrical formulae decreases. Of course for all problems where it is possible, symmetrical formulae should be used. In order to make a more quantitative test of the statistical fine-scaled structure of the orbits resulting from the synthetic mappings we have computed and plotted the Lyapunov Characteristic Number (LCN) (see Benettin et al 1980, Froeschle C, 1984). We consider two orbits starting at $P_0$ and $P'_o$ and denote by $d = \mathrm{dist}(P_0, P'_0)$, the distance between $P_0$ and $P'_0$. After time $\tau$, $P_0$ is in $P_1$ and $P'_0$ in $P'_1$ and we put $d_1 = \mathrm{dist}(P_1, P'_1)$. By a homothesis of the center $P_1$ with a ratio $d/d_1$ we get two new starting points $P_1$ and $P''_1$ such that $\mathrm{dist}(P_1, P''_1) = d$ and we iterate the process. It is well known that the quantity

$$\Gamma(n, p, d, \tau) = \frac{1}{n\tau}\sum_{i=1}^{n}\ln(d_i/d) \tag{15}$$

tends to a limit, independent of $d$ and $\tau$, which is the largest LCN as $n$ goes to infinity. We must keep in mind that strictly speaking the LCN's are limiting values as $t \to \infty$; evidently these cannot in practice be computed. We then define the Lyapunov characteristic indicators (LCIs) as the truncated values of the LCNs for a finite time. The variation of $\lambda = \ln(\Gamma)$ with $n$ for $\tau = 10$ and $d = 10^{-8}$, is shown on fig. 3, for an invariant curve of the Taylor mapping and with the same initial conditions for the different synthetic maps T1, T3, T5 and S3. The grid is such that $N = 100$. For T1 the fit is quite good up to $10^4$ iterations. After that the map shows a strong contracting character and looses its area preserving property. The same occurs but for larger $n$ to the spline approximation S3. The Taylor approximations T3 and T5 are better and show the appearance of divergence from the real mapping after $10^6$ iterations for T3 and $10^7$ for T5. Fig. 4 exhibits the variations of $\lambda$ as a function of $N$ (coarse graining) for a given value of the LCI, $\Gamma(n, p, d, \tau) = \Gamma(10000, p, 10^{-8}, 10)$. Again the Taylor approximations T3 and T5 show the best fits to the real integrable case. The way $\lambda$ varies for a cross section on the $x$-axis at $y_0 = 0$ and $a = -1.3$ is shown as a function of the different mappings (viz. standard map, T1, T3, T5 and S3) in figs. 5a, 5b, 5c for $N = 20$, 50 and 100 respectively. The initial conditions $x_0$, $y_0$ are changed along the $x$-axis ($y_0 = 0$) in steps of 0.1 on the plateaux, and 0.02 at the borders. We again underline the excellence of the fit for the Taylor expansion and also for chaotic orbits the constancy of the values of $\lambda$ independently of both the coarse graining and the order of the approximation.

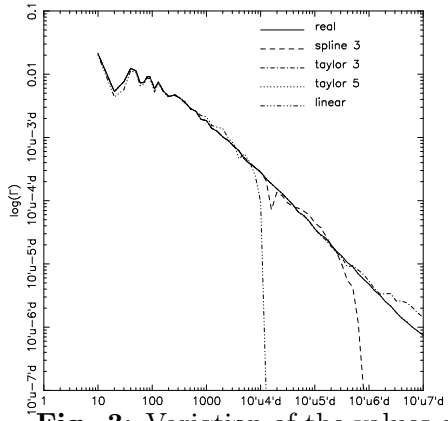



**Fig. 3:** Variation of the values of the Lyapunov Characteristic Indicators (LCI) for initial conditions corresponding to an invariant curve. The standard map and corresponding synthetic maps are considered.

For more clarity, we consider only the real and T5 mappings for three different values of $N = 20$, 50 and 100 and in fig. 6a we plot $\lambda$ versus $x$, as in figs. 5a, 5b, 5c. For $N = 100$, the curves are superimposed and the fit is still very good for $N = 50$. When $N$ is only equal to

20 some slight quantitative differences can be observed but, since the size and location of the stochastic zone remain the same, qualitatively the agreement is excellent. In fig. 6b we present the enlargement of the right hand border of the secondary island which lies to the far left of fig. 6a for $-2.48 \leq x_0 \leq -2.45$ (see region I in fig. 6a). The previous results are confirmed.

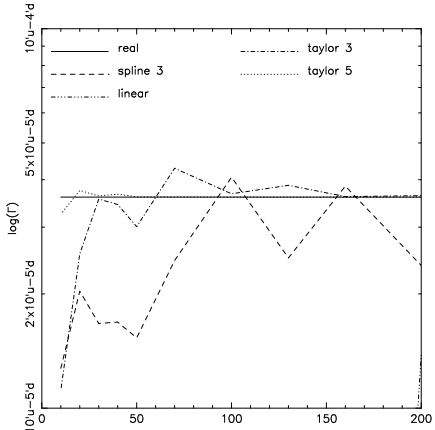



**Fig. 4:** Variation of the values of the Lyapunov Characteristic Indicator (LCI) computed at 100000 iterations against $N$ which characterises the thinness of the grid for different synthetic maps corresponding to the standard map ($a = -1.3$) and for the same initial conditions ($x_0 = 0.5$, $y_0 = 0$) corresponding to an invariant curve of the standard map.

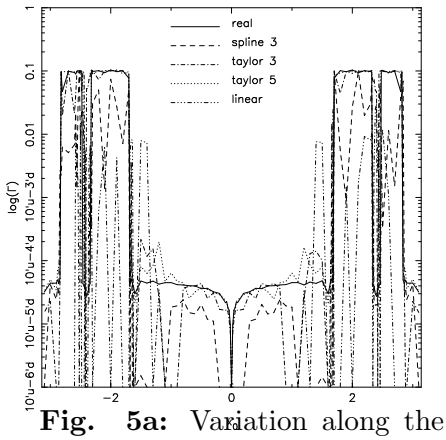



**Fig. 5a:** Variation along the $x$-axis of the values of the Lyapunov Characteristic Indicator (LCI) computed at 100000 iterations at $y_0 = 0$ for a value $N = 20$ of the coarse graining parameter.

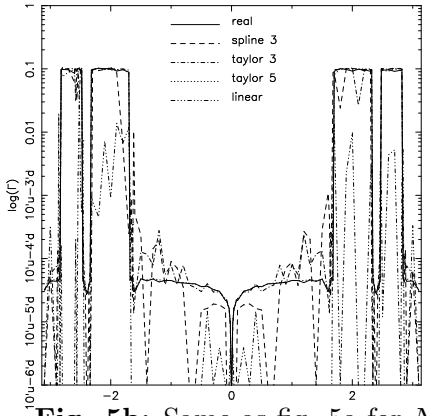



**Fig. 5b:** Same as fig. 5a for $N = 50$

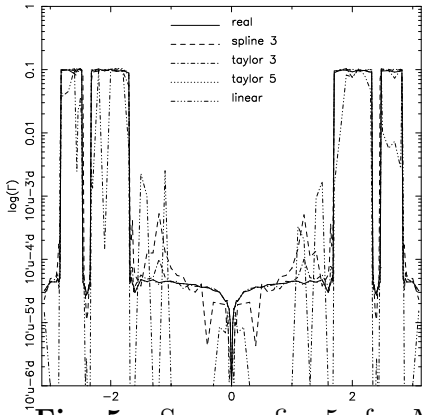



**Fig. 5c:** Same as fig. 5a for $N = 100$

Finally we have tested the method on a special case of the restricted three body problem for which Duncan et al (1989) have developed a special mapping. Fig. 7a exhibits orbits of the Poincare map taking as surface of section the plane defined by the eccentricity and the mean longitude as polar coordinates, when the particle is in conjonction with the planet (i.e. in the rotating frame, when $y = 0$ and $y' > 0$). In figs. 7b, c, d are plotted the corresponding orbits for synthetic mappings T1, T3, T5 for a grid with $N = 100$. If the linear mapping T1 shows only poor qualitative similarities with the Poincare map, the T3 map correctly reproduces the locations, shapes and sizes of zones of regular orbits. Of course the T5 map is even better. It should be noted that in order to obtain the same accuracy we needed to use a smaller grid size (i.e. more points) than for the standard map since the functions which have to be interpolated are less regular. In this case, we needed about 80 minutes to compute the Poincare map by

integrating the equations of motion (fig. 7a) and only 2 seconds for T5 (fig. 7d). However one has to allow for the time needed to compute the grid.

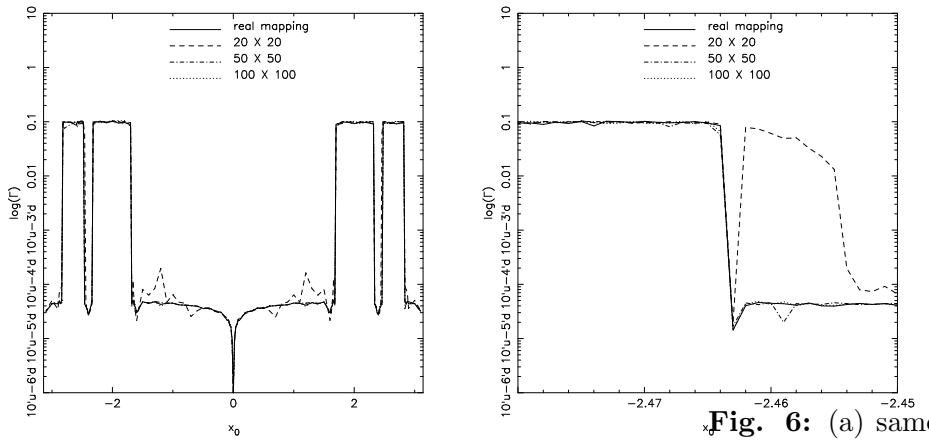



**Fig. 6:** (a) same as fig. 5 for the synthetic mapping T5 and for different values of the parameter grid. (b) enlargement of the border of a secondary island (zone I) for $-2.48 \leq x_0 \leq -2.45$.

Let us remark that the same accuracy as that obtained by the T5 map can be obtained even with the T1 map by drastically decreasing the size of the grid cells (i.e. by increasing $N$). Actually for such an increase in the number of vertices, the price to pay is very small since the additional Poincare return map values at the vertices can be computed using the more precise mapping T5. Figs. 7e,f show the T1 and T3 mappings for a grid with $N = 400$, where 15/16 of the $N^2$ map values at the vertices have been computed using T5 through the already computed grid with $N = 100$. The time needed to compute the 10000 points of the initial grid is about 5 hours 40 minutes and it took only 70 seconds to compute the 150000 news points. In this case T5 is about 300 times faster than the integration. And when looking at the long term evolution, one will still save time if one wants to iterate for 200000 or more steps, since T1 is about 16 times faster than T5. Indeed the accuracy of T1 is greatly improved. Of course the same holds for T3. This combination of the order of interpolation and of different coarse graining, i.e. of different mappings appears very promising for four dimensional mappings where the computational time of T5 is far from negligible as shown in section 2.3

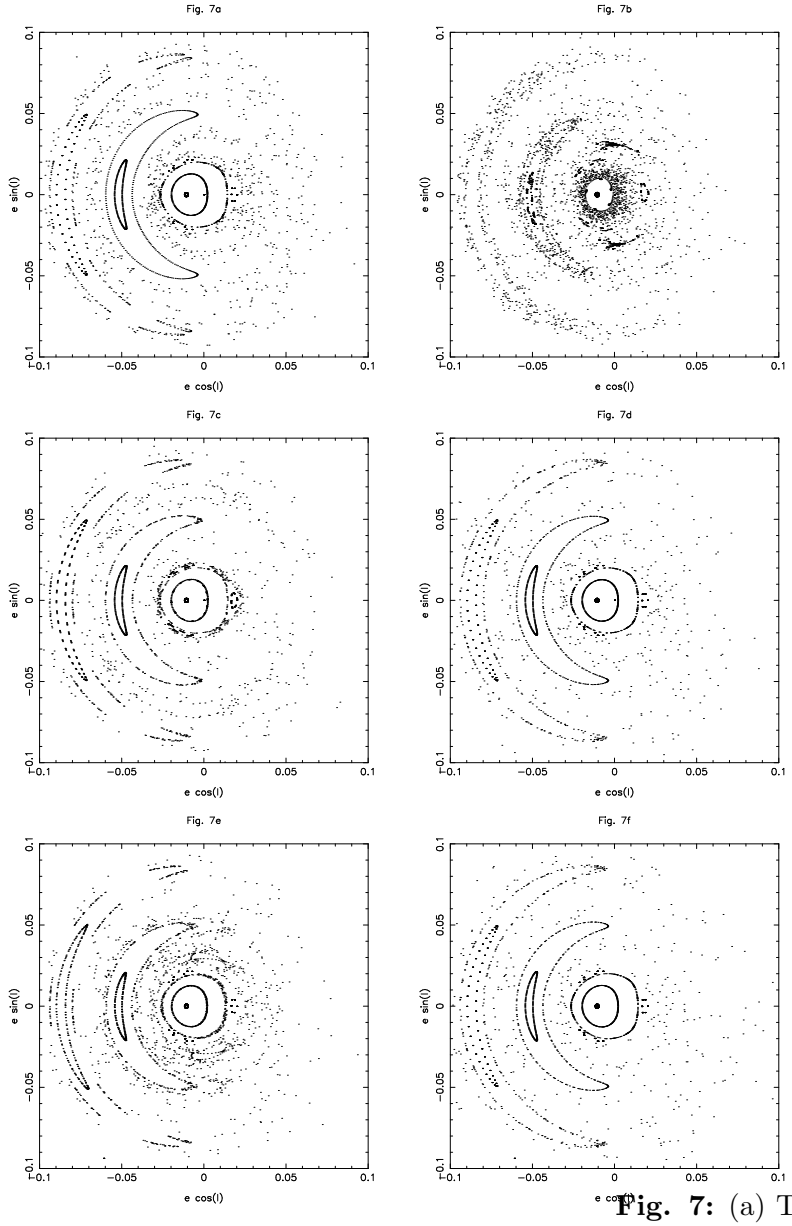



**Fig. 7:** (a) Trajectories of the Poincaré map of the restricted three body problem, in the plane (eccentricity, mean longitude) (polar coordinates) at conjonction with the planet, for the Neptune - Sun mass ratio $m/M_\odot = 5.178\,10^{-5}$ and Jacobi constant 3.0080694. (b, c and d) same as fig. 7a for the synthetic maps T1, T3 and T5. (e and f) Same as fig. 7b,c for a grid with $N = 400$ obtained with T5.

# 4 Conclusion

Synthetic maps appear to be valuable tools for celestial mechanics. However for problems with more than two degrees of freedom the situation is less straightforward. The number of operations required for the Taylor approxmation increases drastically with the dimension of the surface of section. Of course as shown above a lower order map can be used by decreasing the grid size but a second difficulty is the task of storing and recalling the values of the computed images at the vertices. In any case some compromises have to be made between the number of cells and the order of the polynomial approximation. All of these problems will be developed in a forthcoming paper.

# Appendix A

The interpolating formula for the border cells are of the same form as before:

$$f(x) = \sum \Phi_n(\frac{x - x'}{h}) f(x_{i+n}). \tag{16}$$

We first give the formulae for mapping T5. For $x_0 \leq x \leq x_1$, the $\Phi_n(X)$ are:

$$\begin{aligned}
\Phi_0(X) &= \frac{1}{3840}[945 - 3378X + 3800X^2 \\
&\quad -1840X^3 + 400X^4 - 32X^5], \\
\Phi_1(X) &= \frac{1}{3840}[4725 + 2010X - 10760X^2 \\
&\quad +7280X^3 - 1840X^4 + 160X^5], \\
\Phi_2(X) &= \frac{1}{3840}[-3150 + 2860X + 11760X^2 \\
&\quad -11360X^3 + 3360X^4 - 320X^5], \\
\Phi_3(X) &= \frac{1}{3840}[1890 - 2220X - 6800X^2 \\
&\quad +8800X^3 - 3040X^4 + 320X^5], \\
\Phi_4(X) &= \frac{1}{3840}[-675 + 870X + 2360X^2 \\
&\quad -3440X^3 + 1360X^4 - 160X^5], \\
\Phi_5(X) &= \frac{1}{3840}[105 - 142X - 360X^2 \\
&\quad +560X^3 - 240X^4 + 32X^5].
\end{aligned} \tag{17}$$

For $x_1 \leq x \leq x_2$, we have:

$$\begin{aligned}
\Phi_{-1}(X) &= \frac{1}{3840}[-105+142X+360X^2 \\
&\quad -560X^3+240X^4-32X^5], \\
\Phi_0(X) &= \frac{1}{3840}[1575-4230X+1640X^2 \\
&\quad +1520X^3-1030X^4+160X^5], \\
\Phi_1(X) &= \frac{1}{3840}[3150+4140X-5360X^2 \\
&\quad -1120X^3+1760X^4-320X^5], \\
\Phi_2(X) &= \frac{1}{3840}[-1050+20X+4560X^2 \\
&\quad -160X^3-1440X^4+320X^5], \\
\Phi_3(X) &= \frac{1}{3840}[315-90X-1400X^2 \\
&\quad +400X^3+560X^4-160X^5], \\
\Phi_4(X) &= \frac{1}{3840}[-45+18X+200X^2 \\
&\quad -80X^3-80X^4+32X^5].
\end{aligned} \tag{18}$$

By symmetry, we have for $x_{n-2} \leq x \leq x_{n-1}$:

$$\begin{aligned}
\Phi_{-4}(X) &= \frac{1}{3840}[-45-18X+200X^2 \\
&\quad +80X^3-80X^4-32X^5], \\
\Phi_{-3}(X) &= \frac{1}{3840}[315+90X-1400X^2 \\
&\quad -400X^3+560X^4+160X^5], \\
\Phi_{-2}(X) &= \frac{1}{3840}[-1050-20X+4560X^2 \\
&\quad +160X^3-1440X^4-320X^5], \\
\Phi_{-1}(X) &= \frac{1}{3840}[3150-4140X-5360X^2 \\
&\quad +1120X^3+1760X^4+320X^5], \\
\Phi_0(X) &= \frac{1}{3840}[1575+4230X+1640X^2
\end{aligned} \tag{19}$$

$$
\begin{aligned}
& -1520X^3 - 1030X^4 - 160X^5], \\
\Phi_1(X) &= \frac{1}{3840}[-105 - 142X + 360X^2 \\
& +560X^3 + 240X^4 + 32X^5].
\end{aligned}
$$

Finally, for $x_{n-1} \leq x \leq x_n$, the formulae are as follow:

$$
\begin{aligned}
\Phi_{-5}(X) &= \frac{1}{3840}[105 + 142X - 360X^2 \\
& -560X^3 - 240X^4 - 32X^5], \\
\Phi_{-4}(X) &= \frac{1}{3840}[-675 - 870X + 2360X^2 \\
& +3440X^3 + 1360X^4 + 160X^5], \\
\Phi_{-3}(X) &= \frac{1}{3840}[1890 + 2220X - 6800X^2 \\
& -8800X^3 - 3040X^4 - 320X^5], \\
\Phi_{-2}(X) &= \frac{1}{3840}[-3150 - 2860X + 11760X^2 \\
& +11360X^3 + 3360X^4 + 320X^5], \\
\Phi_{-1}(X) &= \frac{1}{3840}[4725 - 2010X - 10760X^2 \\
& -7280X^3 - 1840X^4 - 160X^5], \\
\Phi_0(X) &= \frac{1}{3840}[945 + 3378X + 3800X^2 \\
& +1840X^3 + 400X^4 + 32X^5].
\end{aligned}
\tag{20}
$$

We now give the formulae for mapping T3. Firstly for $x_0 \leq x \leq x_1$:

$$
\begin{aligned}
\Phi_0(X) &= \frac{1}{48}\left[15 - 46X + 36X^2 - 8X^3\right], \\
\Phi_1(X) &= \frac{1}{48}\left[45 + 42X - 84X^2 + 24X^3\right], \\
\Phi_2(X) &= \frac{1}{48}\left[-15 + 6X + 60X^2 - 24X^3\right], \\
\Phi_3(X) &= \frac{1}{48}\left[3 - 2X - 12X^2 + 8X^3\right].
\end{aligned}
\tag{21}
$$

And for $x_{n-1} \leq x \leq x_n$:

$$
\Phi_{-3}(X) = \frac{1}{48}\left[3 + 2X - 12X^2 - 8X^3\right],
$$

$$
\begin{aligned}
\Phi_{-2}(X) &= \frac{1}{48}\left[-15 - 6X + 60X^2 + 24X^3\right], \\
\Phi_{-1}(X) &= \frac{1}{48}\left[45 - 42X - 84X^2 - 24X^3\right], \\
\Phi_{0}(X) &= \frac{1}{48}\left[15 + 46X + 36X^2 + 8X^3\right].
\end{aligned} \tag{22}
$$

# References


[1] Arnold, V.I., 1976. in *Méthodes mathématiques de la mécanique céleste,* Eds. de Moscou,

[2] Baille, Ph., & Froeschle, C., 1989. Astron. Astrophys., in press,

[3] Benettin, G.L., Giorgilli, A., & Strelcyn, J.M., 1980. Meccanica **15**, p. 10.

[4] Chirikov, B.V., 1979. *Phys.Rev.*. **52**, p. 263.

[5] Duncan, M., Quinn, T., & Tremaine S., 1989. Icarus **82**, p. 402.

[6] Froeschle, C., 1970. Astron. Astrophys. **9**, p. 15.

[7] Froeschle, C., 1984. Numéro special du Journal de Mécanique théorique et appliquée p. 101.

[8] Froeschlé, C., & Rickman, H., 1980. *Astron.Astophys.*. **82**, p. 183.

[9] Froeschle, C., & Scholl, H., 1981. Astron. Astrophys. **93**, p. 62.

[10] Hadjidemetriou, J.D., 1986. Journal of applied Mathematics and Physic **37**, p. 776.

[11] Hénon, M., 1981. in *Cours des Houches,* **XXXVI**, North Holland, Amsterdam, p. 57.

[12] Hénon, M., & Petit, J.M., 1986. Celestial Mechanics **38**, p. 67.

[13] Lichtenberg, A.J., & Lieberman, M.A., 1983. in *Regular and Stochastic motion,* Springer Verlag,

[14] Milani, A., Carpino, M.,Hahn, G., & Nobili, A.M., 1989. Icarus **78**, p. 212.

[15] Murray, C.D., 1986. Icarus **65**, p. 70.

[16] Petrosky, T.Y., & Broucke, R., 1988. *CelestialMechanics*. **42**, p. 53.

[17] Rickman, H., & Vaghi, S., 1976. Astron. Astrophys. **51**, p. 327.

[18] Sagdeev, R.Z., & Zaslavsky, G.M., 1987. Il nuovo cimento **97b**, 2.

[19] Schubart, J., 1968. Astron. J. **73**, p. 99.

[20] Sidlichovsky, M., & Melenda, B., 1986. Bull. Astron. Inst. Czechosl. **37**, p. 65.

[21] Varosi, F., Grebogi, C., & Yorke, J.A., 1987. Physics Letters A **124**, p. 59.

[22] Wisdom, J., 1982. Astron. J. **87**, p. 577.

**[12]**

# "Polynomial approximations of Poincaré maps for Hamiltonian systems, II."

PETIT, J-M. & FROESCHLE, C.

*Astron. Astrophys.* **282** (1994)

# Polynomial approximations of Poincaré maps for Hamiltonian systems, II.

Petit J-M.
C.N.R.S., Observatoire de Nice, France
petit@obs-nice.fr

Froeschlé C.
C.N.R.S., Observatoire de Nice, France
claude@obs-nice.fr

**Abstract**

In paper I polynomial interpolating formulae of order 3 and 5 have been proposed and tested for transforming a non–linear differential Hamiltonian system into a map without having to integrate whole orbits as in the well known Poincaré return map technique. The precision of the computations increases drastically with the order of the polynomial fit which requires an extended amount of local information, i.e. information about neighbouring points. The first part of the paper deals with another type of interpolation where the information, within the same accuracy, refers only to the nearest neighbours but takes into account gradient information. The results are in very good agreement with those obtained using an order 3 symmetrical interpolation formula well inside the phase space. Moreover the new method is more effective at the border of the phase space when compared with asymmetrical interpolation. The second part of the paper deals with higher dimensional mappings, i.e. mappings for hamiltonian systems with 3 degrees of freedom.

## 1 Introduction

Poincaré maps are widely used for studying the qualitative behaviour of differential equations (see Henon 1981). In order to study stability problems, many authors have sought explicit algebraic mappings which approximate, at least qualitatively, the Poincaré maps obtained from the original Newton equations. Froeschlé and Petit (1990, paper I) have reviewed some of these mappings. They rely on some approximations made in order to separate either — in the case of

deterministic mappings — an integrable part and some instantaneous perturbations, or — for stochastic mappings — a source of endogeneous/exogeneous stochasticity (see Froeschlé and Rickman, 1988). These mappings are all *ad hoc* and are reliable only in a limited region of phase space where the approximations are valid. In paper I we built a mapping valid everywhere in phase space, following an idea already used by Varosi et al. (1987) but in the framework of non–Hamiltonian systems (i.e., systems where attractors do exist). The method consists of coarse–graining the phase–space surface of section and then interpolating the value of the image of a point. Linear interpolation requires a rather fine graining of the phase space, hence it is necessary to compute a lot of points on the grid. However, Taylor expansions of order 3 and 5 can provide very good results as long as symmetrical interpolation formulae are applied, for which it is necessary to use an extended grid, i.e. a grid extended over the region of phase space one wants to explore. Since there are cases where one cannot cross a given limit, asymmetrical interpolation formulae have been tested, but their accuracy was found to be inferior. Therefore we have developed another type of interpolation, where the information, including that relative to the gradients, is stored only for the nearest–neighbour vertices. Thus, not only images of vertices are computed, but also tangential mappings at each vertex of the grid. The second part of the paper deals with higher dimensional mappings i.e. Poincaré maps for Hamiltonian systems with three degrees of freedom.

# 2 The synthetic mappings

The basic idea of the synthetic mapping is to interpolate the image of a point, when those of a set of points located at the vertices of a grid are given.

In paper I we derived the interpolating formulae for accuracy of order 3 and 5. These formulae require the use of respectively 4 and 6 points along each dimension of the mapping. In other words, we need a certain amount of non–local information. In the following, we use more local information, i.e. we look only at the nearest neighbours but we take more information at these points in order to have the same accuracy.

## 2.1 Taylor interpolation

Assume that we know the values of the functions $x^{(n+1)} = f(x^{(n)}, y^{(n)})$ and $y^{(n+1)} = g(x^{(n)}, y^{(n)})$ at discrete points $x_i$ and $y_j$ and also their derivatives with respect to $x$, $y$ and the cross derivatives with respect to $x$ and $y$. The points are equally spaced in $x$ and $y$. We want to interpolate the values of $f$ and $g$ at any point. We first derive the interpolation formula in one dimension. We write the Taylor expansion for $f$ about a point in the middle of a given interval,

$x' = (x_i + x_{i+1})/2$:

$$\begin{aligned} f(x) &= f(x') + (x-x')f'(x') + \frac{(x-x')^2}{2}f''(x') \\ &\quad + \frac{(x-x')^3}{6}f'''(x') + \mathrm{O}\left((x-x')^4\right), \\ f'(x) &= f'(x') + (x-x')f''(x') \\ &\quad + \frac{(x-x')^2}{2}f'''(x') + \mathrm{O}\left((x-x')^3\right). \end{aligned} \tag{1}$$

We use the middle point to ensure the symmetry of the formula. We need to know the value of $f$ and its derivatives at $x'$. For this we write (**??**) for $x_i$:

$$\begin{aligned} f(x_i) &= f(x') - \frac{h}{2}f'(x') + \frac{h^2}{8}f''(x') \\ &\quad - \frac{h^3}{48}f'''(x') + \mathrm{O}(h^4), \\ f'(x_i) &= f'(x') - \frac{h}{2}f''(x') + \frac{h^2}{8}f'''(x') + \mathrm{O}(h^3), \end{aligned} \tag{2}$$

where $h = x_{i+1} - x_i$. The corresponding formulae for $f(x_{i+1})$ and $f'(x_{i+1})$ are obtained from (**??**) respectively by replacing the minus signs with plus signs. Neglecting terms of order 4 or higher, one can invert (**??**) and write $f(x')$, $f'(x')$, $f''(x')$ and $f'''(x')$ as functions of $f(x_i)$, $f(x_{i+1})$, $f'(x_i)$ and $f'(x_{i+1})$.

$$\begin{aligned} f(x') &= \frac{1}{8}[4f(x_i) + 4f(x_{i+1}) + hf'(x_i) - hf'(x_{i+1})], \\ hf'(x') &= \frac{1}{4}[-6f(x_i) + 6f(x_{i+1}) - hf'(x_i) - hf'(x_{i+1})], \\ h^2f''(x') &= -hf'(x_i) + hf'(x_{i+1}) \\ h^3f'''(x') &= 6[2f(x_i) - 2f(x_{i+1}) + hf'(x_i) + hf'(x_{i+1})]. \end{aligned} \tag{3}$$

By substituting into (**??**) we obtain:

$$\begin{aligned} f(x) &= \frac{f(x_i)}{2}[1 - 3h + 4h^3] \\ &\quad + \frac{f(x_{i+1})}{2}[1 + 3h - 4h^3] \end{aligned} \tag{4}$$

$$+\frac{f'(x_i)}{8}[1-2h-4h^2+8h^3]$$
$$+\frac{f'(x_{i+1})}{8}[-1-2h+4h^2+8h^3].$$

To interpolate a two dimensional function, we iterate the one dimensional interpolation. We want to know the image of point $(x, y)$. Assume that we know the images and gradients at points $(x, y_i)$ and $(x, y_{i+1})$. Then we evaluate the image using eq. (**??**). This means that we must first interpolate $f$ and $\frac{\partial f}{\partial y}$ at these two points. $f$ at $(x, y_i)$ (resp. $(x, y_{i+1})$), is obtained from $f$ and $\frac{\partial f}{\partial x}$ at $(x_j, y_i)$ and $(x_{j+1}, y_i)$ (resp. $(x_j, y_{i+1})$ and $(x_{j+1}, y_{i+1})$). For the value of $\frac{\partial f}{\partial y}$ at $(x, y_i)$, we interpolate using $\frac{\partial f}{\partial y}$ and $\frac{\partial^2 f}{\partial x \partial y}$ at the vertices.

## 2.2 A special 2D case

When using the previous algorithm, we reproduce the Taylor expansion up to order 3 in every direction. So in $n$ dimensions, the $n$-dimensional Taylor expansion is reproduced up to global order 3 and we also have terms of the expansion up to global order $3n$. In the case of a two-dimensional mapping, we need to know 16 values to interpolate one point. But if we want to have the interpolation only up to global order 3, we need only to determine 10 constants. We need to use values only at the four nearest neighbours and give them the same importance. So we have 12 conditions: $f$, $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ at the four points. Here again we expand around the middle point $(x' = (x_i + x_{i+1})/2, y' = (y_j + y_{j+1})/2)$ with all the terms up to global order 3 and two terms of global order 4, and neglecting the other terms of order 4 and higher:

$$\begin{aligned}
f(x,y) &= f(x',y') + (x-x')f_x(x',y') + (y-y')f_y(x',y') \\
&\quad +\frac{(x-x')^2}{2}f_{xx}(x',y') + \frac{(y-y')^2}{2}f_{yy}(x',y') \\
&\quad +(x-x')(y-y')f_{xy}(x',y') + \frac{(x-x')^3}{6}f_{xxx}(x',y') \\
&\quad +\frac{(x-x')^2(y-y')}{2}f_{xxy}(x',y') + \frac{(x-x')(y-y')^2}{2}f_{xyy}(x',y') \\
&\quad +\frac{(y-y')^3}{6}f_{yyy}(x',y') + \frac{(x-x')^3(y-y')}{6}f_{xxxy}(x',y') \\
&\quad +\frac{(x-x')(y-y')^3}{6}f_{xyyy}(x',y'), \qquad (5) \\
f_x(x,y) &= f_x(x',y') + (x-x')f_{xx}(x',y') + (y-y')f_{xy}(x',y')
\end{aligned}$$

$$
\begin{aligned}
& + \frac{(x-x')^2}{2} f_{xxx}(x',y') + \frac{(y-y')^2}{2} f_{xyy}(x',y') \\
& + (x-x')(y-y') f_{xxy}(x',y') + \frac{(x-x')^2(y-y')}{2} f_{xxxy}(x',y') \\
& + \frac{(y-y')^3}{6} f_{xyyy}(x',y'), \\
f_y(x,y) \;=\; & f_y(x',y') + (x-x') f_{xy}(x',y') + (y-y') f_{yy}(x',y') \\
& + \frac{(x-x')^2}{2} f_{xxy}(x',y') + \frac{(y-y')^2}{2} f_{yyy}(x',y') \\
& + (x-x')(y-y') f_{xyy}(x',y') + \frac{(x-x')^3}{6} f_{xxxy}(x',y') \\
& + \frac{(x-x')(y-y')^2}{2} f_{xyyy}(x',y').
\end{aligned}
$$

Writing that these equations must be satisfied with the 12 values we know at the four corners of the cell, we find:

$$
\begin{aligned}
f(x,y) \;=\; & \frac{f(x_i,y_i)}{4}[1+3h+3k+8hk-4h^3(1+2k)-4k^3(1+2h)] \\
& + \frac{f(x_{i-1},y_i)}{4}[1-3h+3k-8hk+4h^3(1+2k)-4k^3(1-2h)] \\
& + \frac{f(x_i,y_{i-1})}{4}[1+3h-3k-8hk-4h^3(1-2k)+4k^3(1+2h)] \\
& + \frac{f(x_{i-1},y_{i-1})}{4}[1-3h-3k+8hk+4h^3(1-2k)+4k^3(1-2h)] \\
& + \frac{f_x(x_i,y_i)}{16}[(1+2k)(-1-2h+4h^2+8h^3)] \\
& + \frac{f_x(x_{i-1},y_i)}{16}[(1+2k)(1-2h-4h^2+8h^3)] \\
& + \frac{f_x(x_i,y_{i-1})}{16}[(1-2k)(-1-2h+4h^2+8h^3)] \\
& + \frac{f_x(x_{i-1},y_{i-1})}{16}[(1-2k)(1-2h-4h^2+8h^3)] \\
& + \frac{f_y(x_i,y_i)}{16}[(1+2h)(-1-2k+4k^2+8k^3)] \\
& + \frac{f_y(x_{i-1},y_i)}{16}[(1-2h)(-1-2k+4k^2+8k^3)]
\end{aligned}
\tag{6}
$$

$$+\frac{f_y(x_i, y_{i-1})}{16}[(1+2h)(1-2k-4k^2+8k^3)]$$
$$+\frac{f_y(x_{i-1}, y_{i-1})}{16}[(1-2h)(1-2k-4k^2+8k^3)]$$

The great advantage of this formula over the one derived from (**??**) is that it does not require knowing the second order cross derivative of $f$. Suppose the mapping is given by numerical integration of differential equations. It is conceptually easy to obtain the first order derivatives of the mapping by a similar integration (see below). But the determination of the second order cross derivative requires a numerical differentiation which is not always easy to do. The problem is that it is not possible to generalize to higher dimensions. In dimension 3, we have 32 conditions (8 nearest neighbours, each with 4 values), and the terms of global order less or equal to 3 give only 20 unknown coefficients. Therefore, we have to choose arbitrarily 12 terms of global order larger than 3. We made a few choices which all lead to singular matrices. However there must exist a choice that gives a non singular matrix, but it might prove very difficult to find. And when increasing the number of dimensions, we increase even more the number of terms we have to choose arbitrarily.

# 3 Derivative evaluation

In the previous algorithms, we need to know the derivatives of the mapping. We have to compute the first order derivatives with the same order of precision as the orbit. Consider the differential equation:

$$\frac{dX}{dt}(t) = F(X, t). \tag{7}$$

Since we want to know a derivative, we integrate this equation for two nearby points: $X$ and $X + \epsilon$. The evolution of the second point is governed by equation:

$$\frac{dX}{dt}(t) + \frac{d\epsilon}{dt}(t) = F(X + \epsilon, t) = F(X, t) + L_{(X,t)}.\epsilon + \mathrm{O}(\epsilon^2), \tag{8}$$

where $L$ is the Jacobian matrix. Since $\epsilon$ is an infinitely small vector, its evolution is governed by the well known variation equations:

$$\frac{d\epsilon}{dt}(t) = L_{(X,t)}.\epsilon \tag{9}$$

This is a linear equation, so we can normalize $\epsilon$ as we wish.

Let us introduce the following notation. $\Gamma$ denotes the intersection of the surface of section $S$ and the constant energy surface $E$. Then we call $T\Gamma_P$ the space tangent to $\Gamma$ at point $P$. We consider a point $X_0 \in \Gamma$ and a basis $\{e_i^{(0)}\}$ of $T\Gamma_{X_0}$ (in our case, the basis consists of two linearly independant vectors). We need to know the image $X_1 \in \Gamma$ of $X_0$ and the image $\{e_i^{(1)}\} \in T\Gamma_{X_1}$ of $\{e_i^{(0)}\}$. For a Hamiltonian system with two degrees of freedom, we have four dimensional phase space $(x_1, x_2, x_3, x_4)$. We define the surface of section by $x_2 = 0$ and $x_4 > 0$ (in our particular case, $x_2 = y$ and $x_4 = dy/dt$) and the surface of constant energy by $E = E_0$. The coordinates $(\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4)$ of the vector $\epsilon$ in $T\Gamma_{X_0}$ are constrained to

$$\begin{aligned} \epsilon_2 = &\ 0, \\ \nabla E.\epsilon = &\ 0. \end{aligned} \tag{10}$$

A problem arises when integrating simultaneously equations (**??**) and (**??**), because when the orbit crosses $\Gamma$, that is in $X_1$, the image of $\epsilon$ does not necessarily belong to $T\Gamma_{X_1}$. We come back to the classical variation method and consider the neighbouring point $X_1 + \epsilon$ where $\epsilon$ is again infinitely small. Using an Euler integration scheme with time step $h$ we integrate until time $t$ when the point reaches $\Gamma$:

$$X(t+h) + \epsilon(t+h) = X(t) + \epsilon(t) + h\frac{d(X+\epsilon)}{dt} \ \in \ \Gamma. \tag{11}$$

We have to solve equation (**??**) for $h$ in order to have $x_2(t+h) + \epsilon_2(t+h) = 0$, where $x_2$ and $\epsilon_2$ are the second components of $X$ and $\epsilon$. We simplify the equation by neglecting the term $h\, d\epsilon/dt$ which is a small quantity of order 2. Since $X_1 = X(t)$ is in $\Gamma$, we have $x_2(t) = 0$. So the time step $h$ is given by:

$$h = -\frac{\epsilon_2(t)}{\frac{dx_2(t)}{dt}}, \tag{12}$$

and the final vector $\epsilon(t+h)$ is:

$$\epsilon(t+h) = \left(\epsilon_i(t) - \frac{\epsilon_2(t)\frac{dx_i(t)}{dt}}{\frac{dx_2(t)}{dt}}\right), \quad i = 1 \ldots 4. \tag{13}$$

Since we are interested in the derivatives, the final vectors will be divided by the modulus of the initial vectors and therefore, we can use unit length vectors to perform the computations, since each operation is linear with respect to $\epsilon$.

We are now able to write explicitly a synthetic mapping from $\Gamma$ into $\Gamma$ using the gradient method. We consider the Hill's problem in a rotating frame (cf Duncan et al., 1989, Froeschlé

and Petit, 1990). The coordinates are $x_1 = x$, $x_2 = y$, $x_3 = dx/dt = \dot{x}$ and $x_4 = dy/dt = \dot{y}$. The surface of section is defined by $y = 0$, $\dot{y} > 0$ and $E = const.$ (Jacobi constant). We do not plot $x, \dot{x}$ directly since for a given value of $x$, there is a maximum value of $|\dot{x}|$ allowed and we do not get a rectangular grid. So we plot $z = x, \lambda = \dot{x}/|\dot{x}|_{max}$, where $\lambda$ belongs to $[-1, 1]$. The limiting value of $|\dot{x}|$ is:

$$|\dot{x}|_{max} = \sqrt{x^2 + \frac{2(1-\mu)}{\mu - x} + \frac{2\mu}{\mu - x - 1} - E}. \tag{14}$$

In the general interpolation scheme (derived from equation **??**), we need the second order cross derivative. This has to be computed by mean of numerical differenciation. We know the images of a point $X_0$ and two vectors $e_1$ and $e_2$. For the cross derivative, we integrate equation (**??**) for vector $e_1$ but with the Jacobian matrix evaluated along the orbit of point $X_0 + \epsilon\, e_2$. This second orbit (very close to the first one) is not actually computed but we approximate it by taking $X(t) + \epsilon\, e_2(t)$.

# 4 Memory management

So far we have always supposed that we know the value of the mapping and eventually its derivatives on a discrete grid. But this can be very memory consuming. In the case of a four dimensional mapping, with 50 points in each dimension, we need to store information for 6 250 000 points. Each point accounts for at least 4 values (when we do not use the derivatives, 16 values for each component if we use derivatives, which means 64 values for each point) needing 8 bytes each (we are using double precision arythmetic). This amount to $2.\,10^8$ bytes of storage.

But a typical orbit will visit only a small part of the total space, especially if it is an invariant orbit. Therefore in order to reduce storage requirements, we only store images of those grid points which are needed to compute the orbit. The storage scheme is based on the hashing technique: to compute the image of a point, we need to know the images of the neighbouring points of the grid. These images are computed and then labelled for the purpose of identification. The labels are just integer indices (at a point of the grid we can make a one to one correspondance with integers). The images are stored in a list which is ordered according to the label index. To compute the next image, we again need to compute the images of the new neighbouring points. Some of them are already computed and we have to locate them in the list, others have to be computed and inserted in the list. These two problems of location and insertion can be time consuming hence we use the double trick of a linked-list (for fast insertion) and a *hash function* (for fast location) (see Appendix A). In fact this hash function

allows one to split a large list into a collection of shorter lists. Actually, applying the hash function to a label gives the entry point of the list where that particular label is stored. In practice, the hash function is simply the remainder modulo $M$:

$$h(k) = 1 + k \bmod M, \tag{15}$$

where $k$ is the labelling index and $M$ an arbitrary integer.

Since the amount of computer memory is limited, we may need, for those orbits that visit a large region, to delete some previous points before storing new ones. These cases are, of course, more time consuming. For technical details, see Appendix A.

# 5 Results

## 5.1 The two dimensional case

An extensive study of the point set Taylor interpolation scheme for the two dimensional case can be found in paper I. We present first a sample of results using point set and tangential Taylor interpolation of order 3. As a model problem we adopt the standard mapping:

$$\begin{aligned} x^{(n+1)} &= x^{(n)} + a\sin(x^{(n)} + y^{(n)}), \qquad (\mathrm{mod}\ 2\pi) \\ y^{(n+1)} &= x^{(n)} + y^{(n)}. \qquad (\mathrm{mod}\ 2\pi) \end{aligned}$$

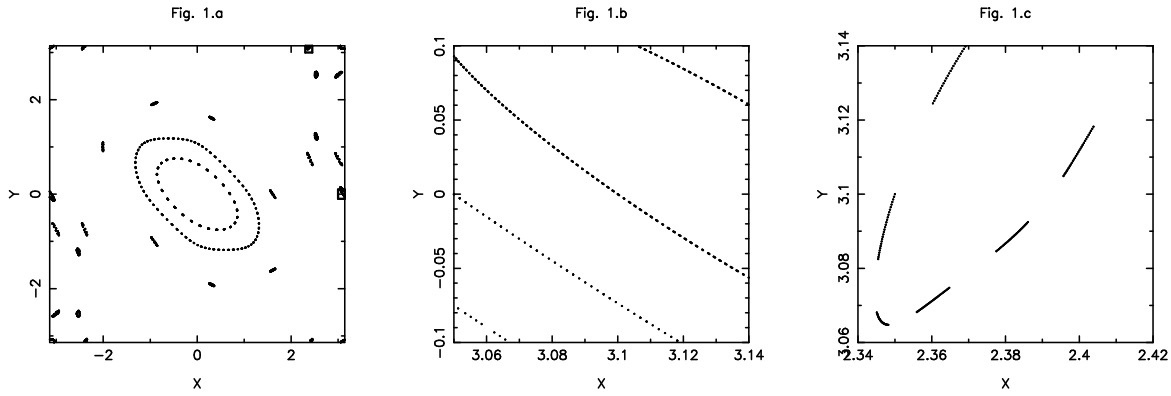
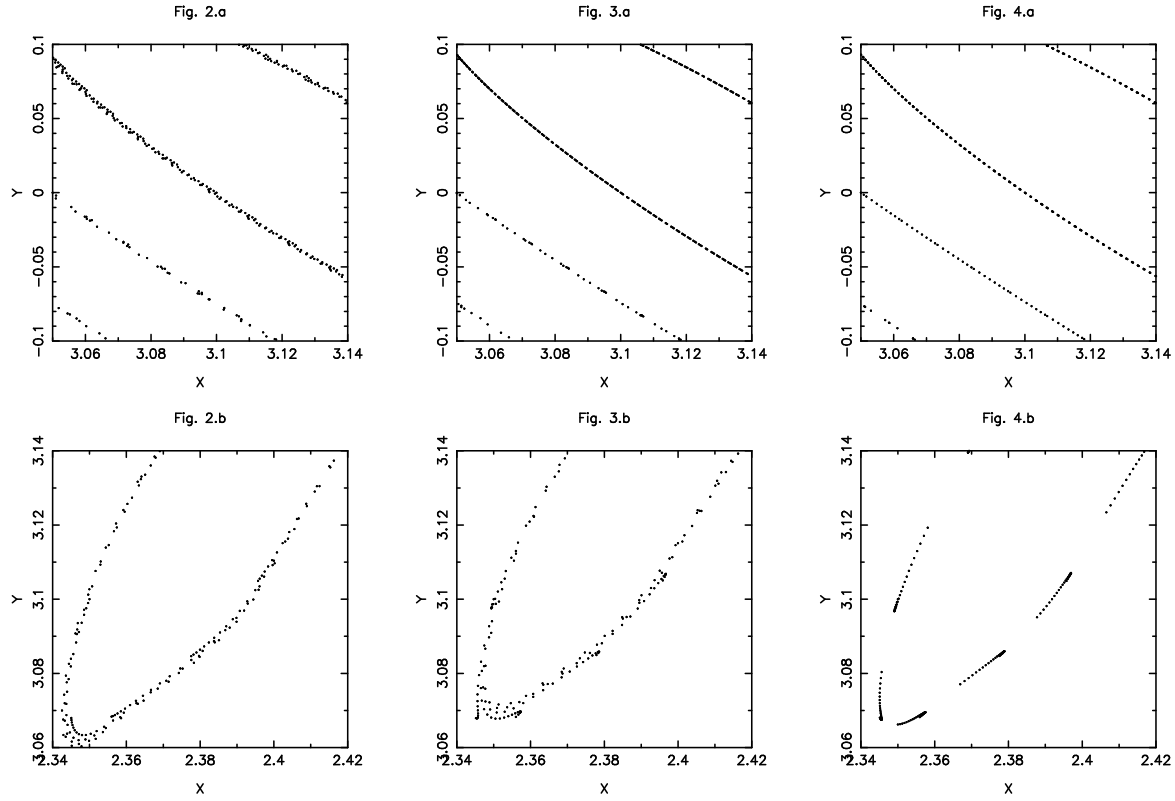



**Fig. 1:** (a) Plots of the standard mapping for $a = -1.3$. (b) and (c) are enlargements of the small boxes shown in (a), respectively at the right border and at the top border.

This has been selected since the computation of orbits is very fast and allows us to calculate many orbits and perform enough iterations for a meaningful comparison. Moreover, for the

tangential Taylor interpolation, the computation of the derivatives of the mapping does not present any problem. Fig. 1 shows orbits of the standard mapping for $a = -1.3$. These mappings exhibit all the well–known typical features of problems with two degrees of freedom, such as invariant curves, "islands", and stochastic zones where the points wander in a chaotic way. Figs. 1b and 1c are magnifications of the small boxes indicated in Fig. 1a. At this magnification level, details like second–order islands become evident and the approximation levels of the synthetic maps are easily visualized. Figs. 2a and 2b correspond to the same orbits and the same magnifications as Figs. 1b–c, but using the Taylor interpolation mapping of order $M = 3$ with decentered formulae on the edges of the mapping; here the grid was regular




**Figs. 2a–b:** Same as in Figs. 1b–c but using a Taylor approximation of order 3 with a regular grid and decentered formulae. **Figs. 3a–b:** The same as in Figs. 1b–c but using a Taylor approximation of order 3 with a non–regular grid and decentered formulae. **Figs. 4a–b:** The same as in Figs. 1b–c but using a Taylor approximation of order 3 with a regular grid and the gradient method.

and characterized by 40 cells. The results are very similar to those of the original map, except for the box close to the border. Figs. 3a–b correspond to the same formulae, but with cells having half the size of the previous ones close to the edges. Apparently this does not greatly improve the quality of the mapping in the border boxes. On the other hand, a definite improvement is obtained using the gradient formulae, as shown in figs. 4a–b. One can see that fig. 4b is much closer to fig. 1c than fig. 2b and even more than fig. 3b. And fig. 4a is closer to fig. 1b than fig. 2a.

We also tested the method on a restricted three body problem for which Duncan et al. (1989) have developed a special mapping. Fig. 5a exhibits orbits of the Poincaré map taking as surface of section the plane defined by the eccentricity and the mean longitude as polar coordinates, when the particle is in conjunction with the planet (i.e, in the rotating frame, when $y = 0$ and $y' > 0$). In fig. 5b are plotted the corresponding orbits for synthetic mapping T3 ($N = 100$). Fig. 5c shows the corresponding orbits for the tangential Taylor interpolation of order 3. In this case the tangential derivatives have been estimated using the method described in $ 3. The agreement is as good as with the corresponding point set interpolation. We must emphasize that all three figures correspond to the same initial conditions. One of these is at the very border of a stable region, which explains why there is one more invariant curve in fig. 5c. On the other hand, in the chaotic regions, the points do not all fit on the grid, and we have to stop the integration very early in the synthetic cases, although not for the real mapping. This explains the different number of points in the chaotic regions.

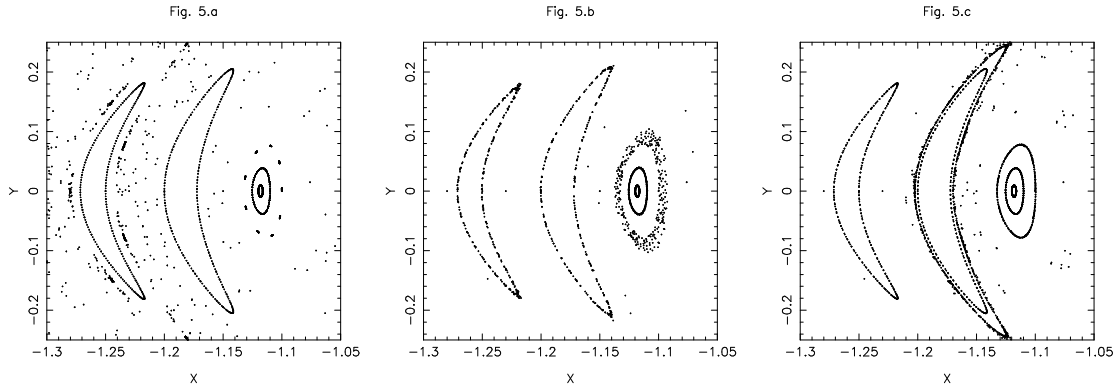



**Fig. 5a:** Trajectories of the Poincaré map of the restricted three–body problem, in the plane giving (as polar coordinates) eccentricity and mean longitude at conjunction with the planet, for the Neptune–to–Sun mass ratio $m/M_0 = 5.178\,10^{-5}$ and a Jacobi constant of 3.0080694. **Fig. 5b:** Same as Fig. 5a but for the point set Taylor 3 interpolation with $N = 100$. **Fig. 5c:** Same as Fig. 5b but for the tangential Taylor 3 interpolation.

## 5.2 Poincaré maps for Hamiltonian systems with three degrees of freedom

### 5.2.1 Graphical display of symplectic four dimensional mappings

When the number of degrees of freedom is equal to 3, the surface of section has 4 dimensions and interpretations of the results are more difficult. Again, using a symplectic four-dimensional mapping appears suitable as a model problem in order to test both the different synthetic mappings and the hash function method. Furthermore a four-dimensional surface of section is populated by points and not by a continuous trajectory, even if they lie on an invariant surface. Therefore, since we compute a finite number of points, if we perform other sections, the probability of a point falling on an arbitrarily chosen surface vanishes. Therefore, we must take not a true section but rather a *slice* having a small yet finite thickness in order to catch some points. Figs. 6a and 7a are examples taken from Froeschlé (1972) for the four-dimensional mapping:

$$\begin{aligned}
x^{(n+1)} &= x^{(n)} + a_1 \sin(x^{(n)} + y^{(n)}) + b \sin(x^{(n)} + y^{(n)} + z^{(n)} + w^{(n)}), \qquad (\mathrm{mod}\ 2\pi) \\
y^{(n+1)} &= x^{(n)} + y^{(n)}, \qquad (\mathrm{mod}\ 2\pi) \\
z^{(n+1)} &= z^{(n)} + a_2 \sin(z^{(n)} + w^{(n)}) + b \sin(x^{(n)} + y^{(n)} + z^{(n)} + w^{(n)}), \qquad (\mathrm{mod}\ 2\pi) \\
w^{(n+1)} &= z^{(n)} + w^{(n)}. \qquad (\mathrm{mod}\ 2\pi)
\end{aligned}$$

A trajectory is a set of points in the $(x, y, z, w)$ space. It is first projected on the three-dimensional $(x, y, z)$ space; then a series of nine slices is taken, defined by $|z - z_0| < 0.01$ with nine regularly spaced values for $z_0$. Parameter values are $a_1 = -1.3$, $a_2 = -1$ for both figures; $b = 0.075$ for fig. 6 and $b = 0.275$ for fig 7. In the first case, the points fall on well defined curves in each slice; this suggests that they fall on a two-dimensional surface in $(x, y, z)$ space. In the second case, the situation is rather different: the points appear to fill a three-dimensional region in $(x, y, z)$ space. Both figures have been obtained by the tangential Taylor 3 ($\tau 3$) synthetic mapping. Figures for the point-set Taylor 3 (T3) mapping are identical. For figures with real mapping, see Figs. 11.d and 11.f from Froeschlé (1972). But again the figures are identical at least in the integrable case ($b = 0.075$).

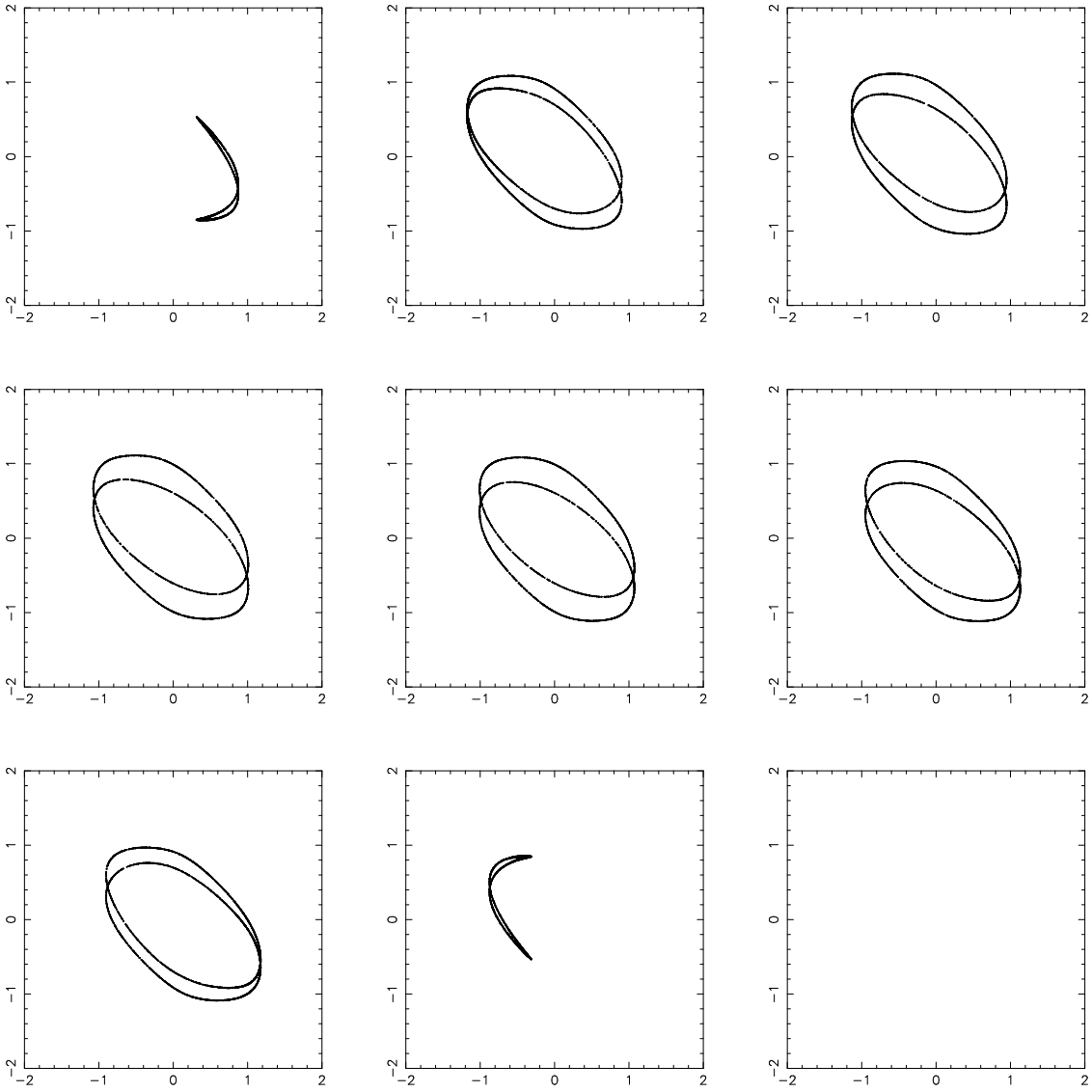


**Fig. 6:** Slices of the subspace $(x, y, z)$ for the tangential Taylor 3 interpolation of the four-dimensional standard map, for $a_1 = -1.3$, $a_2 = -1$. and $b = 0.075$, the starting point being $P = (1, 0, 0.5, 0.5)$. Then slices are defined by $|z - z_0| < 0.01$ with $z_0 = -1.5 + i/3$, where $i$ varies from 1 to 9 from left to right and top to bottom.

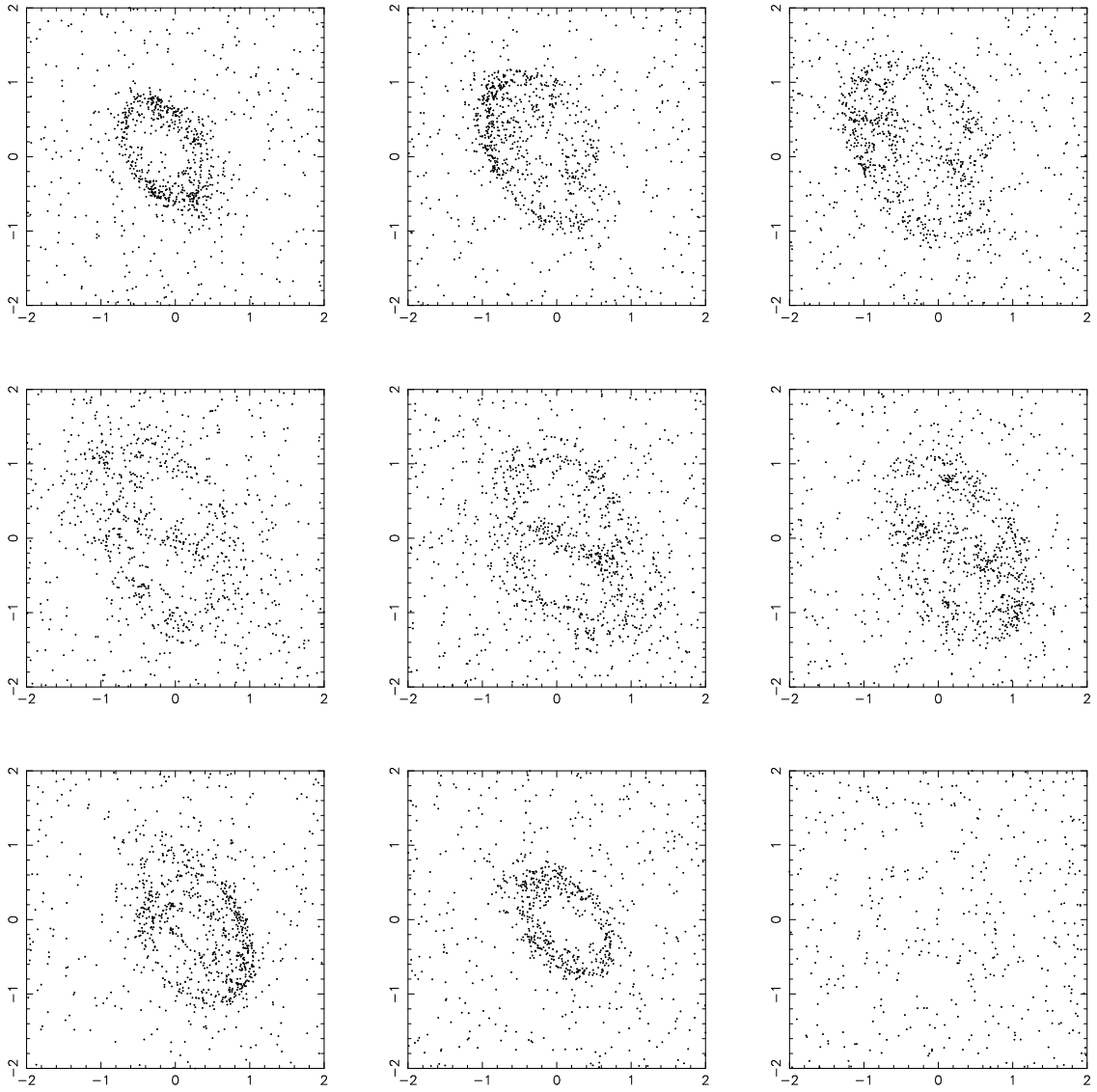


**Fig. 7:** Same as fig. 6, but with $b = 0.275$.

### 5.2.2 The Lyapounov characteristic numbers (LCNs)

A rather different technique consists in determining numerically the Lyapounov characteristic exponents of the trajectories. This allows one to distinguish between quasi-periodic and chaotic

orbits and to map the regular and chaotic regions in phase space. The question raised by Froeschlé (1970) about the number of quasi-integrals for a one parameter set of orbits of the non planar restricted three-body problem has been solved more easily, using LCEs, by Gonczi and Froeschlé (1981). The combination of the two methods (slice cutting and LCEs) has been very useful for discovering and illustrating a sticking-exploding phenomenon. Of course, this technique does not provide any information about the shape of an orbit in phase space, nor about the set of points in the space of section since it gives just two numbers for each trajectory.

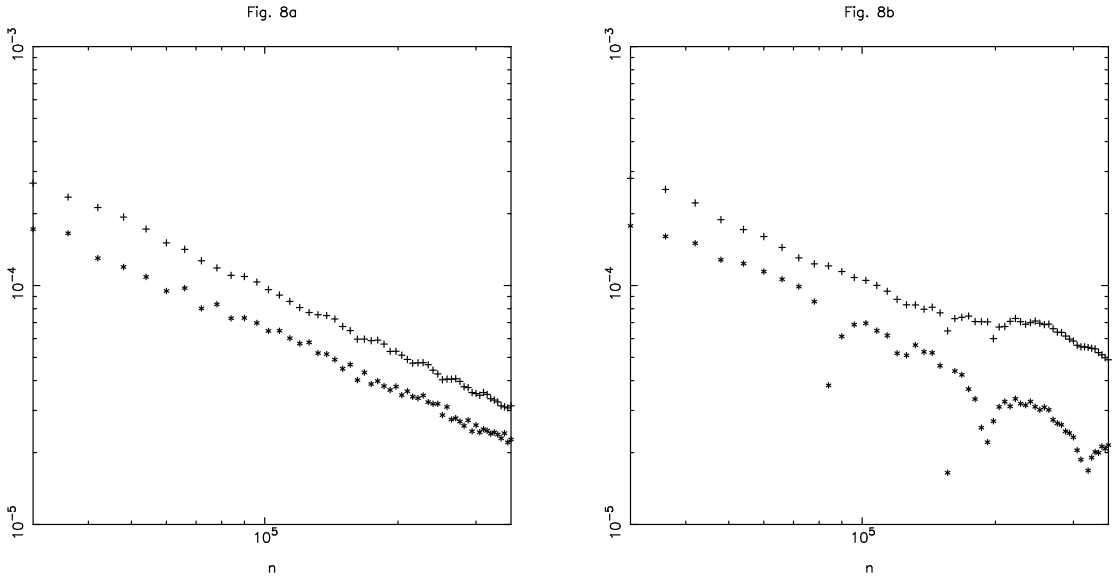



**Fig. 8a:** Positive Lyapounov characteristic indicators for the same orbit as in fig. 6 with $b = 0.075$, for the real mapping. **Fig. 8b:** Same as fig. 8a for the synthetic mapping $\tau 3$.

Fig. 8a shows, on a log-log scale, the behaviour of the two positive Lyapounov characteristic indicators $\gamma_i^n$ (whose limits are identified with the Lyapounov characteristic numbers) as a function of time (i.e. the number $n$ of iterations), for initial conditions taken in the ordered region ($P = (1, 0, 0.5, 0.5)$), with $b = 0.075$. Both $\gamma_i^n$ appear to be decreasing functions of time. The largest shows very little departure from the expected linear behaviour.

Fig. 8b shows the same numbers but computed from the synthetic mappings $\tau 3$. The general behaviour is similar to the previous case, albeit the slope is slightly smaller and the secondary Lyapounov number presents some pathologic patterns. This discrepency is due to the fact that some grid points we use belong to the chaotic zone and therefore influence the computed numbers. This has been tested first by using a finer grid ($N = 100$ instead of $N = 50$). The discrepency between the real and the synthetic mapping almost disappears. Then we have

computed the Lyapounov characteristic indicators with the former grid ($N = 50$) for an orbit closer to the center, that is well within the stable region. This allowed to use only points in the stable region. In this case again, the descrepency disappears.

The agreement between the real mapping and the T3 synthetic mapping is even more dependent of the size of the grid. Since we use grid points further away from the actual point we are looking at, we have more chance to mix stable and choatic regions. This can easily be seen on figs. 9a and 9b which represent, respectively, the dominant and secondary Lyapounov characteristic numbers for different grid sizes.

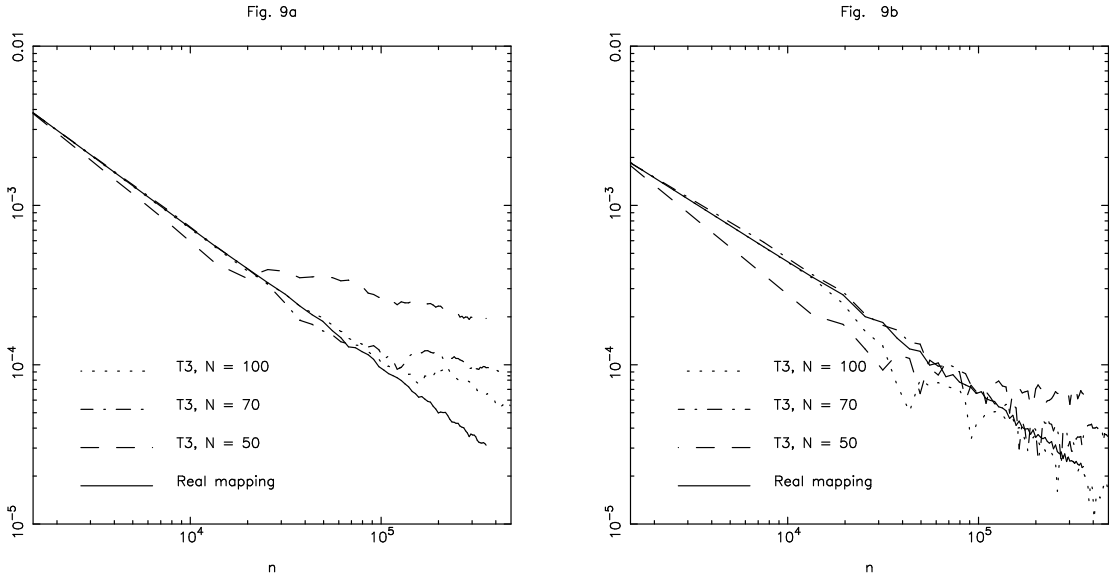



**Fig. 9a:** Dominant Lyapounov characteristic numbers for the same orbit as in fig. 6 with $b = 0.075$, for the real mapping, the synthetic mapping T3 with $N = 50$, $N = 70$ and $N = 100$. **Fig. 9b:** Same as fig. 9a for the secondary Lyapounov characteristic number.

In order to visualize the evolution in time and show in more concrete terms the sticking–exploding phenomenon we have plotted on figs. 10a-c the median slices ($|z| < 0.01$) of the projections along the $w$ axes of the points $P_n$ for $(i-1) \times 40000 \leq n < i \times 40000$, with $i$ varying from 1 to 9 from left to right and top to bottom, and $b = 0.275$, for the real, T3 and $\tau 3$ mappings respectively. Figs. 11a-c show the corresponding variations of the two positive $\gamma_i^n$. Note that one should not pay too much attention to the fact that the times of explosion coincide, since this time is distributed according to a Poisson distribution (Froeschlé and Petit, 1992). We found slightly worse agreement for the explosion time when using the real mapping on two different computers that handle the last bit in floating point operations differently, and

that put the operations in different order due to differences in the compilers.

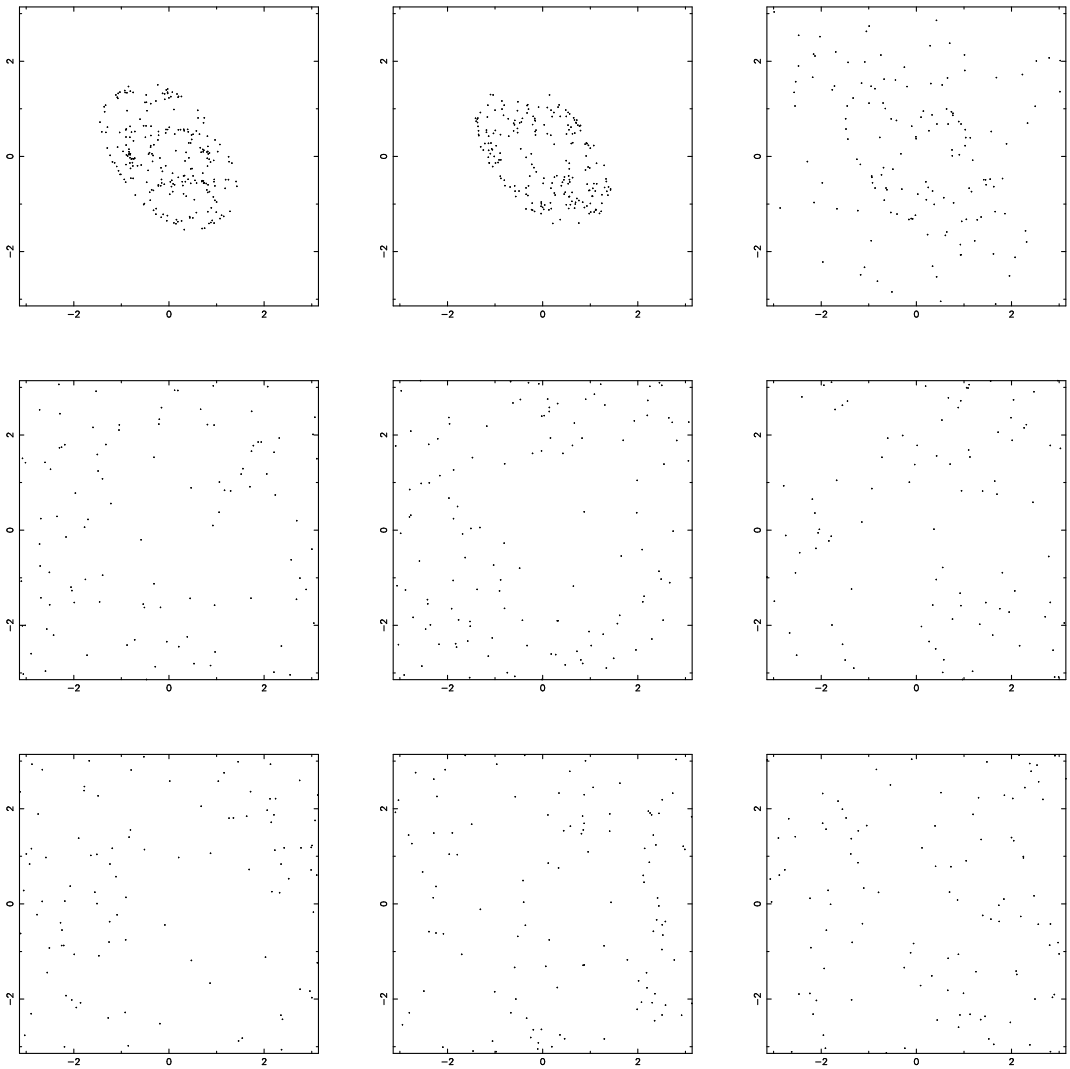


**Fig. 10a:** Median slice ($|z| < 0.01$) for the orbit of fig. 6 with $b = 0.275$ for the real mapping. In each square are plotted the points of 40000 successive iterations.

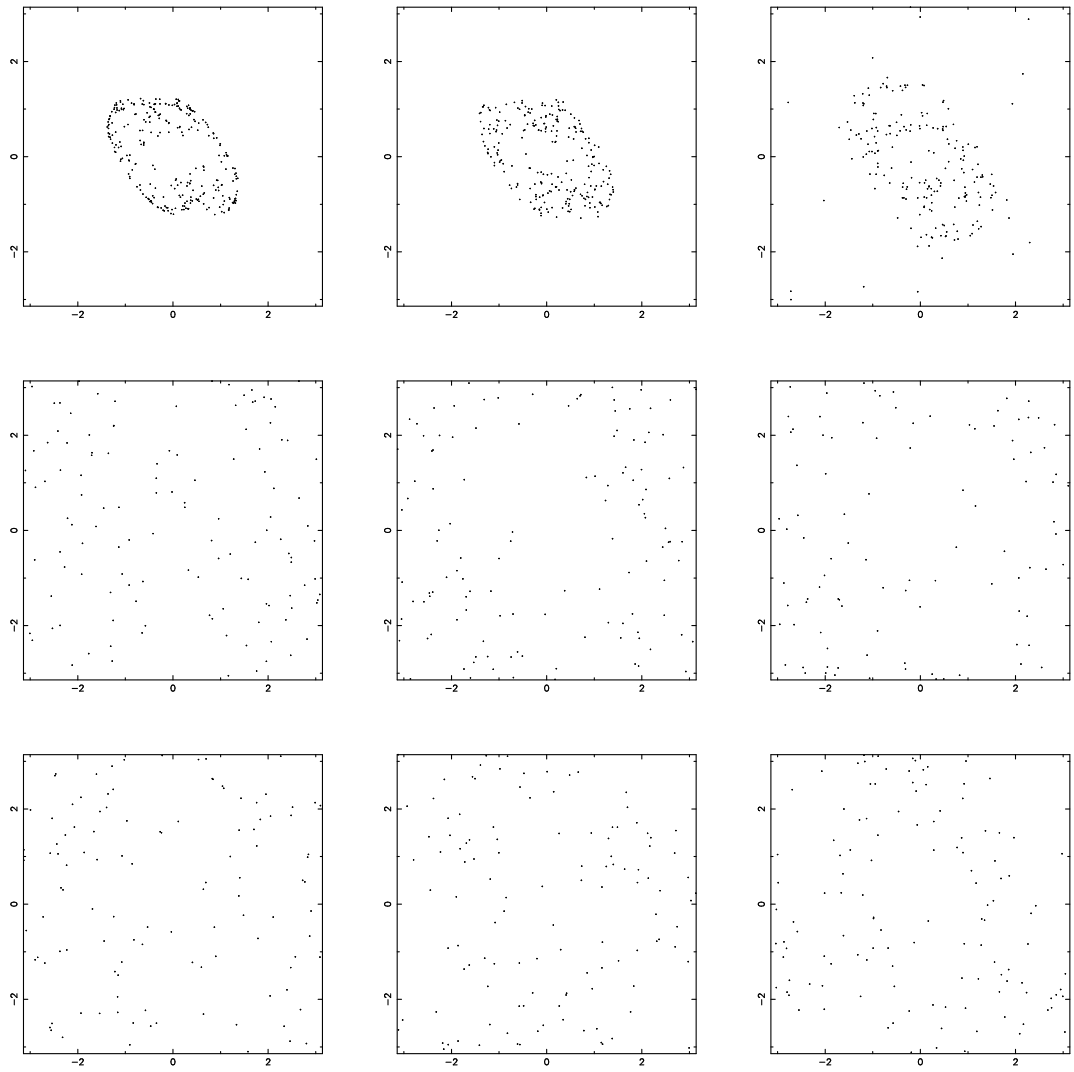


**Fig. 10b:** Same as fig. 10a for the synthetic mapping T3.

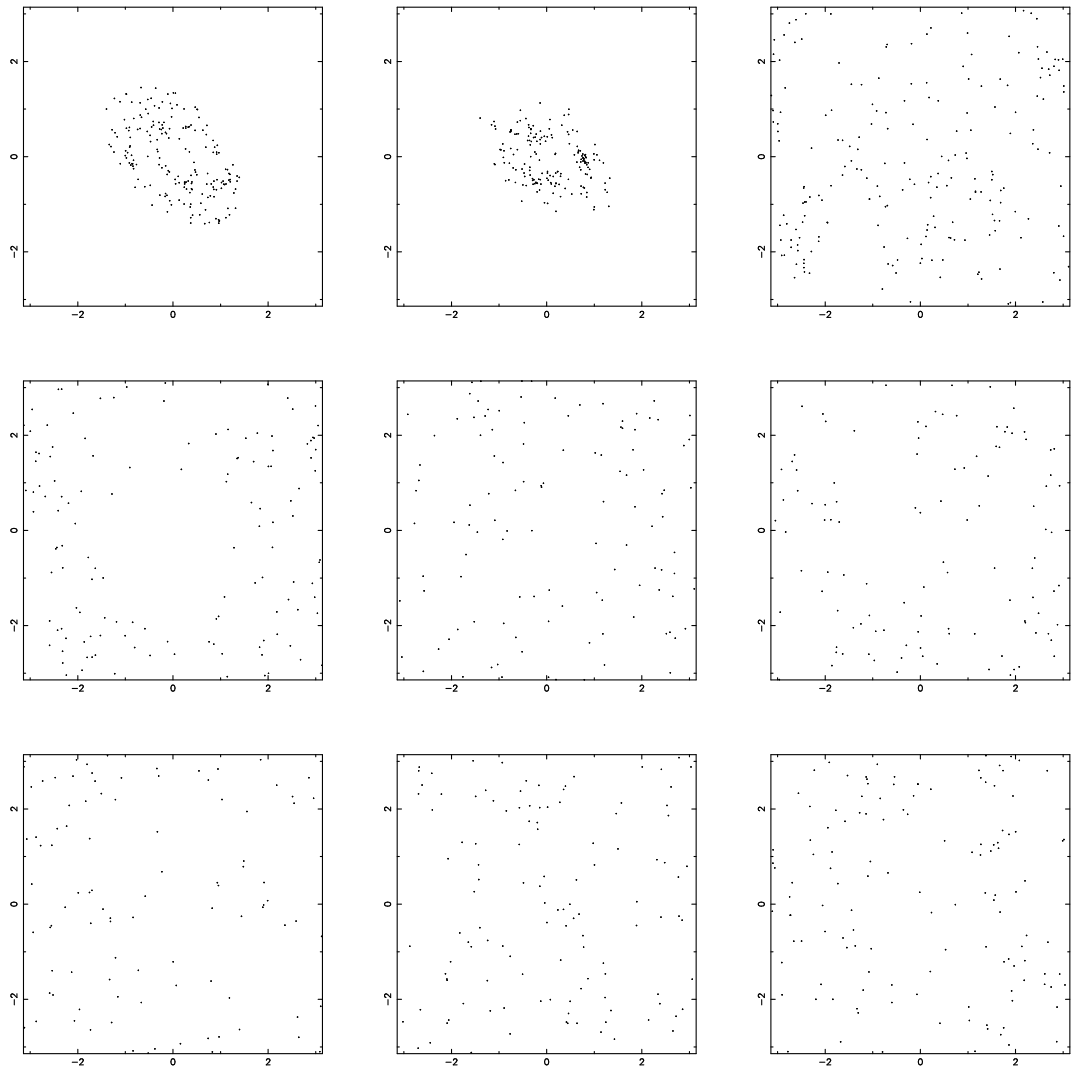


**Fig. 10c:** Same as fig. 10a for the synthetic mapping $\tau 3$.

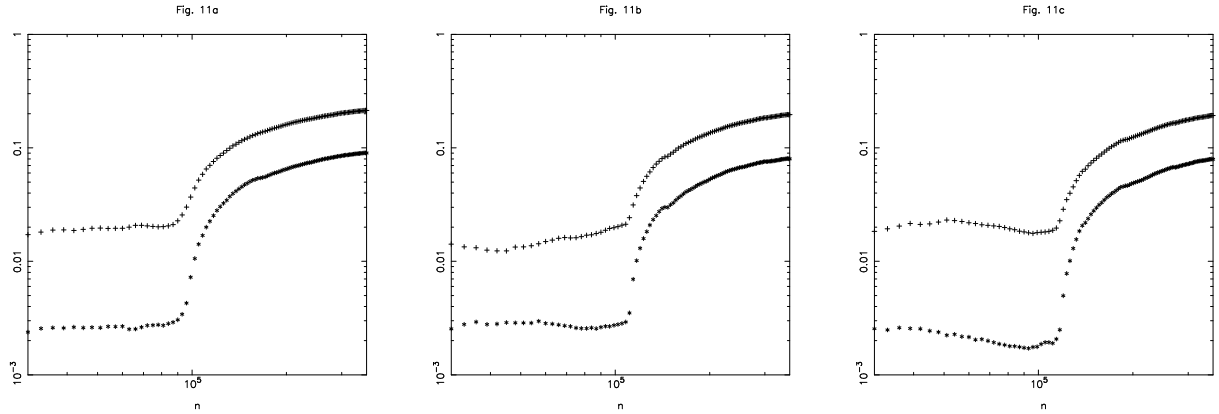



**Fig. 11a:** Positive Lyapounov characteristic numbers for the same orbit as in fig. 10a, for the real mapping. **Fig. 11b:** Same as fig. 11a for the synthetic mapping T3. **Fig. 11c:** Same as fig. 11a for the synthetic mapping $\tau 3$.

# 6 Conclusions

Synthetic maps appear to be valuable tools for celestial mechanics. Depending on the dynamical constraints of the problem like border zones, different interpolations can be used. Even for problems with more than two degrees of freedom, where the situation is less straightforward the use of a hash function has allowed us to overcome difficulties associated with the task of storing and recalling the values of the computed images at the vertices.

# Acknowledgments

The authors want to thank Dr. R. Grauer for stimulating discussions at the outset of this research. They also thank Dr. G. Contopoulos for useful comments during the writing of the final version of this paper.

# Appendix A

We first describe the algorithm for finding a value in an ordered linked-list. We have records composed of labelled data. We store and retrieve data according to the value of the label. The basic idea is to directly reference locations in a table by doing arithmetic transformations of labels into table addresses.

The first step in a search using hashing is to compute a *hash function* which transforms the search label into a table address. Almost no hash function is perfect, and two or more labels might hash to the same table address: the second part of a hashing search is a *collision resolution* process which deals with such labels. The collision resolution method we describe here uses linked lists, and is appropriate in a highly dynamic situation where the number of search labels cannot be predicted.

Hashing is a good example of a "time–space tradeoff". If there were no memory limitations, then we could do any search with only one memory access by simply using the label as a memory address. If there were no time limitations, then we could get by with only a minimum amount of memory by using a sequential search method. Hashing provides a way of using a reasonable amount of memory and time to strike a balance between these two extremes. Efficient use of available memory and fast access to the memory are prime concerns of any hashing method.

*Hash functions*

The first problem we must address is the computation of the hash function which transforms labels into table addresses. This is an arithmetic computation with properties similar to those of random number generators. What is needed is a function which transforms labels (here integers derived from the indices of points in the grid) into integers in the range $[1, M]$ where $M$ is the number of entries in the table. An ideal hash function is one which is easy to compute and approximates a "random" function: for each input, every output should be "equally likely". The method we use here is, for any label $k$, to compute $h(k) = 1 + k \bmod M$. This is a straightforward method which is easy to compute in many environments and spreads the label values out well.

*Separate chaining*

The hash function above will convert labels into table addresses: we still need to decide how to handle the case when two labels hash to the same address. The most straightforward method is to simply build a linked list, for each table address, of the records whose labels hash to that address. Since the labels which hash to the same table position are kept in a linked list, they might as well be kept in order. This leads directly to a generalization of the elementary list searching method. Rather than maintaining a single list with a single list header node, we maintain $M$ lists with $M$ list header nodes. Obviously, the amount of time required for a search depends on the length of the lists (and the relative positions of the labels in them).

Details can be provided by the authors on request.

# References

[1] Duncan, M., Quinn, T., & Tremaine S., 1989. Icarus **82**, p. 402.

[2] Froeschlé, C., 1972. Astron. Astrophys. **16**, p. 172.

[3] Froeschlé, C., & Petit, J.M., 1990. Astron. Astrophys. **238**, pp. 413–423.

[4] Froeschlé, C., & Petit, J.M., 1992. Celestial Mechanics **57**, pp. 123–130.

[5] Froeschlé, C., & Rickman, H., 1988. *CelestialMechanics*. **43**, p. 265.

[6] Gonczi, R., & Froeschlé, C., 1981. Celestial Mechanics **25**, p. 271.

[7] Hénon, M., 1981. in *Cours des Houches,* **XXXVI**, North Holland, Amsterdam, p. 57.

[8] Knuth, D., 1975. in *The art of computer programming: sorting and searching,* **3**, Addison Wesley, p. 506.

[9] Varosi, F., Grebogi, C., & Yorke, J.A., 1987. Physics Letters A **124**, p. 59.

# Chapitre 4

# SUR LES ANNEAUX PLANETAIRES

## 4.1 Analyse de données de Voyager

[8]

"The gravitationnal influence of satellite Pan on the radial distribution of ring-particles in the region of the Encke-division in Saturn's A ring"

SPAHN, F., PETIT, J-M. & BENDJOYA, Ph.

*Celestial Mechanics* **57** (1993)

# The gravitational influence of satellite Pan on the radial distribution of ring–particles in the region of the Encke–division in Saturn's A ring

F. Spahn
Working group "Nonlinear Dynamics" (MPG)
University Potsdam; D-O-1571 Potsdam; Germany

J.-M. Petit
C.N.R.S.
Observatoire de Nice, B.P. 229, 06304, Nice cedex 04, France

Ph. Bendjoya
Observatoire de Nice, B.P. 229, 06304, Nice cedex 04, France

**Abstract**

Here we summarize the models and their similarities with respect to the gravitational influence of small satellites (moonlets) embedded in planetary rings, which have been developed by Petit and Hénon (1987a,b; 1988) and by Spahn (1987; Spahn et al. 1989; 1992). The results and their applicabilities are shown in the case of the Encke–gap Moon Pan located in the outer A–ring of Saturn's rings. To compare the model–profiles with the experimental ones (optical depth profile measured by Voyager–photopolarimeter), the structures embedded in the latter have been confirmed statistically using the "wavelet" analysis method. This analysis points to a size of Pan of about 13 km and an eccentricity of $10^{-4}$ of its orbit.

# 1 Motivation

Much work has been done with respect to the gravitational action of satellites - embedded and outer ones - on planetary rings, which have culminated in the prediction and detection of the satellite Pan in the Encke division of Saturn's rings (Showalter et al. 1986; Showalter 1991).

In this paper we will give a brief summary of the work done concerning the question: " What is the gravitational influence of a small satellite (moonlet) on the radial density distribution of ring–material in the close radial surrounding of its orbit ?"

In Section 2 we summarize the models developed by Petit and Hénon (1987a; b; 1988; hereafter PH 1; 2; 3), and by Spahn (1987; Spahn and Wiebicke 1989; Spahn et al. 1992; hereafter S 1; 2; 3). We enlighten the similarities of the models and the good agreement of a few of their results, which then will be taken together to look for moonlet-like structures in the Voyager occultation data. This is especially done in Section 3 where we compare the results of the models with the radial density- profiles of Saturn's rings measured by the Photopolarimeter (PPS) of Voyager 2 at the Encke division of Saturn's ring where satellite Pan has been discovered recently. This gives a possibility to prove directly the value of the results of the models. Therefore the information about the real structure in the rings has to be separated from the large noise of the PPS-data. This has been realized with the wavelet analysis method which yields a noise-reduced optical depth profile.

In Section 4 the results of the comparison between model and experiment will be discussed and also a small overview of still remaining work can be found there.

# 2 The models

The main common assumptions of either approaches (PH 1; 2; 3; S 1; 2; 3) are:

1. plane motion of all bodies (planet, moonlet, ring-particles);
2. a narrow radial region is considered (compared to the whole radial extension of the rings);
3. the standard-model of spherical, undestructible ring-particles has been applied.

With these assumptions, both models has been applied in order to solve the problem of the gravitational influence of the embedded moonlet on the time-evolution of the radial distribution of ring-particles in the close vicinity (region of overlapping resonances) of the orbit of the perturbing moon. That distribution is given by the function $f(m, r, t)$, $f.dm.dr$ being the

number of particles in the radial interval $[r, r+dr]$ and in the mass-range $[m, m+dm]$ at a given time $t$. Then the evolution of $f(m,r,t)$ can be written by mean of a continuity equation

$$\partial_t f(m,r,t) = \hat{\mathbf{A}}(f,r,t) \cdot f \tag{1}$$

in a very general form.

For more details concerning the meaning of Eq. (1) we refer to the derivation of equation (44) in PH2 and of equation (3) in S2.

The operator $\hat{\mathbf{A}}$ governs the radial transport of the ring–matter and is generally dependent on the gravity of the perturbing moonlet, as well as the interactions between the ring–particles - gravity and direct physical collisions - represented by the dependency of $\hat{\mathbf{A}}$ on $f$ itself (PH2; 3). The integro-differential equation (1) is of a fairly complicated form which makes it impossible to get a general analytical solution. Therefore numerical approaches have been used (PH; S) in order to solve this equation for special cases, e.g. single and bimodal size distributions.

The advantage of the model developed by Petit and Hénon (1987a; b; 1988) is the wide physical generality taken into account in equation (1), which is presented by including:

**a)** arbitrary size distribution of the ring-particles;

**b)** particle interactions $\Rightarrow$ gravity and direct inelastic collisions (fixed coefficient of restitution $\epsilon = constant < 1$; see PH1).

The price for this wide generality is a relatively low number of particles - $N = 500 \cdots 1000$ - taken for the numerical treatment of Eq. (1), which gives a fairly rough resolution of the final radial density profiles.

Such resulting profiles are shown in the fig. 1 for the special case of a bimodal size (mass) distribution for the ring- particles (one moonlet and many much smaller particles). The time evolution of the radial particle distribution is given for a comparatively massive moonlet. In this case the moonlet is able to create an empty gap which is divided by a faint ringlet, covering the orbit of the moonlet (fig. 1f). The ringlet consists of particles in "tadpole" and "horseshoe" orbits (Dermott and Murray 1981). Of high importance is the fact that Petit and Hénon (1988) have been able to obtain an equilibrium width of the gap created by the moonlet, that is a result of the balance between the torque exerted from the moonlet to the ring–material and the torque caused by the particle interactions itself (viscous torque; see also Lissauer et al. 1981). They obtained for that width $W$:

$$W \approx \begin{cases} M_m^{1/3}, & \text{for} \quad M_m < M_0, \\ M_m^{2/3}, & \text{for} \quad M_m \geq M_0, \end{cases} \tag{2}$$

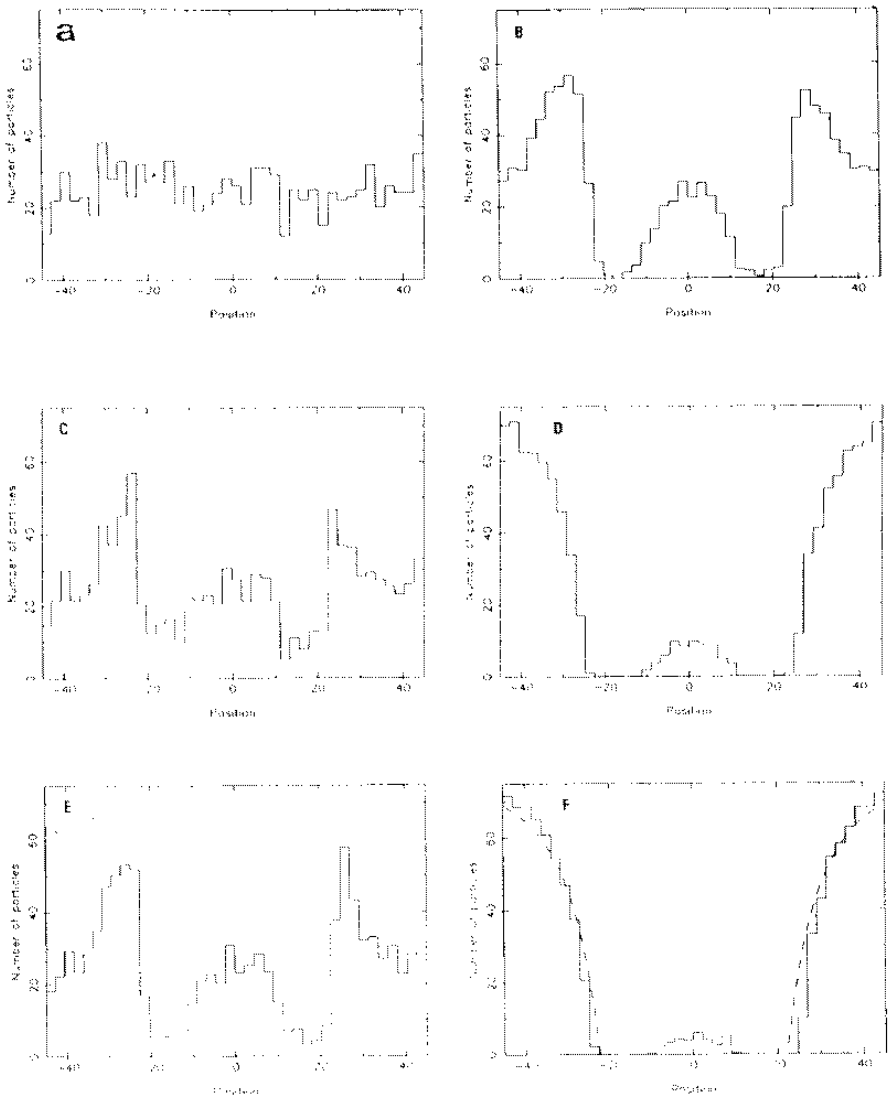



**Fig. 1:** Time evolution of the radial distribution of very small particles ($m \ll M_m$) around the orbit of a moonlet at the "position"= 0.0. The different parts belong to snapshots at different elapsed time of the computer-experiment. Clearly visible is the formation of a gap which is separated by a faint ringlet around the "position"= 0.0 (Figure taken from PH3).

where $M_0$ is the mass threshold depending on the different model parameter (PH3). The order of magnitude is $M_0 \approx 10^{13} kg$ which corresponds to a size of an icy moonlet of a diameter $D \approx 2.7 km$.

In the model of Spahn (1987; S2; S3) particle interactions have been neglected and only the gravitational influence of the moonlet on the ring–material near its orbit has been taken into account. This simplification has enabled to consider (S3):

1. $N = 5000 \cdots 50,000$ particles, giving a high radial resolution of the density profiles;
2. $n$ (up to three) co-orbital moonlets of different masses;
3. eccentric orbits of the moonlets.

However in that model no equilibrium width of the created gap can be expected because of the neglect of the particle interactions - no viscous torque is counteracting the torque exerted by the moon. Here the advantage is to study in detail the dependency of the shape (radial profile) and the width of the inner ringlet on the orbital motion of a massive moonlet ($M_m > M_0$). Concerning the mass-dependence of the radial density profile Spahn and Wiebicke (1989) have obtained results comparable to Petit and Hénon (1988). Assuming a given random motion of the ring–particles (expressed by a mean orbital eccentricity $\overline{e}$; S2), caused by particle interactions and additional outer perturbations (Goldreich and Tremaine 1978; PH1), the inner ringlet is vanishing with decreasing mass of the moonlet $M_m$, while the gap-width scales as "$W \propto M_m^{1/3}$". For very massive moonlets ($h \approx [M_m/3M_p]^{1/3} \approx R_m/[km] \cdot 10^{-4} \gg \overline{e}, \quad \overline{e} \leq 10^{-6}$, $M_p$, $R_m$ and $h$ are the mass of the planet, the radius and the Hill-scale of the moonlet, respectively), the inner ringlet is surprisingly stable and exhibits a detailed structure.

Furthermore, because of the comparable large number of particles $N = 5000 \cdots 50,000$ it has been found (S3), that the density–profile of the inner ringlet shows a detailed structure which contains information about orbital parameters of the moonlet. This is shown in fig. 2, where the different radial density–profiles - taken after a few synodic periods when the structure changes only slightly in time - correspond to different eccentricities $e_m$ of the orbit of the moonlet. Obviously, the larger $e_m$ the narrower the width of the ringlet becomes, which is quite easily seen in fig. 2d where the ringlets caused by moonlet in a circular orbit (solid line), and originated by a moonlet in an eccentric orbit ($e_m = 1.5h$; dashed line), are compared directly. The width $W_r$ of the ringlet can be written

$$W_r \approx C(e_m) \cdot a_m \cdot h(M_m), \tag{3}$$

where $a_m$ is the semi–major axis of the orbit of the moonlet. The factor $C(e_m)$ is a nearly linearly decreasing function of $e_m$ and takes values $C(e_m = 0) \approx 3.0 \cdots C(e_m = 1.5h) \approx 1.2$.

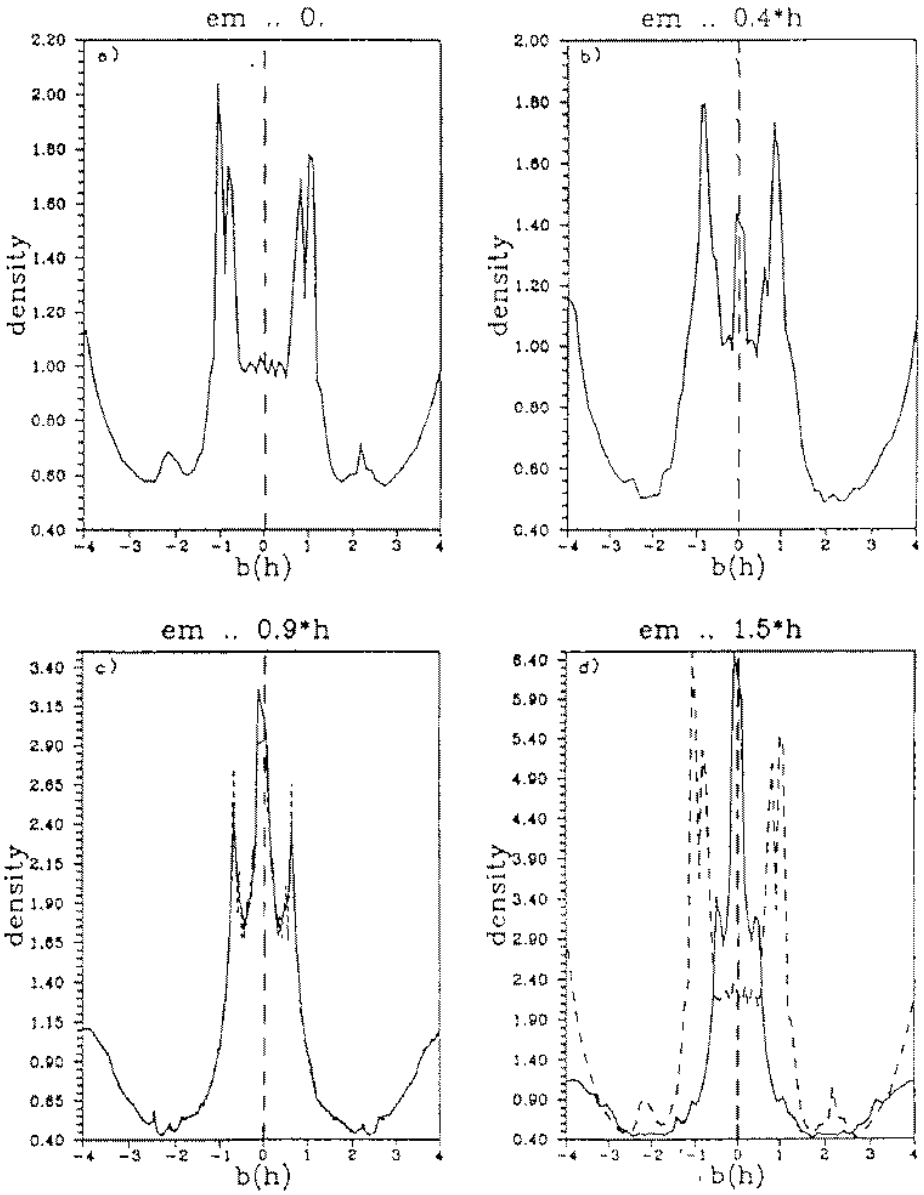



**Fig. 2:** Radial profiles of $\sigma(b, t = constant)$ ($b$ - impact parameter) for different eccentricities $e_m$ of the orbit of the moonlet (case of a "cold" ring $\Rightarrow$ mean-eccentricity of the ring-particles $\overline{e} \to 0$). A narrowing of the ringlet-width is really obvious, and furthermore the shape of the density-profile is turning from a double peaked to a triple peaked one with increasing $e_m$. All profiles show a quasi-symmetry $\sigma(b) \approx \sigma(-b)$ with respect to the orbit of the moonlet at $b = 0.0$.

At the same time the double-peaked shape of the ringlet-profile turns to a triple peaked one with increasing eccentricity $e_m$ .

Summarizing, the results for

1. the gap-formation and its equilibrium–width (Eq. (2)) of Petit and Hénon (1988);
2. the information about the ringlet found by Spahn et al. (1992) (the ringlet-width and the shape of its density-profile);
3. as well as the fact that all profiles obtained (PH 3; S2; 3) are quite symmetric with respect to the mean radial location of the moonlet-orbit, according to $\sigma(b) \approx \sigma(-b)$;

could now be used to search for small embedded satellites. Before doing this in a future work, the value of the above findings will be tested in the case of satellite Pan, embedded in the Encke–gap of Saturn's rings. This is presented in the next section.

# 3 The action of satellite Pan in the Encke-gap

If the results of our models are of any value for the reality, they should be visible also in the optical depth-profile of such a ring-region where a satellite is moving. This is the case for the Encke-division in the A-ring of Saturn, where Showalter (1991) has discovered recently a new satellite - named Pan. Fig. 3 shows the optical depth profile $\tau(r)$ of the Encke–division, which has been obtained with the Photopolarimeter (PPS–data) of Voyager 2. At first one recognizes a wide gap - of 315km width - around the orbit of Pan, which is one of the predictions mentioned above (Section 2). Using this value for the width and applying the relation (2) of PH3 we find that one needs a moonlet of mass $M_m \approx 1.2\ 10^{15} kg$ to create such a gap. For an icy moon the corresponding size would be about $D_m \approx 13km$ . This satellite would be able to confine a ringlet of a width of $W_r \approx 25km$, if the moon is assumed to move in a circular orbit. The narrowing of the ringlet–width (we see only a 15 km wide ringlet) can be explained by an eccentric orbit of Pan, with an eccentricity of $e_{Pan} \approx h_{Pan} \approx 10^{-4}$ (according to relation (3)). Then the radial profile of the ringlet should be triple–peaked (see fig. 2).

The bottom parts of fig. 3 show enlargements of that ringlet region which correspond to the resolutions of $450m$ and $900m$ (processed by binning the PPS-data), respectively. They seem to have the required symmetry, $\tau(b) \approx \tau(-b)$ and furthermore, the $\tau$-profile of the ringlet seems to have at all resolutions the expected triple–peaked shape. However, the shape of the profile of the ringlet is in fact not sure because of the large fluctuations of the PPS–data in the Encke–division.

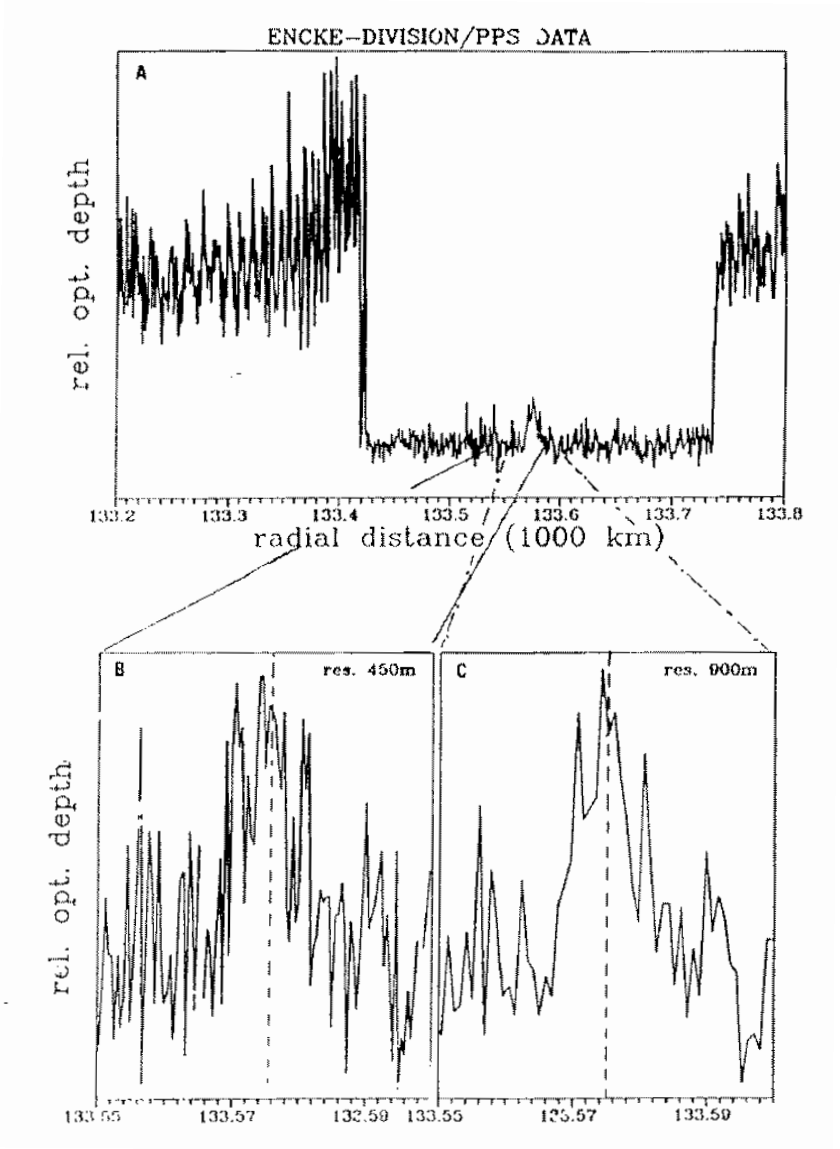



**Fig. 3:** a) Optical depth profile $\tau(r)$ near the Encke-division in Saturn's A-ring, measured by the Photopolarimeter (PPS) of Voyager 2. Clearly visible: a gap with a faint ringlet at its center, where the orbit of Pan is located. Both features - the gap and the ringlet - have been predicted by the models (see Section 2; PH3, and S1). b-c) Enlargements of the region around the faint ringlet at $450m$ and $900m$ resolution, respectively. Both examples of the ringlet data seem to have the symmetry with respect to the orbit of Pan (vertical dashed lines), and appear to be triple peaked.

To weight the structure seen in the ringlet-profile we carried out a modified t-test, which seemed to favour the significance of the triple-peaked structure. However, this was not absolutely convincing.

Consequently, to be more sure about the nature of the structure, we have carried out an independent test by applying the wavelet-analysis method.

The wavelet transform of a signal is the decomposition of this signal onto a basis of functions obtained by dilatations and translations of a unique function called the mother wavelet. It allows to analyse a signal with respect of both a scale parameter and a location parameter, so that it is possible to focus the analysis of the signal at a given location and for a given scale.

Let $s(x)$ be a signal depending only on a space-variable, $\Psi(x)$ the mother wavelet. The wavelet transform of $s(x)$ is then a two variable function $C(a,b)$ depending on a scale variable $a$ and on a location variable $b$:

$$C(a,b) = \langle s(x), \psi_{a,b}(x) \rangle \qquad \text{with} \qquad \psi_{a,b} = \frac{1}{a}\Psi(\frac{x-b}{a}) \tag{4}$$

where $\langle h(x), g(x) \rangle = \int_{-\infty}^{+\infty} h(x) g^*(x) dx$ and where $g^*(x)$ denotes the complex conjugate of $g(x)$. $C(a_0, b_0)$ can be seen as the details present in $s(x)$ at the scale $a_0$ in the vicinity of $b_0$. A wavelet is a zero mean function and rapidly vanishing, which insures from a practical point of view a good localization both in space and scale. From the wavelet transform it is then possible to reconstruct the signal through a reconstruction formula (see Grossmann and Morlet 1985). Our aim has been to use the wavelet transform as a signal processing tool to extract the behavior of the signal in the vicinity of the ringlet, the noise having been removed. The range of considered scales has been chosen in a geometrical progression:

$$a_{i+1} = a_i * 2^{1/12}$$

which insures a practically continuous analysis in scale (see Grossmann et al. 1987), $a_0$ and $a_{max}$ being defined in order to consider the signal from its step of sampling to its mean. In order to make cross comparisons, two analysing wavelets have been chosen: the mexican hat and the french hat (see Bendjoya and Slezak 1992) which are well suited for hole or structure detections. As preliminary encouraging results we will only present here the reconstruction obtained with the mexican hat wavelet. It is to be noticed that the french hat wavelet analysis gives quite comparable results and will be published later (see Bendjoya et al. 1992).

For each scale $a_i$, a wavelet coefficient is computed at each point of the sampled signal and to reach our goal we have to eliminate the non significant coefficients and also the scales corresponding to the statistically estimated noise. One disposes then of a set of coefficients on a range of scale which allows to reconstruct without noise the behavior of the signal on this range of scales.

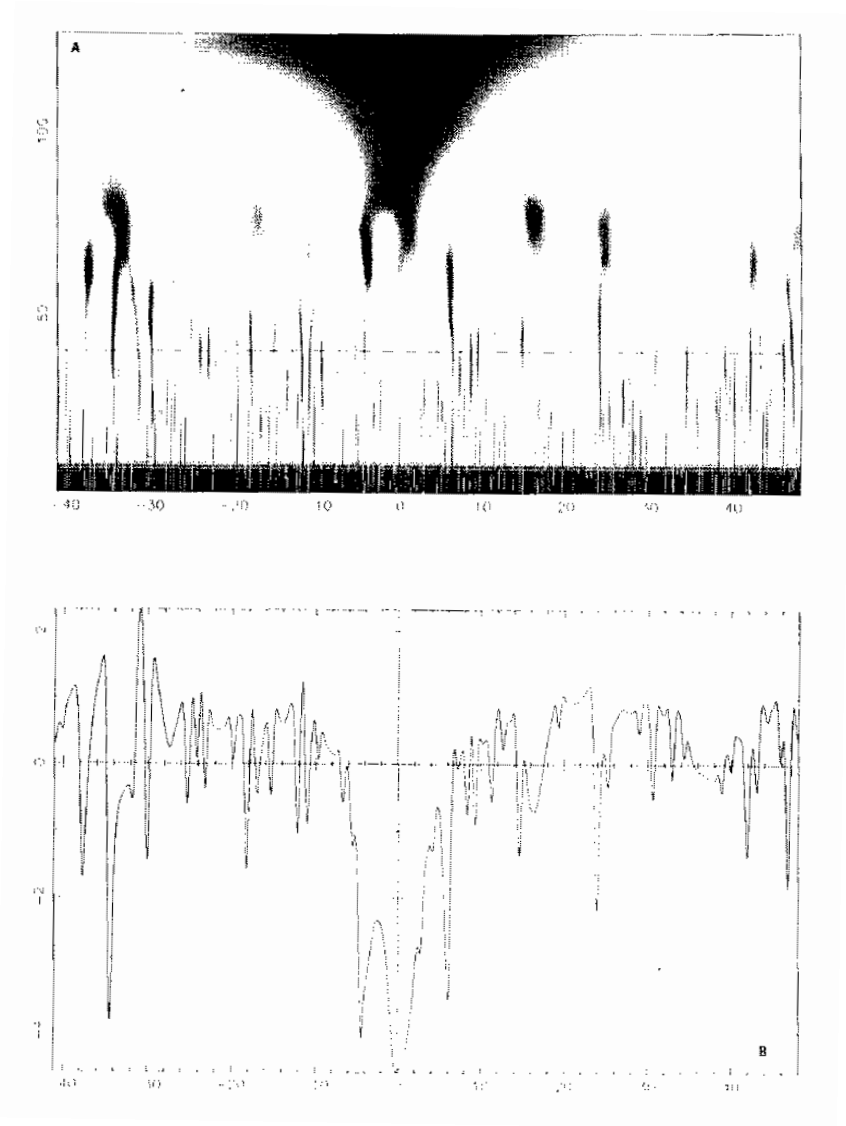



**Fig. 4:** a) A gray level representation of the wavelet transform of the signal. The x-axis gives the radial distance from the center of the Encke–ringlet (orbital radius of 133 576 km), and the y-axis the scale axis. The coefficient less than the threshold are in gray (black for the smallest ones). The gray levels are renormalized for each scale. The black stripe at the very small scale is due to the sampling effect. The dashed line indicates the limiting scale from which the reconstruction is done. b) The reconstructed photon–count signal. The x-axis is the same as above. The y-axis gives the deviation from the mean photon–count in the Encke–division (no extinction of the star signal). Negative value means extinction, that is presence of ring–material.

Therefore, we determine the scales corresponding to the noise then we define a threshold for the coefficients scale by scale. Then we have a set of wavelet coefficients on a range of scales not taking into account the noise, and moreover only the statistically significant coefficients are considered for each scale. The scale below which everything is considered as noise, and also the thresholds of the wavelet coefficients, are derived from the wavelet analysis of a random signal generated with the constraint to fit the distributions of the values of the PPS signal. For each scale of the wavelet transform of the generated signal we compute the distribution of the coefficients. This strategy allows to decrease significantly and quantify the statistical fluctuations in the wavelet coefficients. The minimum scale to consider for the reconstruction without noise has been evaluated to be $a_{40}$ which corresponds to 250 m. Fig. 4a shows the wavelet transform of this signal displayed in a space-scale plane by a gray level representation. This figure is derived from the PPS–photon–count profile $I$, which is related to the optical depth by $\tau \propto \ln(I_0/I)$ ($I_0$ is the photon flux of the occulted star). Since we wanted to detect holes, for each scales the coefficients lower than the threshold are in gray (black for the smallest ones) and the higher ones in white. The dashed line visualizes the scale from which the reconstruction is made. Fig. 4b shows the reconstructed signal from the set of retained coefficients. This reconstruction is made with a threshold at 1- $\sigma$. The structure in three peaks of the ringlet appears obviously. It is to be noticed that this peculiar structure persists for more restrictive thresholds. The existence and the symmetry of the ringlet–peaks around the moonlet position as well as the width of the ringlet are in fairly good accordance with the expected results from the model recalled in the second section of this paper. Another ringlet seems also be pointed out by this analysis around the location 133 542 km from the center of Saturn.

# 4 Discussion

In this paper we have been able to confirm results of theoretical work concerning the gravitational influence of a small satellite, embedded in the rings, on the ring material around its orbit, by analysing the PPS–data of Saturn's rings in the vicinity of the orbit of the satellite Pan. We have been able to estimate the size of Pan to be about $13km$ deduced from the width of the Encke–division, by assuming that the division is swept free by the satellite Pan itself (PH3).

Furthermore, we have compared the shape of PPS–photon–count profile of the Encke–gap–ringlet, as well as its width with theoretical predictions (S3). To do this it has been necessary to remove the noise from the PPS–data in the Encke division, which has been carried out using the wavelet–analysis method. The noise-reduced photon count profile shows three significant depletions inside the ringlet region, pointing to a triple–peaked density profile of the ringlet.

This is in accordance with the results of the model and points to an eccentric orbit of Pan - eccentricity $e_m \approx 10^{-4}$. The width of the ringlet confirms the size estimation of Petit and Hénon (1988).

However, the data analysis presented here is only a first step in analysing the radial density profiles of planetary rings. Future work should address the following points:

1. application of different wavelet-functions to the same ring region;
2. refinement of the estimate of the wavelet-coefficient threshold which separates the noise from "real" structures;
3. applying the wavelet-analysis to other rings (or ring regions), where also small satellites are expected to exist.

With respect to the theoretical modelling the following outstanding questions concerning moonlets in planetary rings are of interest:

1. What about the stability of the inner ringlet, if viscous transport processes, sources (the surface of the moonlets) and sinks (meteorite bombardement) of ring matter is taken into consideration?
2. How does the azimuthal distribution of the matter in the ringlet looks like (2D - simulations)?

Both sets of problems will be investigated in subsequent papers.

# Acknowledgements

We would like to thank Prof. Larry Esposito, Dr. Karen Simmons, and Dr. Joshua Colwell for providing us with the Voyager-2 PPS-data of planetary rings.

# References

[1] Bendjoya, Ph., Petit, J.-M., & Spahn, F., 1993. *Wavelets analysis of the Voyager data on planetary rings. I. Description of the method.* in preparation

[2] Bendjoya, Ph., & Slezak, E., 1992. *Wavelet analysis and applications to some dynamical system.* Celestial Mechanics , in press,

[3] Dermott, S.F., & Murray, C.D., 1981. *The Dynamics of Tadpole and Horseshoe Orbits I. Theory.* Icarus **48**, pp. 1–11.

[4] Goldreich, P., & Tremaine, S., 1978. *The velocity-dispersion in Saturn's rings.* Icarus **34**, pp. 227–239.

[5] Grossmann, A., & Morlet, J., 1985. in *Mathematics+Physics, Lectures on recent results,* ed. L.Streit, World Scientific Publishing,

[6] Grossmann, A., Kronland-Martinet, R., & Morlet, J., 1987. *Reading and understanding continuous wavelet transform.* in *Wavelets, Time-Frequency Methods in Phase Space,* eds. . Combes J.M., Grossmann A. and Tchamitchian Ph., Springer-Verlag, pp. 2–20.

[7] Lissauer, J.J., Shu, F.H., & Cuzzi, J.N., 1981. *Moonlets in planetary rings ?.* Nature **292**, pp. 707–711.

[8] Petit, J.M., & Hénon, M., 1987*a*. *A numerical simulation of planetary rings I. Binary encounters.* Astron. Astrophys. **173**, pp. 389–404.

[9] Petit, J.M., & Hénon, M., 1987*b*. *A numerical simulation of planetary rings II. Monte Carlo model.* Astron. Astrophys. **188**, pp. 198–205.

[10] Petit, J.M., & Hénon, M., 1988. *A numerical simulation of planetary rings III. Mass segregation, ring confinement, and gap formation.* Astron. Astrophys. **199**, pp. 343–356.

[11] Showalter, M., Cuzzi, J.N., Marouf, E.A., & Esposito, L.W., 1986. *Satellite 'wakes' and the orbit of the Encke gap moonlet.* Icarus **66**, pp. 297–323.

[12] Showalter, M., 1991. *Visual detection of 1981 S13, Saturn's eighteenth satellite, and its role in the Encke gap.* Nature **351**, pp. 709–713.

[13] Spahn, F., 1987. *Scattering Properties of a Moonlet (Satellite) Embedded in a Particle Ring: Application to the Rings of Saturn.* Icarus **71**, pp. 69–77.

[14] Spahn, F., & Wiebicke, H.-J., 1989. *Long-Term Gravitational Influence of Moonlets in Planetary Rings.* Icarus **77**, pp. 124–134.

[15] Spahn, F., Saar, A., Schmidt, S. & Schwarz, U., 1992. *The influence of various moonlets on the optical depth profile in planetary rings.* Icarus **100**, pp. 143–153.

**[11]**

# "Wavelet analysis of the Voyager data on planetary rings. I. Description of the method."

BENDJOYA, Ph., PETIT, J-M. *&* SPAHN, F.

*Icarus* **105** (1993)

# Wavelet analysis of the Voyager data on planetary rings. I. Description of the method

Ph. Bendjoya
Observatoire de Nice, B.P. 229, 06304, Nice cedex 04, France
e-mail: benchoc@franck.obs-nice.fr

J.-M. Petit
C.N.R.S.
Observatoire de Nice, B.P. 229, 06304, Nice cedex 04, France
e-mail: petit@bach.obs-nice.fr

F. Spahn
Working group "Nonlinear Dynamics" (MPG)
University Potsdam; D-O-1571 Potsdam; Germany
e-mail: spa@babel.aip.de

**Abstract**

In this paper we present a new method which allows us to extract signal from low (of the order of 1) signal to noise ratio data. The noise is almost completely removed, and the significance of the resulting signal is quantified against pure chance fluctuations. The technique is based on a wavelet transform, and uses a crude pattern–recognition algorithm in the space–scale plane to extract the interesting information. This method has been applied to simulated signals with signal to noise ratio ranging from 0.4 to 2. and the reconstructed signals are shown. From the wavelet transform coefficients, it is possible to gain some quantitative information on the size of the detected structures. As a first real case application, we present the analysis of the Encke gap ringlet in the Saturn's rings. We show that its detailed shape contain at least two peaks, and that its total width is about 8 km, while the width of the left peak is about 1.5 km.

# 1 Introduction

The Voyager 2 PPS data, obtained from the occultation of the star $\delta$ *Scorpii* by the Saturn's rings, have provided a spatial resolution in the radial direction better than 100 m (Lane *et al.* 1982 ; Esposito *et al.* 1983a). Many kinds of structures, at every scale length, have been discovered from these data in Saturn's rings. A certain number of papers have been devoted to the analyses of these data pointing out different kind of features such as density waves (Esposito *et al.* 1983a, 1983b; Brophy and Rosen, 1992), gaps (Petit and Hénon, 1988), sharp edges (Borderies *et al.* 1989),and wakes (Showalter *et al.* 1986). A global catalog of 216 features has been established by Esposito *et al.* 1987. The main philosophy of this catalog was to present the significant features, not to study the detailed shape of low signal-to-noise ratio features. It appears interesting to study such fine details in order to compare the observations with some theoretical models. Indeed, in what concerns a ringlet due to the presence of an embedded moonlet, different models lead to different shapes of the ringlet. This shape is strongly related to the orbital elements of the moonlet and its mass and size (Spahn *et al.* 1992). Another interest is to study structures like small variations in the optical depth as an indirect detection of large particles (from 100 m to a few kilometers in size) that are too small to be directly detected.

The purpose of this paper is to present a new method of signal processing based on the wavelet transform. This method allows us to detect, count and reconstruct localized bumps and/or depletions, with a drastic decrease of noise. Since this method is essentially a local one, the main limitation of this method occurs when one is trying to detect weak structures close to strong ones. An iterative process of analysis is then needed. The method is not based on any modelling of the noise (Gaussian, Poissonian ...), and gives, as it will be shown, satisfying results even for very low signal-to-noise ratio signals (of order of 0.5). In section 2 we briefly recall the principles of the wavelet transform and present the strategy we have applied to use this transform in detection of significant structures. In section 3 we applied this method on simulated signals. We show the details of the algorithm concerning : the noise elimination, the quantification of the significance of detected structures, and the behavior of the reconstructed signal as function of the level of significance. As a first real case application we present the reconstruction of the Encke gap ringlet.

# 2 The method

This section is divided in two parts. The first part gives a summary of the main results on the wavelet transform that will be used in this paper. More details can be found in Daubechies,

1988, Mallat, 1989 and Meyer, 1989. The second part describes the method we have built to extract information and remove noise from the wavelet transform coefficients.

## 2.1 Wavelets transform

We want to study a non–periodic signal $S(x)$ and obtain information on both the location and the size of the characteristic features (later on called structures). We first recall a few basics about the well–known Fourier transform.

The Fourier transform is good at giving the "frequency" (size). And it is a projection on an orthogonal basis. Hence, there exists a unique decomposition and reconstruction formula for a given function. But there is no relationship between the local behavior of $S(x)$ and the amplitude of the Fourier coefficients. This information is so deeply buried in the phases of the coefficients that it is not useful. Another problem is that the functions $\sin(x)$ and $\cos(x)$ are not compactly supported (they are not identically null outside of a finite interval). Therefore we need to know the behavior of $S(x)$ at infinity in order to compute the coefficients, unless we are dealing with a periodic signal. The Fourier coefficients are simply the inner products of $S(x)$ and a set of functions derived from two functions $\sin(x)$ and $\cos(x)$ by a scaling:

$$\sin(x) \to \sin(\omega x), \qquad \cos(x) \to \cos(\omega x) \tag{1}$$

The wavelet transform of $S(x)$ is the decomposition on a basis of functions $w_{a,b}(x)$, all derived from a unique function $w(x)$, called the "mother wavelet", by translation and scaling:

$$w_{a,b}(x) = a^{-1/2} w(a^{-1}(x-b)) \tag{2}$$

The coefficient $a^{-1/2}$ is not mandatory. One can choose any other exponent for $a$, depending on the kind of normalization one is interested in. The exponent $1/2$ is useful for normalizing the energy (the square of the modulus of the coefficients). The function $w(x)$ must satisfy a "compatibility condition":

$$\int_0^\infty \frac{|\hat{w}(\omega)|^2}{|\omega|} d\omega < \infty, \tag{3}$$

where $\hat{w}(\omega)$ is the Fourier transform of $w(x)$, in order to be a wavelet. This means, in particular, that the integral of the wavelet should vanish. It must also be well localized both in space and in scale. Mathematically speaking, we choose an integer $r$ ($r$ is called the order of the wavelet), and we impose the properties:

- For all $x \in R$ (set of reals), all $m \in N$ (set of natural integers), and $0 \leq q \leq r$, there exist $C_m$ such that:

$$\left| \frac{d^q w(x)}{dx^q} \right| \leq C_m (1 + |x|)^{-m}. \tag{4}$$

- The collection
$$2^{j/2}w(2^j x - k), \quad j \in Z, \quad k \in Z \tag{5}$$
is an orthonormal basis of $\mathcal{L}^2(R)$.

The first of these conditions insures the smoothness of $w(x)$, and both conditions together imply that the moments of $w(x)$, up to order $r$ vanish. S. Mallat (1989) discovered a systematic approach to find solutions to this problem. I. Daubechies (1988) showed that, for any $r$, there exist compactly supported functions $w(x)$ satisfying the above conditions.

The wavelet coefficients of $S(x)$ are simply:
$$d_{j,k} = \int_{-\infty}^{\infty} S(x) w_{j,k}(x) dx, \tag{6}$$
and the unique way of recovering $S(x)$ from the coefficients is:
$$S(x) = \sum_j \sum_k d_{j,k} w_{j,k}(x). \tag{7}$$

The previous formula gives $S(x)$ as a sum of details from the finest scale to the largest one. One can also choose another orthonormal basis consisting of the collection $w_{j,k}(x)$, $j \in N$, $k \in Z$ (set of relative integers), together with the collection $q(x-k)$, $k \in Z$, where $q(x)$ is a new smooth function with a rapid decay at infinity. Obviously, the two functions $q(x)$ and $w(x)$ cannot be chosen independently. In this basis, $S(x)$ can uniquely be written as:
$$S(x) = \sum_{k \in Z} c_k q(x-k) + \sum_{j \in N} \sum_{k \in Z} d_{j,k} w_{j,k}(x) = E_0(S) + \sum_{j \in N} D_j(S) \tag{8}$$
where $c_k = \int_{-\infty}^{\infty} S(x) q_k(x) dx = \int_{-\infty}^{\infty} S(x) q(x-k) dx$. This means that $S(x)$ is the sum of a smooth part $E_0(S)$ and a sequence of finer and finer fluctuations $D_j(S)$. The function $q(x)$ is a smoothing function that allows us to cut the doubly infinite sum of eq. (7). The $q(x-k)$'s account for the large scale fluctuations, the $w_{j,k}$'s for the small ones. Obviously, the resolution corresponding to the smooth part can be chosen arbitrarily.

Consider now a real sampled signal. The ranges of variation of $j$ and $k$ are now finite. The scale range is defined by the sampling step for the smallest one (finest resolution) and the length of the signal for the largest (coarsest resolution). The range of variation of $k$ is also defined by the length of the signal. Among the many possibilities for the two functions $q(x)$ and $w(x)$, we choose:
$$\begin{aligned} q(x) &= 1 - |x| \quad \text{for} \quad |x| \leq 1, \\ &= 0 \quad \text{otherwise}, \end{aligned} \tag{9}$$

and:

$$\begin{aligned} w(x) &= \frac{|x|}{4} - \frac{1}{2} \quad \text{for} \quad 1 < |x| \leq 2, \\ &= \frac{1}{2} - \frac{3|x|}{4} \quad \text{for} \quad |x| \leq 1, \\ &= 0 \quad \text{otherwise}, \end{aligned} \tag{10}$$

The process of computing the wavelet coefficients can be done iteratively. The first trivial step is to take the function $q(x)$ with the same resolution as the sampled signal. We only have to compute the $c_k^0$'s (the superscript refers to the number of the step in the iterative process). From the formula for $q(x)$, one can see that the $c_k^0$'s are simply the sampled values of $S(x)$ (the integrals are changed into summations since we are dealing with a discrete set of points). Then we consider a resolution twice as large. We choose the normalizing coefficients such that we replace $q(x)$ by $q(x/2)/2$. There is only one level of fine fluctuations to add to the smoothed part to recover the signal, and it is defined by the set of $d_{1,k} = d_k^1$. We can now repeat the same procedure to the smoothed part of level 1. This will give us two new sets of coefficients $c_k^2$ and $d_k^2$. With our choice of $q(x)$ and $w(x)$, the formulae to obtain the $c_k^j$'s and the $d_k^j$'s are very simple:

$$c_k^j = \frac{1}{2} c_k^{j-1} + \frac{1}{4} \left( c_{k-2^{j-1}}^{j-1} + c_{k+2^{j-1}}^{j-1} \right), \quad j \geq 1, \tag{11}$$

and:

$$d_k^j = c_k^{j-1} - c_k^j. \tag{12}$$

One can see that the $c_k^j$'s are obtained by applying a low–pass filter to the $c_k^{j-1}$'s. The $d_k^j$'s being obtained by a difference between two levels of $c_k^j$'s are in fact the result of a band–pass filter applied to the signal. Due to the shape of the function $q(x)$, the associated wavelet $w(x)$ is often called the "linear wavelet".

For the reconstruction formula, things are quite simple. We are interested in reconstructing the sampled values, not the signal $S(x)$. When adding the $c_k^j$'s and the $d_k^j$'s, we get the $c_k^{j-1}$. Hence, the reconstruction formula is:

$$S(x_k) = c_k^{j_{max}} + \sum_{j=1}^{j_{max}} d_k^j, \tag{13}$$

where $j_{max}$ is the maximum scale corresponding to the length of the signal.

A more general description of this method has been given by S. Mallat (1989) in the framework of the multiresolution analysis.

The wavelet transform described above is the one that allows us to work with the coefficients in order to remove the noise in a real signal (we are working with coefficients in an orthogonal basis).

## 2.2 Signal processing

Our aim is to use the wavelet transform as a signal processing tool to extract the behavior of the signal while removing the noise. We also want to have the possibility to compute the statistical significance of detections. The main philosophy of the procedure consists in isolating, in the space–scale plane where the wavelet transform of the signal is unfolded, the patterns made by coefficients pointing out structures with a level of significance versus the noise fluctuations. Then the signal is reconstructed from the selected coefficients. The patterns of significant wavelet coefficients are exhibited by means of a double thresholding. A first one is very restrictive, leading to the detection of very significant structures within the signal, with a drastic decrease of the noise, and a second one which allows the addition of more and more local details as it is chosen more and more permissive. These thresholdings are made scale by scale by comparing the wavelet coefficients derived from the studied signal to the distribution of coefficients computed from a pseudo–random signal. The pseudo–random signal is generated from the signal. For each point $x_k$ of the signal, we draw at random a value $S_{pse}(x_k)$ of the pseudo–random signal, according to the probability distribution computed from the real signal. This strategy has already been applied in asteroid family detection (Bendjoya *et al.* , 1991) and allows us to take into account the intrinsic distribution of the noise instead of having to model it.

The wavelet transform of a white noise signal gives fluctuating coefficients, with no noticeable pattern in the space–scale plane (see Fig. 1). Also of importance is the fact that almost no coefficients deviate greatly from the mean. Fig. 1b displays a representation by a gray scale coding of the wavelet transform coefficients of the signal performed by means of the linear wavelet. The y-axis stands for the scales and the x-axis for the positions. For a given scale the value of a coefficient is coded by a gray level which varies linearly from black to white with the value of the coefficient in the range the mean of the coefficients plus or minus four times the standard deviation of the coefficients at the given scale. The normalization is made scale by scale. For a noiseless signal, there is a definite pattern in the space–scale plane (see Fig. 2). Particularly, some coefficients deviate by a large amount from the mean, and they tend to the mean on each side of the maximum for each scale (this is due to the absence of structures on each side), with a width proportional to the scale. In the following, in order to determine the mean and quantify a "large deviation", we evaluate the mean and standard deviation, for each scale, of the coefficients of the wavelet transform of a pseudo–signal $S_{ps}(x)$. As mentioned

above, this pseudo–signal is a random data set having the same cumulative distribution as the

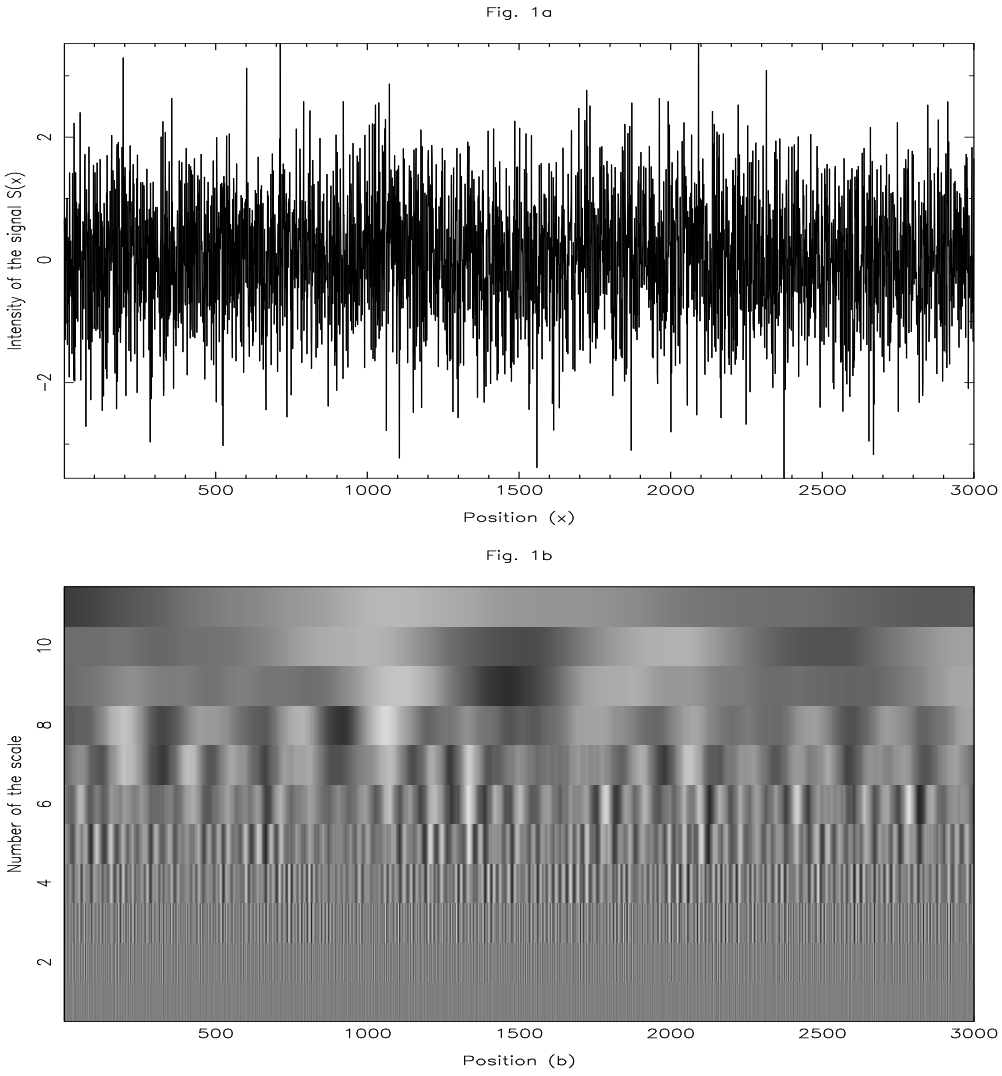



**Fig. 1:** a) A white noise signal. The x–axis is the position, and the y–axis is the signal intensity $S(x)$ b) Gray scale representation of the coefficients of the wavelet transform of the noise signal of Fig. 1a, using the linear wavelet (see section 2). The x–axis is the position, and the y–axis is the number of the scale. For a given scale the value of a coefficient is coded by a gray level which varies linearly from black to white with the value of the coefficient in the range the mean of the coefficients plus or minus four times the standard deviation of the coefficients at the given scale. The correspondence between value and gray level is made scale by scale.

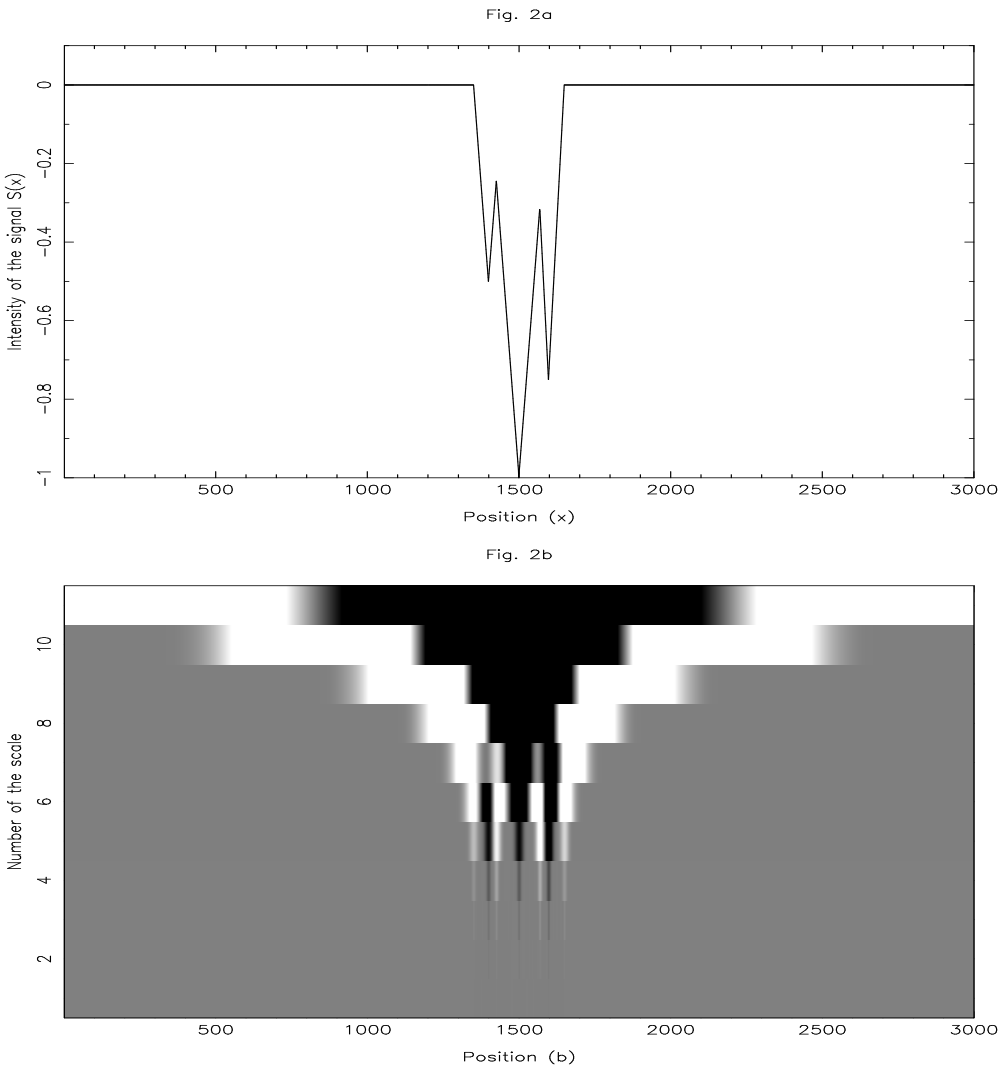



**Fig. 2:** a) A three peak noiseless signal artificially generated. b)Same as Fig. 1b for the signal of Fig. 2a.

signal under study $S(x)$.

We use a pattern–recognition algorithm in order to isolate the significant patterns in the space–scale plane. This is done by a double thresholding. First, we search for points in the space–scale plane with a very high level of significance (the seeds). Those seeds are the points where the coefficients deviate by a large amount, i.e. by more than a given number times the

standard deviation, from the mean. This defines a first threshold, the seed–threshold $T_{se}(a_i)$. Usually we take $T_{se}(a_i) = 5\sigma_{S_{ps}}(a_i)$, where $\sigma_{S_{ps}}(a_i)$ is the standard deviation of the coefficients of scale $a_i$ of the pseudo–random signal. The thresholding is done by retaining the points where the coefficients minus the mean, defined above, are larger (resp. smaller) than $T_{se}$ (resp. $-T_{se}$) for a bump (resp. hole) detection. The seeds give a rough estimate of the size, position and shape of significant structures in $S(x)$. In order to add details to these detected structures, we need to take into account some coefficients in the vicinity of the seeds.

We define a second threshold, the skeleton–threshold $T_{sk}$ which determines the points that might be associated with the seeds. Usually $1.5\sigma_{S_{ps}}(a_i) \leq T_{sk}(a_i) \leq 2\sigma_{S_{ps}}(a_i)$. Among the points that are selected by this second weaker threshold, those that are connected by nearest neighbor relation to a seed actually belong to a pattern. In other words, a wavelet coefficient is included if the absolute value of its difference from the mean value is above the threshold, and if it is adjacent to a "seed" either in the location index or the scale index direction. These new points are new seeds. The procedure of extension is iterated until no new point is added. As seen before (Fig. 2) a structure in $S(x)$ gives significant coefficients on a range of scales. The most significant ones, corresponding to the seeds which belong to restricted range of scales (see Fig. 5a). Therefore we have to extent the seed along the scale axis. Furthermore, at a given scale, only the coefficients in the middle of the structure are selected by the seed–threshold. However the structure contributes to coefficients that are not selected by $T_{se}$ but are significant according to $T_{sk}$ (see Fig. 5b). Then we must consider them. At this stage of the procedure we have a set of coefficients, with a level of significance defined by $T_{sk}$ but the cores of the patterns thus defined are significant at the level defined by $T_{se}$.

Once we have determined the skeletons, we use the reconstruction formulae with the coefficients of the skeletons unchanged, and the other ones put scale by scale to the mean. Unfortunately, if we do so, the reconstruction of a bump (resp. hole) gives rise to a hole (resp. bump) on each side of the signal. Due to the shape of the analysing wavelet, a bump, for example, gives coefficients that are larger than the mean in its central part and smaller than the mean on its edges (see Fig. 3, the wavelet coefficients of a Dirac). Therefore, in order to avoid this reconstruction bias, we must include also the coefficients that are smaller than the mean in the vicinity of the skeleton of the bump. The size of the vicinity is the extent of the wavelet at each scale. A precise description of the algorithm to avoid the bias is done in Appendix A.

The different steps of the procedure can be summarized as follow : (i) a first restrictive thresholding $T_{se}$ allows to localize the strongly significant structures within the signal, and eliminate nearly all noise; (ii) the seeds formed by the coefficients selected by $T_{se}$ are then enriched along both the scale and the space axis by related coefficients coming from a second more permissive thresholding $T_{sk}$; (iii) in order to avoid some bias of reconstruction, some coefficients are considered in the vicinity of the skeletons defined by the enriched seeds; (iv) the

non selected coefficients are changed to the mean of each scale; (v) the signal is reconstructed from the coefficients obtained after step (iv).

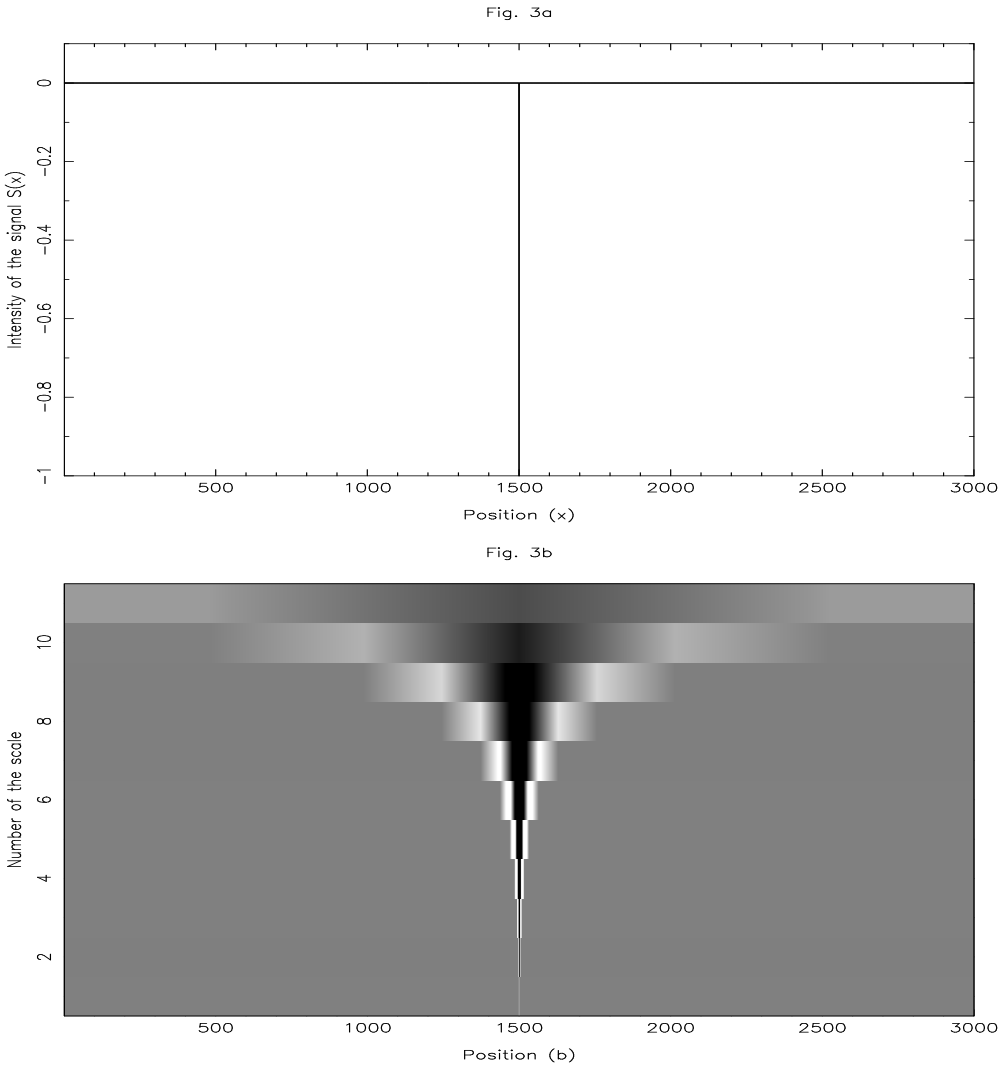



**Fig. 3:** a) A Dirac signal. b) Same as Fig. 1b for the signal of Fig. 3a.

# 3 Applications

## 3.1 Simulated signal

The models of gravitational interactions between the particles of a planetary ring and a moonlet orbiting within this ring has led Petit and Henon (1988) to exhibit the formation of a stable ringlet from numerical simulations. Collisions between particles were taken into account. This restricted the number of particles used in the simulations, which in turn prevented the analysis of fine details within the ringlet. The hydrodynamic approach of Spahn *et al.* (1992) allows us to propose as a stable solution a ringlet the density of which is inhomogeneous and presents a symmetric feature composed by three "peaks" in the distribution of matter.

By analyzing the Voyager data in the Encke gap of Saturn's A ring, Showalter (1991) detected the moonlet Pan which may be responsible for a ringlet in this area. The diameter of this moonlet was estimated by Showalter to be around 20 km. An analysis made by Spahn on the Voyager 2/PPS data by means of a *t-test* method seemed to show the presence of a three-peak structure (private communication). Unfortunately the very low signal to noise ratio (of order of 1) of the data prevented him from convingcingly detecting this peculiar structure with the methods he used. This signal exhibits different structures on a very large range of characteristic sizes. Therefore the wavelet transform appears well-suited for processing such a signal.

The idea is to create first a signal with a three–peaked structure and add to it different realizations of gaussian noise in order to simulate the Voyager data. The wavelet analysis method has been performed for different shapes of the three–peak structure, different signal to noise ratios and different realizations of the noise.

## 3.2 Wavelet coefficients

Fig. 2a shows a noiseless simulated three-peak structure and Fig. 4a the same with an added noise realization. We define the signal to noise ratio ($S/N$) as the ratio between the amplitude of the structure (here, 1) and the standard deviation of the gaussian noise we have added. This definition gives a local meaning to $S/N$. In our case, we are not interested in large scale structures but rather in local ones. Fig. 4a shows a signal to noise ratio of 1.2, since the standard deviation is 0.833 and the amplitude of the noiseless signal is 1.

Fig. 4b displays the gray level coded coefficients obtained from the wavelets analysis of the signal of Fig. 4a. Fig. 5a shows the coefficients selected by the first thresholding $T_{se}$ (here, $T_{se}(a_i) = 5\sigma_{S_{ps}}(a_i)$). One can easily distinguish the patterns forming the seeds in the space–scale plane of the wavelet coefficients. One can also remark how the noise corresponding to the

small scales is eliminated.

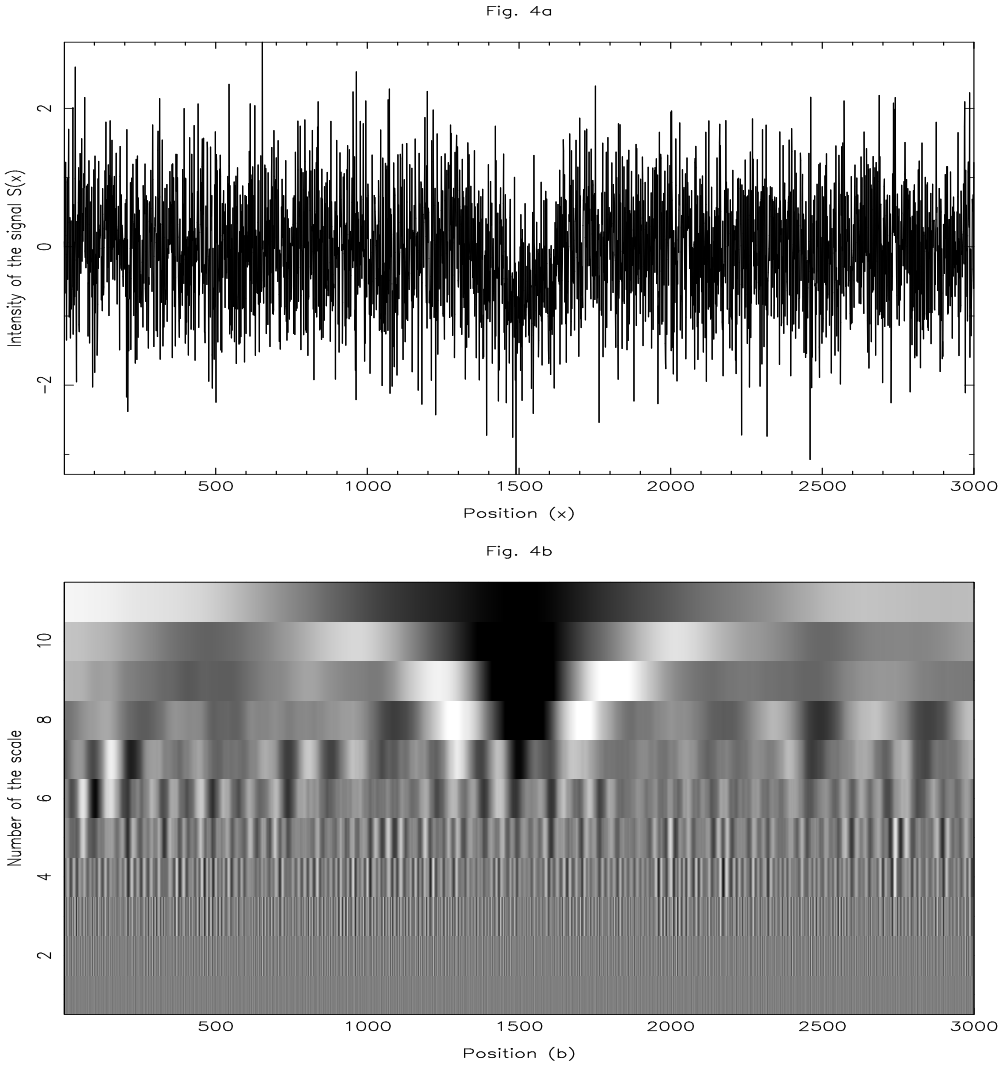



**Fig. 4:** a) Superposition of Fig. 2a and a white noise leading to a signal to noise ratio of 1.2. b)Same as Fig. 1b for the signal of Fig. 4a.

Fig. 5b shows the skeletons of coefficients created by adding to the seeds the connected coefficients coming from the second thresholding ($T_{sk}(a_i) = 1.6\sigma_{S_{ps}}(a_i)$) and also a contour plot of the coefficients added to avoid bias reconstruction (see section 2.2 step (iii)).

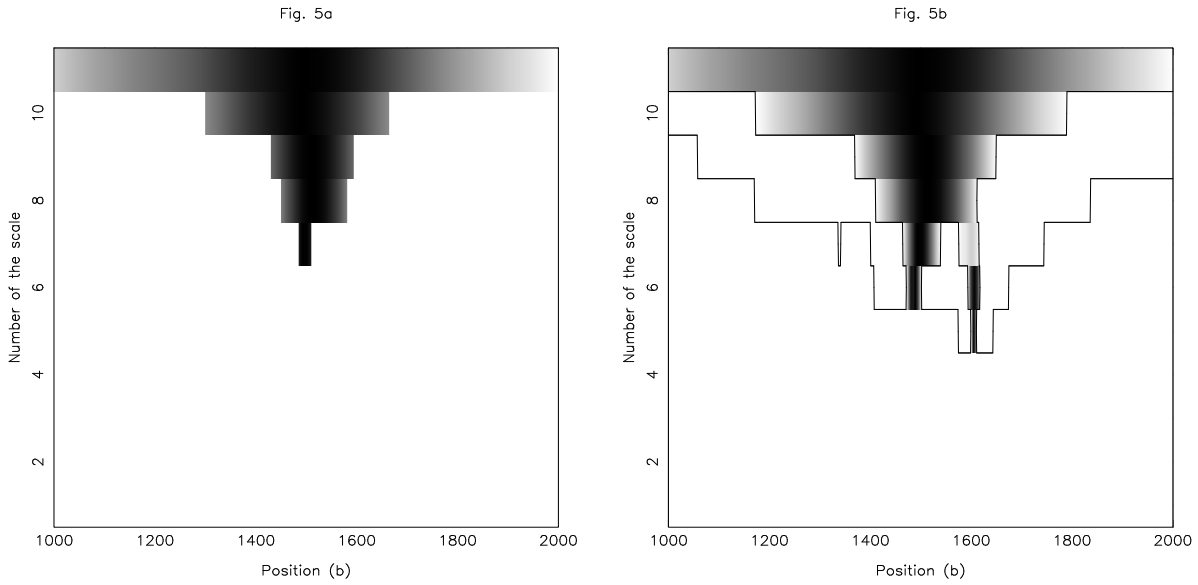



**Fig. 5:** a) Gray scale representation of the coefficients selected by the first and most restrictive thresholding ($T_{se}(a_i) = 5.0\sigma_{S_{ps}}(a_i)$), extracted from the coefficients of Fig. 4b (a zoom is made around the region of interest). b) Coefficients satisfying the second thresholding ($T_{sk}(a_i) = 1.6\sigma_{S_{ps}}(a_i)$) and connected to the previous ones have been added. The contour plot shows the location of positive coefficients added to avoid reconstruction biases.

## 3.3 Reconstructed signals

The main problem at this step of the analysis is to define the "best" second threshold. The "best" $T_{sk}$ (quoted hereafter $T^*_{sk}$) must be understood as the threshold which is permissive enough to add fine details to the localized structures, but also strict enough to avoid the reappearance of localized sharp features, with a relatively large amplitude, and which can be associated with noise. One must keep in mind that there are always finite probabilities that some true features of the signal will be lost by thresholding and some noise spikes will make it through the threshold. The goal is to optimize the trade–off by picking the "best" threshold, avoiding any qualitative criterion. For the choice of $T^*_{sk}$ we have developed a strategy relying on the philosophy of the wavelet transform.

The idea is to get an indicator of the presence of noise in the reconstructed signal and to choose $T^*_{sk}$ as the value of the threshold for which this indicator changes suddenly. As we have already pointed out, the wavelet transform provides a multi-resolution analysis in which small details in the signal are represented at the small scales. More precisely the noise is mainly

detected by the first scale. Therefore we compute the first scale of the wavelet transform on the reconstructed signals obtained for different $T_{sk}$. The indicator of the presence of noise is the standard deviation of the wavelet coefficients at this smallest scale.

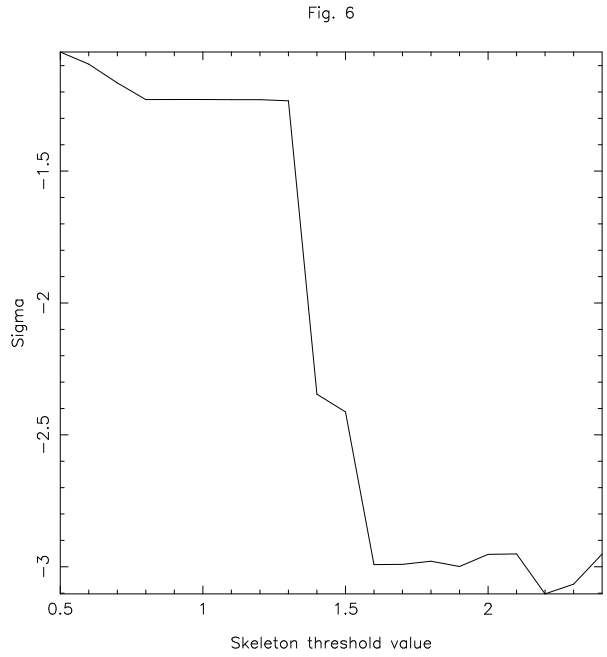



**Fig. 6:** The standard deviation (in a log scale) of the wavelet coefficients at the smallest scale computed from the reconstructed signals , as a function of $T_{sk}$

Fig. 6 shows the plot of this standard deviation as a function of $T_{sk}$ (in a semi-log representation). The curve is typically made of plateaus and sharp changes. In each plateau there is no significant change of the reconstructed signal. Each sharp change (when decreasing $T_{sk}$ ) corresponds to the adding of new details and eventually some noise. Among a set of discret values of $T_{sk}$, from 2.5 downto 0.5 sampled by a step of 0.1, we choose the smallest value of $T_{sk}$ for each plateau. This gives us one or several values of $T^*_{sk}$ for which we reconstruct the signal.

Fig. 7 shows the different reconstructed signals as a function of various choices of $T_{sk}$. One can see how the noise is gradually added as $T_{sk}$ decreases. By a simple look to this series of reconstructed signals (from bottom right to top left), one is tempted to choose as the value of $T^*_{sk}$, the one before the alteration of the smoothness of the reconstructed signal. Practically, in fig. 7, we would choose $T^*_{sk}(a_i) = 1.6\sigma_{S_{ps}}(a_i)$, which correspond to the lower limit of the plateau in fig. 6.

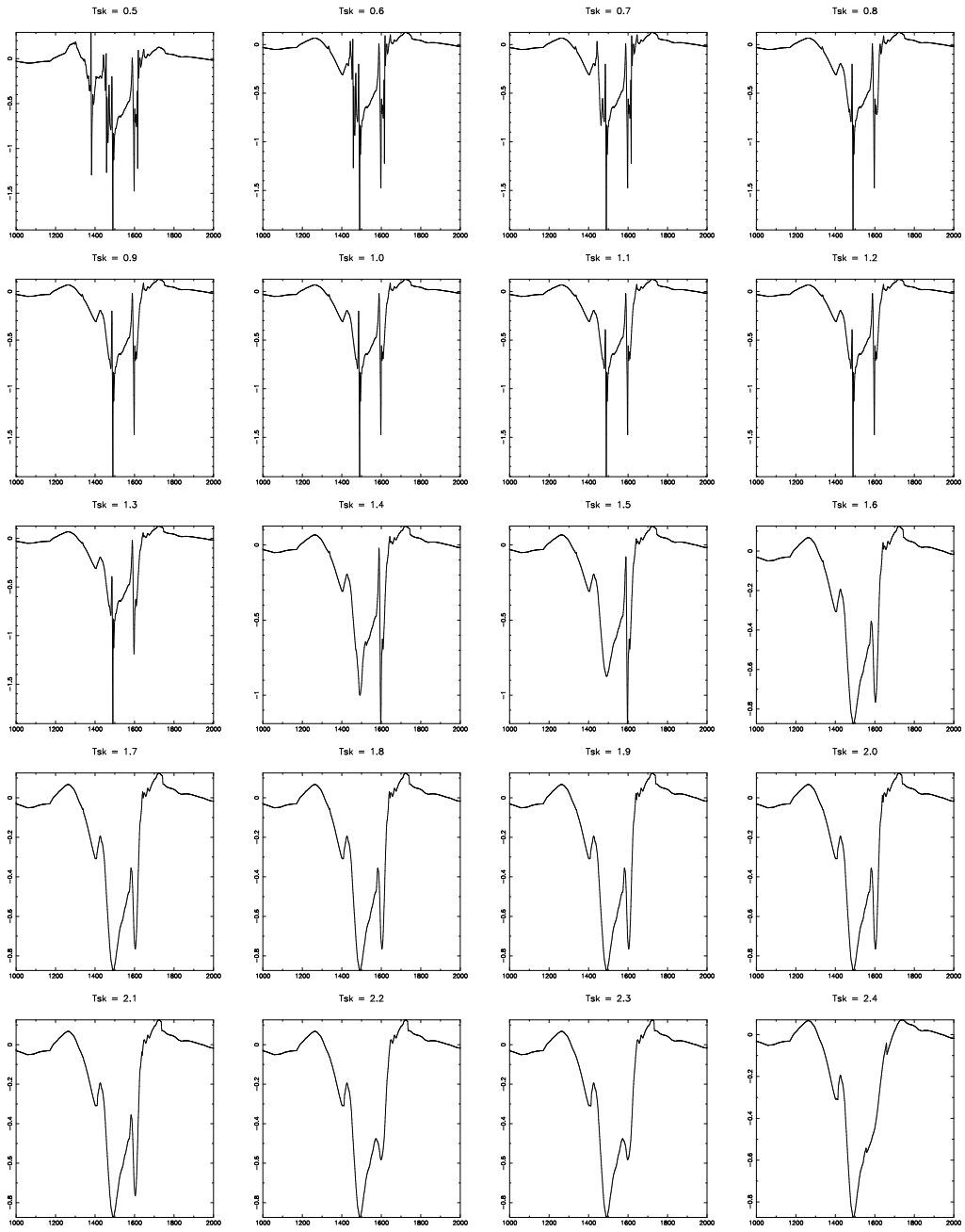



**Fig. 7:** The series of reconstructed signals (three peaks structures with a noise leading to S/N = 1.2), for a set of $T_{sk}$ from 0.5 to 2.4 with a step of 0.1

Figs. 8 shows the reconstructed signal (full line) superimposed on the simulated noiseless structure (dashed line), for $T^*_{sk}(a_i) = 1.6\sigma_{S_{ps}}(a_i)$. One can appreciate the quality of the reconstruction of this signal with a signal to noise ratio of 1.2 The noise is eliminated and the reconstructed structure has a level of significance of, at least, $1.6\sigma_{S_{ps}}$ (for each scale) versus chance fluctuations which corresponds for a gaussian distribution of the wavelet coefficient to a level of confidence of 94.5 %. However, the core of the structure (the part which is associated to the coefficients selected by $T_{se}(a_i)$) has a level of significance of $5.\sigma_{S_{ps}}$, corresponding to a confidence of 99.99997 %. Thus, the details added with $T_{sk}$ are more significant than other coefficients with amplitudes of the same order of magnitude since they are connected by nearest neighbor relation to this particularly robust core.

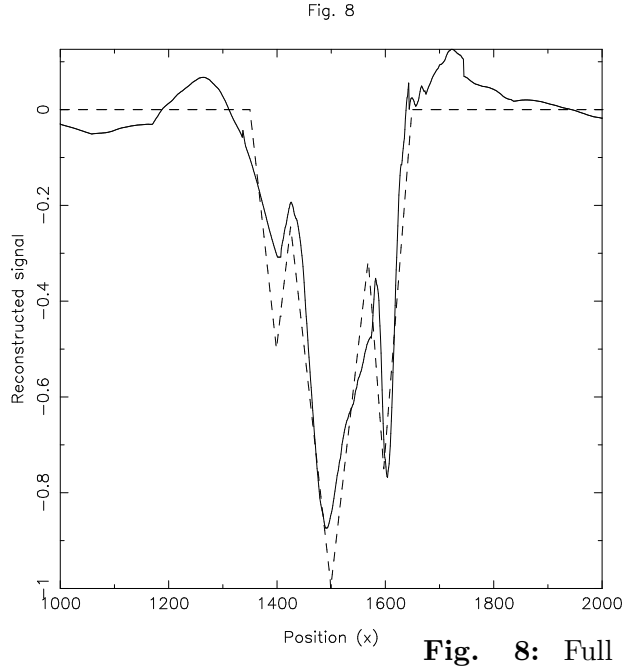



**Fig. 8:** Full line: the signal reconstructed from the coefficients of Fig. 5c. Dashed line: the initial noiseless signal.

This reconstructed signal is typical of ones that we obtained from several noise simulations. No major discrepancies occurs in the reconstructed signals when different noise realizations are performed. Only the depths of the different peaks varies a little but not significantly. When looking at fig. 8 one can remark that the left peak is less completely reconstructed than the central and right ones. This is due to the presence of the larger and deeper central peak which spoils the detection of this smaller and shallower peak. In fact, there are no coefficients selected by the two thresholding that correspond to that peak. But the presence of the peak modifies

the value of the coefficients in the vicinity of the central peak, where we incorporate coefficients to eliminate reconstruction biases. Hence the small branch at scale 7 in Fig. 5b.

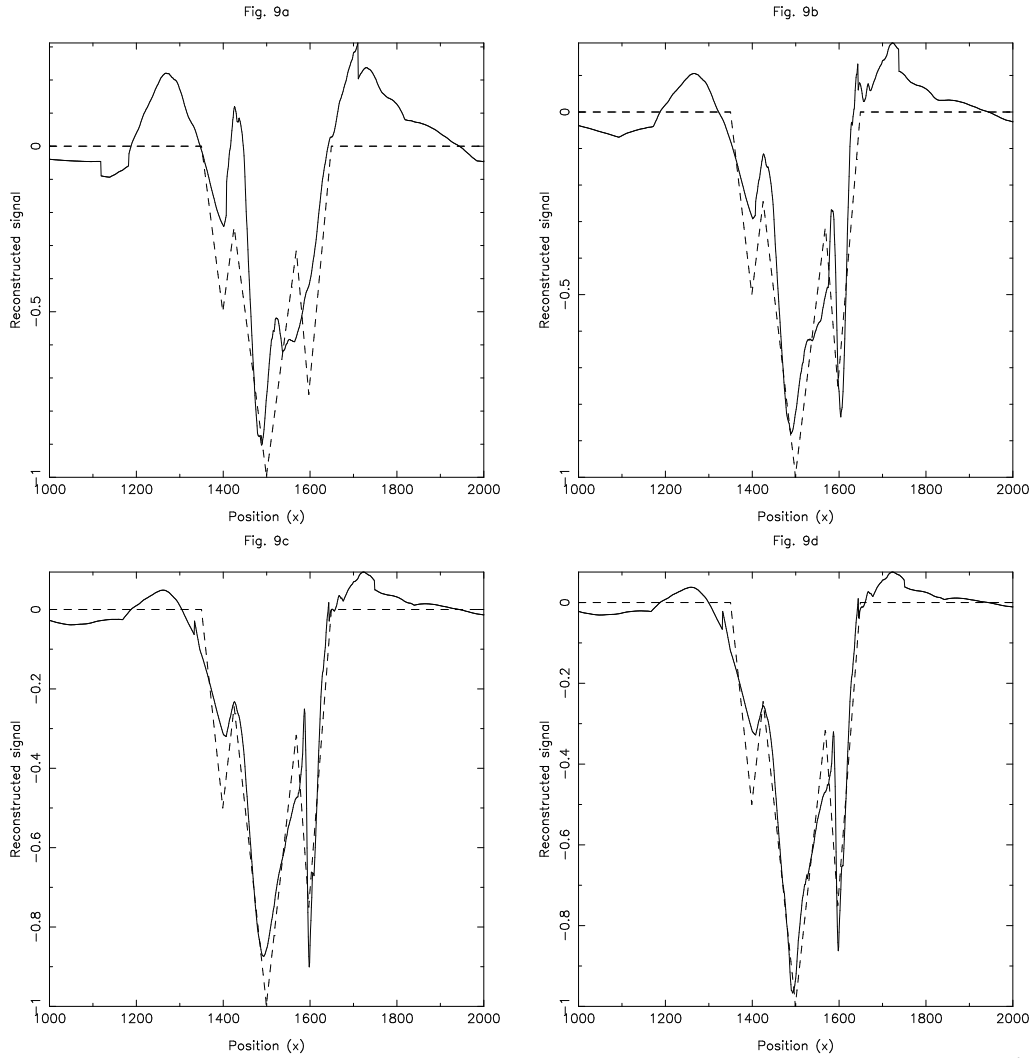



**Fig. 9:** a) Reconstruction (full line) of a signal similar to that of Fig. 4a, with a signal to noise ratio of 0.4. The second threshold $T_{sk}$ is chosen with the procedure described in section 3.3. The dashed line shows the initial signal before adding the noise. b) Same as Fig. 8a with signal to noise ratio of 0.8. c) Same as Fig. 8a with signal to noise ratio of 1.6. d) Same as Fig. 8a with signal to noise ratio of 2.0.

In addition to the shape of the structure, one can also gain some quantitative information.

The depths (or amplitude) of the different peaks and their respective width can be estimated from the reconstructed signal. The estimated widths of the three detected peaks have been found equal to 54, 129 and 38 from left to right which are good estimations of the real values: about 50, 132 and 57 respectively. These widths have been estimated at the base of each peak. The size of the global structure, measured at the base of the left peak, is found to be 249 to be compared to 256 for the real structure. In what concerns the depths we find respectively 0.11, 0.41 and 0.51 to be compared with 0.25, 0.43 and 0.68. The depth of the overall structure, estimated relative to the mean position of the plateau on each side, is 0.88 (true value: 1.0).

In order to have an idea of how the reconstruction process is influenced by the signal to noise ratio, we show in Figs. 9a-d the reconstructed signals for $S/N = 0.4$, 0.8, 1.6 and 2.0 respectively. Obviously, for very small $S/N$ the algorithm is not able to resolve the internal details of the structure, but it is still able to find it, with the rather good estimate of the size. It appears from these simulations that this method gives satisfying results on signals with signal-to-noise ratio greater than about 0.4.

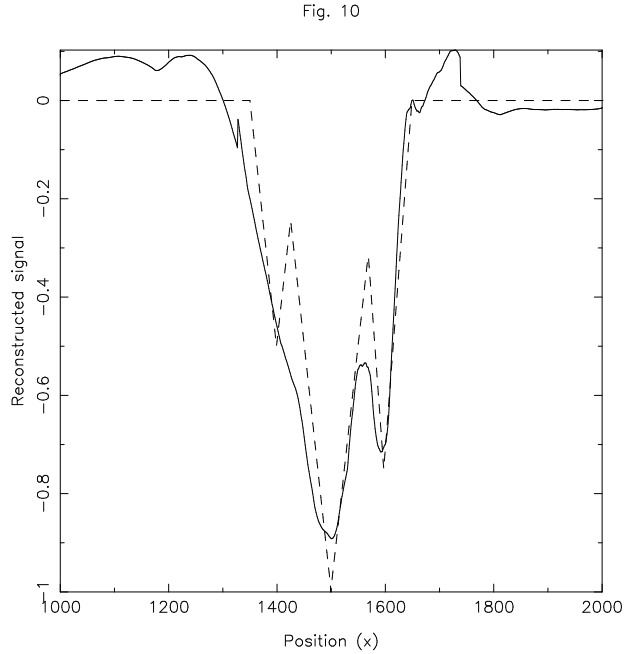



**Fig. 10:** Reconstruction (full line) of a signal with a signal to noise ratio of 1.2 and a different noise realization from that of Fig. 8. The worst case found in our study.

All the signals we showed up to now have been created with the same realization of the

noise, only changing its amplitude. This noise realization was typical of the ensemble. Fig. 10 shows the worst case we have found among different generations of noise and pseudo–random signal for a signal to noise ratio of 1.2. This was the only case nearly as bad, the other ones being closely similar to the reconstruction displayed in fig. 8

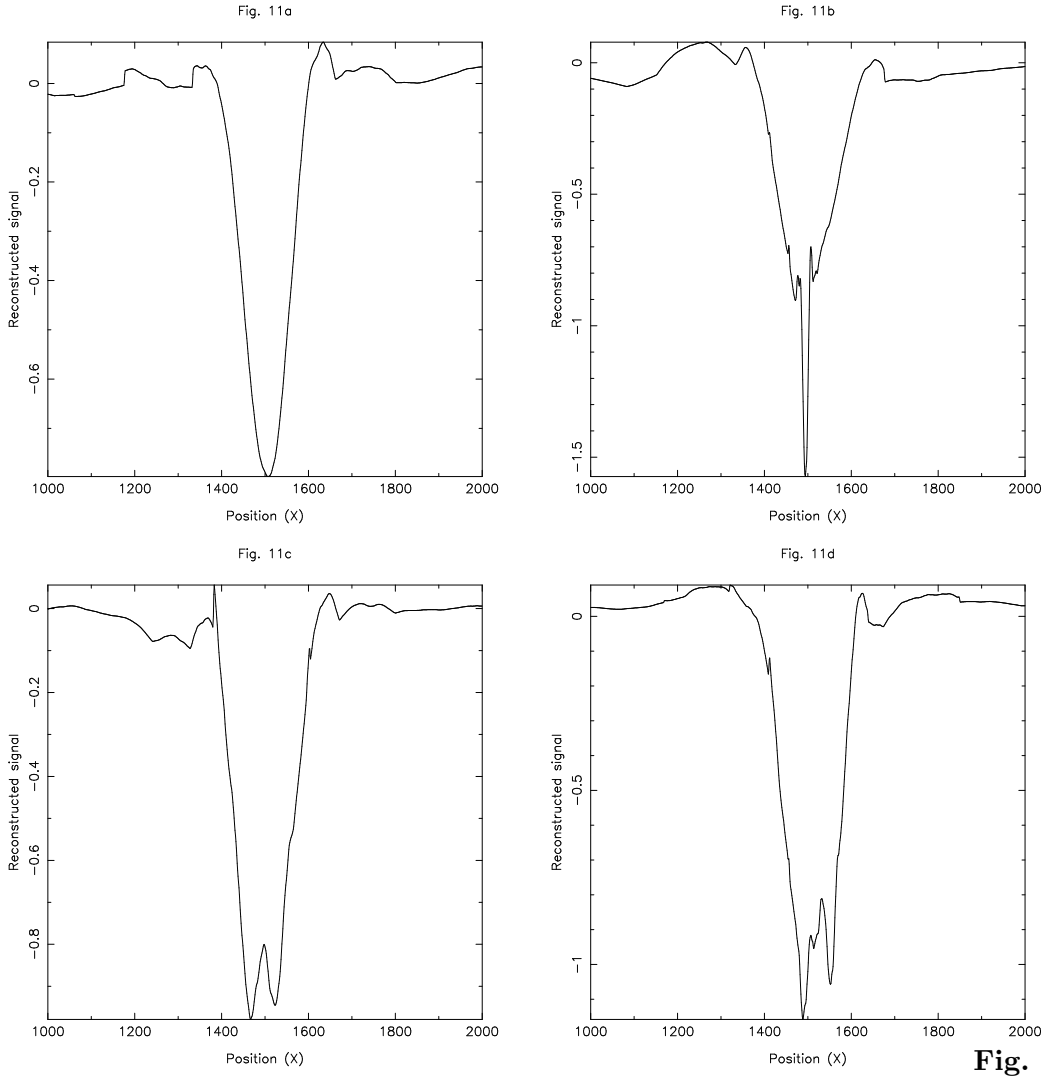



**Fig. 11:** The four typical kinds of the reconstruction of a single gaussian peak with a signal-to-noise ratio of 1.2. a) smooth mono peak ; b) mono-peak with a very sharp and strong feature ; c) double peak ; d) triple peak.

We have also applied our method to a signal with three Gaussian peaks to test the effect of the shape of the peaks. No significant effect of the structure shape has been found. We have also simulated a single Gaussian peak in order to estimate the percentage and the importance (in a qualitative sense) of false detections of weak features. For this purpose we have generated 100 mono-peak signals with different realizations of the noise, but keeping the same signal-to-noise ratio fixed at 1.2. The structure has been automatically reconstructed following the above described procedure. Four types of reconstructed signals appeared with the following percentages : 1) smooth mono peak (59 %) ; 2) mono-peak with a very sharp and strong feature (12 %) on the side. This kind of feature can unambiguously be interpreted as being due to the presence of a strong localized noise ; 3) a double peak (21 %) 4) a triple peak (8 %). These two later cases are to be qualitatively compared to the simulations of the three peak structures (fig. 8 ). The four typical kinds of reconstructed peaks are displayed in figs. 11 (a to d respectively for cases 1 to 4)

Summarizing the results of the different reconstructions, we have been able to : (i) recover more than 70 % of mono peaks, (ii) differentiate the false detections of two and three peaks from the real ones (see figs. 11 c-d and 9 a-d), (iii) to detect and reconstruct the three peaks structures even with low signal-to-noise ratio. This have led us to apply the whole procedure, first in the debated, as previously mentioned, Encke gap of the Saturn A ring.

## 3.4 Data of Voyager in the Encke gap in Saturn's A ring

In the previous section, we have shown the possibilities and also the limitations of the method. Particularly, we have seen that this method tends to eliminate the small peaks rather than adding peaks due to the noise. We will then apply our technique to the PPS Voyager data in order to examine the detailed structure of the ringlet found in the Encke division. Our aim is to give some new and independent information which will validate or invalidate the theory of Spahn (1992).

The PPS data of Voyager 2 have been obtained by counting photons coming from a star of reference and occulted by the different parts of Saturn's rings. Fig. 12 shows the analyzed data in the Encke gap of the Saturn A ring. Since we are counting the number of photons coming through the rings, the smaller the density of ring material, the larger the number plotted. Hence a decrease in the level of the signal (what we have called a depletion) corresponds to an increase of the ring density. The data used here are raw data. This choice has been driven by the wish to avoid any preliminary smoothing as it is done when dealing with the optical depth for which a background must be estimated (see Esposito *et al.* 1987). The overall goal is merely to assess the efficiency of this method.

Figs. 13 a-b display the reconstructed signals with $T_{se}(a_i) = 5\sigma_{S_{ps}}(a_i)$ and $T_{sk}(a_i)/\sigma_{S_{ps}}(a_i) =$

1.3, and 0.8 respectively. The structures have been automatically reconstructed for these two values of $T_{sk}$, following the criterion we have imposed for the reconstruction. In fig. 13a two peaks are well identified and there is also the trace of a third one (on the right side of the central peak). This later appears when reconstructing from the weaker $T_{sk}$. Its level of confidence, versus noise fluctuations, is therefore smaller and one can see in fig. 13b how noise is added on the left peak at this level of detection. This means that that this third smallest peak is very close to the level of noise.

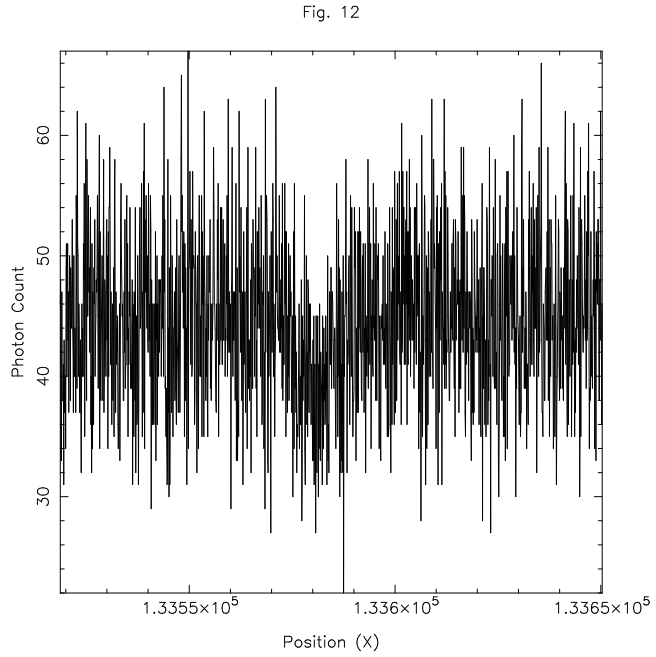



**Fig. 12:** Voyager 2 PPS data of the Encke gap region in Saturn's rings. The x–axis gives the radial distance from the center of Saturn in kilometers, the y–axis gives the photon counts.

In order to get quantitative information from the reconstructed signal we consider both figures. The global structure is located at 133580.7 km from the center of Saturn and has a typical width of 14.2 km. This value is estimated in fig 13a just above the right peak detected in fig 13b. The left peak is centered at 133576.8 km and has a width of 3.3 km (from fig. 13a). The central one is 8.75 km wide and finally the right one is 3.3 km wide, located at 133586.0 km from the center of Saturn. From the the top of each peak we estimate the depths at 2.2, 2.8, 1.5 units respectively, for the left central and right peaks (keeping in mind that we deal with photon counts, and so, with integers). Finally from the mean position of the plateau on

each outer side of the structure the depth of the global feature is measured at 6.5 units.

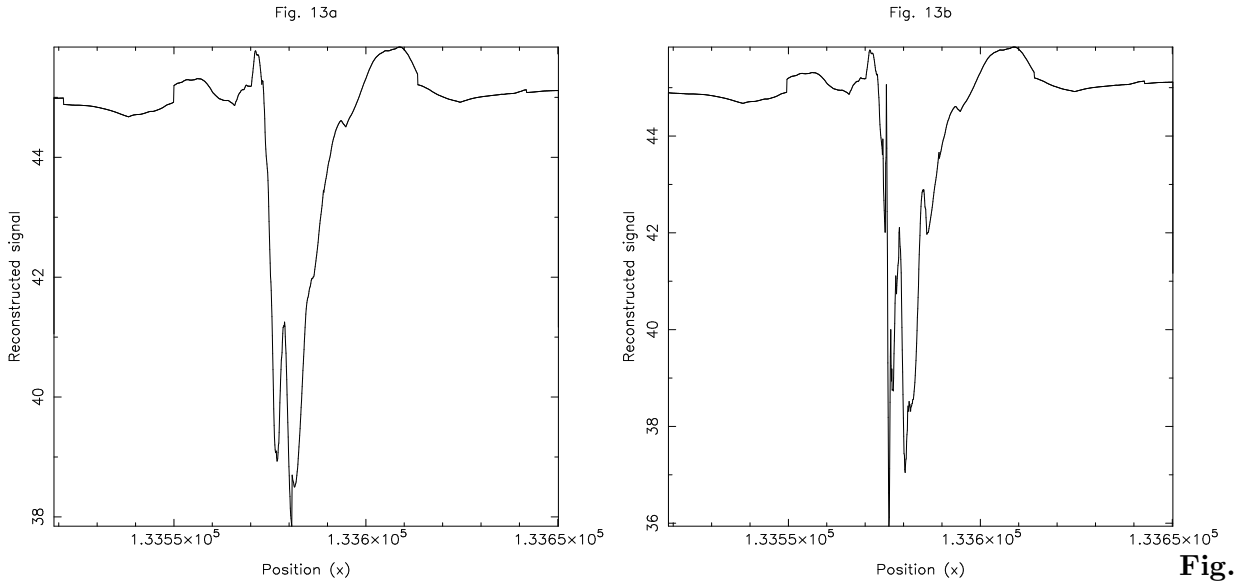



**Fig. 13:** a) Reconstruction of the signal of Fig. 12 with $T_{sk}(a_i) = 1.3\sigma_{S_{ps}}(a_i)$. b) with $T_{sk}(a_i) = 0.8\sigma_{S_{ps}}(a_i)$.

In order to make some conclusions about the physical parameters of satellite Pan, which is possibly responsible to the ringlet, we plan to apply our method on the results of numerical simulations modelling the interactions between the moonlet and the ring's particles.

In the theory of Spahn, the azimuthally integrated radial density profile was symmetric about the position of the moonlet. But the real ringlet has no reason to be symmetric due to the asymmetry in the position of outer satellites. Moreover, recent numerical simulations show that the radial density profile varies from three to two peaks depending on the azimuth relative to the position of the moonlet.

# 4 Comparison with other methods

In order to make a rough comparison between our method of structure detection and some already existing techniques, we present in fig. 14 the result of another analysis. This later is a classical one which relies on the use of a low pass filter applied to the Fourier Transform of the signal. The low pass filter is a smooth window with a cut-off frequency from which the higher frequencies are removed from the spectrum of the signal. Theses high frequency components,

are mostly noise, and do not contribute to the signal as reconstructed by the inverse Fourier transform. Fig. 14a and 14b display respectively the reconstructed signal for two different widths of the low pass filter.

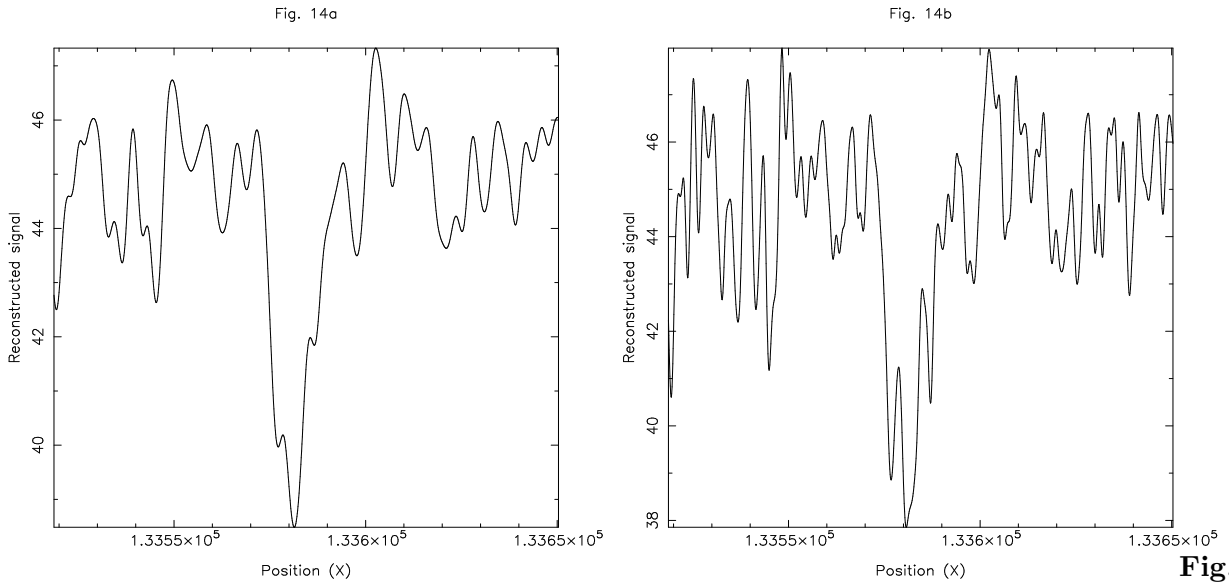



**Fig. 14:** The same data as in fig.12 analyzed by means of the low pass filter Fourier analysis for 2 different value of the cut-off frequency.

In this method, the noise is globally decreased instead of locally. Hence, it clearly appears that, even if a three peak structure seems to be extracted from the signal the amplitude of the fine features is of the same order of the fluctuations due to the noise. Moreover there is no way to quantify the level of detection of these features and hence to distinguish them quantitatively from the noise.

# 5 Conclusion

We now have a powerful signal processing tool which is able to detect structures even in a strongly noisy signal and which allows one to quantify the level of significance of the detected structures versus the noise fluctuations. Moreover the quality of the reconstructed signal and the information available through wavelet analyses performed by different wavelets lead to a fairly good estimation of the physical parameters of the detected structures. We have been able to remove the noise from the raw data of the Voyager 2/PPS experiment in the region of the

Encke gap. The Encke gap ringlet exhibits, with a high level of confidence, a two peak profile. According to Spahn *et al.* (1992), this points to a nearly circular orbit of Pan. However, the reconstructed profile shows only qualitative agreement with their model. They obtained very sharp and narrow peaks at the edges, while the right edge of the reconstructed structure is rather smooth (Fig. 13a). On the other hand, taking the size of Pan found by Showalter (1991) of about $20km$ in diameter (icy composition assumed), the satellite should be able to confine a ringlet of width $W > 30km$. But we see only $W \approx 14km$. However, Spahn *et al.* .(1992) showed that the higher the eccentricity of the orbit of the embedded moonlet – here Pan – the smaller the width $W$ of the ringlet (Eq. 8 and Fig. 2 in Spahn *et al.* , 1992). According to their findings the satellite Pan should have an eccentricity of the order of its Hill-scale: $e_{Pan} \approx 10^{-4}$. If we assume that this is correct, then the ringlet should show a three peak structure similar to the artificial signal presented in Fig. 2a. However, there is a hint for the existence of a third peak (albeit of very small amplitude), but with a lower level of confidence.

Recent simulations (Spahn *et al.* , 1993) point at deficiencies of the model of Spahn *et al.* (1992). These new investigations, where the assumption of axisymmetry has been dropped, show that the previous results are correct only if averaged azimuthally, while the actual radial density profile can depend on the azimuthal longitude in the case of eccentric moonlet orbits. So it might happen that one observes a single, or double or also a triple–peaked ringlet profile for the same model-run. Thus the existence of the central and a companion peak cannot be excluded, where the third peak appears even at another azimuthal longitude.

However this first application on the Encke gap ringlet has been performed with raw data (photon count). Since all the theoretical models use either surface mass density or optical depth, some parameters cannot be reached by analyzing raw data. We need to transform these data in optical depth profile in order to obtain the values of these parameters. We are now performing analyses in a systematic way on the Voyager data of Saturn and Uranus rings. The results will be presented in a forthcoming paper.

# Appendix A

We describe here the algorithm to add coefficients in order to avoid reconstruction biases. As previously mentioned, a depletion in the signal gives negative coefficients at its location and positive coefficients on each side (see section 2.2 and Fig. 2b). If we don't keep the positive coefficients, the reconstruction will give bumps on each side of the reconstructed depletion. Therefore, we want to keep these coefficients. The algorithm is as follows. For each scale, do:

i find the sets of contiguous coefficients selected by the thresholding,

ii on each side of each contiguous set (neighboring in both the scale and location directions), determine a range the size of which is the extent of the wavelet at that scale,

iii in each range, keep the coefficients that are greater [resp. smaller] than the mean for the detection of a depletion [resp. bump],

iv add the coefficients that are smaller [resp. greater] than the mean, but not selected by the thresholding and lie between a set of contiguous selected coefficients and the coefficients added in step (iii) (this avoids discontinuities on the edges of the reconstructed structures).

# Acknowledgements

We would like to thank Larry Esposito, Karen Simmons, and Joshua Colwell for providing us with the Voyager-2 PPS-data of planetary rings. We are also very grateful to Jeffrey D. Scargle for his very helpful and meaningful comments. We finally thank Michel Hénon for his suggestions

[1] Bendjoya, Ph., Slezak, E. & Froeschlé, Cl., 1991. *The wavelet transform: a new tool for asteroid family determination.* Astron. Astrophys., in press, **251**, pp. 312–330.

[2] Borderies, N., Goldreich, P. & Tremaine, S., 1989. *The formation of sharp edges in planetary rings by nearby satellites.* Icarus **80**, pp. 344–360.

[3] Brophy, T.G. & Rosen, P.A., 1992. *Density waves in Saturn's rings probed by radio and optical occultation: Observational tests of theory.* Icarus **99**, pp. 448–467.

[4] Daubechies, I., 1988. *Orthogonal bases of compactly supported wavelets.* Comm. Pure Appl. Math. **41**, pp. 909–996.

[5] Esposito, L.W, O'Callahan, M., Simmons, K.E., Hord, C.W., West, R.A., Lane, A.L., Pomphrey, R.B., Coffeen & D.L., Sato, M., 1983*a*. *Voyager photopolarimeter stellar occultation of Saturn's rings.* J. Geophys. Res. **88**, pp. 8643–8649.

[6] Esposito, L.W, O'Callahan, M. & West, R.A., 1983*b*. *The structure of Saturn's rings: implications from the Voyager stellar occultations.* Icarus **56**, pp. 439–452.

[7] Esposito, L.W., Harris, C.C & Simmons, K.E., 1987. *Features in Saturn's rings.* The Astroph. J. Supp. **63**, pp. 749–770.

[8] Grossmann, A., Kronland-Martinet, R. & Morlet, J., 1989. *Reading and understanding continuous wavelet transform.* in *Wavelets, Time-Frequency Methods in Phase Space,* eds. Combes J.M., Grossmann A. and Tchamitchian Ph., Springer-Verlag, pp. 2–20.

[9] Lane, A.L., Hord, C.W., West, R.A., Esposito, L.W., Coffeen, D.L., Sato, M., Simmons, K.E., Pomphrey, R.B. & Morris, R.B., 1982. *Photopolarimetry from Voyager 2: preliminary results on Saturn, Titan and the rings.* Science **215**, pp. 537–543.

[10] Mallat, S.G., 1989. *A Theory for Multiresolution Signal Decomposition: The Wavelet Representation.* IEEE Transactions on Pattern Analysis and Machine Intelligence **11**, 7, pp. 674–693.

[11] Meyer, Y., 1989. *Orthogonal wavelets.* in *Wavelets, Time-Frequency Methods in Phase Space,* eds. Combes J.M., Grossmann A. and Tchamitchian Ph., Springer-Verlag, pp. 21–37.

[12] Petit, J.M. & Hénon, M., 1988. *A numerical simulation of planetary rings III. Mass segregation, ring confinement, and gap formation.* Astron. Astrophys. **199**, pp. 343–356.

[13] Showalter, M., Cuzzi, J.N., Marouf, E.A. & Esposito, L.W., 1986. *Satellite 'wakes' and the orbit of the Encke gap moonlet.* Icarus **66**, pp. 297–323.

[14] Showalter, M., 1991. *Visual detection of 1981 S13, Saturn's eighteenth satellite, and its role in the Encke gap.* Nature **351**, pp. 709–713.

[15] Spahn, F., Saar, A., Schmidt, S. & Schwarz, U., 1992. *The influence of various moonlets on the optical depth profile in planetary rings.* Icarus **100**, pp. 143–153.

[16] Spahn, F. & Scholl, H., 1993. in preparation

**[36]**

# "A new insight in Uranus rings: a wavelet analysis of the Voyager 2 data."

PETIT, J-M. *&* BENDJOYA, Ph.

in *Proceedings of the Pacific Astronomical Society* (1994)

# A new insight in Uranus rings: a wavelet analysis of the Voyager 2 data

J.M. Petit

Ph. Bendjoya

*Observatoire de la Côte d'Azur BP 229 06304 Nice cedex 4 France*

**Abstract.** A new signal processing analysis, based on the wavelet transform has been developed. It allows the detection and the reconstruction of fine structures in a very noisy signal. It removes the noise and gives a quantified level of detection of the structures against chance fluctuations. This powerful method has been applied on the PPS Voyager 2 data on the Uranus rings. A preliminary catalog of structures found in the $\sigma$ Sagitarii occultation experiment, is proposed here.

## 1. Introduction

The data obtained in the PPS (recording of brightness of star $\delta$ Sco using a photopolarimeter) experiment of Voyager 2 have revealed a whole hierarchy of structures in planetary rings, at all scales and with a wide range of amplitudes. In order to determine physical parameters of the rings and embedded (or shepherding) satellites we have to be able to determine precisely the density profile of features with a signal to noise ratio of the order of 1. In this paper, we briefly describe a new method of signal processing based on the wavelet transform. The method has be fully described in Bendjoya *et al.* 1993. We will recall here the main philosophy of the analysis and we will propose, as an example of application, an inventory of the fine structures detected in the $\sigma$ Sagitarii occultation experiment of Voyager 2.

## 2. Philosophy of the method

The signal processing tool presented here is particularly well suited for the detection and reconstruction of structures (bumps or depletions) embedded in a very noisy signal. Simulations have been performed and have shown that the method remains powerful for signals with a signal-to-noise ratio of order of one. The method gives a reconstruction of the underlying structures precise enough to derive physical parameters with a relative uncertainty of order of 10 %. Moreover the detection of these structures can be quantified against chance fluctuations by giving a level of significance.

The method relies on the following model: a signal is a sum of details at different scales i.e. different characteristic sizes. A natural assumption is that the significant information is not equivalently displayed at all scales. For example,

the small scales are generally related to the noise while a structure lives on a set of larger scales, determined by its shape, depth and width.

The idea is then to decompose the signal onto a peculiar band pass filter basis which allows one to separate the details at the different scales. This basis is called the wavelet basis and the decomposition of the signal onto this basis is called the wavelet transform. The wavelet transform acts like a local mathematical zoom which allows detection and reconstruction of a structure without noise by taking into account only the details corresponding to that structure.

A signal is defined with a cutoff frequency $f$: the sampling step. The procedure of the wavelet transform consists in applying a low pass filter with a cutoff frequency of $f/2$. We thus obtain a smoothed function of the signal. A "good" choice (Daubechie, Mallat) of the low pass filter allows the derivation of a band pass filter just by taking the difference point by point between the signal and the smoothed function. This band pass filter extracts the details contained in the frequency band $[f/2, f]$. By iterating the same filtering on the smoothed function, we get on one hand a new smoothed function with a cutoff of $f/4$ and on the other hand the details we lost when passing from a resolution to the next one. Fig. 1 illustrates the whole procedure.

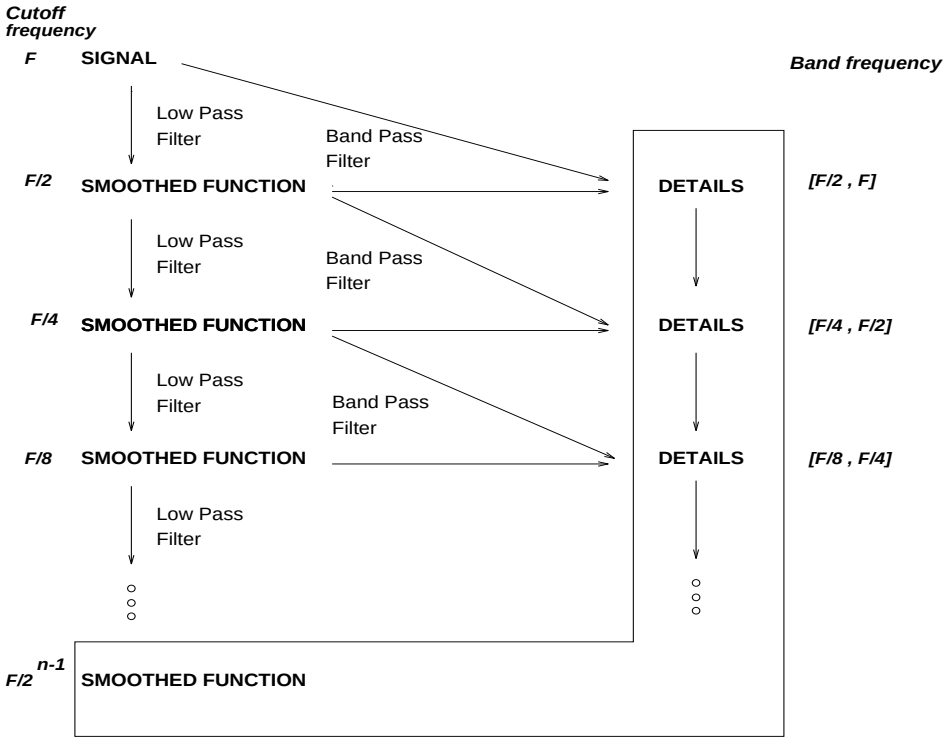



Figure 1. Scheme of the wavelet transform algorithm.

The band pass filter bank provides a basis onto which the signal is decomposed. This basis is the wavelet basis and is chosen so that each element is a

zero mean function. Since the successive steps are obtained by difference we can reconstruct the signal by simply adding the last smoothed function with all the details. This is the scheme of a multi resolution analysis. Fig. 2 shows concretely how it works: starting from the signal one can see the successive smoothed functions and the details at different scales. Since the wavelet basis is composed by zero mean functions, the details (also known as the wavelet coefficients) have a zero mean.

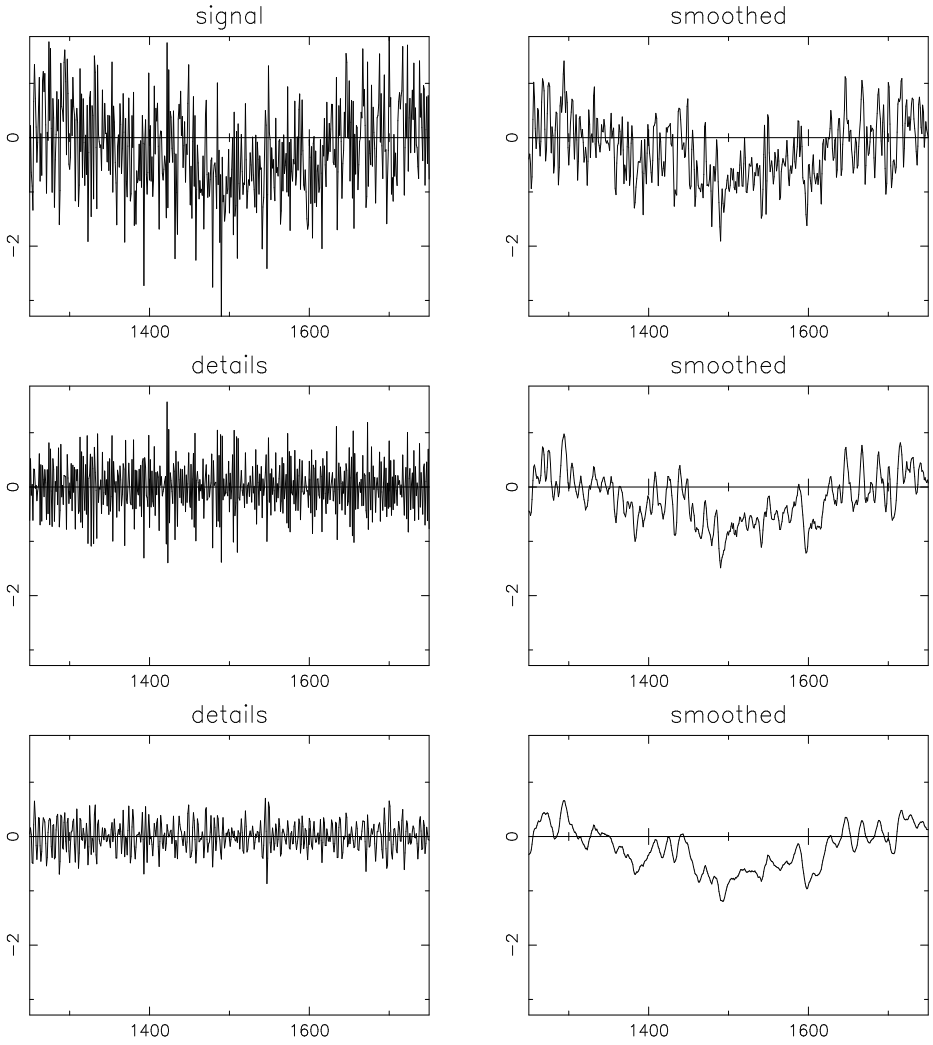



Figure 2. Left: the original signal sampled at frequency $f$ (top), the details in the frequency range $[f/2; f]$ (middle) and the details in the frequency range $[f/4; f/2]$ (bottom). Right: the successive smoothed signals at frequencies $f/2$ (top), $f/4$ (middle) and $f/8$ (bottom).

Since the noise is generally associated with the smallest scales the idea is to reconstruct the signal by adding the last smoothed function and a set of selected coefficients discarding those corresponding to the noise. In order to select these coefficients, we do a thresholding in the wavelet plane. The thresholding of the

wavelet coefficients is done by comparing for each scale the wavelet coefficients derived from the signal with those derived from a pseudo random signal. This later is built randomly from the signal following the same histogram of values. In such a way the noise is not derived from a model and no assumption is made on it. Due to the shape of the wavelet (zero mean function) a bump of the size of the studied scale will give high positive coefficients. Similarly, a depletion will give large amplitude negative coefficients. the threshold is computed from the distribution of the wavelet coefficients of the pseudo random signal. When looking for bumps, the threshold is the mean value of the considered scale plus a given number $n$ times the r.m.s. When looking for depletions, the threshold is the mean value of the considered scale minus a given number $n$ times the r.m.s. Among the wavelet coefficients computed from the real signal, only those greater (in absolute value) than the threshold will be kept for the reconstruction. By construction the pseudo random signal and the signal itself are similar at the small scales (scales of the noise). A sufficiently high threshold ($n$ of order of 4) will discard the wavelet coefficients of these smallest scales and hence the noise. On the other hand at larger scales the pseudo random signal will not provide high (in absolute value) coefficients but the true signal will. If there is a structure, it will be conserved for the reconstruction. Fig. 3 displays a scheme of the thresholding procedure.

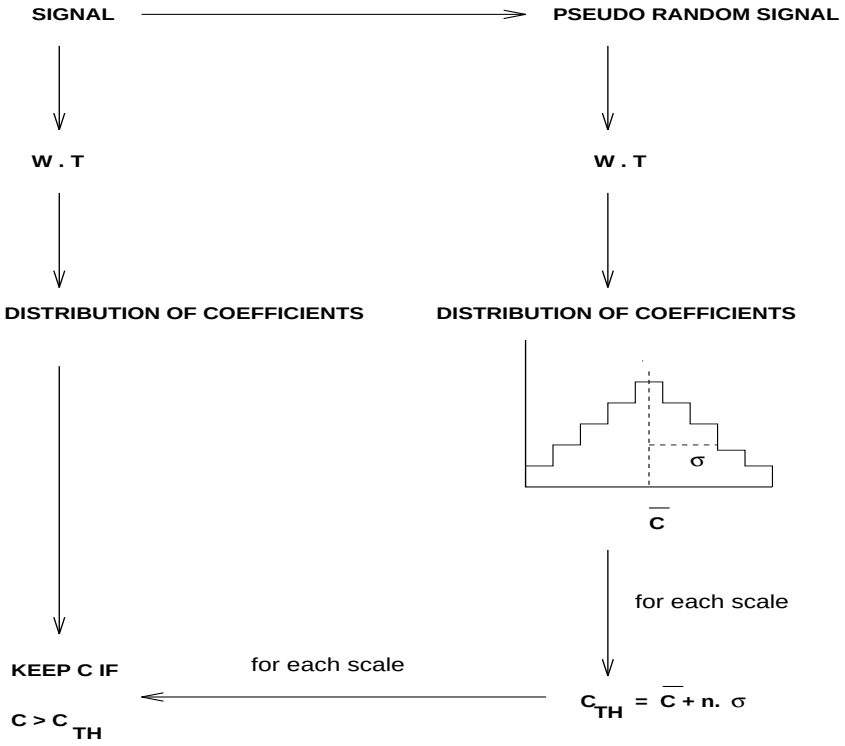



Figure 3. Scheme of the thresholding procedure.

In fact the reconstruction is made after a second thresholding more permissive than the first, whose role is to remove the noise. This second thresholding

(typically $n$ between 1 and 2) allows to add locally details to the structures detected after the first thresholding. Finally the level of detection (or of confidence against chance fluctuations) is given by the second thresholding for the finest details and by the first thresholding for the structure as a whole.

This is the main difference with the more classical filtering methods such as Fourier Transform or binning which perform a global filtering and are unable to add details (small scales or high frequency) only where it is needed.

## 3. Results

### 3.1. A simulated structure

In order to show the efficiency of the method we have built an artificial structure with a three peak shape (fig. 4a) and plunged it in a gaussian noise in order to get a signal to noise ratio ($S/N$) of 1.2 (fig. 4b). We define the signal to noise ratio ($S/N$) as the ratio between the amplitude of the structure (here, 1) and the standard deviation of the gaussian noise we add. This definition gives a local meaning to $S/N$. In our case, we are not interested in large scale structures but rather in local ones. The signal shown in fig. 4b has a signal to noise ratio of 1.2, since the standard deviation is 0.833 and the amplitude of the noiseless signal is 1. Fig. 4c displays the reconstructed signal (solid line) superimposed to the original structure (dashed line).

The choice of such a shape for the simulated structure has been suggested by the models of gravitational interactions. The rings are known to be maintained by the gravitational interaction with satellites (Goldreich and Tremaine 1978, Shu 1984, Shu *et al.* 1985 a,b). The effect of an embedded satellite is to clear a gap (provided its mass is large enough) and retain a faint ringlet on the same orbit as itself. The profile of such a ringlet can have two or three peaks (Spahn *et al.* 1992).

The first thresholding of the wavelet coefficients is 4 times the r.m.s of coefficient distribution from the pseudo random signal. We recall that this first thresholding remove the noise and detect the structure. The second thresholding has been automatically chosen by the procedure equal to 1.5. This second pass add locally details to the detected structure. The value of this second thresholding corresponds to a level of significance against chance fluctuation of: 93.3%. One can see the great ability of the method to remove noise, to detect and to reconstruct precisely fine structures. The two bumps on each side of the reconstructed three peak depletion are due to the property of the wavelets to be zero mean functions. This phenomenon is all the more so marked since the underlying structure is thinner.

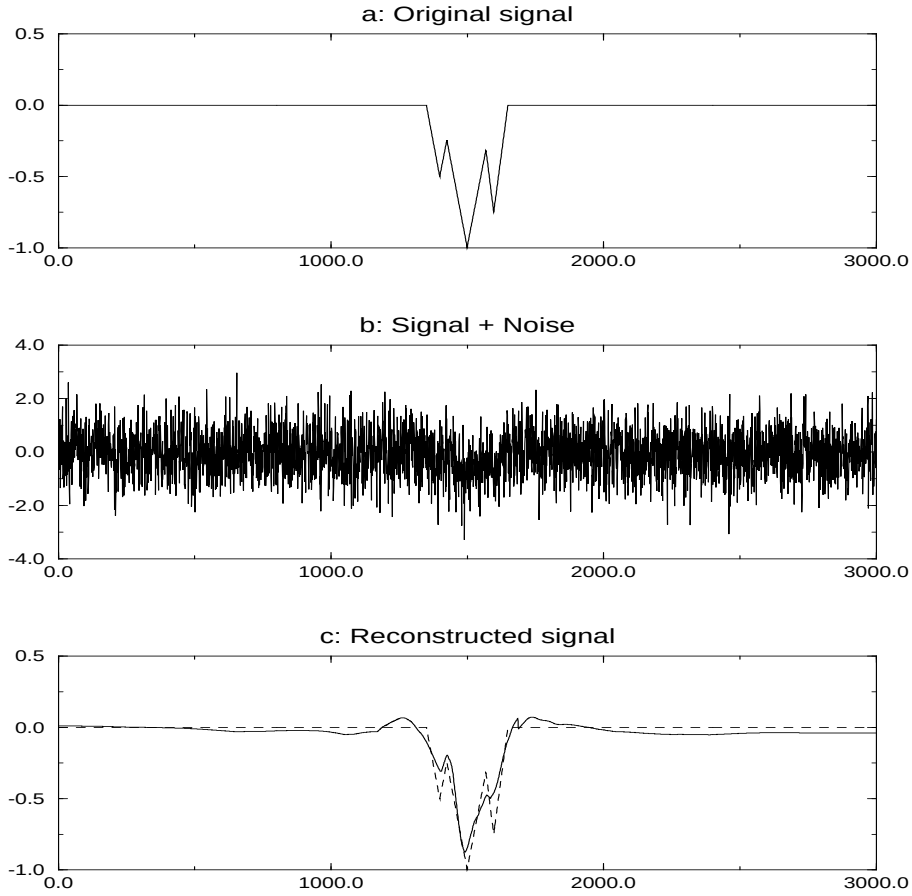



Figure 4. a) the simulated three peaks structure; b) the analysed signal with $S/N = 1.2$; c) the superposition of the reconstructed signal and the initial structure (dashed line).

The wavelet analysis method has been performed for different shapes of the three–peak structure, different signal to noise ratios and different realizations of the noise (see Bendjoya *et al.* 1993) leading to very good detections and reconstructions.

### 3.2. The Voyager 2 PPS data

The Voyager 2 PPS data analysed here were obtained from the occultation of the star $\sigma$ *Sagitarii* by the Uranus' rings. During the occultations the spacecraft passed behind the $\epsilon$ and $\delta$ rings in a nearly tangential occultation as viewed from Voyager as it approached Uranus. Every 10 msec the PPS recorded the brightness of the occulted star, with successive data points separated by 10 m

in distance from Uranus' center for $\epsilon$ ring and 3.8 m for $\delta$ ring. These data have been provided to us by Mark Showalter from the PDS.

We have made the analysis on subsets of 32,000 data points with an overlapping of 16,000 points. This overlapping enables us to avoid the edge effects such as false detection or missing detection. The detection and reconstruction have been performed in a fully automated manner. Local analyses have been performed around the detected structures. The data are provided with the number of the record and the photon counts. The conversion to the radii (from the center of the planet) is made assuming an uninclined equatorial ring (French *et al.* 1988).

### 3.3. Inventory of the Uranus rings structures in the $\sigma$ Sagitarii occultation experiment

Results are displayed in fig. 5 and in table 1. In this table the widths have been computed at the widest location of the reconstructed structure. We recall that all these reconstructed structures have been detected with a threshold of 4 $\sigma$. The thresholds given in column 4 of table 1 correspond to the second thresholding and give the level of confidence of the details added to the global structure.

We only present here the structures found both in ingress and egress. We will not discuss here the possible origin of the fine structures especially inside the $\epsilon$ and $\delta$ rings. A forthcoming paper will be devoted to a comparison between theoretical models and the detection of these fine details. Nevertheless one can obviously see how the noise is removed and how the quality of the reconstruction can bring clues to the modelling of the formation and the stability of the rings. Moreover the detection and reconstruction of these fine structures can be indicators of the presence of embedded moonlets.

Among the 1,400,000 data points, the analysis has provided the detection and reconstruction of the well known $\epsilon$ and $\delta$ rings plus three new ringlets. We have called these ringlets respectively $\mathrm{I}_{wv}$, $\mathrm{II}_{wv}$ and $\mathrm{III}_{wv}$. The "$_{wv}$" subscript stands for "wavelet" and its role is to avoid any confusion with possible previous classifications.

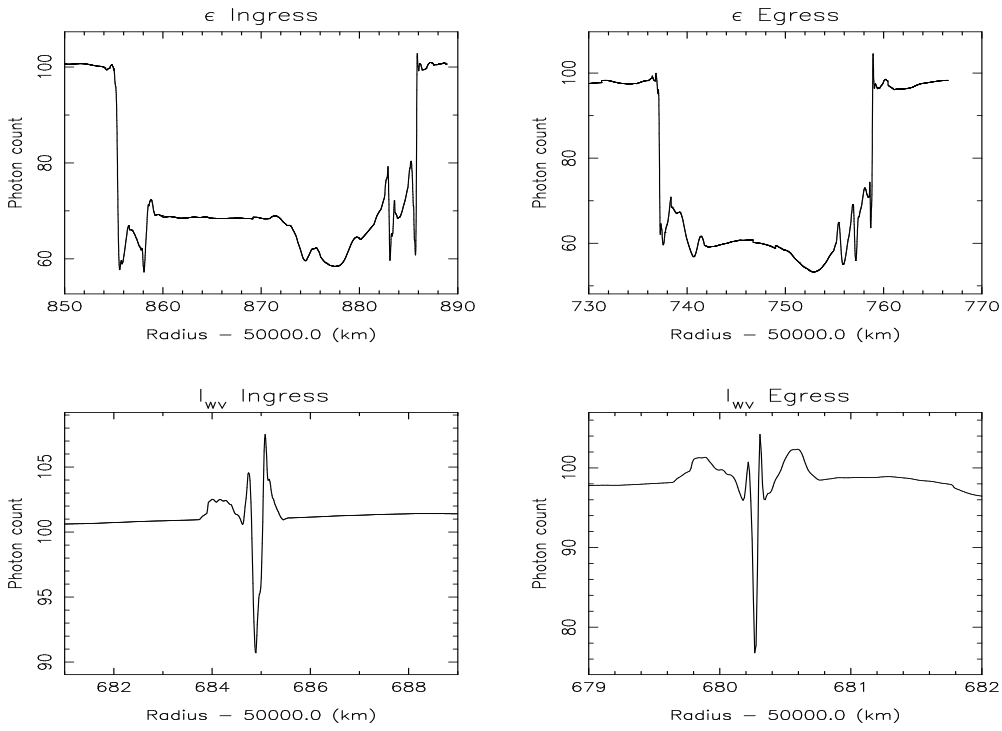
The ringlet $\mathrm{I}_{wv}$, located at 50685 km (ingress) and 50680 km (egress) from the center of Uranus could be associated with Ring I detected by Lane *et al.* (1986) from the occultation of $\beta$ Persei. However they found a width of 16 km where we get a width of the order of a few hundred meters. But they did not detect the ring in the $\sigma$ Sagitarii occultation, due to their method of analysis based on binning. This would be consistent with a narrowing of the thick part of the ring. In such a case our method allows detection of features not accessible to older methods.

The newly detected ringlet $\mathrm{II}_{wv}$, located at 50,023 km is quite wide since its wider part is almost 6 km at ingress and a little more than 2 km at egress. The very good agreement in position between ingress and egress is a factor in favor of the detection of a real feature, in addition to the rather large value of the second threshold which give a high confidence in the reality of the details of the profiles (94.5% at ingress and 98.6% at egress). As for the optical depth of this ringlet, it is comparable at egress to the optical depth of ring $\mathrm{I}_{wv}$ at ingress.

Although the last new detected ringlet, located at 47,977.4 km, is very narrow, it has been found both at ingress and egress and it correspond to a rather large number of data point since the precision in radial distance at that location is 1.4 m. And even if the level of confidence in all the fine details is not extremely high (90.3% and 91.9%), the structure itself has been detected at $4\sigma$, i.e. with a confidence of 99.997%.

Table 1. List of detected features

| Name | Midpoint range | Width (km) | Threshold |
|---|---|---|---|
| $\epsilon$ (I) | 50,870 | 30.7 | 1.4 |
| $\epsilon$ (E) | 50,748 | 21.7 | 1.4 |
| $I_{wv}$ (I) | 50,684.9 | 0.25 | 1.6 |
| $I_{wv}$ (E) | 50,680.3 | 0.075 | 1.5 |
| $II_{wv}$ (I) | 50,023.4 | 5.8 | 1.6 |
| $II_{wv}$ (E) | 50,023.1 | 2.26 | 2.2 |
| $\delta$ (I) | 48,300.5 | 7.6 | 1.4 |
| $\delta$ (E) | 48,301.7 | 7.6 | 1.1 |
| $III_{wv}$ (I) | 47,977.4 | 0.10 | 1.3 |
| $III_{wv}$ (E) | 47,972.7 | 0.11 | 1.4 |



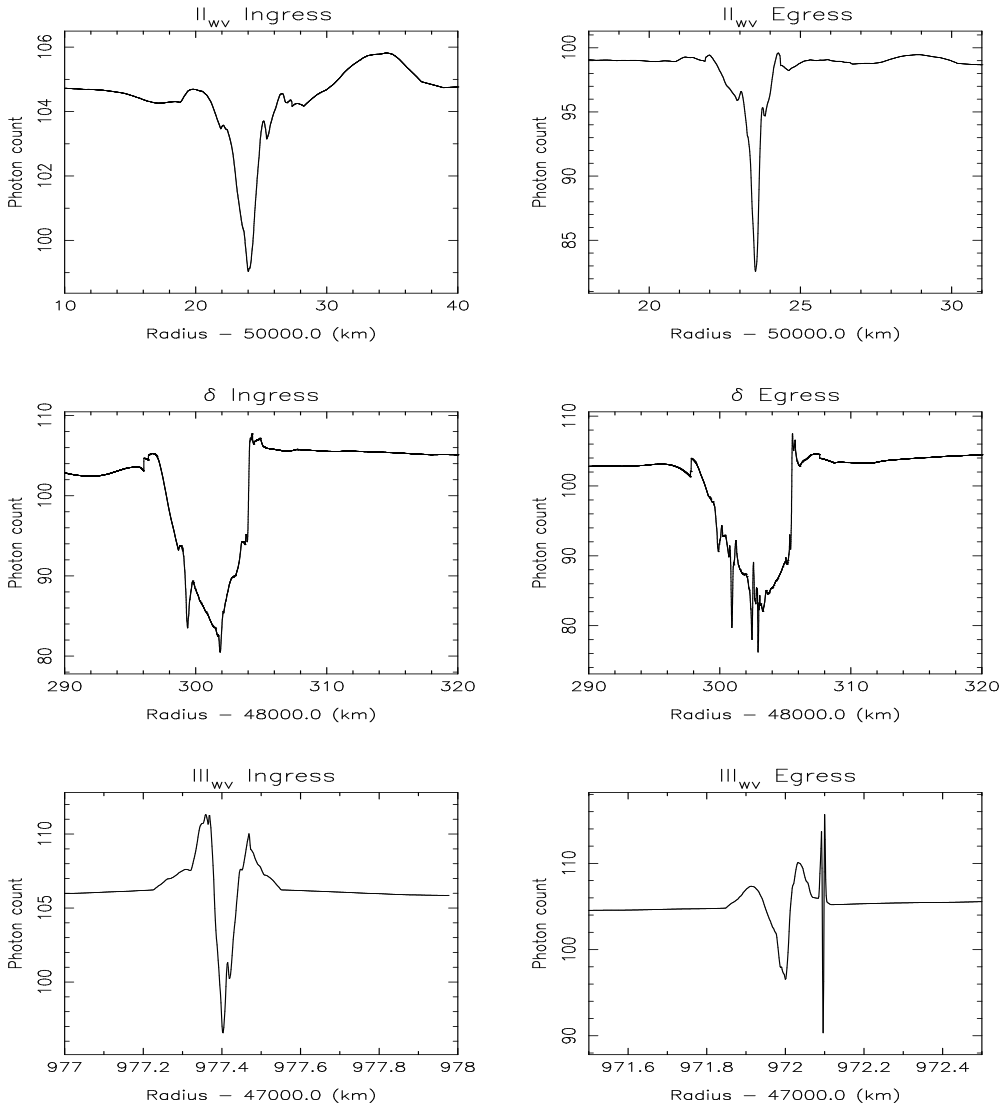



Figure 5. Reconstruction of structures detected in the $\sigma$ Sagitarii occultation data both for ingress (left) and egress (right).

## 4. Conclusion

We presented here a powerful signal processing tool for detection of faint structures in 1 dimension data set. It allowed us to analyse the $\sigma$ Sagitarii occultation observations performed by Voyager 2. In addition to the well known $\epsilon$ and $\delta$ rings, we also found 3 new probable rings. One of these (ring $\mathrm{I}_{wv}$) is probably the same as Ring I detected by Lane *et al.* (1986) in the $\beta$ Persei. In the $\epsilon$ ring, we are able to separate a plateau in the center of the ring and peaks on each side. By comparison, the standard methods exhibit a noisy reconstructed signal

which makes it difficult to separate real peaks. The detection of the new features may need to be confronted to the analysis of $\beta$ Persei occultation observations. This analysis, which we are undertaking now, will allow us to explore the Uranus ring system closer to the planet. The present analysis was performed on a data set extending only down to 47900 km from the center of Uranus. The analysis of both data set will allow us to compare observation with the predictions of theoretical models.

**Acknowledgments.** We would like to thank Mark Showalter for providing us with the data used in this study. We are grateful to Franck Spahn for continuous interesting discussions.

## References

Bendjoya, Ph., Petit, J. M., Spahn, F. 1993, *Icarus* 105, 385

Daubechie, I. 1988 *Comm. Pure Appl. Math.* , 41, 448

French, R. G., *et al.* 1988 *Icarus*, 73, 349

Goldreich, P. & Tremaine, S. 1978 *Icarus*, 34, 240

Lane, A. L. *et al.* 1986 *Science*, 233, 65

Mallat, S. G. 1989 in *IEEE Transactions on Pattern Analysis and Machine Intelligence*

Shu, F. H. 1984, in Planetary Rings, 513

Shu, F. H., Yuan C. & Lissauer J. J. 1985 ApJ291, 356 a

Shu, F. H.,Dones, L., Lissauer J. J., Yuan C., & Cuzzi 1985 ApJ299, 542 b

Spahn, F., Saar, A., Schmidt, S., & Schwarz, U. 1992 *Icarus*, 100, 143

## 4.2 Viscosité des anneaux planétaires

**[14]**

"Viscosity in Keplerian Disks: Steady-State Velocity Distribution and Non-Local Collision Effects."

PETIT, J-M. & GREENBERG, R.

*Icarus*, in press

# Viscosity in Keplerian Disks: Steady-State Velocity Distribution and Non-Local Collision Effects

J.M. Petit
C.N.R.S., Observatoire de Nice, France
&
Lunar and Planetary Laboratory
University of Arizona
Tucson, AZ 85721
petit@bach.obs-nice.fr

Richard Greenberg
Lunar and Planetary Laboratory
University of Arizona
Tucson, AZ 85721
Phone: 602-621-6940
FAX: 602-621-5133
greenberg@lpl.arizona.edu

**Abstract**

A kinetic model based on a numerical algorithm, rather than Boltzman's equation, yields the viscosity and velocity distribution for colliding, finite-size particles in a planetary ring. Results are similar to those of many-particle simulations, and show that non-local effects due to the finite size are dominant in many cases of interest. Only for small particles does the viscosity decrease with increasing optical thickness sufficiently for the standard ringlet instability model to apply. This numerical kinetic theory will allow study of multi-size particle distributions, as well as scattering due to gravitational interactions or alternative collision models.

# 1 Introduction

Angular momentum transport is a key process driving the large-scale structure of planetary rings. It determines the time-scale for radial spreading, which is remarkably short in the case of Saturn's rings, and the behavior of wave structures. Analytical studies of momentum transport generally follow traditional formulations of statistical mechanics, with modified versions of Boltzman's equation representing the dynamics in a keplerian swarm of particles (Goldreich and Tremaine 1978, Shu and Stewart 1985, Araki and Tremaine 1986). Such approaches often involve complicated mathematics that can obscure the underlying physical processes and make the modeling of critical collisional effects difficult to evaluate.

Greenberg (1988) took a more heuristic and physical approach which led to results similar to earlier analyses. Greenberg's approach was used by Ojakangas and Greenberg (1990) to examine momentum transport in the case of a disk or planetary ring with a radial surface density gradient. The general approach involves estimating the concurrent redistribution of velocity due to collisions and due to keplerian motion. The collisions tend to reduce orbital eccentricities due to the effect of impact energy loss on relative velocities; collisions may also tend to increase some eccentricities as scattering creates a more isotropic relative-velocity distribution relative to the elliptical distribution associated with swarms of orbiting particles. At the same time keplerian motion tends to drive an isotropic distribution toward a distribution with an excess of radial velocity components. Modeling these competing effects on velocities gives a steady-state velocity distribution, which in turn can be integrated to find momentum transport and its coefficient, the viscosity $\nu$.

Another approach to study of transport processes has been to construct numerical many-particle simulations in which the actual motion of individual particles is computed and tracked. Simulations by Wisdom and Tremaine (1988) and by Salo (1991) have provided considerable insight into the dynamics of such systems.

In computing viscosity, two components of momentum transport must be taken into account. One component is found by determining (to some level of approximation) the probability distribution of the velocity of particles' centers at a representative point in a model planetary ring. Then the momentum transport is evaluated by integrating over the velocity distribution, weighted by the probability, the angular momentum carried as particles move radially through the test locality. Because this calculation ignores transport by interactions between particles centered at different locations, it yields the so-called "local" component of the viscosity.

However, "non-local" transport must be considered if the particles have significant finite size. Particles centered on opposite sides of a shear plane may collide and transport momentum without their centers actually crossing over. This contribution to the viscosity is potentially important if the particles' sizes approach the scale of their eccentric motion, $ae$ (alternatively

called their random or thermal motion by some authors). Similarly, non-local effects must be considered if mutual gravitational effects near encounter are to be included, as well as collisions, in a momentum transport model.

The evaluation of viscosity $\nu$ by Goldreich and Tremaine (1978) included only the local component. Their work showed a decrease in $\nu$ as the optical thickness $\tau$ of the ring increased above unity. That result inspired explanations for the ringlet structure of Saturn's rings, because it was shown to lead to potential ringlet-forming instabilities (Ward 1981, Lin and Bodenheimer 1981). Greenberg's (1988) analysis also only considered the local component of $\nu$.

Araki and Tremaine (1986) did include the non-local transport due to finite-size particles, but they invoked a statistical approach that cannot be readily interpreted in terms of the physics of individual particle collisions. An important result was that $\nu$ increases with $\tau$, even for $\tau > 1$. This work showed that non-local effects can be very important, and that they seemed to eliminate the possibility of ringlet instabilities.

The many-particle simulations by Wisdom and Tremaine (1988) and by Salo (1991) yielded similar results, with ring particle radii of 1m. Salo did suggest that reduction of the particle size in the model to 1cm might give the dependence of $\nu$ on $\tau$ that is required for the ringlet instability.

In this paper, we study the effects of the collisional process using a numerical algorithm that tracks the evolution of the velocity distribution, by statistically taking into account the continual processes of mutual collision among particles and of keplerian motion. The model includes finite-size particles and accounts for both local and non-local transport. The algorithm is based on the physical description of the mechanics by Greenberg (1988). This approach provides great flexiblity in incorporating physical models of the outcomes of collisions, without the constraints of fitting a Boltzman formulation. Moreover, because the Boltzman equation generally has to be solved numerically in any case, it does not really provide any advantage over our algorithmic computation. Our results are complementary to the many-particle simulations, providing statistical results rather than the experiment-specific results of the simulations.

Our formalism is structured to allow eventual inclusion of the effect of gravitation between particles, in cases where Hill's approximation is valid. In principle, a "close encounter" may involve collisions alone in the case without gravitation, or both collisions and gravitation when Hill's approximation applies. Results with gravitational effects will be described in future work; the computational results obtained in this paper involve only collisions. First we describe a two-dimensional ("flat ring") model. Then in Sec. IV, we describe corrections that approximate a three-dimensional model. The results confirm the essential results of the earlier work, and lay the groundwork for inclusion of gravitational encounters and alternative collision models.

## 2 Method of Computation

The local viscosity is a parameter describing the transport of momentum across a boundary, due to particles crossing that boundary. In the case of disks, the relevant boundary is a circle around the central planet. The radial and tangential components of the velocity of particles relative to the circular orbital velocity at a point on that circle are given by $V_r$ and $V_t$. The momentum transported is proportional to $V_t$ and the rate of transport is proportional to $V_r$. Hence the rate of transport of momentum is proportional to $\langle V_r V_t \rangle$, which is the key quantity to be evaluated. Viscosity is readily calculated by scaling to the density and systematic velocity shear (e.g. Greenberg 1988).

The non-local transport is due to exchange of angular momentum between pairs of colliding particles, each with one particle centered inside the shear boundary and the other outside. We must then sum over all possible collisions involving such pairs and compute the transfer of angular momentum, which is not included in the calculation of local viscosity.

For both cases (local and non-local), we first need to know the equilibrium velocity distribution in order to determine the frequencies of all the possible collisions and of all boundary crossings. This is done by integrating the equation of evolution of the velocity distribution function which accounts for the keplerian shift and the scattering due to collisions.

We have developed a numerical model that computes the evolution of the velocity distribution in $(V_r, V_t)$ space by taking a series of small time steps, and in each one moving the particles along keplerian paths (the 2/1 ellipses described by Greenberg 1988) and also redistributing them according to collisional statistics. For the keplerian motion we use the epicyclic approximation, i.e. the calculations are first-order in orbital eccentricity. In each time step, the probable number of collisions between every pair of points in $(V_r, V_t)$ space is computed for each time step (see Eq. 2 below), and particles are then redistributed according to a collisional scattering model.

For the cases considered here, in our scattering model the particles are assumed to be solid spheres, and the collisions occur with no tangential friction (i.e. perfectly slippery particles so that relative tangential velocity at the contact point is unchanged). A coefficient of restitution $\epsilon$ with a value between 0 and 1 describes the change in the radial component of the relative velocity due to energy loss at impact. In most of our numerical experiments, $\epsilon$ is assumed to be a function of impact velocity, according to the data of Bridges et al. (1984). This collisional model is fairly conventional in ring dynamics studies, although some models (unlike ours) compute $\epsilon$ from the mean velocity of the system, rather than separately for each impact geometry.

In order to compute the contribution to momentum transport due to collisions between one particle centered on one side of the boundary and another particle centered on the other side

(non-local), we account for all possible collisions that transport momentum in this way. We assume initially a two-dimensional model in which every particle is on the same orbital plane. In effect, all collisional dynamics are identical to that of hockey pucks on a plane, or of infinitely long, vertically oriented cylinders. After discussing the results of this two-dimensional case, we will describe modifications that simulate a three-dimensional disk, i.e. one that is thicker than the particle size so that particles can pass above and below one another.

Consider Fig. 1, which shows a particle #1 with its center barely on one side of a shear boundary, being hit by a second particle on the other side. We let $r_0$ be the distance from the center of the primary to the circular shear boundary under consideration, and $R_1$ and $R_2$ are the radii of the particles. The particles' surface number density is given by

$$N_{surf} = \int_e \int_\theta n(e,\theta) de d\theta, \tag{1}$$

where $\theta$ is the mean anomaly of the particle, $e$ is orbital eccentricity, and $n$ the number of particles (per area of the disk) with $e, \theta$. (At any position in the disk, the velocity is characterized by $e, \theta$.) The number of particles #2 hitting #1 during a time $dt$ at an orientation within $d\beta$ of the position $\beta$ is

$$N_{2,hit} = n(e_2,\theta_2) de_2 d\theta_2 (R_1 + R_2) d\beta |V_v| dt, \tag{2}$$

where $V_v$ is the component of the relative velocity perpendicular to the surface of the particles at the point of contact. From the epicyclic approximation (first order in $e$) of keplerian motion, $V_v$ is given by:

$$\begin{aligned} V_v &= \Omega_0 r_0 \Bigg[ \left(e_2 \sin(\theta_2) - e_1 \sin(\theta_1)\right) \sin(\beta) \\ &\quad + \left( e_2 \cos(\theta_2) - e_1 \cos(\theta_1) - 3\frac{R_1 + R_2}{r_0} \sin(\beta) \right) \frac{\cos(\beta)}{2} \Bigg], \end{aligned} \tag{3}$$

where $\Omega_0$ is the keplerian circular angular velocity at $r_0$.

For any given contact orientation $\beta$, a more general momentum-transfer impact might have the two particles at any distance from the primary such that the shear boundary lies between their centers. In Fig. 1, for example, the two particles could be shifted leftward, keeping $\beta$ constant. The radial distance of particle #1 from the primary can be expressed as:

$$r(\lambda) = r_0 - \lambda (R_1 + R_2) sin(\beta) \tag{4}$$

where the coefficient $(0 > \lambda > 1)$ expresses the position relative to the shear boundary. In Fig 1, $\lambda$ is very small.

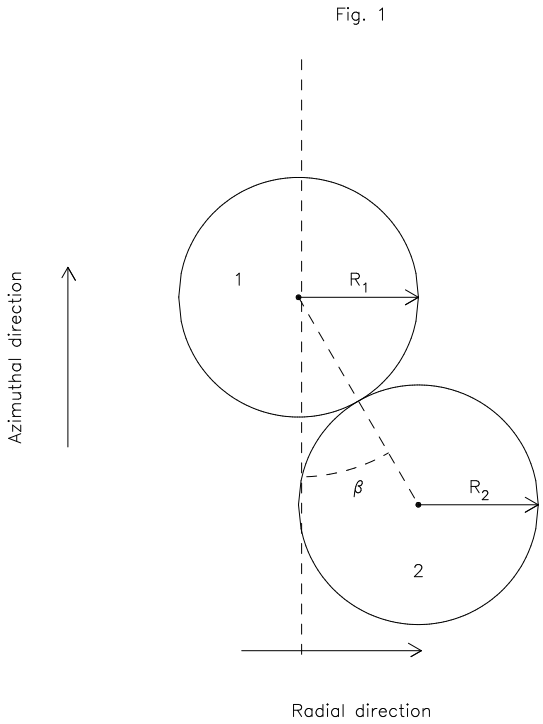



**Fig. 1:** Geometry of impact between two finite-sized particles centered at different distances from a planet. The vertical dashed line just to the right of the center of particle #1 is the shear boundary under consideration and lies along a constant radial distance from the planet.

Next consider the change in angular momentum during a collision. Before the impact, the angular momentum of particle #1 is:

$$A_1(r) \;=\; \Omega(r) r^2 (1 + \frac{e_1}{2}\cos(\theta_1)). \tag{5}$$

Expanding to first order in the small quantities $e$ and $(r - r_0)/r_0$, we obtain:

$$\begin{aligned} \Omega(r) &= \Omega_0 + (r - r_0)\frac{d\Omega}{dr}(r = r_0) \\ &= \Omega_0 \left[1 + \frac{3}{2}\lambda \sin(\beta)\frac{R_1 + R_2}{r_0}\right] \qquad (6) \\ r^2 &= r_0^2 \left[1 - 2\lambda \sin(\beta)\frac{R_1 + R_2}{r_0}\right] \qquad (7) \end{aligned}$$

$$A_1(r) \;=\; \Omega_0 r_0^2(1-\frac{\lambda}{2}\sin(\beta)\frac{R_1+R_2}{r_0}+\frac{e_1}{2}\cos(\theta_1)). \tag{8}$$

Let us call $\delta A$ the outward transport of angular momentum for one collision. Taking into account that particle #1 may be closer to the primary than #2 ($sin(\beta) > 0$ as in Fig. 1) or farther out ($sin(\beta) < 0$), we have:

$$\delta A \;=\; -sign(sin(\beta))(A_1' - A_1) \tag{9}$$

where primes denote after–collision quantities. The transport of angular momentum per unit of time is:

$$\begin{aligned}
\frac{dA}{dt} &= \int_{e_2} de_2 \int_{\theta_2} d\theta_2 \int_{e_1} de_1 \int_{d\theta_1} d\theta_1 n(e_2,\theta_2) n(e_1,\theta_1) \\
&\quad \times \int d\beta (R_1+R_2)|V_v| \int_{\min(r(0),r(1))}^{\max(r(0),r(1))} \delta A(r) 2\pi r dr \\
&= \int_{e_2} de_2 \int_{\theta_2} d\theta_2 \int_{e_1} de_1 \int_{d\theta_1} d\theta_1 n(e_2,\theta_2) n(e_1,\theta_1) \\
&\quad \times \int d\beta (R_1+R_2)|V_v| 2\pi \int_0^1 d\lambda |\sin(\beta)| (R_1+R_2) r(\lambda) \delta A(\lambda) \\
&= \int_{e_2} de_2 \int_{\theta_2} d\theta_2 \int_{e_1} de_1 \int_{d\theta_1} d\theta_1 n(e_2,\theta_2) n(e_1,\theta_1) \\
&\quad \times \int d\beta (R_1+R_2)^2 \Omega_0 r_0^2 2\pi r_0 (-\sin(\beta)) |V_v| (\frac{e_1'}{2}\cos(\theta_1') - \frac{e_1}{2}\cos(\theta_1)).
\end{aligned} \tag{10}$$

Note that to first order in $e$, $r(\lambda)\delta A(\lambda)$ is actually independent of $\lambda$.

From numerical evaluation of (10), we can calculate the non–local angular momentum transport and hence the viscosity.

# 3 Results for single-size particles in a flat ring

Our first numerical experiments were applied to a two-dimensional ring with particles all of a single size. Initial conditions were selected with a distribution of orbital eccentricities at a distance of $10^5$ km from a Saturn-mass primary, such that there is a linear decrease of the number of bodies with each $e$ value, from a maximum at $e = 0$ to zero at some selected value. Usually the initial maximum $e$ is about $2 \times 10^{-7}$, corresponding to $er_0$ being about 20m. The initial conditions also have the particles at any point being uniformly distributed in mean anomaly $\theta$, i.e. the longitudinal distance from their own pericenters. Each of the cases

explored in this paper reached a steady state after about 2 hours of CPU time on an HP 735/99 computer independent of the optical depth.

As discussed by Greenberg (1988), the distribution in $(e, \theta)$ maps directly into $(V_r, V_t)$ space. The coordinates eccentricity $e$ vs. mean anomaly $\theta$ are essentially polar coordinates equivalent to local velocity coordinates, as defined by Ojakangas and Greenberg. A population of particles with a single $e$ value and a uniform distribution in $\theta$ would be distributed in $(V_r, V_t)$ space along ellipses with a 2:1 axis ratio, elongated in the $V_r$ direction. Particles with $\theta = 0$ lie on the $+V_t$ axis, those with $\theta = \pi/2$ lie twice as far from the origin on the $+V_r$ axis, etc.

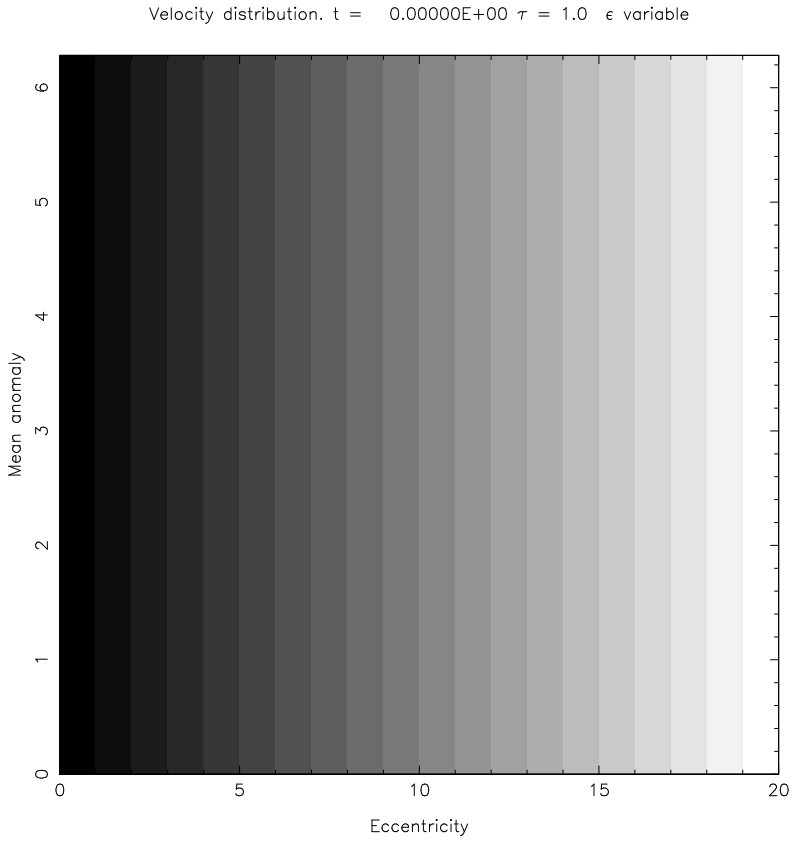



**Fig. 2a:** Initial velocity distribution for a case with two-meter- radius particles in a two-dimensional ring with optical thickness $\tau = 1$, represented by eccentricity and mean anomaly, which map directly into $(V_r, V_t)$.

Fig. 2a shows the initial distribution for a case with two-meter radius particles in a disk with optical thickness $\tau = 1$. This case is similar to conditions in Saturn's A and C rings

(Marouf et al. 1983), where most of the mass is due to particles several meters across with optical thickness of those particles approaching unity (although most of the total optical depth is due to much smaller particles that play a less important role in the dynamics).

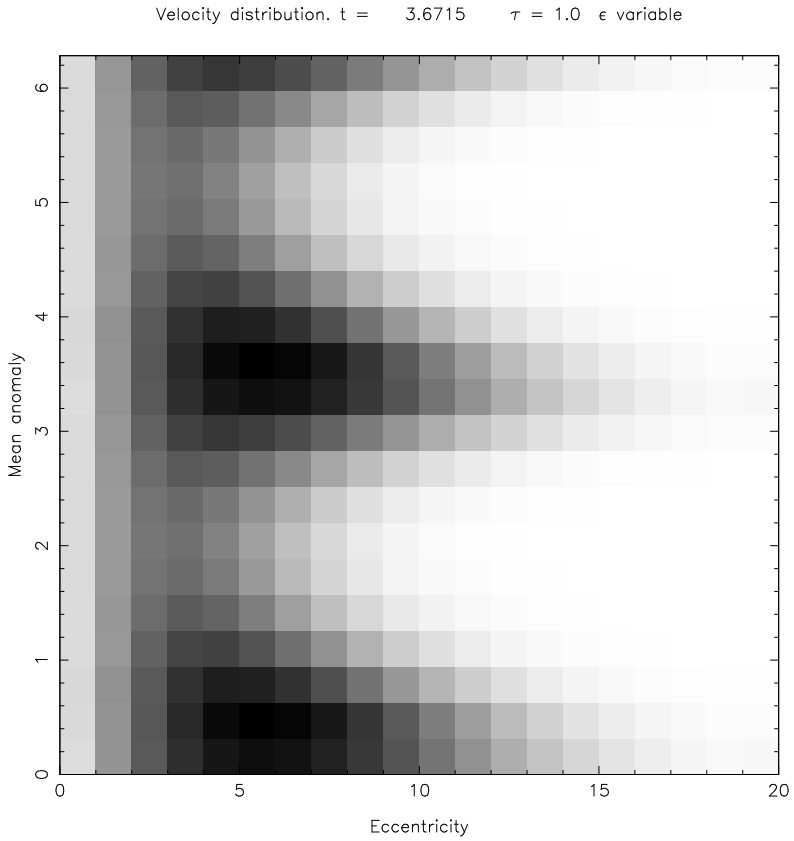



**Fig. 2b:** The steady-state distribution reached after only half an orbital period.

Fig. 2b shows the distribution after about half an orbital period. The particle distribution has clustered at values of $\theta$ that correspond to positive values of $V_r V_t$, consistent with previous evaluations of this distribution. Remarkably, the distribution shown in Fig. 2b is already almost a steady-state; Several orbits later there is negligible further change. This rapid evolution to the steady state is consistent with analytical estimates of collision rates that carefully account for collision cross-sections (eg. Ojakangas and Greenberg, 1990). (A conventional rule of thumb that collision frequencies are roughly $2\tau$ per orbit gives too low a rate.)

The average velocity and angular momentum transport in this steady state are summarized

in Table 1. This first case appears on the first line with $\tau = 1$ (the fifth line of data in Table 1.). The thermal velocity ($9 \times 10^{-4} m/s$) is almost five times larger than that required for a steady-state in Goldreich and Tremaine's (GT) model ($2 \times 10^{-4} m/s$). By itself, this change in $V$ inserted into the GT theory would yield an increase in the kinematic viscosity by a factor proportional to $V^2$, i.e. about twenty, as seen in comparing column $\nu_{GT}$ ($1.7 \times 10^{-5} m^2/s$) with $\nu'_{GT}$ ($3.3 \times 10^{-4} m^2/s$). When we compute the viscosity from the local mass transport in our steady-state population (Fig. 2b), we get a similar, but somewhat larger, value $\nu_{local} = 4.1 \times 10^{-4} m^2/s$. The non-local viscosity is yet an order of magnitude higher, $2.8 \times 10^{-3} m^2/s$.

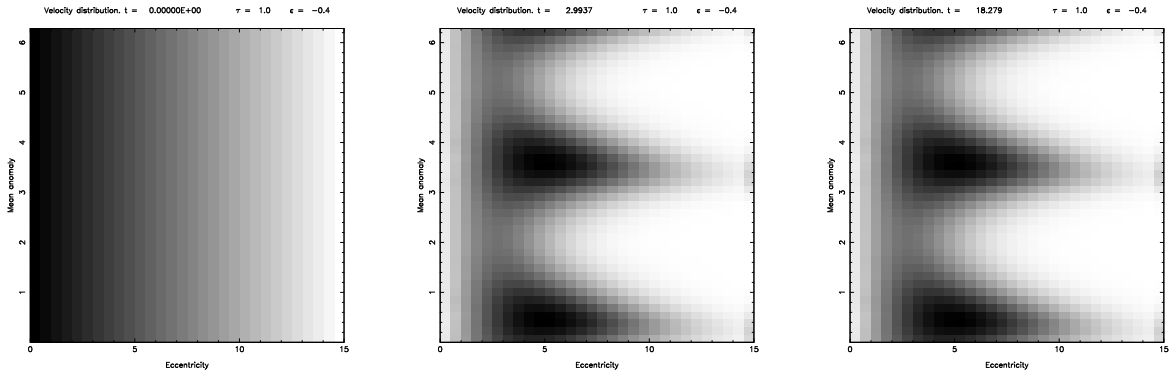



**Fig. 3:** A time sequence (a,b,c) showing the evolution of the velocity distribution in a case where the coefficient of restitution $\epsilon$ is less than the critical value for equilibrium. Velocities damp to a steady-state governed by particle size.

It is certainly reasonable that inclusion of the non-local component of viscosity would increase the total viscosity. However, non-local effects are also responsible for the increase in $V$, because particles can reach across the shear boundary, collide, and convert shear velocity into eccentric motion. Thus the steady-state velocity distribution itself is modified by the non-local effects. As a consequence, even what we have called the local viscosity is modified by non-local effects, to the extent that the velocity distribution is affected.

An important caveat regarding these results is that they are based on the two-dimensional model. As shown in Table 1, in this case $ae$ is almost four times larger than the particle radius. In a real planetary ring, partitioning of collisional energy out of the plane would produce ring thickness comparable to $ae$, so the validity of the 2D model becomes marginal. However, as shown in Sec. IV, when modifications are made to the model to account for ring thickness, the trends found in this first case are confirmed, even more strongly.

Next consider a similar case except that $\epsilon$ is assumed to be fixed at a value of 0.4, rather

than dependent on $V$. According to the GT model, for $\tau = 1$, a fixed $\epsilon$ of that value would result in indefinite decrease in average $V$, with no steady-state viscosity value. Figs. 3a,b,c show a time sequence in such a case. From the initial state in Fig. 3a, after half an orbital period (t=3 in Fig. 3b) the distribution is nearly at the steady state (Fig. 3c). Results are summarized on line 11 of Table 1. The mean $V$ has only dropped to a value slightly smaller that in the case of variable $\epsilon$. Here $ae$ is very close to the particle diameter, so it is already outside the range of validity of the GT model. The finite size of the particles sets a lower limit on the damping of random velocity, which is evident because we have included non-local effects. Viscosity values are similar to those computed in the previous case; again the non-local component of viscosity is dominant.

As shown in Table 1, cases with $\epsilon$ fixed at lower values, 0.1 and 0.3, lead to very similar distributions. Again, $V$ damps to a value controlled by the particle size such that $ae$ is comparable to $R$. $V$ and computed viscosity values are similar to those obtained for $\epsilon = 0.4$.

Changing the fixed value of $\epsilon$ to 0.5 yields the sequence shown in Figs. 4a,b,c. Here random velocities quickly increase to high values. By 5% of an orbital period (Fig. 4b), most particles, which originally had $ae$ less than $\sim$ 4m, have $ae >$ 10m. During the next several periods, the remainder of the particles have $ae$ pumped up to $>$ 10m (Fig. 4c). According to these results, the critical value of epsilon for a steady state is between 0.4 and 0.5, considerably lower than the value $\sim 0.8$ found by earlier studies.

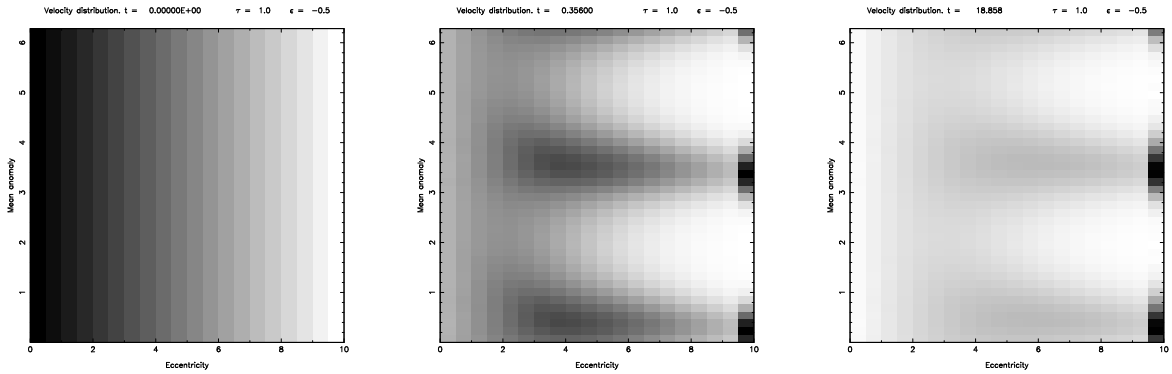



**Fig. 4:** A time sequence (a,b,c) showing the evolution of the velocity distribution in a case where the coefficient of restitution $\epsilon$ is greater than the critical value for equilibrium. Velocities increase indefinitely.

Steady state results for other cases with $R =$ 2m are shown in Figs. 5 and 6, where $\epsilon$ depends on $V$ as in Fig. 2a-b, but here $\tau$ is 0.1 and 10, respectively. Numerical results for these

cases are given in Table 1, including also the case with $\tau = 0.01$. In all these cases, viscosity is considerably larger than the local component. For larger $\tau$, the non-local component of viscosity becomes overwhelming. For smaller $\tau$, collisions are relatively infrequent, so a large part of the momentum transport is expected to be due to particles crossing the shear boundary; Table 1 shows that the non-local part is not dominant (although still very important) for small $\tau$. Even in the case of low $\tau$, we are finding total kinematic viscosity 50 times larger than given by the GT model, for example.

| $\tau$ | $\epsilon$ | $R$ | $ae$ | our $V$ ($\times 10^4$) | $V_{GT}$ ($\times 10^4$) | $\nu_{GT}$ ($\times 10^5$) | $\nu'_{GT}$ ($\times 10^5$) | $\nu_{local}$ ($\times 10^5$) | $\nu_{non-local}$ ($\times 10^5$) |
|---|---|---|---|---|---|---|---|---|---|
| .01 | V | 2. | 11. | 13.6 | 5 | .2 | 1.52 | 4.55 | 3.24 |
| .1 | V | 2. | 11. | 13.5 | 5 | 2. | 14.8 | 37.3 | 36.5 |
| .1 | V | .2 | 2.4 | 3.02 | 5 | 2. | .742 | 3.12 | 1.27 |
| .1 | V | .01 | .54 | .665 | 5 | 2. | .036 | .134 | .0346 |
| 1. | V | 2. | 7.2 | 8.95 | 2 | 1.7 | 32.9 | 41.4 | 281. |
| 1. | V | .5 | 2.5 | 3.12 | 2 | 1.7 | 4.01 | 5.29 | 28.9 |
| 1. | V | .2 | 1.2 | 1.54 | 2 | 1.7 | .976 | 1.07 | 6.11 |
| 1. | V | .01 | .34 | .423 | 2 | 1.7 | .0737 | .0221 | .109 |
| 1. | .1 | 2. | 4.1 | 5.13 | | | 10.8 | 14.2 | 119. |
| 1. | .3 | 2. | 5.4 | 6.67 | | | 18.6 | 26.6 | 185. |
| 1. | .4 | 2. | 6.3 | 7.79 | | | 25.0 | 36.4 | 235. |
| 10. | V | 2. | 6.0 | 7.41 | 1 | .08 | 4.47 | 24.6 | 2270. |
| 10. | V | .2 | 1.0 | 1.28 | 1 | .08 | .134 | .596 | 47.1 |
| 10. | V | .01 | .30 | .378 | 1 | .08 | .0116 | .0118 | .823 |

**Table 1:** Summary of results for the two-dimensional model in mks units (viscosities are in $m^2/s$). Column "$\tau$" shows the optical thickness; "$\epsilon$" shows the coefficient of restitution, with V indicating a velocity-dependent value based on the experimental results of Bridges et al. (1984); "$R$" shows the particle radius in meters; "our $V$" shows the RMS magnitude of velocity vector $(V_r, V_t)$; "$ae$" shows the value of $ae$ corresponding to "our $V$", for comparison with $R$; "$V_{GT}$" shows the RMS velocity according to the theory of Goldreich and Tremaine (1978, "GT"), such that the corresponding value of $\epsilon$ (from Bridges et al., 1984) gives a steady-state; "$\nu_{GT}$" shows the corresponding kinematic viscosity according to GT's theory. "$\nu'_{GT}$" is found by inserting the value of "our $V$" into the GT theory (i.e. their Fig. 4), instead of the value of $V$ used in the previous column; "$\nu_{local}$" is based on our evaluation of the mass transport due to particles crossing the shear boundary in the steady-state population that evolves in our numerical model; "$\nu_{non-local}$" is based on our evaluation of the non-local transport as calculated according to the theory described in Section II.

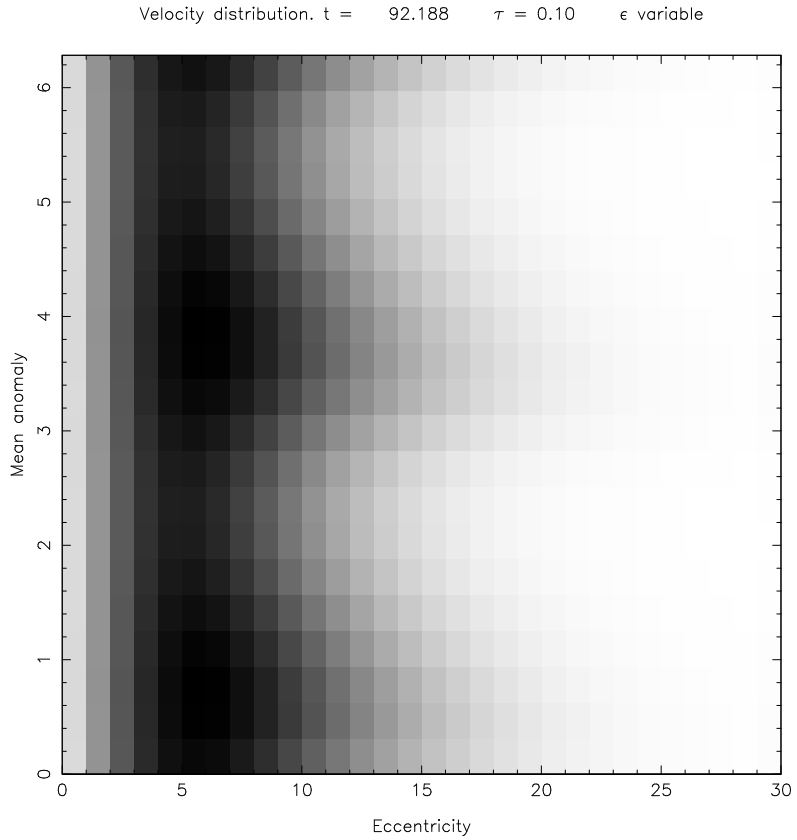



**Fig. 5:** Steady-state result for $\tau = 0.1$ shows expected clustering near $\theta = 45°$ and 225°, as discussed by Greenberg (1988).

Another meaningful trend is seen in the comparison of Figs. 5 and 6, showing why local viscosity decreases both for very large and very small $\tau$ (a result first found by GT). For low $\tau$, Greenberg (1988) showed that particles should be widely distributed in $\theta$ with modest concentrations near $\theta = 45°$ and 225°, such that local viscosity (recall proportional to $\langle V_r V_t \rangle$) is small. This result is clearly demonstrated in Fig. 5. At the opposite extreme of large $\tau$, concentrations in the $\theta$ distribution are stronger, but viscosity is limited because the concentrations are only slightly above $\theta = 0°$ and 180°.

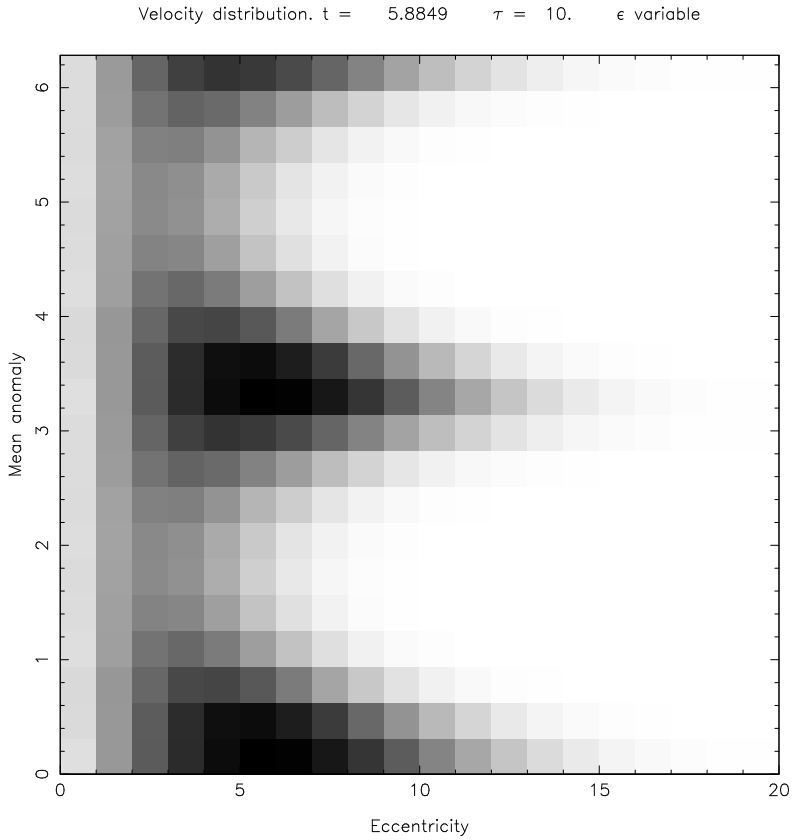



**Fig. 6:** Steady-state result for $\tau = 10$ shows expected clustering near $\theta = 0°$ and 180°.

Results discussed so far all assumed particle radii of 2m. Table 1 also shows results for smaller particles. For $R = 20$cm, $V$ is reduced by a factor $\sim 5$ relative to the 2m cases for the same $\tau$. These values, as well as the local viscosity, are fairly close to the GT values, which are independent of particle size. This closer agreement with GT for smaller bodies probably reflects the diminished importance of particles' physical extent in converting keplerian shear to random motion.

For still smaller particles (see the cases with $R = 1$cm in Table 1), several of the trends discussed above continue to hold. For $\tau > 1$ the non-local viscosity remains very important, but for smaller $\tau$, the non-local viscosity becomes less than the local component, the only place on Table 1 where that occurs. Note however that for these small particles, the steady state random motion is such that $ae$ is typically $> 30$ times the particle size. Because random motion

would be partitioned to include out-of-plane motion, the two-dimensional model is probably only relevant for real planetary rings where particles with $R > 20$cm dominate dynamical processes. Moreover, it must be recognized that any results for the case $\tau > 1$ are artificial in a two-dimensional case.

# 4 Three-dimensional cases

In order to evaluate the effect of significant ring thickness, we have modified our numerical model in the following way. We account for collisional partitioning of random motion between in-plane and out-of-plane components by taking the ring thickness $h$ to be equal to $2ae$. We again model each particle as a cylinder, with its axis perpendicular to the orbital plane, but now we assume that the length of the axis is equal to the diameter $2R$ of the particle. This allows us to reduce the collision frequency by a factor $2R/h$. In all other ways the algorithm remains as in the 2D model. Although the vertically aligned cylinders do not allow for modelling of exchange between in-plane and out-of-plane motion, the assumption of equiparition is reasonable and the cylinder do allow us to account for the less well-understood in-plane transport processes.

Results using this three-dimensional model are shown in Table 2. In addition, the viscosity values are shown graphically in Fig. 7. We consider the cases in the same order as they were discussed in Section III. For $\tau = 1$ and $R = 2$m, with the velocity-dependent $\epsilon$, we find that the steady-state random velocity is nearly twice what we found with the 2-D algorithm, giving a comparable increase in the local viscosity (both versions, $\nu'_{GT}$ and $\nu_{local}$). However, the non-local viscosity decreases by a factor of 5 relative to the 2-D model, because of the reduced impact rate in the thicker disk. The local viscosity is about 40 times that given by the GT theory, and the non-local component doubles the mass transport above the local component alone.

Next we considered the effect of using a fixed value of epsilon, still with $\tau = 1$ and $R = 2$m. As in the 2D case, we confirm that sufficiently large $\epsilon$ values lead to indefinite increase in random velocity and in corresponding values of $ae$ (e.g. for $\epsilon = 0.6$); smaller $\epsilon$ values (e.g. the case $\epsilon = 0.5$ shown in Table 2) lead to values of $ae$ comparable to the particle size. The critical $\epsilon$ value for a steady state is between 0.5 and 0.6, larger than given by the 2D calculations, but still less than the equilibrium value $\epsilon = 0.8$ found by GT in their model (with local-viscosity only).

Exploring the effects of varying $\tau$ for a fixed particle size $R = 2$m, we find similar trends to the 2D models. Viscosity is considerably larger than $\nu_{GT}$, especially for larger $\tau$. However, the values are not as extreme as in the 2D case. For small $\tau$ the total viscosity is only a few times larger than $\nu_{GT}$, while for $\tau = 10$, it is still several orders of magnitude larger.

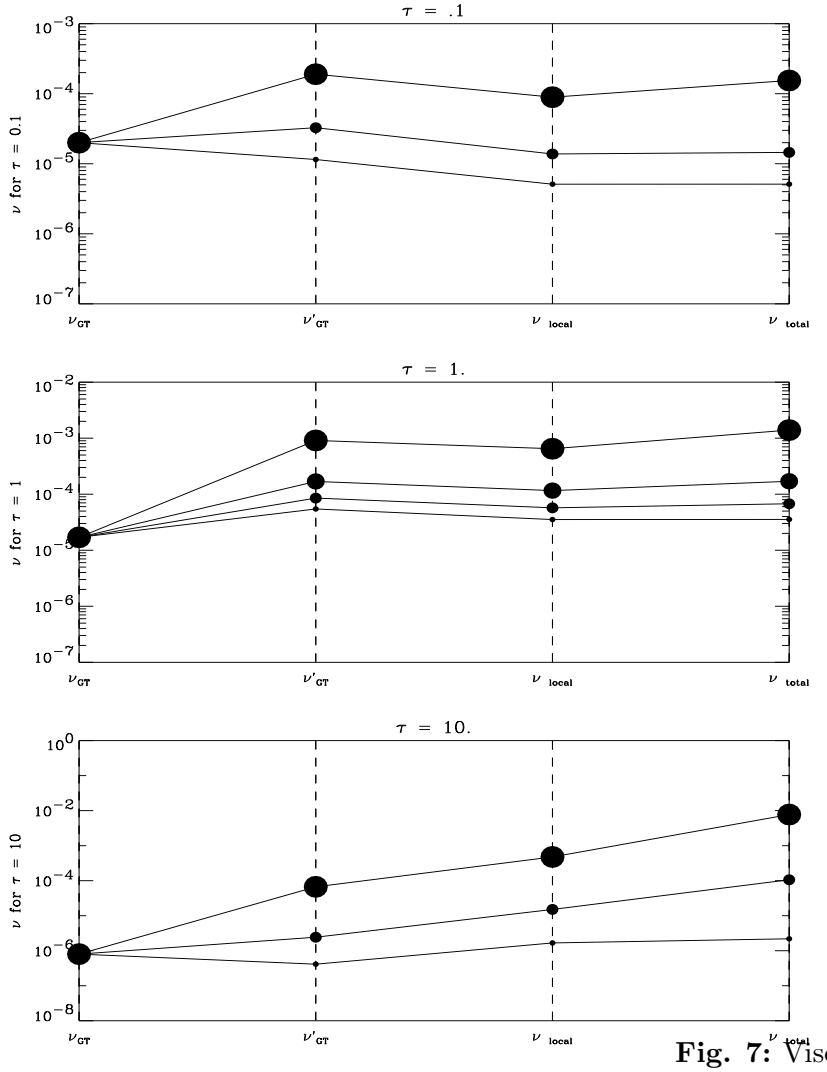



**Fig. 7:** Viscosity values for cases shown in Table 2. All cases shown are for the velocity-dependent $\epsilon$, and the dot sizes correspond to the particle radii in Table 2 (smallest dot is for R = 1cm, largest is for R = 2m). Note the decrease in $\nu_{GT}$ as $\tau$ increases above 1, which would imply the possibility of ringlet instability. The difference in values between $\nu_{GT}$ and $\nu\prime_{GT}$ indicates the effect of the different velocities found in our calculations relative to Goldreich and Tremaine's theory. The GT theory yields the local component of viscosity only, and can be compared with our results in the column labeled $\nu_{local}$. The total viscosity for each case is shown in the last column, which is the sum of the last two columns in Table 2 ($\nu_{local}$ plus $\nu_{non-local}$). The right hand column of this figure ($\nu_{total}$) shows that only for the cm-scale particles does $\nu$ decrease with increasing $\tau$, so as to drive ringlet instabilities.

| $\tau$ | $\epsilon$ | $R$ | $ae$ | our $V$ ($\times 10^4$) | $V_{GT}$ ($\times 10^4$) | $\nu_{GT}$ ($\times 10^5$) | $\nu'_{GT}$ ($\times 10^5$) | $\nu_{local}$ ($\times 10^5$) | $\nu_{non-local}$ ($\times 10^5$) |
|---|---|---|---|---|---|---|---|---|---|
| .01 | V | 2. | 8.6 | 15.4 | 5 | .2 | 1.95 | .799 | .637 |
| .01 | V | .01 | 3.2 | 5.85 | 5 | .2 | .281 | .214 | .000021 |
| .1 | V | 2. | 8.6 | 15.3 | 5 | 2. | 19.0 | 8.92 | 6.53 |
| .1 | V | .2 | 3.5 | 6.34 | 5 | 2. | 3.27 | 1.38 | .0729 |
| .1 | V | .01 | 2.1 | 3.76 | 5 | 2. | 1.15 | .511 | .00038 |
| 1. | V | 2. | 8.6 | 14.9 | 2 | 1.7 | 91.1 | 64.8 | 74.6 |
| 1. | V | .5 | 3.7 | 6.41 | 2 | 1.7 | 16.9 | 11.6 | 5.42 |
| 1. | V | .2 | 2.6 | 4.56 | 2 | 1.7 | 8.52 | 5.71 | 1.04 |
| 1. | V | .01 | 2.1 | 3.64 | 2 | 1.7 | 5.45 | 3.53 | .0162 |
| 1. | .5 | 2. | 10.2 | 17.8 | | | 130. | 69.5 | 74.5 |
| 10. | V | 2. | 6.3 | 9.06 | 1 | .08 | 6.68 | 47.6 | 726. |
| 10. | V | .2 | 1.2 | 1.73 | 1 | .08 | .243 | 1.51 | 9.13 |
| 10. | V | .01 | .48 | .713 | 1 | .08 | .0413 | .167 | .0521 |

**Table 2:** Results for the three-dimensional model. See caption of Table 1 for definitions of quantities shown.

For $R = 20$cm, the local viscosity values are comparable to the 2D results. The non-local component is much smaller than found with the 2D model: For $\tau \leq 1$, it is negligible compared to the local component, but for $\tau = 10$ it is much larger than the local component and 100 times larger than $\nu_{GT}$.

For cm-size particles, where $ae >> R$, we expected the 2D results to be invalid. The 3D model does give very different results. In general the non-local viscosity is small compared with the local component, and the local component is comparable to $\nu_{GT}$.

# 5 Conclusions

Our computations of angular momentum transport in planetary rings, taking into account the role of the finite sizes of particles, confirms significant effects on the relative velocity distribution among colliding particles and on transport processes, due to collisions between particles whose centers are at different distances from the central planet. The latter effect is responsible for the non-local component of the viscosity. The various trends of interest are apparent in Fig. 7.

Generally, the rms random velocity appears to be quite close to the value based on Goldreich and Tremaine's original theory for cases with the smallest particle sizes (cm-scale). This result can be seen comparing the velocity values in Table 2, or by comparing $\nu_{GT}$ and $\nu'_{GT}$ in Table

2 or Fig. 7. (The difference between those two viscosity values is due only to the difference in random velocity.) On the other hand, the velocity values we find for multi-meter particles is a few times larger than the GT value. This result is reasonable: A model that accounts for the finite size of particles would in fact be expected to be most relevant for larger particles rather than for small particles, which can be approximated by infinitesimal sizes.

Similarly, computations of momentum transport, as characterized by the kinematic viscosity components $\nu_{local}$ and $\nu_{non-local}$, show greatest deviation from the GT results for the larger bodies. Evidently the ability to reach radially and transmit momentum during a collision is most effective for larger bodies. Even the local component is affected by the finite particle size because of the effect on the random velocity. There is also a strong and not unexpected correlation with optical thickness $\tau$: the greater frequency of collisions with larger $\tau$ means that a proportionately greater contribution to momentum transport comes during collisions than during the free radial motion of particles, which was assumed dominant process in the GT theory.

Fig. 7 shows that, although $\nu_{total}$ increases with $\tau$ for larger particles, it decreases strongly for cm-size particles as $\tau$ increases above 1. That result seems to support Salo's (1991) suggestion that the ringlet instability mechanism may be possible for rings that are predominantly composed of the smaller particles.

Our results can also be compared with the many-particle simulations of Wisdom and Tremaine (1988) and of Salo (1991), which gave results very similar to one another. For the cases that they considered, with particle radii of 1m, we have computed and compared our own values of $\nu/\Omega R^2$ (c.f. Salo's Fig. 5c) with those given by Salo (Fig. 8). Note that the non-local component is in fairly good agreement, generally about 10% higher than the many-particle simulations. The results are remarkably consistent given the completely different calculational process. The local component also agrees very well for values of $\tau < 0.5$, but diverges substantially for larger $\tau$. In general for all of these cases with $R = 1m$, the steady-state rms relative velocity among particles is about 20% higher in our calculations than in the many- particle simulations.

There are several possible reasons for the disagreement between our results especially for the larger values of $\tau$. First, our three-dimensional model is really a corrected version of a two- dimensional calculation. More precise results will require a full three-dimensional calculation, which could still use our statistical approach. Second, our model does not include the effects of physical packing among particles which can occur when the filling factor (the fraction of space filled with the volume of particles) becomes high. This effect can in principle be included by modifying the statistical treatment, including restricting the range of possible collision geometries in a tightly packed system.

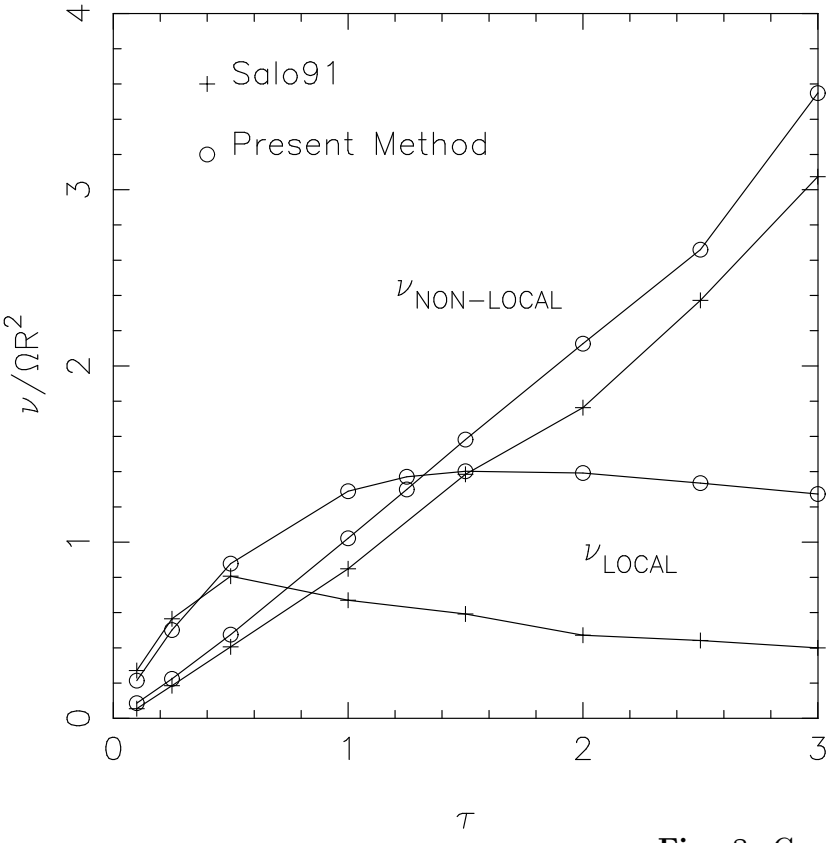



**Fig. 8:** Comparison of viscosity values found for a set of cases explored in the many-body simulations of Salo (1991) and recomputed by our statistical method. The local component differs significantly for $\tau > 0.5$, but the non-local component agrees remarkably well.

The computational technique introduced here provides insight into the statistical processes that determine viscosity, supplementing the results of direct numerical many-body simulations. Moreover, it has the potential for application to other interesting models. We plan to consider momentum transport in systems with two or more different sizes of particles for example. It will also be possible to examine the effects of gravitational interactions among particles, which may introduce a significant non-local component to the viscosity under some circumstances. Such results may be applicable to other dynamical systems, such as planetesimal populations during planet formation, as well as to known planetary rings. Such conditions are ideally suited for application of our technique, because the added features will involve very little additional computation time, and the low $\tau$ conditions are in the regime where our approach is strongest and where direct numerical simulations may be impractical. Non-local interactions, whether

due to finite size of colliding bodies or mutual gravitation clearly play an important role in momentum transport in a variety of kind of astronomical disks.

# Acknowledgments

We are grateful for help and comments by Dan Durda, Joe Plassmann, Jyrki Hanninen, and Michel Henon. Thoughtful reviews by Glen Stewart and Heikki Salo were also valuable. The research reported here was done during Jean-Marc Petit's extended visit to the Lunar and Planetary Laboratory, on leave from the Observatoire de Nice. This work was supported by NASA's Planetary Geology and Geophysics program.

# References

[1] Araki S., & Tremaine S., 1986. *The dynamics of dense particle disks.* Icarus **65**, pp. 83–109.

[2] Bridges, F.G., Hatzes, A., & Lin, D.N.C., 1984. *Structure, stability, and evolution of Saturn's rings.* Nature **309**, pp. 333–335.

[3] Goldreich P., & Tremaine S., 1978. *The velocity dispersion in Saturn's rings.* Icarus **34**, pp. 227–239.

[4] Greenberg, R., 1988. *Particle properties and the large-scale structure of planetary rings: Rebound characteristics and viscosity.* Icarus **75**, pp. 527–539.

[5] Ojakangas G.W., & Greenberg R., 1990. *Viscosity and mass transport in nonuniform keplerian disks.* Icarus **88**, pp. 146–171.

[6] Salo, H., 1991. *Numerical simulations of dense collisional systems.* Icarus **90**, pp. 254–270.

[7] Shu, F.H., & G.R. Stewart, 1985. *The collisional dynamics of particulate disks.* Icarus **62**, pp. 360–383.

[8] Wisdom, J. and S. Tremaine, 1988. *Local simulations of planetary rings.* Astron. J. **95**, pp. 925–940.

# Chapitre 5

# SUR LES ASTEROIDES

## 5.1 Evolution physique des petits corps

[**9**]

"Modelling the outcomes of high-velocity impacts between small solar system bodies"

PETIT, J-M. & FARINELLA, P.

*Celestial Mechanics* **57** (1993)

# Modelling the outcomes of high–velocity impacts between small solar system bodies

Jean-Marc Petit

CNRS, Observatoire de Nice, B.P. 229, 06304 Nice Cedex 04, France

Paolo Farinella

Observatoire de Nice, B.P. 229, 06304 Nice Cedex 04, France

Dipartimento di Matematica, Università di Pisa, Via Buonarroti 2, 56127 Pisa, Italy

**Abstract**

We present a self–consistent numerical algorithm aimed at predicting the outcomes of high–velocity impacts between asteroids (or other small bodies of the solar system), based on a set of model input parameters which can be estimated from the available experimental evidence, and including the possible gravitational reaccumulation of ejected fragments whose velocity is less than a suitably defined escape velocity. All the fragment mass distributions are modelled by truncated power laws, and a possible correlation between fragment ejection velocity and mass is taken into account in different ways, including a probabilistic one. We analyze in particular the effectiveness of the gravitational reaccumulation process in terms of different choices of the collisional parameters and the assumed relationship between fragment speed and mass. Both the transition size beyond which solid targets are likely to reaccumulate a large fraction of the fragment mass and the collision energy needed to disperse most of the fragments are sensitive functions of the assumed fragment velocity versus mass relationship. We also give some examples of how our algorithm can be applied to study the origin and collisional history of small solar system bodies, including the asteroid 951 Gaspra (recently imaged by the Galileo probe) and the asteroid families.

# 1 Introduction

As a consequence of their non–negligible eccentricities and mutual inclinations, the orbits of most asteroids can occasionally intersect each other, opening the way to the occurrence of high–velocity impacts. A typical main–belt asteroid undergoes collisions with other (normally, much smaller) asteroids at an average velocity of 5.8 $km/s$ and at a rate of the order of $3 \times 10^{-18}$ $(R/km)^2 N_{pr}$ $yr^{-1}$, where $R$ and $N_{pr}$ are the assumed target radius and number of projectiles (Farinella and Davis, 1992). As a consequence, mutual impacts — whose outcomes range from small–scale cratering events to catastrophic target break–up — have played a crucial role in the evolution of the asteroid belt (the pioneering work in this field dates back to the 50s, e.g. Öpik, 1951, and Piotrowski, 1953; for a recent a review, we refer to Davis *et al.*, 1989). In the 80s it has been also realized that disruptive impacts may have been important in the origin and evolution of other types of small solar system bodies, like some icy satellites of the outer planets (Farinella *et al.*, 1983), the planetary rings (Harris, 1984), and the interplanetary dust complex (Sykes *et al.*, 1989).

In order to model the outcomes of collisional events occurring in the solar system, hypervelocity impact experiments have been carried out in the laboratory at a much smaller scale (for reviews, see Davis *et al.*, 1986; Fijiwara *et al.*, 1989), and from the observed results some simple empirical relationships have been derived to describe the collisional outcomes in terms of several basic *input parameters*, such as the projectile and target masses, the impact velocity, some coefficients describing the impact response of the target's material, and sometimes the incidence angle (when it is relevant to distinguish between "central" and "grazing" impacts). For the sake of simplicity, both the target and the projectile are normally considered as nearly homogeneous and spherical bodies (see Chapman *et al.*, 1989, for a discussion of possible complications). It is important to keep in mind that at least one physically fundamental difference exists between small laboratory targets and real celestial bodies, which are tens or hundreds of km across: in the latter case, self–gravity is not negligible, and this has two important consequences. First, self–compression can strengthen the interior of the bodies making more difficult to fracture and shatter them, and second — if fragmentation occurs — mutual gravity can cause the partial reaccumulation of the ejected fragments into a so–called *pile of rubble*.

From the theoretical point of view, our understanding of the complex physical processes which take place after an energetic impact is still quite limited. Research work on this subject is currently in progress, based either on semiempirical working hypotheses about the fracture criteria and the geometry of fragment velocity fields (Paolicchi *et al.*, 1989, 1992), or on the integration of suitable hydrodynamical equations in two or three dimensions (Ryan *et al.*, 1990; Ryan and Melosh, 1992; Nolan *et al.*, 1992; Asphaug *et al.*, 1992). However, the corresponding numerical experiments are still so complicated and demanding (in terms of CPU time and

memory) that they cannot be used to explore the full range of the relevant parameters, in order to interpolate and extrapolate to actual solar system conditions, nor to simulate the long–term collisional evolution of systems like the asteroid belt or the planetary rings. Therefore, the results of both the laboratory and the numerical experiments need to be incorporated into suitable scaling algorithms, requiring a number of input parameters (which depend on the "initial conditions" of the impact and the detailed physical processes following it) and capable of predicting in a fast and effective manner the outcome of any possible impact occurring in the system under scrutiny.

In this paper (Secs. 2 to 5) we present such a numerical algorithm, which is physically simple (but still fully self–consistent) and at the same time fairly realistic, as it includes models for the effects of self–gravity and fragment reaccumulation. This algorithm is a follow–up and a refinement of those used in the past by Greenberg *et al.* (1978), Davis *et al.* (1979, 1985, 1989), and Farinella *et al.* (1992), as it includes some additional physical effects. In Sec. 6 we shall discuss how much the predictions of our algorithm are sensitive to different choices of some collisional parameters, and how the effectiveness of the gravitational reaccumulation process depends upon the assumed relationship between the ejection velocity and the mass of the fragments. We shall also give some examples of how this algorithm can be applied to model the origin of asteroid families and the collisional evolution of different types of small solar system bodies.

# 2 The fragmentation threshold

In this and the next three sections we are going to describe our algorithm to predict the outcome of a collision between two (spherical and homogeneous) bodies of masses $M_i$, densities $\rho_i$, diameters $D_i = (6M_i/\pi\rho_i)^{1/3}$, impact strenghts $S_i$ (the index $i$ is always equal to either 1 or 2), when the relative velocity is $V_{rel}$ and the relative kinetic energy is $E_{rel} = M_1 M_2 V_{rel}^2/2(M_1 + M_2)$. The impact strength $S$ is defined as the mimimum specific energy delivered to each body to cause catastrophic fragmentation of it. We are implicitly assuming that an abrupt transition always occurs between the cratering and shattering regimes. This is supported by the experimental results, which show that the transition from localized target damage to widespread fragmentation is associated with a modest increase of the impact energy (see e.g. Fujiwara *et al.*, 1977, Fig. 10). For targets having a significant self–gravity, recent hydrocode modelling work by Nolan *et al.* (1992) and Greenberg *et al.* (1992) suggests that the transition may be less distinct than indicated by laboratory experiments, in the sense that a very large crater can form and remain on the surface even though a significant fraction of the target material is shattered and reaccumulated. While we shall keep our "catastrophic transition" assumption,

we shall treat the formation of giant craters in a way that does not conflict with the above results (see Sec. 4).

We further assume that in the shattering case the available energy is partitioned in equal parts to the two colliding bodies (Hartmann, 1988), so for each of them fragmentation occurs if $E_{rel} > 2SM/\rho x_{cr}$ (here $x_{cr}$ is a numerical coefficient which will be justified later). The value of $S$ is obtained from that observed in laboratory impact fragmentation experiments ($S_0$) through the scaling formula:

$$S_i = \left( S_{0,i} + \frac{\pi\gamma}{15} G\rho_i^2 D_i^2 \right) \left( \frac{D_i}{0.2\ m} \right)^{-1/4} \quad , \tag{1}$$

where the term proportional to $D^2$ is due to self-compressional strengthening of the interior of each body (see Davis *et al.*, 1985), while the $D^{-1/4}$ factor is consistent with strain–rate scaling with size (Housen and Holsapple, 1990). The numerical coefficient $\gamma$ is possibly material–dependent and has to be determined by experiments (see e.g. Housen *et al.*, 1991). For solid materials like basalt or concrete, abundant experimental evidence indicates that $S_0 \approx 3 \times 10^6\ J/m^3$.

Given $S_i$ and the other impact parameters, our goal is to compute the mass distribution of the bodies generated by the impact. Since (i) this is is possible only for a finite number of target–projectile pairs $(M_1, M_2)$, and (ii) it is normally useless to keep track of all the masses of the individual fragments, in our numerical computations we normally grouped the fragments into a limited number (ranging between $\approx 20$ and 100) of discrete, logarithmic mass bins, having their limits (and central values) scaled by a constant amount (typically, a factor $g = 2$). In other words, we considered a geometric sequence of $N_{bin}$ "central" mass values $m_{bin,k} = g^{k-1} m_{bin,1}$ $(k = 1, ..., N_{bin})$, with every bin spanning the range $(m_{bin,k}/\sqrt{g}, m_{bin,k}\sqrt{g})$. In order to compute the amount of mass reaccumulated by self–gravity after an impact, we also considered the fate of the low–mass "tail" of the distribution, ranging between the assumed lower cutoff of the fragment mass ($M_{dust}$, see later) and the lower limit of the first bin $(m_{bin,1}/\sqrt{g})$.

# 3 Shattering events

When one body is shattered, the mass of the largest remaining fragment is given by $M_{max,i} = M_i f_{l,i}$, where

$$f_{l,i} = \frac{1}{2} \left( \frac{S_i M_i}{\rho_i E_{rel}/2} \right)^{1.24} \quad , \tag{2}$$

in agreement with the experimental relationship derived by Fujiwara *et al.* (1977), which predicts that when the critical energy threshold for fragmentation is just reached, a "core" half the mass of the original body survives, while this fraction decreases rapidly for increasing impact energies. Actually, the previous formula is valid for head–on impacts; if we take into account that oblique impacts are on the average less effective in shattering the involved bodies, $f_{l,i}$ can be averaged over the incidence angles leading to

$$\overline{f_{l,i}} = 3 f_{l,i}^{2/3} - 2 f_{l,i} \tag{3}$$

(Davis *et al.*, 1985). Hereinafter by $f_{l,i}$ we will always mean the averaged value $\overline{f_{l,i}}$ given by Eq.(3). In order to keep the definition of threshold collision as that generating a largest fragment with half the mass of the target, the condition for fragmentation given in Sec. 2 includes a numerical factor $x_{cr} = 4^{-1/1.24} \cong 0.327$, derived by noting that $\overline{f_{l,i}} = 1/2$ when $f_{l,i} = 1/8$. This implies that the effective strength of the targets versus impacts from random directions is about 3 times higher than the value found in the laboratory for head–on impacts.

Next, we need to model the mass distribution of the fragments. Following many previous investigations (e.g., Hartmann, 1969; Kresàk, 1977; Greenberg *et al.*, 1978; Zappalà *et al.*, 1984; Fujiwara *et al.*, 1989), we shall assume here a simple, albeit somewhat artificial, mathematical model: a *Pareto* power–law distribution truncated at a mimimum cut–off mass $M_{dust}$ and consistent with the fact that $M_{max,i} = f_{l,i} M_i$ is the mass of the largest fragment. More precisely, we define $N_i(\geq m)$ to be the number of fragments of body $i$ having mass larger than $m$ and $M_i(\leq m)$ and $M_i(> m)$ the total mass of fragments of body $i$ with an individual mass respectively smaller and larger than $m$; we assume that $N_i(\geq m)$ has a discontinuity at $m = M_{max,i}$ and that we know that there is just one fragment of mass $M_{max,i}$. If $\Theta(x)$ is the Heaviside step function ($= 0$ for $x < 0$ and $= 1$ for $x \geq 0$) and $\delta(x)$ is the Dirac delta function ($\delta(x) = d\Theta(x)/dx$), then we have

$$N_i(\geq m) = B_i m^{-b_i} \Theta(M_{max,i} - m) \Theta(m - M_{dust}) + B_i (M_{dust})^{-b_i} [1 - \Theta(m - M_{dust})] \ , \tag{4}$$

where $b_i$ is the characteristic exponent of the cumulative size distribution, and $B_i = (M_{max,i})^{b_i}$, so that $N_i(\geq M_{max,i}) = 1$. The differential (or incremental) fragment size distribution is then given by:

$$\begin{aligned} n_i(m)\, dm &= -dN_i(\geq m) \\ &= \Big[ b_i B_i m^{-b_i - 1} \Theta(M_{max,i} - m) + \delta(m - M_{max,i}) \Big] \Theta(m - M_{dust}) dm \ , \end{aligned} \tag{5}$$

while for the total mass accounted by fragments of mass smaller than any given value $m$ we obtain (provided $b_i \neq 1$):

$$\begin{aligned} M_i(\leq m) &= \int_0^m m\ n_i(m)\ dm \\ &= \frac{b_i B_i}{1-b_i}\left[m^{1-b_i} - (M_{dust})^{1-b_i}\right][1-\Theta(m-M_{max,i})] \\ &+\left[\frac{b_i B_i}{1-b_i}\left[(M_{max,i})^{1-b_i} - (M_{dust})^{1-b_i}\right] + M_{max,i}\right]\Theta(m-M_{max,i})\ . \end{aligned} \tag{6}$$

Of course, the cumulative mass accounted by fragments larger than $m$ will be $M_i(>m) = M_i - M_i(\leq m)$. Since mass conservation implies $M_i(\leq M_{max,i}) = M_i$, from Eq. (6) we get the condition

$$M_i(1-f_{l,i}) = \frac{b_i B_i}{1-b_i}[(M_{max,i})^{1-b_i} - (M_{dust})^{1-b_i}]\ , \tag{7}$$

which allows us to rewrite Eq.(6) as

$$\begin{aligned} M_i(\leq m) &= \frac{b_i (M_{max,i})^{b_i}}{1-b_i}\left[m^{1-b_i} - (M_{dust})^{1-b_i}\right][1-\Theta(m-M_{max,i})] \\ &+M_i\Theta(m-M_{max,i})\ . \end{aligned} \tag{8}$$

Eq.(7) is important because from it, once $M_i$, $M_{dust}$ (which are "input parameters") and $f_{l,i}$ (or $M_{max,i}$, to be derived from Eqs.(2,3)) are known, the exponent $b$ can be derived in a unique way. For this purpose, Eq.(7) can be rewritten as

$$b = \frac{\frac{M}{M_{dust}}(1-f_l)}{\frac{M}{M_{dust}} - \left(\frac{M_{max}}{M_{dust}}\right)^b} \tag{9}$$

and solved numerically by an iterative method. As pointed out with simplified formulae (ignoring e.g. the lower cutoff $M_{dust}$) by Greenberg *et al.* (1978) and Farinella *et al.* (1982), the assumption that the fragment mass distribution is given by a single–exponent power law allows us to uniquely specify its "free parameters" $B_i$ and $b_i$. However, we have to stress here that this assumption is certainly an oversimplification of reality. Indeed, many experiments show that a much better fit is obtained by using different exponents in different mass ranges (Fujiwara *et al.*, 1989); the same property is shared by the real asteroid population (Cellino *et al.*, 1991).

On the other hand, the physical meaning of the transition masses (or sizes) for which the best–fitting exponents change is not currently well understood; the corresponding values probably depend both on the material properties and on the impact parameters. As a consequence, we did not feel justified at this stage to introduce in our algorithm such a complicating feature as a two– or three–exponents power law to represent the fragment mass distribution — implying several more free parameters to be given in input, whose choice would be almost arbitrary.

It is easy to see that the critical value $b_i = 1$ separates the case in which most of the mass is in the largest bodies ($b_i < 1$) from the opposite one. Actually, note that when $M_{dust} \Rightarrow 0$, $b_i \Rightarrow 1 - f_{l,i}$ (rather than the commonly used relationship $b = (1 + f_l)^{-1}$, see e.g. Greenberg *et al.*, 1978). For $M_{dust} > 0$, the "supercatastrophic" cases when $f_l \ll 1$ can yield $b > 1$.

How many fragments are going to escape from the gravitational well of the two colliding bodies? A detailed answer would require the knowledge of the original position of the fragments inside their parent bodies, and of their motion immediately after the ejection. Although this problem could be tackled with models such as that of Paolicchi *et al.* (1989), its treatment is well beyond the scopes of the present work. Instead, we are going to derive an "effective escape velocity" $V_{esc}$ defined by analogy with the usual definition of surface escape velocity of a celestial body, and to assume that only the fragments ejected with speeds exceeding $V_{esc}$ will escape "to infinity" (i.e., will achieve independent heliocentric orbits). We use the energy balance equation

$$\frac{1}{2}(M_1 - M_{max,1} + M_2 - M_{max,2})V_{esc}^2 + W_{tot} = W_{fr,1} + W_{fr,2} + W_h \ , \tag{10}$$

where we have assumed that after the impact both the colliding bodies are shattered (if this is not the case for body $i$, $W_{fr,i}$ will have to be changed into $W_{crat,i}$, to be derived in Sec. 4) and that the largest fragment from either of them has negligible kinetic energy (this is consistent with experimental results by Fujiwara and Tsukamoto, 1980, Nakamura and Fujiwara, 1991, and others). The various $W$ terms in Eq.(10) stand for the following potential energies: $W_{tot}$ is the total potential energy of the two bodies just before fragmentation, that is

$$W_{tot} = -\frac{3GM_1^{5/3}}{5Q_1} - \frac{3GM_2^{5/3}}{5Q_2} - \frac{GM_1M_2}{Q_1M_1^{1/3} + Q_2M_2^{1/3}} \ , \tag{11}$$

where $Q_i = (4\pi\rho_i/3)^{-1/3}$; $W_{fr,i}$ is the sum of the self-gravitation potential energy of each fragment of body $i$, derived form the mass distribution (5):

$$W_{fr,i} = \frac{3G}{5Q_i}\int_{m=0}^{m=\infty} m^{5/3}\ n_i(m)\ dm = -\frac{3G}{5Q_i}\frac{5M_{max,i}^{5/3} - 3b_iM_{max,i}^{b_i}M_{dust}^{5/3-b_i}}{5 - 3b_i} \tag{12}$$

and $W_h$ is an estimate of the potential energy when the escaped fragments are separated by a distance of the order of the Hill's radius of the total colliding mass in the gravitational field of the central mass (for the asteroids, the Sun) of mass $M_0$ and orbital distance $R_0$:

$$W_h = -\frac{3G(M_1+M_2)^{5/3}}{5}\frac{(3M_0)^{1/3}}{R_0} \ . \tag{13}$$

# 4 Cratering events

Assume now that for body $i$, $E_{rel} < 2S_iM_i/\rho_i x_{cr}$, so that it is not shattered by the collision. In this case a crater is formed on it. If the case of a small–scale collision, resulting into a crater of mass smaller than 1% of the target mass, we assume that the mass $M_{crat}$ excavated from the crater (and converted into ejected fragments) is just proportional to the impact energy, through a *crater excavation coefficient* $\alpha$ to be given in input and depending on the material properties (for "soft" and "hard" materials, laboratory data yield $\alpha \approx 4 \times 10^{-4}\ s^2/m^2$ and $\alpha \approx 10^{-5}\ s^2/m^2$, respectively; see Stöffler *et al.*, 1975, and Dobrovolskis and Burns, 1984, p. 467). For larger craters, we still assume a linear dependence of the excavated mass on $E_{rel}$, but with coefficients such that the largest possible crater — created when $E_{rel}$ is just less than the fragmentation threshold $2S_iM_i/\rho_i x_{cr}$ — has a mass of 1/10 that of the target. Note that this is much more than the largest crater predicted by conventional scaling theories, which has a diameter of the order of the target radius and a volume $\approx 3\%$ of the target volume; rather, our assumption is in agreement with the recent results of numerical hydrocode simulations of asteroid impacts, suggesting that such giant craters (or concavities) can be formed without complete break–up of the target, but damaging the material structure enough for generating a substantial amount of regolith (Greenberg *et al.*, 1992).

Imposing continuity for $M_{crat} = M_i/100$, we obtain the relationships

$$M_{crat} = \alpha E_{rel} \qquad \text{for} \quad E_{rel} \le \frac{M_i}{100\alpha} \ , \tag{14}$$

$$M_{crat} = \frac{9\rho_i x_{cr}\alpha}{200S_i\alpha - \rho_i x_{cr}}E_{rel} + \frac{M_i}{10}\frac{\rho_i x_{cr} - 20S_i\alpha}{\rho_i x_{cr} - 200S_i\alpha} \qquad \text{for} \quad \frac{M_i}{100\alpha} < E_{rel} \ , \tag{15}$$

In the cratering case, the largest surviving body has always a mass $M_i - M_{crat}$; we shall again assume a single–exponent power law mass distribution of the ejected fragments, but with an exponent $b_i$ having the constant value 0.8, instead of being related to the impact parameters (as in the shattering case). As a consequence, the largest fragment from the crater has a mass $f_{l,i}M_{crat}$, with $f_{l,i}$ given by the implicit equation

$$f_{l,i} = 1 - b_i + b_i f_{l,i}^{b_i} \left( \frac{M_{dust}}{M_{crat}} \right)^{1-b_i} \quad , \tag{16}$$

which can be easily solved numerically. Also, in the cratering case the post–impact potential energy is given by

$$\begin{aligned} W_{crat,1} &= -\frac{3G}{5Q_i} \int_{m=0}^{m=\infty} m^{5/3}\ n_i(m)\ dm - \frac{3G(M_i - M_{crat})^{5/3}}{5Q_i} \\ &= -\frac{3G}{5Q_i} \frac{5M_{max,i}^{5/3} - 3b_i M_{max,i}^{b_i} M_{dust}^{5/3-b_i}}{5 - 3b_i} - \frac{3G(M_i - M_{crat})^{5/3}}{5Q_i} \ , \end{aligned} \tag{17}$$

which has to be substituted in Eq.(10) for $V_{esc}$ instead of $W_{fr,i}$.

# 5 Ejection velocity vs. mass relationship and reaccumulation

Most models of asteroid collisional evolution (e.g., Davis *et al.*, 1985, 1989) neglect any possible correlation between the ejection velocity and the mass of the fragments (both from shattering and for cratering events), and just assume that the fraction of the total fragment mass ejected with speeds exceeding some given value $V$ is $(V/V_{min})^{-k}$, $V_{min}$ being a fixed lower cutoff to the ejection velocity. Such a relationship was observed for crater ejecta (with $k \approx 9/4$) by Gault *et al.* (1963), and later confirmed by other experiments. On the other hand, recent experimental evidence (Davis and Ryan, 1990; Nakamura and Fujiwara, 1991) indicates that some correlation exists, with the largest fragments normally going somewhat slower; this is also predicted by the semiempirical models of Paolicchi *et al.*, 1989, 1992). It is easy to show (see e.g. Nakamura and Fujiwara, 1991) that in this case the relationship given above for the mass fraction ejected at speeds $> V$ still holds provided the fragment speeds are proportional to the power $-r$ of their mass, with $r = (1 - b)/k$.

Here we shall assume that there is a definite correlation between the mass and the velocity of a fragment, which is modelled either in a deterministic way (e.g., through a definite formula relating the two quantities), or stochastically (see later). We define $E_i = E_{rel}/2$ as the energy received by body $i$ at impact, and $E_{fr,i} = f_{ke} E_i$ as the total kinetic energy of ejecta from body $i$, where $f_{ke}$ is an input parameter specifying which fraction of the impact energy is partitioned into kinetic energy of the fragments (for a detailed discussion about this *anelasticity parameter*, see Davis *et al.*, 1989). In order to avoid any possible divergence of the kinetic energy of the

fragments, we set a high velocity cutoff $V_{max}$ to the distribution (a reasonable phyical motivation for this is that the fragment velocity is unlikely to exceed the sound velocity in the material, of the order of some $km/s$). To give the explicit relationship between the velocity $v$ and the mass $m$ of a fragment, we must distinguish two cases according to the sign of $r_i$. If $r_i > 0$:

$$\begin{aligned} v &= C_i m^{-r_i} \qquad \text{for} \qquad \bar{M}_i \leq m \leq M_{max,i} \\ v &= V_{max} \qquad \text{for} \qquad m < \bar{M}_i \ . \end{aligned} \tag{18}$$

where $\bar{M}_i = (V_{max}/C_i)^{-1/r_i}$. On the other hand, if $r_i < 0$ (this case is treated for consistency, see later):

$$\begin{aligned} v &= C_i m^{-r_i} \qquad \text{for} \qquad M_{dust} \leq m \leq \bar{M}_i \\ v &= V_{max} \qquad \text{for} \qquad \bar{M}_i < m \ . \end{aligned} \tag{19}$$

The constant coefficient $C_i$ is obtained from an energy conservation equation. Consistently with the definition of $V_{esc}$, in the shattering case the largest fragment is assumed to have a negligible kinetic energy in the reference frame of the center of mass. On the other hand, this is not true in the cratering case. Therefore, we will insert the corresponding term in the conservation equation multiplied by a factor $\lambda_i$, with $\lambda_i = 0$ for shattering and $\lambda_i = 1$ for cratering. We also define $\bar{V}_i = C_i \left(M_{max,i}\right)^{-r_i}$. We now have two possibilities: either for all fragments $C_i m^{-r_i}$ is smaller than $V_{max}$, in which case we just solve the equation:

$$\begin{aligned} E_{fr,i} &= \lim_{\epsilon \to 0} \int_{M_{dust}}^{M_{max,i}-\epsilon} \frac{v^2}{2} m n_i(m) dm + \lambda_i \frac{\bar{V}_i^2}{2} M_{max,i} \\ &= \lambda_i C_i^2 \frac{M_{max,i}^{1-2r_i}}{2} + \frac{C_i^2}{2} \frac{b_i M_{max,i}^{b_i}}{1-b_i-2r_i} \left[ M_{max,i}^{1-b_i-2r_i} - M_{dust}^{1-b_i-2r_i} \right] \end{aligned} \tag{20}$$

and obtain

$$C_i = \sqrt{\frac{2E_{fr,i}}{\lambda_i M_{max,i}^{1-2r_i} + \frac{b_i M_{max,i}^{b_i}}{1-b_i-2r_i} \left[ M_{max,i}^{1-b_i-2r_i} - M_{dust}^{1-b_i-2r_i} \right]}} \ , \tag{21}$$

or for some fragments (i.e. those having mass $M_{dust}$ if $r_i > 0$ or $M_{max,i}$ if $r_i < 0$) $C_i m^{-r_i} > V_{max}$, in which case the correct equations are different depending on the sign of $r_i$: for $r_i > 0$ we have

$$E_{fr,i} = \lim_{\epsilon \to 0} \int_{\bar{M}_i}^{M_{max,i}-\epsilon} \frac{v^2}{2} m n_i(m) dm + \frac{V_{max}^2}{2} M(\leq \bar{M}_i) + \lambda_i \frac{\bar{V}_i^2}{2} M_{max,i} \ , \tag{22}$$

$$M(\leq \bar{M}_i) = \frac{b_i}{1-b_i} M_{max,i}^{b_i} \left[ \bar{M}_i^{1-b_i} - M_{dust}^{1-b_i} \right] \tag{23}$$

and

$$\begin{aligned} E_{fr,i} &= \lambda_i C_i^2 \frac{M_{max,i}^{1-2r_i}}{2} + \frac{V_{max}^2}{2} \frac{b_i}{1-b_i} M_{max,i}^{b_i} \left[ \bar{M}_i^{1-b_i} - M_{dust}^{1-b_i} \right] \\ &\quad + \frac{C_i^2}{2} \frac{b_i M_{max,i}^{b_i}}{1-b_i-2r_i} \left[ M_{max,i}^{1-b_i-2r_i} - \bar{M}_i^{1-b_i-2r_i} \right] \\ &= C_i^2 \frac{M_{max,i}^{1-2r_i}}{2} \left[ \lambda_i + \frac{b_i}{1-b_i-2r_i} \right] + C_i^{k_i} b_i \frac{M_{max,i}^{b_i} V_{max}^{2-k_i}}{2} \left[ \frac{1}{1-b_i} - \frac{1}{1-b_i-2r_i} \right] \\ &\quad - \frac{V_{max}^2}{2} \frac{b_i}{1-b_i} M_{max,i}^{b_i} M_{dust}^{1-b_i} \ ; \end{aligned} \tag{24}$$

on the other hand, for $r_i < 0$ we have:

$$E_{fr,i} = \int_{M_{dust}}^{\bar{M}_i} \frac{v^2}{2} m n_i(m) dm + \frac{V_{max}^2}{2} M(> \bar{M}_i) \ , \tag{25}$$

$$M(> \bar{M}_i) = \frac{b_i}{1-b_i} M_{max,i}^{b_i} \left[ M_{max,i}^{1-b_i} - \bar{M}_i^{1-b_i} \right] + \lambda_i M_{max,i} \ , \tag{26}$$

$$\begin{aligned} E_{fr,i} &= \frac{V_{max}^2}{2} \frac{b_i M_{max,i}^{b_i}}{1-b_i} \left[ M_{max,i}^{1-b_i} - \bar{M}_i^{1-b_i} \right] + \lambda_i \frac{V_{max}^2}{2} M_{max,i} \\ &\quad + \frac{C_i^2}{2} \frac{b_i M_{max,i}^{b_i}}{1-b_i-2r_i} \left[ \bar{M}_i^{1-b_i-2r_i} - M_{dust}^{1-b_i-2r_i} \right] \\ &= -C_i^2 \frac{M_{max,i}^{b_i}}{2} \frac{b_i}{1-b_i-2r_i} M_{dust}^{1-b_i-2r_i} + C_i^{k_i} b_i \frac{M_{max,i}^{b_i} V_{max}^{2-k_i}}{2} \left[ \frac{1}{1-b_i-2r_i} - \frac{1}{1-b_i} \right] \\ &\quad + \frac{V_{max}^2}{2} M_{max,i} \left[ \lambda_i + \frac{b_i}{1-b_i} \right] \ . \end{aligned} \tag{27}$$

Thus, $C_i$ is the solution of the equation

$$aC_i^{k_i} + b - C_i^2 = 0 \ , \tag{28}$$

where

$$\begin{aligned} a &= M_{max,i}^{2r_i+b_i-1} V_{max}^{2-k_i} \left[ \frac{2b_i r_i}{[(1-2r_i-b_i)\lambda_i + b_i]\,(1-b_i)} \right] \\ b &= 2M_{max,i}^{2r_i-1} \left[ \frac{1-b_i-2r_i}{(1-2r_i-b_i)\lambda_i + b_i} \right] \left[ E_{fr,i} + \frac{V_{max}^2}{2} \frac{b_i}{1-b_i} M_{max,i}^{b_i} M_{dust}^{1-b_i} \right] \end{aligned} \tag{29}$$

for $r_i > 0$, and:

$$\begin{aligned} a &= M_{dust}^{2r_i+b_i-1} V_{max}^{2-k_i} \left[ \frac{2r_i}{1-b_i} \right] \\ b &= 2M_{dust}^{2r_i+b_i-1} M_{max,i}^{-b_i} \left[ \frac{b_i+2r_i-1}{b_i} \right] \left[ E_{fr,i} - \frac{V_{max}^2}{2} M_{max,i} \left( \lambda_i + \frac{b_i}{1-b_i} \right) \right] \end{aligned} \tag{30}$$

for $r_i < 0$. Our program first tries the explicit formula (19); then, if the $V_{max}$ cutoff turns out to be relevant, it solves numerically Eqs. (22). Also, we keep the freedom of giving as an input parameter either $r_i$ or $k_i$, and then deriving the other one through the simple relationship

$$r_i = \frac{1-b_i}{k_i} \ . \tag{31}$$

This implies that in the "supercatastrophic" case $b_i > 1$, when most of the fragment mass is concentrated at small sizes, we obtain the somewhat counterintuitive result that either $r_i$ or $k_i$ has a negative value.

All the previous formulae have been derived for a *deterministic* relationship between $v$ and $m$, as given by Eqs. (17). A more realistic treatment may assume that the same relationship has instead a probabilistic nature, i.e., it holds for the r.m.s. value $V_0 = \sqrt{\langle v^2 \rangle}$ of a Maxwellian distribution function for $v$:

$$P(v; V_0) = \sqrt{\frac{2}{\pi}} \, \frac{3\sqrt{3} v^2}{V_0^3} \, \exp\left( -\frac{3v^2}{2V_0^2} \right) \ . \tag{32}$$

Note that the mean kinetic energy for bodies of a given mass $m$ is still given by $mV_0^2/2$. Therefore, neglecting statistical fluctuations, we can keep the same energy conservation equations obtained before to compute the constant $C_i$ in this case (provided $V_{max}$ is now interpreted as

the higher cutoff to $V_0$). Also in this *probabilistic* case, we have the choice of fixing in input either $r_i$ or $k_i$.

The self–gravitational reaccumulation process is then treated as follows. The impact is assumed to always generate a remnant body formed by the two largest remnants from the colliding bodies (either the largest shattered fragment(s) or the cratered target(s)), plus the non–escaping fraction of the smaller fragments. Let us define $M_{esc,i} = (C_i/V_{esc})^{1/r_i}$, namely the mass corresponding to the escape velocity. In the deterministic case, if $r_i > 0$, the reaccumulated (non–escaping) fragments are those with mass in the range between $\max(M_{dust}, M_{esc,i})$ and $M_{max,i}$ ; if $r_i < 0$, the reaccumulation range is between $M_{dust}$ and $\min(M_{esc,i}, M_{max,i})$. The corresponding total mass is easily found by integrating between these limits $m.n_i(m)dm$, with $n_i(m)dm$ given by Eq.(5).

In the probabilistic case, the whole range between $M_{dust}$ and $M_{max,i}$ is considered. For the mass bins containing $\leq 100$ bodies, the non–escaping fragments are drawn at random by using the velocity distribution (24) — more precisely, given $V_0$ (derived from $m$ through equations similar to (17)), a number is drawn at random according to the distribution (24), and then compared to $V_{esc}$. For the mass bins containing more than 100 fragments, we just compute the fraction of non–escaping fragments $f$ from

$$f = \int_0^{V_{esc}} \sqrt{\frac{2}{\pi}} \frac{3\sqrt{3}v^2}{V_0^3} \exp -\frac{3v^2}{2V_0^2} dv \ ; \tag{33}$$

calling $\mathrm{erf}(v)$ the error function

$$\mathrm{erf}(v) = \frac{2}{\sqrt{\pi}} \int_0^v \exp -x^2 dx \ , \tag{34}$$

we get

$$f = -\frac{2}{\sqrt{\pi}} \sqrt{\frac{3}{2}} \frac{V_{esc}}{V_0} \exp -\frac{3V_{esc}^2}{2V_0^2} + \mathrm{erf}\left( \sqrt{\frac{3}{2}} \frac{V_{esc}}{V_0} \right) \ . \tag{35}$$

As for the "tail" of the mass distribution (i.e., the fragments having masses between $M_{dust}$ and the lower limit of the smallest mass bin, $m_{bin,1}/\sqrt{g}$), we keep the option of dealing with it "deterministically" or in agreement with the probabilistic procedure. In the former case, the reaccumulated tail mass is the integral of $m.n_i(m)dm$ between the values $\max(M_{dust}, M_{esc,i})$ and $\min(M_{max,i}, m_{bin,1}/\sqrt{g})$ (provided $r_i > 0$), or between $M_{dust}$ and $\min(M_{esc,i}, M_{max,i}, m_{bin,1}/\sqrt{g})$ (if $r_i < 0$).

On the other hand, the "probabilistic" tail treatment is done by taking into account the velocity distribution (24). This is made by evaluating the double integral

$$I = bM_{max}^{b} \int_{M_{dust}}^{m_{bin,1}/\sqrt{g}} \int_{0}^{V_{esc}} P(v; V_0(m)) m^{-b} dv\, dm \ . \tag{36}$$

If $V_0 = V_{max}$, the integral over the mass can be done analytically and that over the velocity reduces to (25); otherwise a numerical calculation is carried out.

# 6 Some results and applications

In this Section we shall illustrate some results of our model by means of appropriate plots. We are not going here to explore systematically the multi–dimensional space of our input model parameters in order to study how the collision outcomes depend on them, but only to touch on some specific issues, which we believe are important from the physical point of view. Therefore, we have assumed a set of *nominal* input values as follows: density $\rho = 2500\ kg/m^3$; impact velocity $V_{rel} = 5.81\ km/s$; lower cutoff of the fragment mass $M_{dust} = 10^{-5}\ kg$; higher cutoff of the fragment ejection velocity $V_{max} = 5\ km/s$; pre–scaling impact strength $S_0 = 3 \times 10^6\ J/m^3$ ($S$ was then derived by means of Eq. (1), with $\gamma = 1$); anelasticity parameter $f_{ke} = 0.1$; crater excavation coefficient $\alpha = 10^{-4}\ s^2/m^2$. For our numerical computations we used 50 logarithmic mass bins spaced by a factor 2 (i.e., a factor 1.26 in size), with central values of the corresponding radii ranging from $1.211 \times 10^{-2}\, km$ to $10^3\, km$.

Figs. 1 and 2 give an idea of the mass distribution predicted by the model for shattered bodies of radii ranging from $1\, km$ to $10^3\, km$ (roughly, the size range where most observable asteroids and natural satellites are found). We have plotted in the same frame the mass ratio $f_l$ between the largest fragment and the target (or the total mass excavated in the crater) and the exponent $b$ of the cumulative fragment mass distribution (see Eqs. (2), (3) and (9)) versus the logarithm of the radius $R$ of the parent body, by taking a fixed value (either $10^{-4}$ or $10^{-3}$) of the projectile–to–target mass ratio $M_2/M_1$. As shown by Fig. 1, both $b$ and $f_l$ approach the limit value 1/2 when $M_2/M_1$ and $R$ are such that the threshold between shattering and cratering impacts is reached (this occurs going towards larger target radii, due to the increase of $S$ with size (see Eq.(1)). $b$ is of course closer to 1 (and $f_l$ to zero) for more disruptive collisions, owing to a lower strength or a larger projectile–to–target mass ratio. For craters, on the other hand, both $b$ and $f_l$ are assumed to have constant values (see Sec. 4). Notice that for the nominal values of $S_0$ and $V_{rel}$, a collision involving a projectile of 1/10 the size of the target normally causes extensive break–up, and generates a swarm of fairly small fragments. For an asteroid $200\, km$ across, such as the parent bodies of several asteroid families, the average collision rate quoted in Sec. 1 implies that such collisions have a probability of order unity of occurring over the solar system lifetime (taking into account that several thousands of bodies of diameter

$\approx 20\,km$ exist in the main asteroid belt, see Cellino *et al.*, 1991).

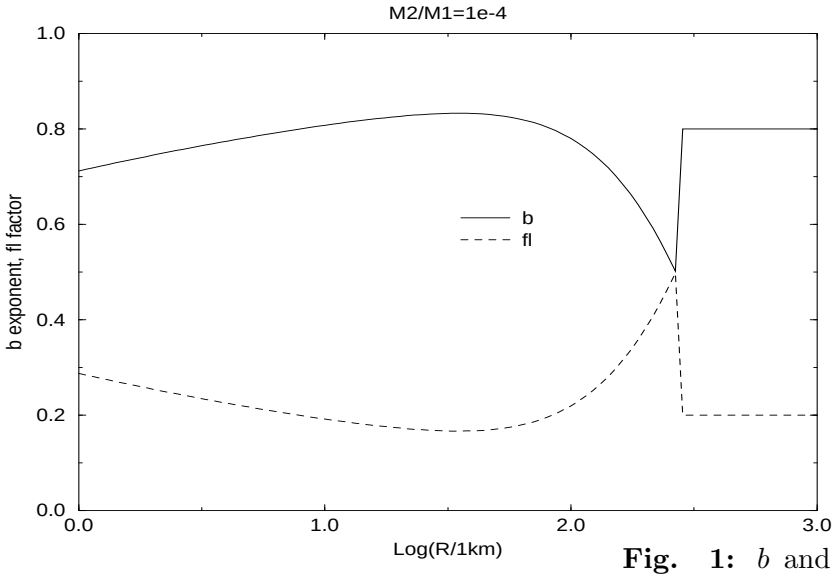



**Fig. 1:** $b$ and $f_l$ versus the logarithm of the radius of the parent body for the nominal values of the impact parameters and a projectile–to–target mass ratio of $10^{-4}$.

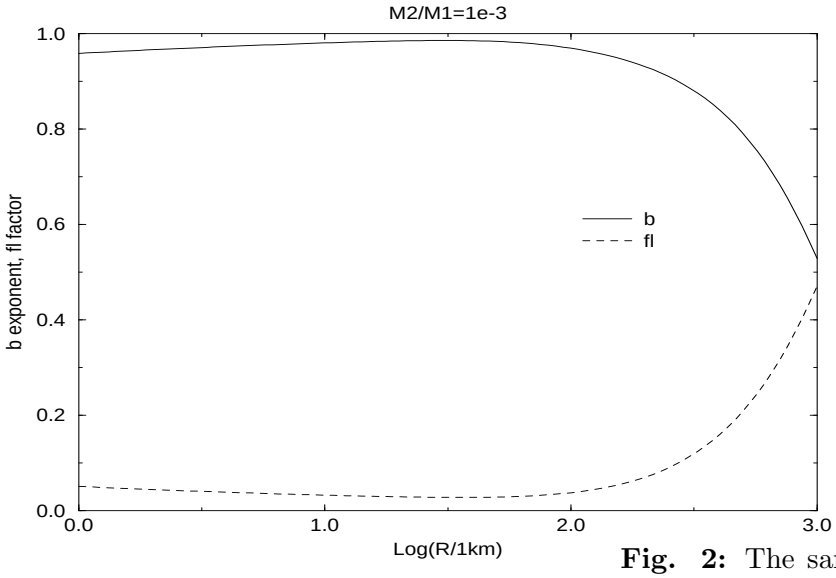



**Fig. 2:** The same as Fig. 1, but for $M_2/M_1 = 10^{-3}$.

The curves shown in Figs. 1 and 2 are not significantly changed when the lower mass cutoff of the fragments $M_{dust}$ is modified within reasonable bounds. Actually, as discussed in Sec. 3, this parameter becomes relevant only for *supercatastrophic* impacts, namely when $b$ is close to 1. This is illustrated in Fig. 3, where the relationship between $f_l$ and $b$ is plotted for three values of $M_{dust}$ (differing from each other by orders of magnitude) in the supercatastrophic case $M_2/M_1 = 1$. Note that for $b$ smaller than unity the approximation $f_l \cong (1 - b)$, which corresponds to the limit $M_{dust} \Rightarrow 0$, is fairly accurate.

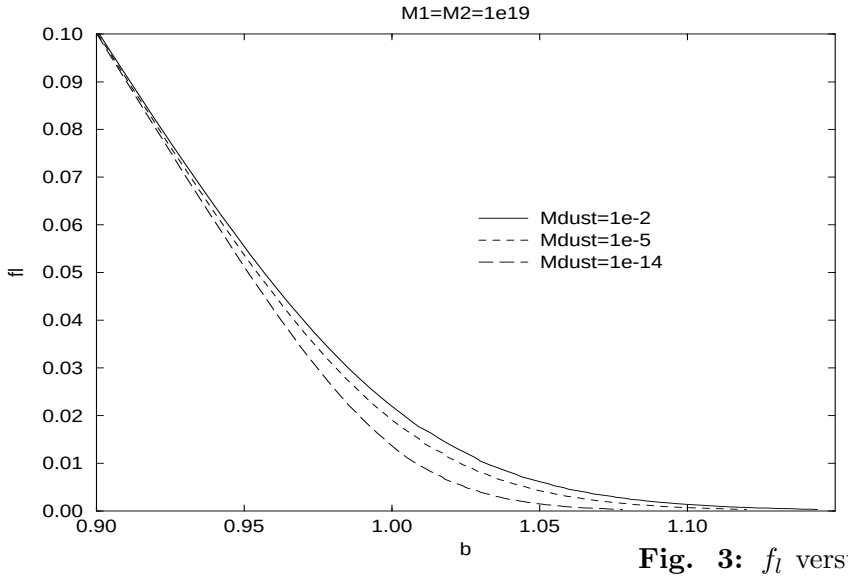



**Fig. 3:** $f_l$ versus $b$ for $M_1 = M_2 = 10^{19}\ kg$ and different values of $M_{dust}$.

Fig. 4 shows, for three different values of the projectile–to–target mass ratio, how the averaged escape velocity $V_{esc}$ defined by Eq.(10) (normalized to the surface escape velocity of $M_1$) varies as a function of $b$. As shown by Farinella *et al.* (1988), when $M_2 \ll M_1$ and $b$ is close to 1, the limiting value of the normalized escape velocity is about 0.77. The curves shown in Fig. 4 are somewhat different from that appearing in Fig. 1 of the above–mentioned paper, because (i) we now take into account the mass of the projectile (so the curves are shifted upward for higher values of $M_2/M_1$), and (ii) we have made the assumption that the largest fragment, whose mass is higher for smaller values of $b$, has always a negligible kinetic energy. Fig. 4 shows that taking a constant value $\approx 0.8$ for the normalized escape velocity after shattering impacts is normally a good approximation. As shown by Fig. 5, using $M_{dust} > 0$ has little influence on the escape velocity.

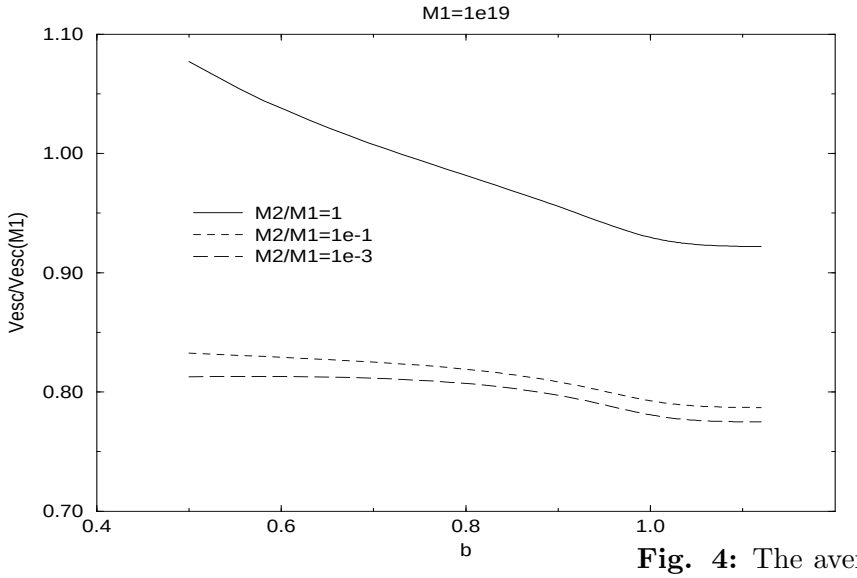



**Fig. 4:** The average fragment escape velocity, normalized to the surface escape velocity of $M_1$, versus $b$ for three different values of the projectile–to–target mass ratio.

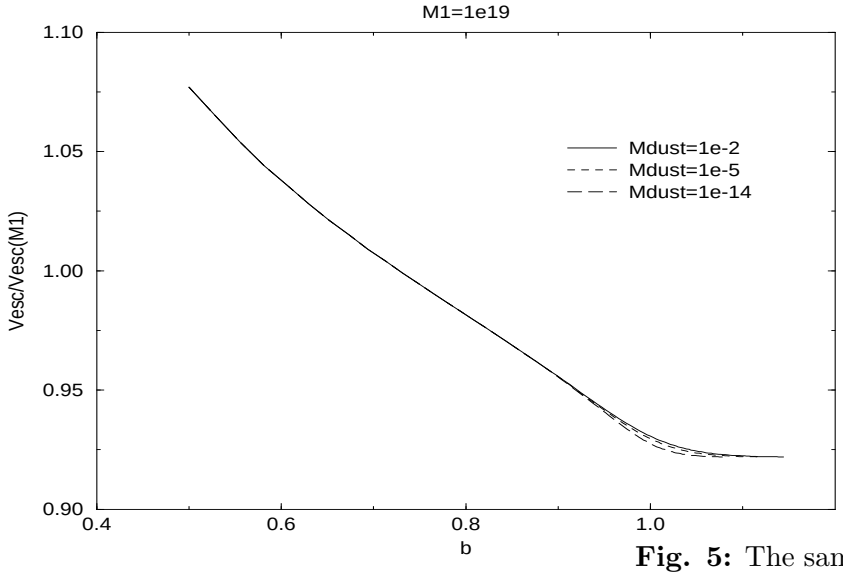



**Fig. 5:** The same as Fig. 4 for $M_2/M_1 = 1$, but for three different values of $M_{dust}$.

One of the most interesting aspects of the algorithm presented in this paper is that it allows us to model in several alternative ways the gravitational reaccumulation process, through the use of different fragment velocity vs. mass relationships, as discussed in Sec. 5. Qualitatively speaking, the reaccumulation process is due to the fact that, for targets having escape velocities of tens to hundreds of $m/s$, the amount of energy sufficient for fragmenting them is normally not enough to accelerate all the fragments to speeds exceeding $V_{esc}$. Thus the slower fragments fall back and reaccumulate onto the largest one, giving rise to a so–called *pile of rubble*. The relevance of this process for asteroid evolution has been discussed in detail more than a decade ago by Davis *et al.* (1979) and Farinella *et al.* (1982). But of course the quantitative aspects of it depend in a critical manner on the conversion efficiency of the projectile's kinetic energy into kinetic energy of fragments (i.e., the *anelasticity coefficient* $f_{ke}$), as well as on the way the available kinetic energy is partitioned among fragments of different masses. We shall focus here on the latter issue.

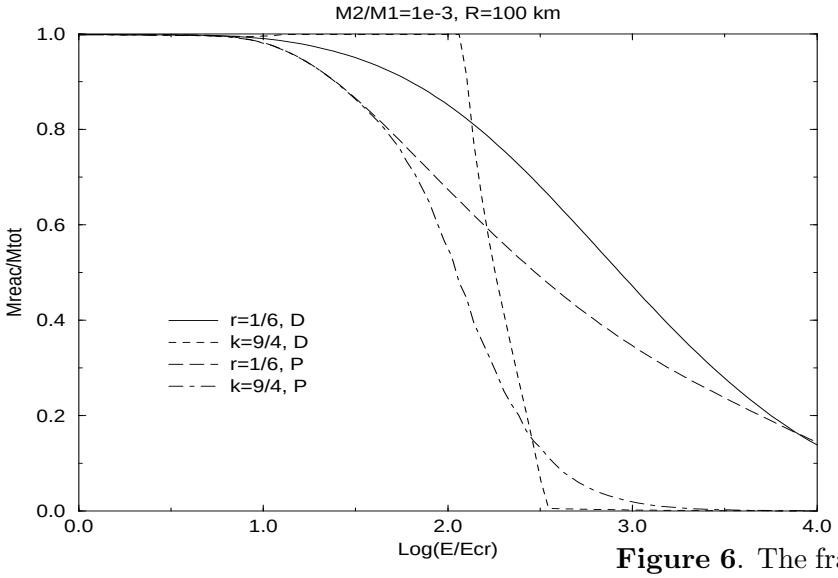



**Figure 6**. The fraction of reaccumulated mass for an impact with $M_2/M_1 = 10^{-3}$ against a target of radius $100\,km$, versus the logarithm of the impact energy. The latter parameter is normalized to the minimum energy required for fragmentation. Different fragment velocity vs. mass relationships are used (see text), with D standing for "deterministic model" and P for "probabilistic model".

Fig. 6 shows, for a target $100\,km$ in radius and $M_2/M_1 = 10^{-3}$, the fraction of reaccumulated mass versus the logarithm of the impact energy, normalized to the critical energy for fragmentation. Clearly, with our nominal parameter choice, there is a gap of more than two

orders of magnitude between the fragmentation threshold and the energy needed to disperse to infinity a significant fraction of the target mass. As a consequence, most asteroids of radius $\approx 100\,km$ have been probably shattered but not *disrupted* (i.e., have reaccumulated) as a consequence of the most energetic collisions they have undergone over the history of the solar system. Fig. 6 also shows that the reaccumulation process is confined to a narrow energy range if we assume $k = 9/4$, i.e. we require that for every velocity $V$ (exceeding a lower cutoff) the total mass ejected with higher speeds is $\propto V^{-9/4}$. The reason for this is that since $r = (1-b)/k$ (see Sec. 5) and $b$ is close to 1 (see Fig. 2), the exponent $r$ in this case is small, implying that the velocity of fragments has a weak dependence on their mass: thus, there is an abrupt transition between the case when nearly all the fragments are slower than $V_{esc}$ and the opposite one. The transition is less sharp in the probabilistic case, since in this case for every mass the fragment velocities have a significant chance variability — and it is much smoother when $r$ (instead of $k$) is fixed to the value 1/6, consistent with experimental data. In this case, even after a supercatastrophic impact a significant fraction of mass is always reaccumulated, because most of the mass resides in small fragments and is forced to move comparatively fast, so that

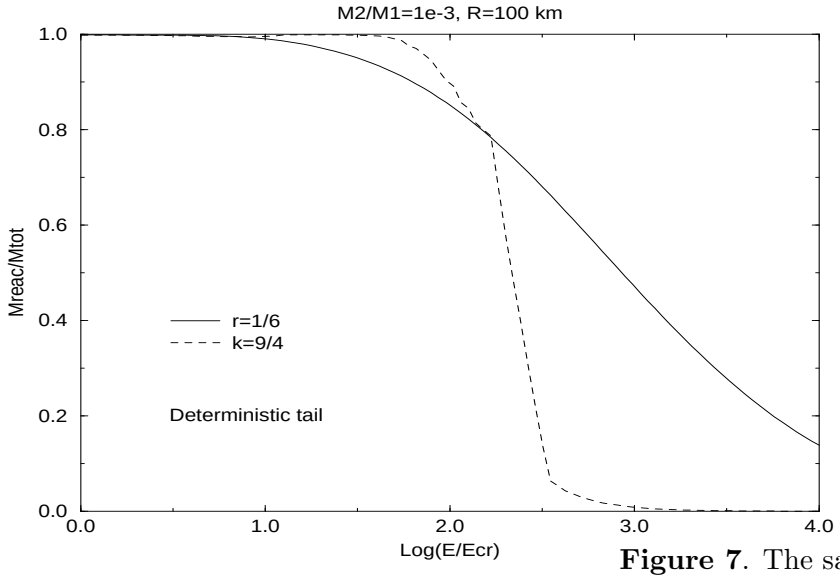



**Figure 7**. The same as Fig. 6, but treating always in a deterministic case the small–size tail of the fragment mass distribution.

conservation of energy causes the largest fragments to be slow and to reaccumulate. Fig. 7 shows what happens if the bodies in the small–size tail of the mass distribution (fragments smaller than about 20 $m$) are treated deterministically, as described in Sec. 5: while the $r = 1/6$ curve is unaffected, because in this case the tail fragments are always fast enough to

escape, there is a significant change in the $k = 9/4$ case.

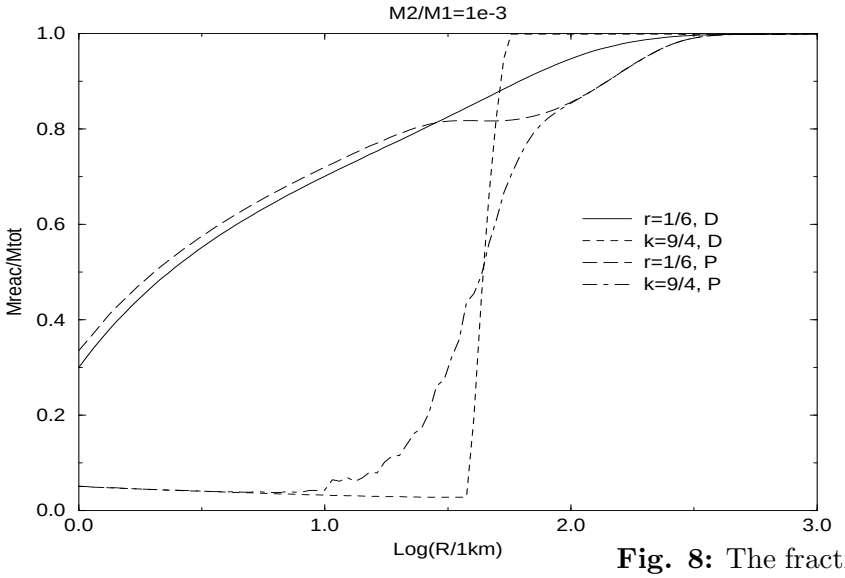



**Fig. 8:** The fraction of reaccumulated mass for a shattering impact with $M_2/M_1 = 10^{-3}$ versus the logarithm of the target radius. The same four models have been used as in Fig. 6.

A physically important consequence of the different behaviors of the $k = 9/4$ and the $r = 1/6$ models is apparent from Figs. 8 and 9, which give, for the same models used to draw Fig. 6 and 7, the fraction of reaccumulated mass versus target size. Again, for the $k = 9/4$ models (but less for the probabilistic ones), there is a comparatively sudden transition from negligible to substantial reaccumulation, which occurs when the velocity of most fragments becomes lower than $V_{esc}$. This would imply comparatively short collisional lifetimes of small asteroids — say, in the 1 to $50\,km$ diameter range. Also, in this case small asteroids would be unlikely to be *rubble piles*, but would mostly look like single, competent fragments. The opposite conclusion would follow from the $r = 1/6$ models, for which the figures show that even small bodies are always reaccumulated; recall, however, that when $V_{esc}$ is small only a small number of massive fragments may become reaggregated. This is possibly an effective mechanism to form $km$–sized contact or nearly–contact binary asteroids, such as recently observed by radar (Ostro *et al.*, 1990).

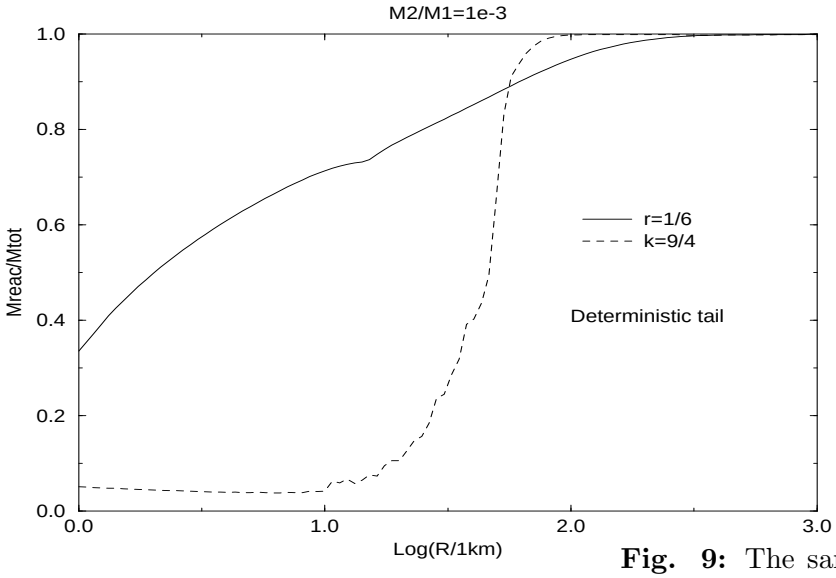



**Fig. 9:** The same as Fig. 8, but treating always in a deterministic case the small–size tail of the fragment mass distribution.

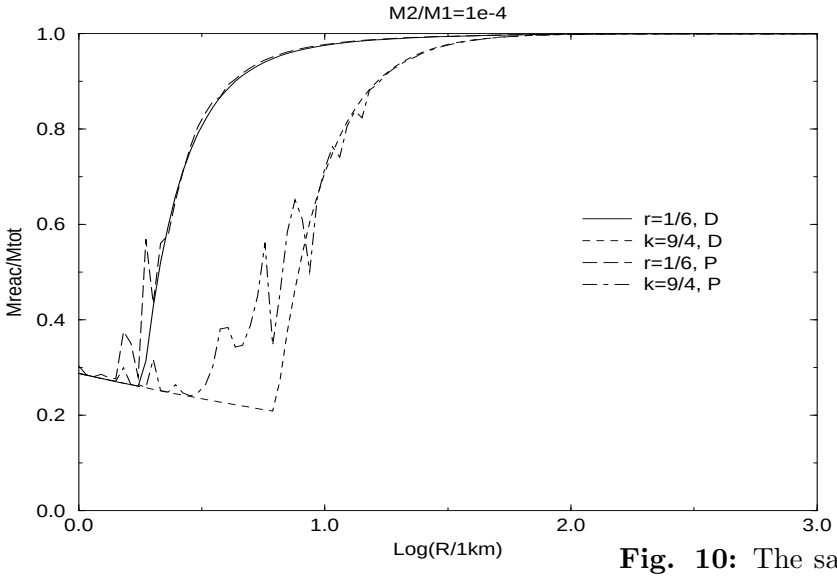



**Fig. 10:** The same as Fig. 8, but for $M_2/M_1 = 10^{-4}$.

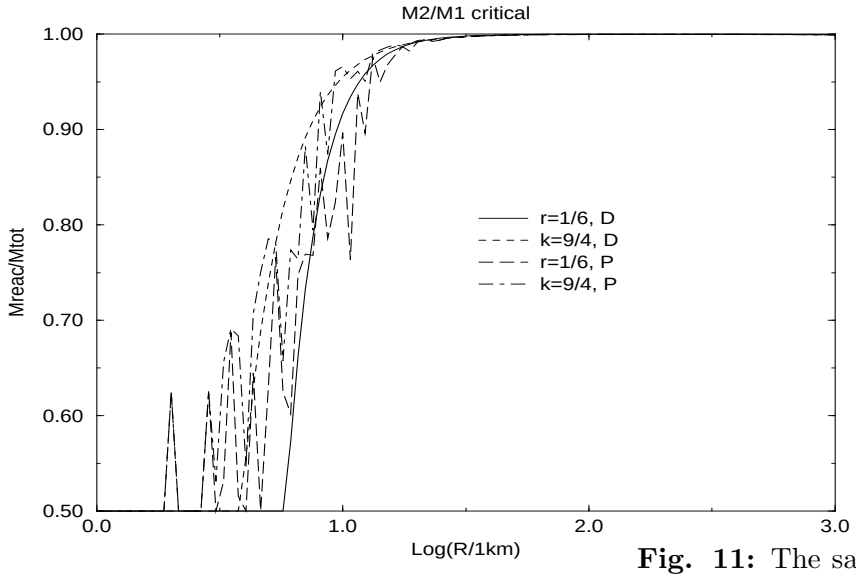



**Fig. 11:** The same as Fig. 8, but for $M_2/M_1$ just exceeding the threshold value for fragmentation.

For a smaller projectile–to–target mass ratio ($10^{-4}$, still sufficient for fragmentation), Fig. 10 shows that both the $r = 1/6$ models and the $k = 9/4$ models display a transition from complete fragment escape (with the chance exception of some large fragments in the probabilistic cases, giving rise to irregular curves) to substantial reaccumulation. Of course here the transition occurs at smaller target sizes (less than $10\,km$ radius), but there is a shift of about a factor 3 between the two types of models, due to the lower velocity of the big fragments in the $r = 1/6$ cases. For a barely shattering impact (Fig. 11), on the other hand, all the models yield similar results, with the transition occurring at a radius of $\approx 5\ km$ (plus or minus a factor 2) and large fluctuations in the probabilistic cases.

When $M_2/M_1 = 10^{-5}$, we are in the cratering case (see Sec. 4), with Eqs.(14) yielding between 2% and 3% of excavated mass (see Fig. 12; the crater mass increases with target size because of the decreasing impact strength in this size range). Again the reaccumulation of crater ejecta is effective at smaller sizes in the $r = 1/6$ models, and is anyway practically complete for bodies exceeding 100 $km$ diameter.

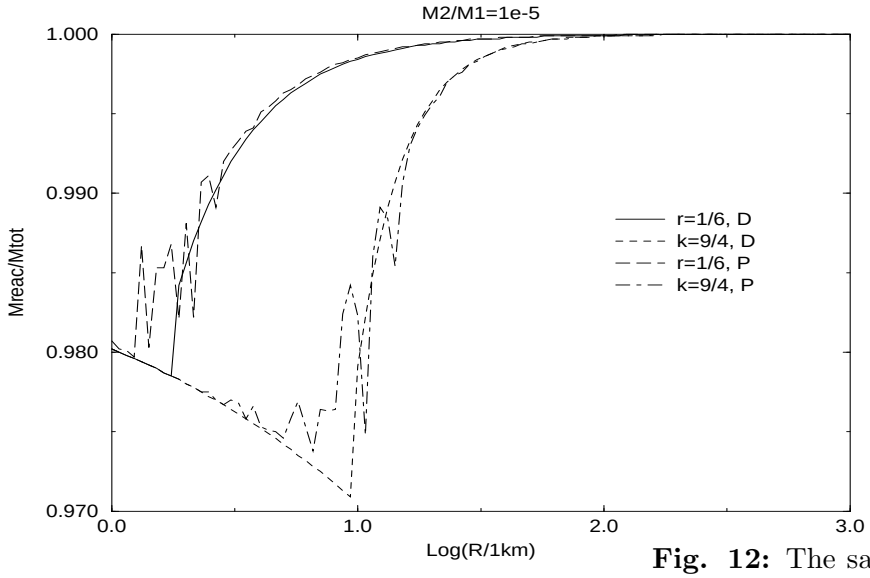



**Fig. 12:** The same as Fig. 8, but for $M_2/M_1 = 10^{-5}$. This mass ratio results in a cratering event.

These results provide some hints to the collisional history of 951 Gaspra, the asteroid about $13\,km$ in mean diameter recently approached and imaged by the Galileo probe (Belton *et al.*, 1992). Gaspra's global shape is strongly irregular, with two very large concavities (about $8\,km$ across) evident on its limb. However, the general appearance of the surface is fairly smooth, with no sharp edges and a relative paucity of $km$–sized craters (while several linear "grooves" are apparent). To explain these features and the distribution of colors and albedo variations, which hints to the presence of surface regolith, Greenberg *et al.* (1992) have suggested that Gaspra's interior is a rubble pile, overlain by a substantial regolith layer; the large concavities would represent giant impact basins, formed by events close to the catastrophic break–up threshold. This interpretation is consistent with our models, in particular those including a significant ejecta velocity vs. mass dependence, as they show that even for bodies as small as Gaspra a large fraction of the material ejected after impacts can be reaccumulated by self–gravity both in the shattering and in the cratering regime (see Figs. 10–12). Thus Gaspra may have been originated in a shattering event followed by the reaggregation of a few large blocks; the subsequent formation of some giant craters may have produced the surface regolith and erased the small–scale topography.

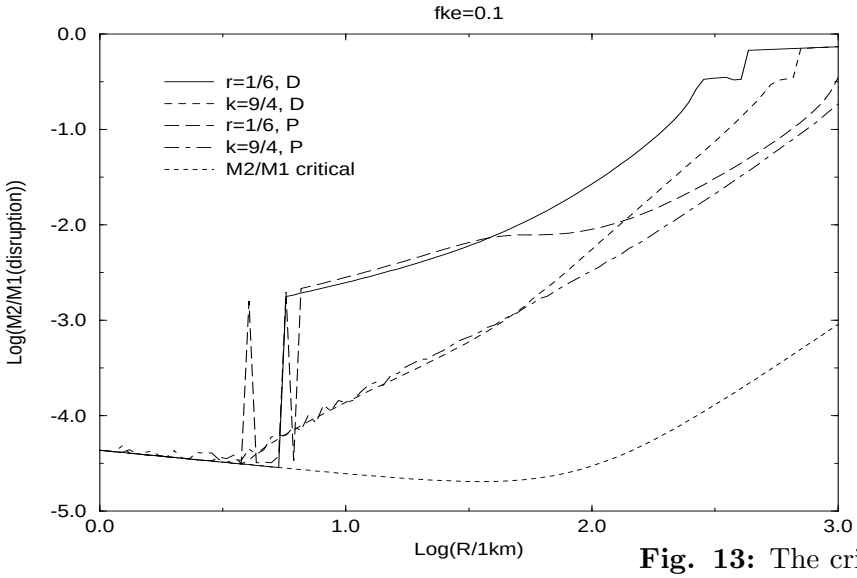



**Fig. 13:** The critical mass ratio for *disruption* versus the logarithm of the target size, for the four models used in the previous figures. The bottom (dashed) line is the threshold mass ratio for *fragmentation*.

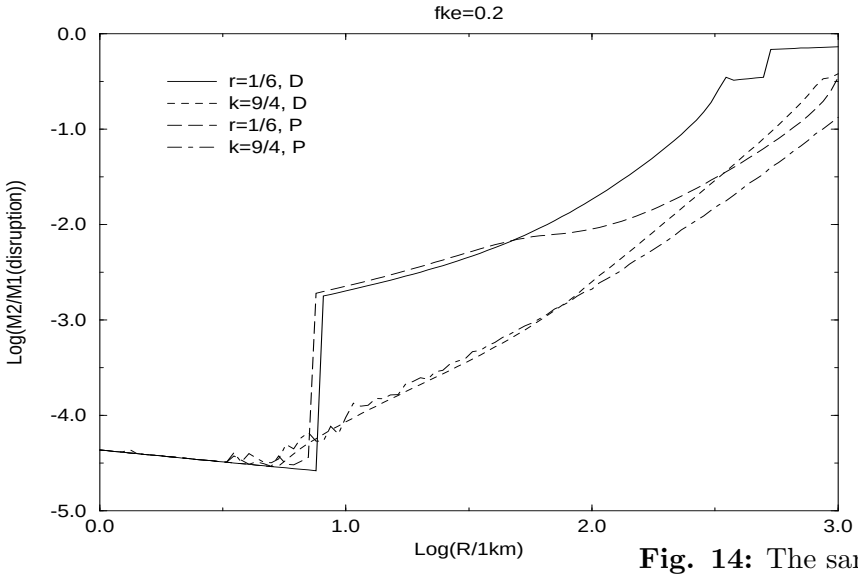



**Fig. 14:** The same as Fig. 13, but with $f_{ke} = 0.2$.

Figs. 13 and 14 show another type of diagram, which should be compared with Fig. 5 of Davis *et al.* (1985) and Fig. 5 of Housen and Holsapple (1990). Here the lower (dashed) curve corresponds to the critical projectile–to–target mass ratio for fragmentation, which is a function of size because of the strength scaling relationship (1). The other curves give the threshold mass ratio for collisional disruption, with one half of the target mass dispersed to infinity. While for small asteroids, whose gravity is not efficient in reaccumulating fragments, the fragmentation and disruption thresholds coincide, for asteroids $\approx 100\,km$ across there is a large gap between the two thresholds, due to the gravitational reaccumulation process. In this case, the curves show that the for the $r = 1/6$ models a sudden discontinuity appears when a critical size is exceeded, while for the $k = 9/4$ case there is a more gradual transition. This behaviour can be explained as follows: in the $k = 9/4$ case, when all the fragment velocities are close to each other, a modest specific energy (i.e., mass ratio) increase may be enough to exceed $V_{esc}$; on the other hand, in the $r = 1/6$ case, after self–gravity has become strong enough to reaccumulate the largest fragments, a moderate increase in the specific energy causes a mass shift to fragments of smaller sizes, which receive a larger fraction of the available energy, thus preventing the largest fragments from starting faster (and being able to escape).

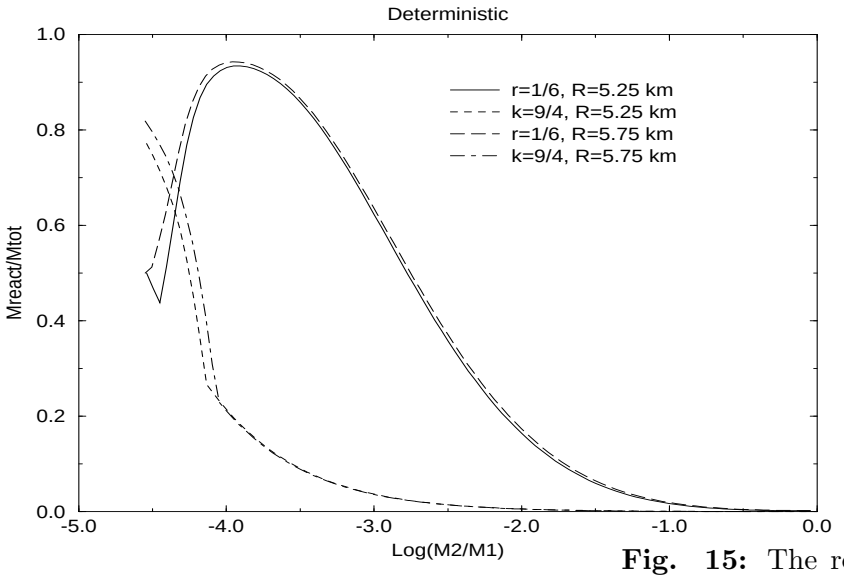



**Fig. 15:** The reaccumulated mass fraction versus $M_2/M_1$ for the deterministic $r = 1/6$ and $k = 9/4$ models and for two values of the target radius, close to each other but on the two sides of the critical value for which reaccumulation starts to be effective.

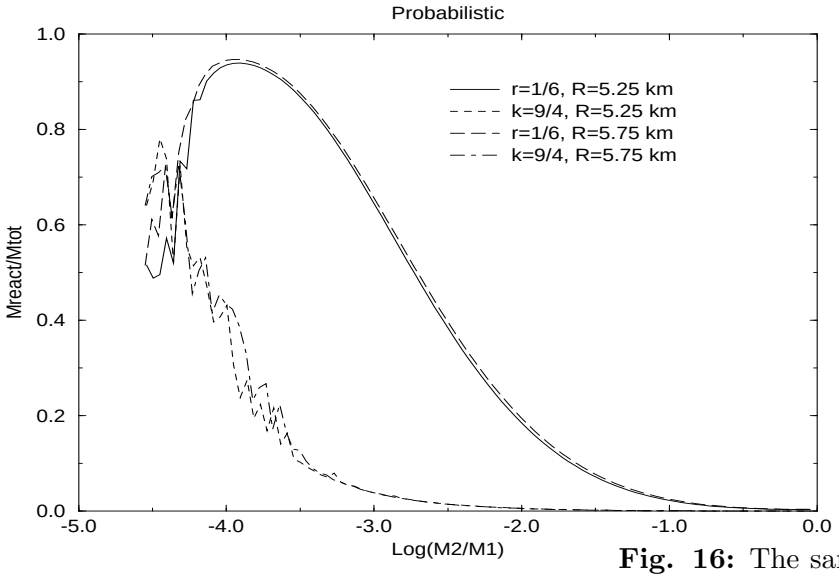



**Fig. 16:** The same as Fig. 15 but for the probabilistic models.

This sudden "runaway reaccumulation" phenomenon in the $r = 1/6$ models is illustrated in Figs. 15 and 16. Here the reaccumulated mass fraction is plotted versus $M_2/M_1$ for two values of the target radius, close to each other but on the two sides of the critical value. While with the $k = 9/4$ model the reaccumulated mass always decreases in a nearly–monotonic way when the impact gets more energetic, the opposite happens with the $r = 1/6$ model: for the lower size, when the fragmentation threshold is exceeded by a small amount, there is still a small drop of the reaccumulated mass fraction to values $< 1/2$, but this drop disappears at the larger size, since in this case more energetic collisions result initially into more effective reaccumulation; only for much higher mass ratios the reaccumulated fraction decreases, falling again below 1/2 only for values of $M_2/M_1$ exceeding the fragmentation threshold by almost two orders of magnitude.

This is just an example of what we believe is an important finding of this work: for asteroidal–sized bodies, even keeping fixed the (poorly known) parameter $f_{ke}$, the predicted collisional outcomes depend in a sensitive way on the adopted fragment velocity vs. mass model. Take for instance the formation of an asteroid family from a target $\approx 200\,km$ across, which requires a collisional event capable of ejecting to infinity a substantial fraction of the mass of the parent body. Fig. 13 shows that, even keeping $f_{ke} = 0.1$ and limiting ourselves to the four models described earlier, the required projectile–to–target mass ratio ranges from about 1/300 ($k = 9/4$, probabilistic) to 1/30 ($r = 1/6$, deterministic), implying a projectile

diameter between 30 $km$ and 65 $km$. This entails about a factor 4 of uncertainty in the available number of projectiles ($\approx$ 1100 versus 300, respectively; see Fig. 3 of Cellino *et al.*, 1991), with a similar uncertainty in the collision lifetimes and therefore in the expected number of families. With the average collision rate derived by Farinella and Davis (1992; see Sec. 1) and the current population of main–belt asteroids, the expected number of disrupted 200 $km$–sized parent bodies may range between 1.5 and 5. The agreement with the results of family searches such as that of Zappalà *et al.* (1990) is good, but reducing the uncertainty would be very important.

# 7 Conclusions and future work

The algorithm described in this paper is simple enough to be applied in a quick and effective way to a variety of studies on the collisional evolution of small solar system bodies. In Sec. 6 we have presented some examples related to the evolution of asteroids and the origin of asteroid families, but similar analyses — with comparable resulting uncertainties — could be performed for other cases of catastrophic or nearly–catastrophic impacts occurred in the solar system. A list might include: the formation of Vesta's family from a giant crater preserving intact most of the basaltic crust of this asteroid (Davis *et al.*, 1985; Rosa *et al.*, 1992); the collisional origin of the dust bands discovered by IRAS (Sykes *et al.*, 1989); the formation of reaccumulated, rapidly spinning (or binary) asteroids (Farinella *et al.*, 1982, 1985; Weidenschilling *et al.*, 1989); the origin of very large craters and impact basins observed on many natural satellites (Chapman and McKinnon, 1986); the collisional disruption of Hyperion (Farinella *et al.*, 1983); the formation of planetary rings from disrupted satellites (Harris, 1984).

Of course, many input parameters of the collisional models — and also the validity of several empirical relationships used in Secs. 2 to to 6 — are still subject to considerable uncertainty. We are confident that this uncertainty will be reduced in the near future as result of the combination between further laboratory experiments and complex numerical simulations of the collisional disruption process. We think that the algorithm presented here is simple and flexible enough to easily incorporate new theoretical and experimental results. Once this is done, the output of our code can be written as a three–dimensional matrix, specifying how many objects end up in the $i$–th mass bin following a collision between two bodies lying in the $j$-th and $k$–th mass bins. This matrix can then be read by a program aimed at simulating the time evolution of an initial asteroid (or planetesimal, or planetary ring particle) population by computing in any time step for every pair of mass bins the expected number of collisions, as in the work of Greenberg *et al.* (1978), Davis *et al.* (1979, 1985, 1989), and Farinella *et al.* (1992). We are going to present in a forthcoming paper some results on the sensitivity of the outcomes of the simulated asteroid collisional evolution (in terms of the observable size distribution), with

respect to the fragment velocity vs. mass correlation models, as discussed in Sec. 6, and other poorly known input parameters. We also plan to develop in the future, along the same lines discussed here, a more comprehensive algorithm capable of computing impact outcomes, not only in terms of sizes but also including spins, i.e., angular momenta of rotation — for the physics involved in this problem see Davis *et al.* (1989), Cellino *et al.* (1990) and Farinella *et al.* (1992).

# Acknowledgments

We are grateful to the reviewer R. Greenberg for pointing to us the recent hydrocode modelling work, and to A. Campo Bagatin, A. Cellino, D.R. Davis, G. Martelli, P. Paolicchi and V. Zappalà for useful discussions and comments. J.–M. Petit's stay in Pisa in 1991 was supported through the CNRS–CNR researcher exchange agreement. P. Farinella acknowledges financial support by the Observatoire de la Côte d'Azur and the EEC research grant no. SC1–0011–C(GDF).

# References

[1] Asphaug, E., Benz, W., & Melosh, H.J., 1992. *Asteroid impact studies with SPH 3D (abstract).* Bull. Amer. Astr. Soc., in press,

[2] Belton, M.J.S., Veverka, J., Thomas, P., Helfenstein, P., Simonelli, D., Chapman, C., Davies, M.E., Greeley, R., Greenberg, R., Head, J., Murchie, S., Klaasen, K., Johnson, T.V., McEwen, A., Morrison, D., Neukum, G., Fanale, F., Anger, C., Carr, M., Pilcher, C., 1992. *Galileo encounter with 951 Gaspra: First pictures of an asteroid.* Science **257**, pp. 1647–−1652.

[3] Cellino, A., Zappalà, V., Davis, D.R., Farinella, P., & Paolicchi, P., 1990. *Asteroid collisional evolution, I. Angular momentum splash: Despinning asteroids through catastrophic collisions.* Icarus **87**, pp. 391–−402.

[4] Cellino, A., Zappalà, V., & Farinella, P., 1991. *The size distribution of main–belt asteroids from IRAS data.* Mon. Not. R. astr. Soc. **253**, pp. 561–−574.

[5] Chapman, C.R., & McKinnon, W.B., 1986. *Cratering on planetary satellites.* in *Satellites,* eds. J.A. Burns and M.S. Matthews, Univ. of Arizona Press, pp. 492–−580.

[6] Chapman, C.R., Paolicchi, P., Zappalà, V., Binzel, R.P, & Bell, J.F., 1989. *Asteroid families: Physical properties and evolution.* in *Asteroids II,* eds. R.P. Binzel, T. Gehrels & M.S. Matthews, Univ. of Arizona Press, pp. 386–−415.

[7] Davis, D.R., 1990. *On social disintegration in the solar system: The disruption of families.* presented at the Third International Workshop on Catastrophic Disruption of Small Solar System Bodies (Kyoto, Japan, September 1990).

[8] Davis, D.R., Chapman, C.R., Greenberg, R., & Harris, A.W., 1979. *Collisional evolution of asteroids: Populations, rotations, and velocities.* in *Asteroids,* ed. T. Gehrels, Univ. of Arizona Press, pp. 528–−557.

[9] Davis, D.R., Chapman, C.R., Weidenschilling, S.J., & Greenberg, R., 1985. *Collisional history of asteroids: Evidence from Vesta and the Hirayama families.* Icarus **62**, pp. 30–−53.

[10] Davis, D.R., Farinella, P., Paolicchi, P., & Zappalà, V. (eds.), 1986. *Catastrophic Disruption of Asteroids and Satellites.* J. Ital. Astron. Soc. **57**, n. 1.

[11] Davis, D.R., Farinella, P., Paolicchi, P., Weidenschilling, S.J., & Binzel, R.P., 1989. *Asteroid collisional history: Effects on sizes and spins.* in *Asteroids II,* eds. R.P. Binzel, T. Gehrels & M.S. Matthews, Univ. of Arizona Press, pp. 805–−826.

[12] Davis, D.R., & Ryan, E.V., 1990. *On collisional disruption: Experimental results and scaling laws.* Icarus **83**, pp. 156–−182.

[13] Dobrovolskis, A., & Burns, J.A., 1984. *Angular momentum drain: A mechanism for despinning asteroids.* Icarus **57**, pp. 464–−476.

[14] Farinella, P., Paolicchi, P., & Zappalà, V., 1982. *The asteroids as outcomes of catastrophic collisions.* Icarus **52**, pp. 409–−433.

[15] Farinella, P., Milani, A., Nobili, A.M., Paolicchi, P., & Zappalà, V., 1983. *Hyperion: Collisional disruption of a resonant satellite.* Icarus **54**, pp. 353–−360.

[16] Farinella, P., Paolicchi, P., & Zappalà, V., 1985. *Original abundance of asteroids from their present rotational properties.* Mon. Not. R. astr. Soc. **216**, pp. 565–−570.

[17] Farinella, P., Paolicchi, P., Cellino, A., & Zappalà, V., 1988. *Escape velocity for fragments of collisionally shattered solar system bodies.* Bull. Obs. Astron. Belgrade **138**, pp. 88–−90.

[18] Farinella, P., & Davis, D.R., 1992. *Collision rates and impact velocities in the main asteroid belt.* Icarus **97**, pp. 111–−123.

[19] Farinella, P., Davis, D.R., Paolicchi, P., Cellino, A., & Zappalà, V., 1992. *Asteroid collisional evolution: An integrated model for the evolution of asteroid rotation rates.* Astron. Astrophys. **253**, pp. 604–−614.

[20] Fujiwara, A., Kamimoto, G., & tsukamoto, A., 1977. *Destruction of basaltic bodies by high–velocity impact.* Icarus **31**, pp. 277–−288.

[21] Fujiwara, A., & Tsukamoto, A., 1980. *Experimental study on the velocity of fragments in collisional breakup.* Icarus **44**, pp. 142–−153.

[22] Fujiwara, A., Cerroni, P., Davis, D.R., Ryan, E., Di Martino, M., Holsapple, K., & Housen, K., 1989. *Experiments and scaling laws on catastrophic collisions.* in *Asteroids II,* eds. R.P. Binzel, T. Gehrels & M.S. Matthews, Univ. of Arizona Press, pp. 240–−265.

[23] Gault, D.E., Shoemaker, E.M., & Moore, H.J., 1963. *Spray ejected from the lunar surface by meteoroid impact.* NASA Tech. Note **D–1767**,

[24] Greenberg, R., Wacker, J.F., Hartmann, W.K., & Chapman, C.R., 1978. *Planetesimals to planets: Numerical simulation of collisional evolution.* Icarus **35**, pp. 1–−26.

[25] Greenberg, R., Nolan, M.C., Bottke, W.F., Kolvoord, R.A., Veverka, J., & the Galileo imaging team, 1992. *Collisional and dynamical evolution of Gaspra (abstract).* Bull. Amer. Astr. Soc., in press,

[26] Harris, A.W., 1984. *The origin and evolution of planetary rings.* in *Planetary Rings,* eds. R. Greenberg & A. Brahic, Univ. of Arizona Press, pp. 641–−659.

[27] Hartmann, W.K., 1969. *Terrestrial, lunar and interplanetary rock fragmentation.* Icarus **10**, pp. 201–−213.

[28] Hartmann, W.K., 1988. *Impact strength and energy partitioning in impacts into finite solid targets.* Lunar Planet. Sci. **XIX**, pp. 451–−452.

[29] Housen, K.R., & Holsapple, K.A., 1990. *On the fragmentation of asteroids and planetary satellites.* Icarus **84**, pp. 226–253.

[30] Housen, K.R., Schmidt, R.M., & Holsappple, K.A., 1991. *Laboratory simulations of large scale fragmentation events.* Icarus **94**, pp. 180–−190.

[31] Kresàk, Ľ., 1977. *Mass content and mass distribution in the asteroid system.* Bull. Astron. Inst. Czech. **28**, pp. 65–−82.

[32] Nakamura, A., & Fujiwara, A., 1991. *Velocity distribution of fragments formed in a simulated collisional disruption.* Icarus **92**, pp. 132–−146.

[33] Nolan, M.C., Asphaug, E., & Greenberg, R., 1992. *Numerical simulation of impacts on small asteroids (abstract).* Bull. Amer. Astr. Soc., in press,

[34] Öpik, E.J., 1951. *Collision probabilities with the planets and the distribution of interplanetary matter.* Proc. Royal Irish Acad. **54A**, pp. 164–−199.

[35] Ostro, S.J., Chandler, J.F., Hine, A.A., Rosema, K.D., Shapiro, L.L., & Yeomans, D.K., 1990. *Radar images of asteroid 1989 PB.* Science **248**, pp. 1523–−1528.

[36] Paolicchi, P., Cellino, A., Farinella, P., & Zappalà, V., 1989. *A semiempirical model of catastrophic breakup processes.* Icarus **77**, pp. 187–−212.

[37] Paolicchi, P., Verlicchi, A., & Cellino, A., 1992. *A new semiempirical model for catastrophic impact processes (abstract).* Bull. Amer. Astr. Soc., in press,

[38] Piotrowski, S.I., 1953. *The collisions of asteroids.* Acta Astron. Ser. A **6**, pp. 115–−138.

[39] Rosa, F., Cellino, A., Zappalà, V., & Farinella, P., 1992. *Crater ejecta from Vesta.* in preparation

[40] Ryan, E.V., Asphaug, E., & Melosh, H.J., 1990. *Hydrocode predictions of collisional outcomes: Effects of target size.* Lunar Planet. Sci. **XXI**, pp. 1155–−1156.

[41] Ryan, E.V., & Melosh, H.J., 1992. *Understanding asteroid collisions: The effect of target size on collisional outcome (abstract).* Bull. Amer. Astr. Soc., in press,

[42] Stöffler, D., Gault, D.E., Wedekind, J., and Polkowski, G., 1975. *Experimental hypervelocity impacts into quartz sand: Distribution and shock metamorphism of ejecta.* J. Geophys. Res. **80**, pp. 4062–−4077.

[43] Sykes, M.V., Greenberg, R., Dermott, S.F., Nicholson, P.D., and Burns, J.A., 1989. *Dust bands in the asteroid belt.* in *Asteroids II,* eds. R.P. Binzel, T.Gehrels *&* M.S.Matthews, Univ. of Arizona Press, pp. 336–−367.

[44] Weidenschilling, S.J., Paolicchi, P., *&* Zappalà, V., 1989. *Do asteroids have satellites?.* in *Asteroids II,* eds. R.P.Binzel, T.Gehrels and M.S.Matthews, Univ. of Arizona Press, pp. 643–−658.

[45] Zappalà, V., Farinella, P., Knežević, Z., *&* Paolicchi, P., 1984. *Collisional origin of asteroid families: Mass and velocity distributions.* Icarus **59**, pp. 261–285.

[46] Zappalà, V., Cellino, A., Farinella, P., *&* Knežević, Z., 1990. *Asteroid families. I. Identification by hierachical clustering and reliability assessment.* Astron. J. **100**, pp. 2030–−2046.

**[13]**

# "Fragment ejection velocities and the collisional evolution of asteroids."

CAMPO-BAGATIN, A., FARINELLA, P. *&* PETIT, J-M.

*Planet. Space Sci.* **42** (1994)

# FRAGMENT EJECTION VELOCITIES AND THE COLLISIONAL EVOLUTION OF ASTEROIDS

A. CAMPO BAGATIN[1], P. FARINELLA[1,2] AND J.–M. PETIT[2]

*1. Gruppo di Meccanica Spaziale, Dipartimento di Matematica, Università di Pisa, Via Buonarroti 2, 56127 Pisa, Italy*

*2. Observatoire de la Côte d'Azur, B.P. 229, 06304 Nice Cedex 04, France*

Submitted to **Planetary and Space Science**: August 1993
Revised version: January 1994

Send editorial communications to:
Paolo Farinella
OCA — Observatoire de Nice
B.P. 229
06304 NICE CEDEX 4, FRANCE
E–mail ZAPPALA@OBS-NICE.FR
Tel. 92003049, Fax 92003033

**Abstract.** We have applied the algorithm developed by Petit and Farinella (1993) to model the outcomes of impacts between asteroids of different sizes, to show that a crucial feature of these models is the assumed relationship between velocity and mass of fragments ejected after a shattering impact. Not only how the mean velocity depends upon mass is important to determine the extent of fragment reaccumulation, but also the distribution of velocities about the mean values. The available experimental evidence on this issue is still sparse, and does not constrain the collisional models well enough to allow us to make reliable predictions on the outcomes of impacts between bodies of size much larger than the laboratory targets. As a consequence, when the collisional outcome models are used as an input for simulations of the asteroid collisional history since the origin of the solar system, the results show a strong sensitivity to the assumed velocity vs. mass relationship. This sensitivity is stronger in the diameter range (a few tens to a few hundreds of km) where the self–gravitational reaccumulation of fragments is most effective, but may extend also to much smaller sizes.

## 1. Introduction

Developing realistic models of the outcomes of asteroidal impacts is a basic prerequisite to understanding the collisional history of asteroid sizes and rotations since the end of the accretion phase $4.5 \times 10^9$ yr ago (Davis et al. 1989). Such impacts occur at velocities of several km/s, and most current knowledge on their outcomes is based on extrapolations of small–scale laboratory experiments (Davis et al. 1986; Fujiwara et al. 1989). Recently, it has become possible to measure at the same time the size and the ejection velocity for large numbers of fragments generated in the impact or explosive break–up of a solid target (Davis and Ryan 1990; Nakamura and Fujiwara 1991; Nakamura et al. 1992; Giblin et al. 1994), and a correlation beween these two quantities has been detected. For large–scale cratering events, a similar correlation had been found by Vickery (1986, 1987), through an analysis of the distribution of secondary craters surrounding some large impact craters on Mercury, the Moon and Mars. This is potentially an important finding, because fragment ejection velocities determine the extent to which material ejected from the target can be reaccumulated by self–gravity, so as to form a regolith layer (for cratering events) or a so–called *pile of rubble* (in the case of break–up; see Davis et al. 1979, Farinella et al. 1982).

Petit and Farinella (1993) have presented a self–consistent numerical algorithm aimed at predicting the collisional outcomes, based on a set of model input parameters which can be estimated from the available experimental evidence, and including the possible gravitational reaccumulation of ejected fragments whose velocity is less than a suitably defined escape velocity. The possible correlation between fragment ejection velocity and mass has been taken into account in different ways, including a probabilistic one. In that paper the effectiveness of the gravitational reaccumulation process has been analyzed for different choices of the collisional parameters and different assumptions on the relationship between fragment speed and mass. Both the transition size beyond which a large fraction of the fragment mass is likely to be reaccumulated and the collision energy needed to disperse "to infinity" most of the fragments were found to be sensitive functions of the assumed fragment velocity vs. mass relationship.

The purpose of this paper is that of assessing how much this sensitivity may affect the final size distributions of asteroids obtained from collisional evolution codes such as those described in Greenberg et al. (1978), Davis et al. (1979, 1985, 1989, 1994), Farinella et al. (1992) and Campo Bagatin et al. (1994). These codes simulate the collisional history of asteroids by evolving in time the numbers of objects residing in a set of discrete size bins. At any given time step, the expected number of collisions between bodies belonging to any pair of bins is computed, and in each case the collisional outcome algorithm provides the size distribution of the objects resulting from the impacts, which are then redistributed in the bins. In this way, given an initial size distribution (presumably resulting from the primordial accretion phase) and a suitable set of impact response parameters, the size distribution resulting from $4.5 \times 10^9$ yr of evolution can be calculated, and then compared with the real one as determined from astronomical observations.

The remainder of this paper is organized as follows. In Sec. 2 we recall briefly the main assumptions embedded in our collisional outcome model and the adopted parameter choice. In Sec. 3 we discuss some results from this model. In Sec. 4 we present several

numerical simulations of the collisional evolution of the entire asteroid population, and comment on the importance of the assumed fragment velocity vs. mass relationship. Some conclusions from this work and the remaining open problems are summarized in Sec. 5.

## 2. The collisional outcome model

The collisional outcome algorithm of Petit and Farinella works in the following way. For each colliding body the available impact energy per unit volume is compared with a fragmentation threshold $S$ (called *impact strength*), which is assumed to vary with size according to the strain–rate scaling theory of Housen and Holsapple (1990) and to include a gravitional self–compressional term in agreement with Davis et al. (1985) and Housen et al. (1991); this term is multiplied by a free input parameter $\gamma$. When a body is shattered, the mass ratio between the largest fragment and the parent body $f_l$ is assumed to be a decreasing function of the impact energy according to an empirical relationship determined by Fujiwara et al. (1977), and the fragment mass distribution is modelled by a truncated power law whose lower cutoff is an input parameter $M_{dust}$ and whose exponent ($-b$ for the cumulative distribution) becomes larger and larger as $f_l$ decreases. On the other hand, in the cratering case the fragment mass distribution is assumed to have a fixed exponent, and the total mass excavated from the crater is determined from an empirical formula relating it to the impact energy (for small craters, this is a simple linear relationship; the ratio $\alpha$ between crater mass and projectile energy is called *crater excavation coefficient*, and is another input parameter of the model).

In order to determine how many fragments escape from the gravitational well of the colliding bodies, an *effective escape velocity* $V_{esc}$ is computed by a suitable energy balance equation, assuming that the largest fragment has always a negligible kinetic energy and taking into account the self–gravitational energy of the fragments. It is assumed that the total kinetic energy of the fragments is given by a fixed fraction of the projectile's kinetic energy (this fraction is the so–called *anelasticity parameter* $f_{KE}$), and then this constraint is used in computing their velocity distribution, as explained below. The fragments whose velocity exceeds $V_{esc}$ are assumed to escape "to infinity", the other ones are reaccumulated onto the largest fragment.

Most models of asteroid collisional evolution (e.g., Davis et al. 1985, 1989) neglect any possible correlation between the ejection velocity and the mass of the fragments (both from shattering and for cratering events), and just assume that the fraction of the total fragment mass ejected with speeds exceeding some given value $v$ is $(v/V_{min})^{-k}$, $V_{min}$ being a fixed lower cutoff to the ejection velocity. Such a relationship was observed for crater ejecta (with $k \approx 9/4$) by Gault et al. (1963); see Housen (1992) for a recent re–analysis of these experiments. On the other hand, recent experimental evidence (Davis and Ryan, 1990; Nakamura and Fujiwara, 1991; Nakamura et al. 1992) indicates that some correlation between fragment velocity and mass does exist, with the largest fragments on average going somewhat slower; this is also predicted by the semiempirical models of Paolicchi et al. (1989, 1992). It is easy to show (see e.g. Nakamura and Fujiwara 1991) that in this case the relationship given above for the mass fraction ejected at speeds $> v$ still holds provided the fragment speeds are proportional to the power $-r$ of their mass $m$, with $r = (1-b)/k$. The experimental results appear to be consistent with $r \approx 1/6$, although the dispersion of

the data is considerable and smaller values of $r$ cannot be ruled out (Takagi et al. 1992).

Petit and Farinella (1993) have used such a power relationship in either a *deterministic* way (i.e., assuming that a fragment of given mass has exactly the velocity value predicted by the empirical formula) or a *probabilistic* way (see below). Either $r$ or $k$ can be chosen as the input parameter specifying the adopted velocity vs. mass relationship, with the fragment velocity always bound to be less than a higher cutoff $V_{max}$.

In the *probabilistic case*, for every fragment mass $m$ the $V_0 \propto m^{-r}$ relationship was assumed to hold for the r.m.s. value $V_0 = \sqrt{\langle v^2 \rangle}$ of a Maxwellian probability distribution function for the fragment velocity $v$:

$$P(v; V_0)dv = \sqrt{\frac{2}{\pi}} \, \frac{3\sqrt{3}v^2}{V_0^3} \, \exp\left( - \frac{3v^2}{2V_0^2} \right) dv \ , \tag{1}$$

The choice of a Maxwellian velocity distribution had the following motivations: (i) it roughly matches the (limited) experimental data available, which show a large dispersion of the velocities measured for every mass, rather than a strong correlation; (ii) a Maxwellian distribution for the magnitude of the velocity vector arises whenever the three components of it follow Gaussian distributions with zero averages and equal sigmas — the simplest assumption for a nearly–isotropic velocity field; (iii) Eq. (1) is consistent with a fairly simple method to compute the individual fragment velocities, as the mean kinetic energy for bodies of mass $m$ is simply given by $mV_0^2/2$, so the energy conservation equation to derive the normalization of the $V_0$ vs. $m$ relationship is the same as in the deterministic case. Thus for each fragment an ejection velocity value is drawn at random using the distribution (1), and is compared to $V_{esc}$ to determine whether the fragment escapes or is reaccumulated (actually, this is done only when mass bins containing no more than 100 fragments are considered; otherwise Eq. (1) is integrated to obtain the fraction of escaping fragments for the given mass). Thus, whereas in the deterministic case when $r > 0$ all the fragments smaller than some threshold mass (corresponding to $v = V_{esc}$) are always reaccumulated, the transition is more gradual in the probabilistic case.

In this work, we have adopted the following "standard" set of input parameter values: density $\rho = 2500$ kg/m$^3$; impact velocity $V_{rel} = 5.85$ km/s; lower cutoff of the fragment mass $M_{dust} = 10^{-5}$ kg; higher cutoff of the fragment ejection velocity $V_{max} = 5$ km/s; pre–scaling impact strength (at laboratory target sizes of 0.2 m) $S_0 = 10^6$ J/m$^3$ ($S$ was then scaled up assuming a value $\gamma = 1$ for the coefficient of the self–compressional term); anelasticity parameter $f_{KE} = 0.1$; crater excavation coefficient $\alpha = 10^{-4}$ s$^2$/m$^2$. This parameter choice is derived from the available astronomical and experimental evidence, as discussed in Fujiwara et al. (1989) and Davis et al. (1989). In particular, the $S_0$ value is within the typical range — spanning about an order of magnitude — observed for "rocky" materials such as basalt, granite, etc. (see Fujiwara et al. 1989, Fig. 2; note that Petit and Farinella 1993 adopted a value 3 times larger), whereas the value of $f_{KE}$ is some 5 times larger than inferred from the experiments, but consistent with the properties of asteroid families (see Davis et al. 1989, Fig. 1). In view of their large uncertainty, we shall discuss in Sec. 4 some results obtained by choosing different values for these two parameters. For our numerical computations we used 50 logarithmic mass bins spaced by a factor 2 (i.e., a factor 1.26 in size), with central values of the corresponding radii ranging from $1.211 \times 10^{-2}$

km to $10^3$ km. As concerns the fragment velocity vs. mass relationship, we have chosen as an input parameter either $k = 9/4$ or $r = 1/6$, using $r = (1 - b)/k$ to determine the missing coefficient, as explained earlier — note that this would give at the same time the "experimental" values $r = 1/6$ and $k = 9/4$ for $b = 5/8$, which is reasonable since it can be shown that as $M_{dust} \Rightarrow 0$, $b \Rightarrow (1 - f_l)$, and $f_l < 1/2$ is the definition of shattering impact (see Petit and Farinella 1993, Sec. 3).

The output of the collisional outcome algorithm can be written as a three–dimensional matrix, specifying how many objects (escaping fragments, cratered targets, reaccumulated rubble piles) end up in the $i$–th mass bin following a collision between two bodies lying in the $j$-th and $k$–th mass bins. This matrix can then be read by a program aimed at simulating the time evolution of an initial asteroid population by computing in any time step for every pair of mass bins the expected number of collisions, as in the work of Greenberg et al. (1978), Davis et al. (1979, 1985, 1989, 1994), Campo Bagatin et al. (1994) and Farinella et al. (1992). This will be further detailed below.

## 3. Results from the collisional outcome model

Let us start by commenting some plots that illustrate the results of the collisional outcome algorithm summarized in the previous Section. Fig. 1 shows, for a target 100 km in radius and a $10^{-3}$ projectile–to–target mass ratio, the fraction of reaccumulated mass vs. the logarithm of the impact energy, normalized to the threshold energy for fragmentation. A gap of two to three orders of magnitude is apparent between the fragmentation threshold and the energy needed to disperse to infinity a significant fraction of the target mass. As a consequence, with this parameter choice most asteroids of radius $\approx 100$ km would have been shattered but not completely disrupted (i.e., they would have partially reaccumulated) as a consequence of the most energetic collisions undergone over the history of the solar system. Fig. 1 also shows that the reaccumulation process is confined to a narrower energy range in the $k = 9/4$ case, i.e. when we require that for every velocity $V$ (exceeding a lower cutoff) the total mass ejected with higher speeds is $\propto V^{-9/4}$. The reason for this is that since $r = (1 - b)/k$ (see Sec. 2) and for such supercatastrophic collisions $b$ is close to 1, the exponent $r$ in this case is small, implying that the velocity of fragments has a weak dependence on their mass: thus, there is an abrupt transition between the case when nearly all the fragments are slower than $V_{esc}$ and the opposite one. The transition is less sharp in the probabilistic case, since in this case for every mass the fragment velocities have a significant dispersion about the mean value — and it is much smoother when $r$ (instead of $k$) is fixed to the value 1/6. In this case, even after a supercatastrophic impact a significant fraction of the target mass is reaccumulated, because most of the mass resides in small fragments and is forced to move comparatively fast, so that conservation of energy causes the largest fragments to be slow and reaccumulate.

The same phenomenon can be seen in Fig. 2, where the fraction of reaccumulated mass is plotted vs. target size. Again, for the $k = 9/4$ cases (but less for the probabilistic one), there is a comparatively sharp transition from negligible to substantial reaccumulation, which occurs when the velocity of most fragments becomes lower than $V_{esc}$. As remarked by Petit and Farinella (1993), this behavior would imply comparatively short collisional lifetimes for small asteroids, which would be unlikely to become rubble piles,

but would mostly look like single, competent fragments. The opposite conclusion would follow from the $r = 1/6$ models, for which Fig. 2 shows that even small bodies are always reaccumulated; recall, however, that when $V_{esc}$ is small only a small number of massive fragments may become reaggregated.

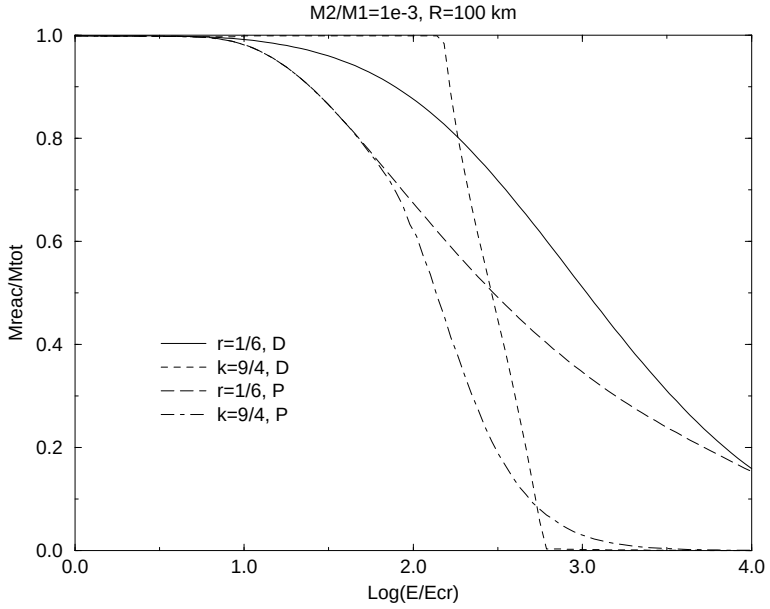



**Figure 1**. The fraction of reaccumulated mass for an impact with $10^{-3}$ projectile–to–target mass ratio against a target of radius 100 km, vs. the logarithm of the impact energy. The latter parameter is normalized to the minimum energy required for fragmentation. Different fragment velocity vs. mass relationships are used (see text), with D standing for "deterministic model" and P for "probabilistic model".

Fig. 3 shows another type of diagram, which should be compared with Fig. 5 of Davis *et al.* (1985) and Fig. 5 of Housen and Holsapple (1990). Here the lower (dashed) curve corresponds to the critical projectile–to–target mass ratio for fragmentation, which is a function of size because of the strength–scaling relationship. The other curves give the threshold mass ratio for collisional disruption, with one half of the target mass dispersed to infinity. While for small asteroids, whose gravity is not efficient in reaccumulating fragments, the fragmentation and disruption thresholds coincide, in agreement with Fig. 1 for asteroids $\approx 100$ km across there is a large gap between the two thresholds, due to the gravitational reaccumulation process. In this case, the curves show that for the $r = 1/6$ models a sudden discontinuity appears when a critical size is exceeded, while for the $k = 9/4$ case there is a more gradual transition. This behaviour can be explained as follows: in the $k = 9/4$ case, when all the fragment velocities are close to each other, a modest specific energy (i.e., projectile–to–target mass ratio) increase may be enough to exceed $V_{esc}$; on the other hand, in the $r = 1/6$ case, after self–gravity has become strong enough to reaccumulate the largest fragments, a moderate increase in the specific energy causes a mass shift to fragments of smaller sizes, which receive a larger fraction of the available energy, thus preventing the largest fragments from starting faster (and being

able to escape).

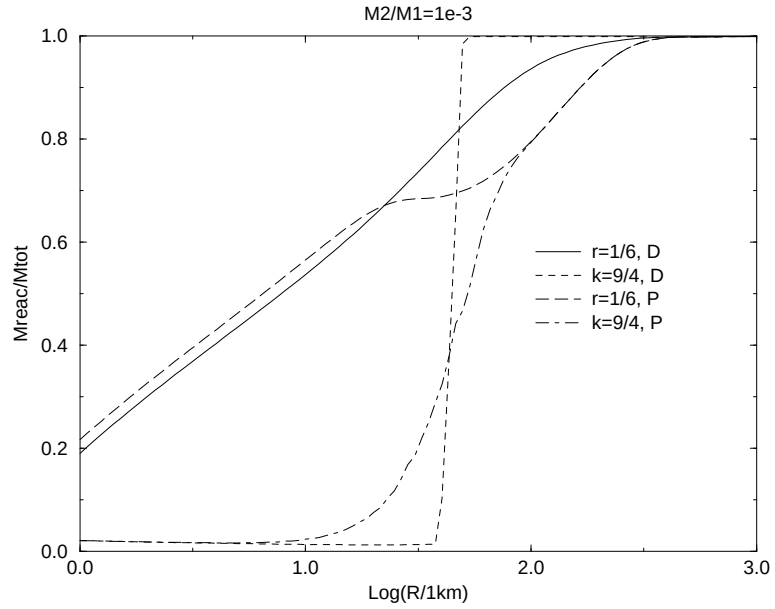



**Figure 2**. The fraction of reaccumulated mass for a shattering impact with $10^{-3}$ projectile–to–target mass ratio vs. the logarithm of the target radius. The same four models have been used as in Fig. 1.

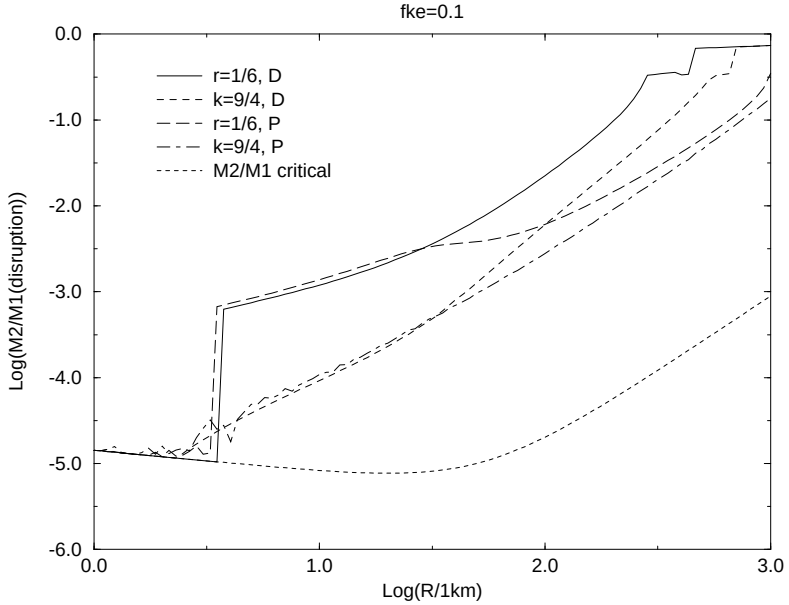



**Figure 3**. The critical mass ratio for *disruption* vs. the logarithm of the target size, for the four models used in Figs. 1 and 2. The bottom (dashed) line shows the projectile–to–target threshold mass ratio for *fragmentation*, derived from scaling laws for the impact strength $S$ as described in the text.

## 4. Simulations of asteroid collisional evolution

We now come to the numerical simulations of the collisional history of the asteroid belt. In summary, we have developed a code which integrates numerically (by a simple variable–stepsize Eulerian technique) the following set of first–order differential equations:

$$\frac{dN_k}{dt} = \sum_{i,j=1}^{n_{bin}} f_{ijk} p_{ij} \frac{(N_i - \delta_{ij}) N_j}{(1 + \delta_{ij})} - B \frac{N_k}{D_k} \,, \tag{2}$$

where $N_k$ are the populations of the $n_{bin}$ (= 50) mass bins; $f_{ijk}$ is the number of fragments injected into the $k$–th bin by an impact between two asteroids belonging to the $i$–th and the $j$–th mass bins, as determined by the collisional outcome algorithm described earlier; $p_{ij}$ is the rate of impacts for a pair of asteroids belonging to the same two bins (derived from their impact cross section and the average *intrinsic collision probability* as estimated by Farinella and Davis 1992); $\delta_{ij}$ (= 1 if $i = j$ and 0 otherwise) accounts for collisions between asteroids residing in the same bin; and the last term in the right–hand member accounts for the removal of small particles due to the Poynting–Robertson drag, whose efficiency is assumed to be inversely proportional to the (mean) particle diameter $D_k$ of the $k$–th bin (for the constant coefficient $B$ the value $1/(4.5 \times 10^{10})$ m/yr has been adopted).

Note that using the same matrix $f_{ijk}$ over the whole evolution time span requires that we assume that the impact response parameters defined in Sec. 2 do not change with time. Actually, one may expect that when an asteroid is converted by an energetic collision into a rubble pile (a weakly bonded aggregate of fragments), its subsequent behaviour becomes very different from the previous one. There is some experimental evidence that this is not the case, however, at least as far as the impact strength $S_0$ is concerned (Ryan et al. 1991). As for $f_{KE}$ and $k$, the available data are too few to justify the development of a specific model for rubble piles.

As for the initial conditions, the starting population was built by joining two power law distributions at a transition diameter of 100 km, with 400 bodies assumed to exceed this size and being distributed according to the cumulative law $N(> D) \propto D^{-3.5}$ (with the only exception of the four largest size bins, containing Ceres, Pallas and Vesta, which were assumed to start already with 1, 0, 0 and 2 bodies). For $D < 100$ km, we used the cumulative distribution $N(> D) \propto D^{-3}$. This choice for the initial population is the same we have made in Davis et al. (1994) and Campo Bagatin et al. (1994), and similar to those made in Davis et al. (1989) and Farinella et al. (1992), and we refer to those papers for a discussion on its motivations. We stress that we did not aim in this work at obtaining a match to the real asteroidal size distribution or at exploring the full space of the model parameters, but only at assessing the importance of making different assumptions about the fragment velocity distribution.

Figs. 4 to 6 show the results of these simulations and their sensitivity to the assumed relationship between fragment velocity and mass. In Fig. 4 the two cases using a "deterministic" relationship (see Sec. 2) are compared to the "standard" case used in all previous work on asteroid collisional evolution, in which no dependence of velocity on mass was assumed (in this case a mass fraction $[1 - (V_{esc}/V_{min})^{-9/4}]$ was simply assumed to reaccumulate for $V_{esc} > V_{min}$, with the lower cutoff velocity $V_{min}$ derived from $f_{KE}$ through

conservation of energy). Of course, the largest difference appears between the standard case and the $r = 1/6$ case, which corresponds to the strongest velocity vs. mass dependence. Whereas this difference is always considerable for asteroids $\approx 100$ km in size, in the $r = 1/6$ case it is still significant at diameters down to $\approx 100$ m. There are two concurring reasons for this behaviour: (i) in the $r = 1/6$ case some gravitational reaccumulation occurs even on very small targets (see Sec. 3, Fig. 2), because the largest fragments are always very slow, hence the response to impacts of small asteroids is affected by the behaviour of fragment velocities; (ii) since most small asteroids are first– or multi–generation fragments escaped after the break–up (and partial reaccumulation) of larger parent asteroids, their final abundance is sensitive to the efficiency of gravitational reaccumulation onto larger bodies. On the other hand, the $k = 9/4$ case differs from the standard one mainly in the size range (between 40 and 200 km in diameter) where reaccumulation is most important.

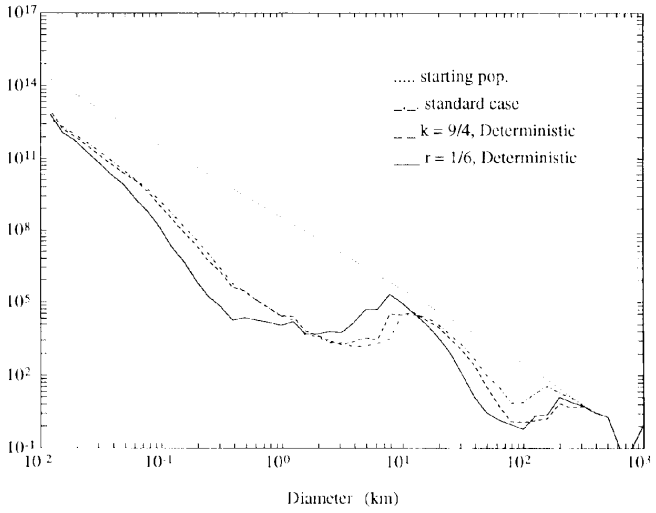



**Figure 4**. Results of numerical simulations of the asteroid collisional history. The dotted line corresponds to the assumed initial population, while the other curves are the final size distribution derived with different assumptions about the fragment velocity vs. mass relationship. Note that in this type of diagram, showing incremental populations of equal logarithmic bins vs. logarithm of their central diameter, a power–law size distribution appears as a straight line of slope coinciding with the index of the corresponding *cumulative* distribution.

Figs. 5 and 6 show the differences between the deterministic and the probabilistic cases for both the $r = 1/6$ and the $k = 9/4$ case. Although some difference propagates to smaller sizes, the largest discrepancy lies in the reaccumulation–effective diameter range around 100 km. These results show that the final size distribution of the observable asteroids depends sensitively not only upon the assumed relationship between *mean* fragment velocity and mass, but also on the dispersion of that velocity about the mean for any given mass. This highlights the crucial importance of obtaining new experimental data on

fragment velocity distributions.

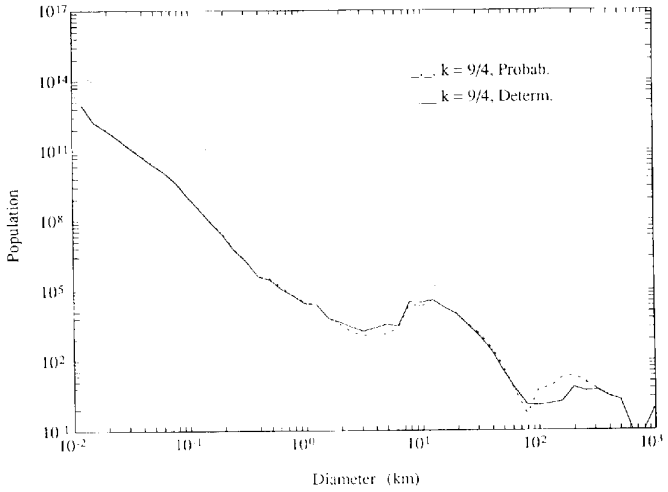



**Figure 5**. The same as Fig. 4, but comparing the results obtained with the deterministic and probabilistic models in the case $k = 9/4$ (see text).

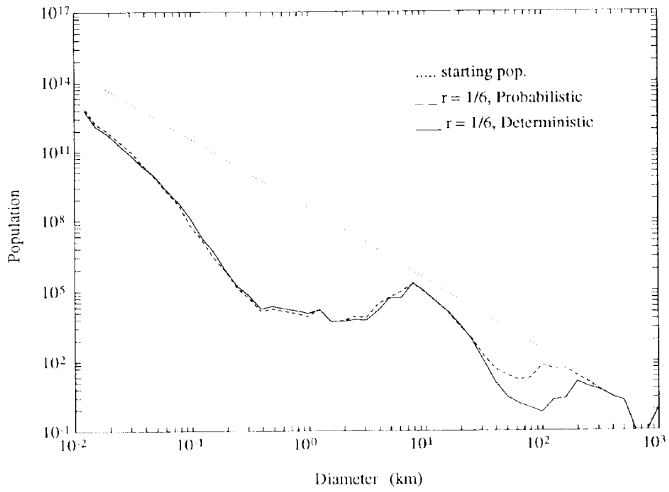



**Figure 6**. The same as Fig. 5, but for the case $r = 1/6$ (see text).

How does the sensitivity to the velocity vs. mass relationship discussed in this paper compare with the sensitivity to plausible changes of other input parameters which are needed to run the collisional outcome model? Figs. 7 and 8 show how the results of the

"standard" case already displayed in Fig. 4 are affected when $S_0$ is changed by plus or minus an order of magnitude (Fig. 7), and $f_{KE}$ is changed by plus or minus a factor of two (Fig. 8). Changing $S_0$ has very strong effects, as the final asteroid populations at some sizes (e.g., a few km) are changed by orders of magnitude. Note, however, that the resulting effects are not obvious: for instance, in the diameter range between about 10 and 50 km, the final populations are *more* depleted when the asteroids are assumed to be stronger. This is caused by the fact that these bodies are more frequently disrupted by km–sized projectiles, which on the contrary are much more abundant. This type of effect, together with the presence of a small–size cutoff in the evolving size distribution, is responsible for the "wavy" appearance of the final distribution, with the typical wavelength inversely correlated to $S_0$ (Campo Bagatin et al. 1994). Even if the results of the simulations are very sensitive to $S_0$, however, we recall that for any given material this parameter can be measured in the laboratory much more easily than the fragment velocities, as such measurements do not need complex fast–framing imaging devices, but only an *a posteriori* assessment of the impact outcome.

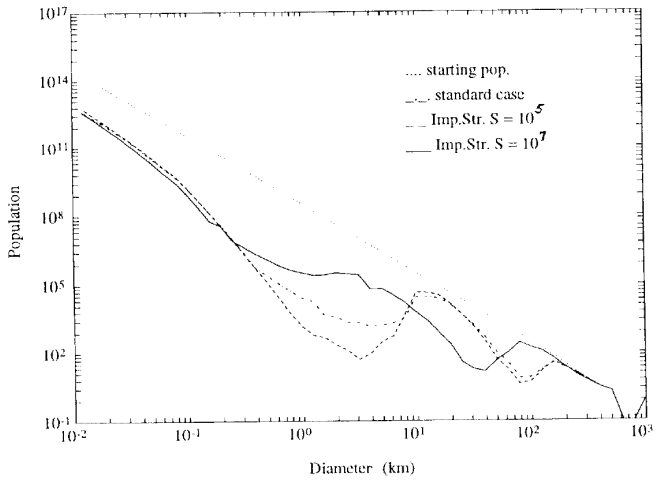



**Figure 7**. The same as Fig. 4, but comparing the standard case (with $S_0 = 10^6$ J/m$^3$) to the results obtained by changing $S_0$ by plus or minus an order of magnitude.

The results illustrated in Fig. 8 are even more interesting in the present context. The changes resulting when $f_{KE}$ varies by a factor of four (i.e., typical fragment velocities vary by a factor of two, as $f_{KE}$ is related to the total kinetic energy) are not negligible, and they can be interpreted easily. The final populations are almost insensitive to the value of $f_{KE}$ at small sizes, where self–gravitational effects are negligible, whereas the location of the "bump" caused by the onset of reaccumulation (at diameters $\approx$ 10 km) is somewhat shifted. Also, for the largest asteroids (diameters $>$ 100 km) higher fragment velocities result into a less effective reaccumulation, hence in a stronger depletion. However, it is worth remarking that the changes shown in Fig. 8 are smaller than those displayed in

Fig. 4 and 6, and similar to those displayed in Fig. 5. Thus, knowing how the fragment velocities are related to their mass and also how they are distributed for any given mass about the mean value, is at least as important as determining the average magnitude of these velocities (i.e., measuring $f_{KE}$).

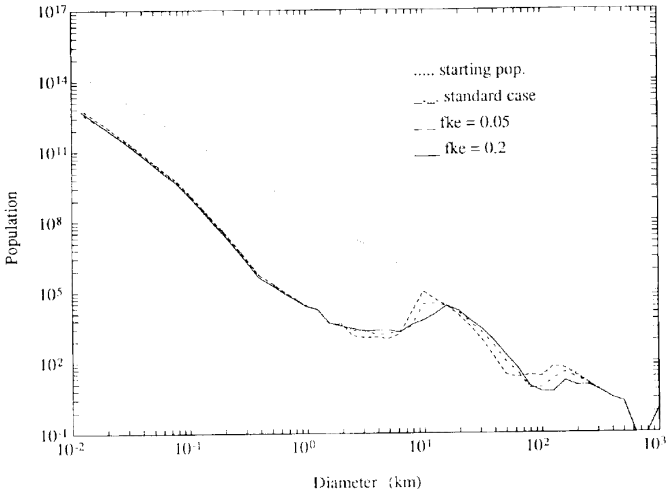



**Figure 8**. The same as Fig. 4, but comparing the standard case (with $f_{KE} = 0.1$) to the results obtained by changing $f_{KE}$ by plus or minus a factor of two.

## 5. Conclusions

Numerical simulations of the asteroid collisional evolution are motivated by the desire to understand how the current asteroid population has been processed starting from an original population of newly accreted proto–asteroids, which would have been the building blocks for an asteroidal planet if the accumulation process had not been interrupted. However, to derive information about this primordial population, we need to use realistic physical models of how collisions have affected asteroids of different sizes, in particular how frequently they are disrupted and how many new smaller (fragment) asteroids are generated in this case.

In this paper we have shown that a crucial feature of these models is the relationship between fragment ejection velocity and mass. Not only how the mean velocity depends upon mass is important, but also the distribution of velocities about the mean values. The available experimental evidence on this issue is still sparse, and does not constrain the collisional models well enough to allow us to make reliable predictions on the outcomes of impacts between bodies tens to hundreds of km in size. This adds to the current uncertainty on the values of collisional parameters such as $S_0$ and $f_{KE}$ and the appropriate scaling laws for them, and as a consequence we are still unable to simulate the evolution of the asteroid size distribution over billions of years having full confidence that the results will

tell us the history of the real belt. Hopefully, further experimental and modelling work will improve this situation.

**Acknowledgments**. We are grateful to D.R. Davis, A. Nakamura and P. Paolicchi for useful discussions. Helpful comments by I. Giblin and an anonymous referee are also acknowledged. J.–M. Petit's stay in Pisa in 1992 was supported through the CNRS–CNR researcher exchange agreement. P.F. worked on this project in 1992–93 while staying at the Nice Observatory thanks to the "G. Colombo" fellowship of the European Space Agency.

## References

**Campo Bagatin, A., A. Cellino, D.R. Davis, P. Farinella, and P. Paolicchi,** Wavy size distributions for collisional systems with a small–size cutoff, *Planetary Space Sci.*, submitted, 1994.

**Davis, D.R., and E.V. Ryan,** On collisional disruption: Experimental results and scaling laws, *Icarus* 83, 156–182, 1990.

**Davis, D.R., C.R. Chapman, R. Greenberg, and A.W. Harris,** Collisional evolution of asteroids: Populations, rotations, and velocities, in *Asteroids*, ed. T. Gehrels, Univ. of Arizona Press, pp. 528–557, 1979.

**Davis, D.R., C.R. Chapman, S.J. Weidenschilling, and R. Greenberg,** Collisional history of asteroids: Evidence from Vesta and the Hirayama families, *Icarus* 62, 30–53, 1985.

**Davis, D.R., P. Farinella, P. Paolicchi, and V. Zappalà (eds.),** *Catastrophic Disruption of Asteroids and Satellites*, *J. Ital. Astron. Soc.*, vol. 57, n. 1, 1986.

**Davis, D.R., P. Farinella, P. Paolicchi, S.J. Weidenschilling, and R.P. Binzel,** Asteroid collisional history: Effects on sizes and spins, in *Asteroids II*, eds. R.P. Binzel, T. Gehrels and M.S. Matthews, Univ. of Arizona Press, pp. 805–826, 1989.

**Davis, D.R., E.V. Ryan, and P. Farinella,** Asteroid collisional evolution: Results from current scaling algorithms, *Planet. Space Sci.*, in press, 1994.

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

**Farinella, P., P. Paolicchi, and V. Zappalà,** The asteroids as outcomes of catastrophic collisions, *Icarus* 52, 409–433, 1982.

**Farinella, P., D.R. Davis, P. Paolicchi, A. Cellino, and V. Zappalà,** Asteroid collisional evolution: An integrated model for the evolution of asteroid rotation rates, *Astron. Astrophys.* 253, 604–614, 1992.

**Fujiwara, A., G. Kamimoto, and A. Tsukamoto,** Destruction of basaltic bodies by

high–velocity impact. *Icarus* 31, 277–288, 1977.

**Fujiwara, A., P. Cerroni, D.R. Davis, E. Ryan, M. Di Martino, K. Holsapple, and K. Housen, K.,** Experiments and scaling laws on catastrophic collisions, in *Asteroids II*, eds. R.P. Binzel, T. Gehrels and M.S. Matthews, Univ. of Arizona Press, pp. 240–265, 1989.

**Gault, D.E., E.M. Shoemaker, and H.J. Moore,** Spray ejected from the lunar surface by meteoroid impact, *NASA Tech. Note D–1767*, 1963.

**Giblin, I., G. Martelli, P.N. Smith, A. Cellino, M. Di Martino, P. Farinella, and P. Paolicchi,** Field fragmentation of macroscopic targets simulating asteroidal catastrophic collisions, *Icarus*, submitted, 1994.

**Greenberg, R., J.F. Wacker, W.K. Hartmann, and C.R. Chapman,** Planetesimals to planets: Numerical simulation of collisional evolution, *Icarus* 35, 1–26, 1978.

**Housen, K.R.,** Crater ejecta velocities for impacts on rocky bodies, *Lunar Planet. Sci. Conf. XXIII*, 555–556, 1992.

**Housen, K.R., and K.A. Holsapple,** On the fragmentation of asteroids and planetary satellites, *Icarus* 84, 226-253, 1990.

**Housen, K.R., R.M. Schmidt, and K.A. Holsappple,** Laboratory simulations of large scale fragmentation events, *Icarus* 94, 180–190, 1991.

**Nakamura, A., and A. Fujiwara,** Velocity distribution of fragments formed in a simulated collisional disruption, *Icarus* 92, 132–146, 1991.

**Nakamura, A., K. Suguiyama, and A. Fujiwara,** Velocity and spin of fragments from impact disruptions, I. An experimental approach to a general law between mass and velocity. *Icarus* 100, 127–135, 1992.

**Paolicchi, P., A. Cellino, P. Farinella, and V. Zappalà,** A semiempirical model of catastrophic breakup processes, *Icarus* 77, 187–212, 1989.

**Paolicchi, P., A. Verlicchi, and A. Cellino,** A new semiempirical model for catastrophic impact processes (abstract), *Bull. Amer. Astr. Soc.* 24, 960, 1992.

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

**Ryan, E.V., W.K. Hartmann, and D.R. Davis,** Impact experiments 3: Catastrophic fragmentation of aggragate targets and relation to asteroids, *Icarus* 94, 283–298, 1991.

**Takagi, Y., M. Kato, and H. Mizutani,** Velocity distribution of fragments of catastrophic impacts, in *Asteroids, Comets, Meteors 1991*, eds. A.W. Harris and E. Bowell, Lunar and Planetary Institute, pp. 597–600, 1992.

**Vickery, A.M.,** Size–velocity distribution of large ejecta fragments, *Icarus* 67, 224–236, 1986.

**Vickery, A.M.,** Variation in ejecta size with ejection velocity, *Geophys. Res. Lett.* 14, 726–729, 1987.

## 5.2 Evolution de l'astéroïde Ida et son satellite Dactyl

**[16]**

"Collisional and Dynamical History of Ida."

GREENBERG, R., BOTTKE, W., NOLAN, M., GEISSLER, P., PETIT, J-M., DURDA, D., ASPHAUG, E. & HEAD, J.

*Icarus* **120** (1996)

**[17]**

# "Erosion and Ejecta Redistribution on 243 Ida and its Moon."

GEISSLER, P., PETIT, J-M., DURDA, D., GREENBERG, R., BOTTKE, W. & NOLAN, M.

*Icarus* **120** (1996)

**[18]**

# "The discovery and Orbit of 1993 (243)1 Dactyl."

BELTON, M., MUELLER, B., D'AMARIO, L., BYRNES, D., KLAASEN, K., SYNNOTT, S., BRENEMAN, H., JOHNSON, T., THOMAS, P., VEVERKA, J., HARCH, A., DAVIES, M., MERLINE, W., CHAPMAN, C., DAVIS, D., DENK, T., PETIT, J-M., GREENBERG, R., MCEWEN, A., STORRS, A. *&* ZELLNER, B.

*Icarus* **120** (1996)

**[20]**

# "The long term dynamics of Dactyl's orbit."

PETIT, J-M., DURDA, D., GREENBERG, R., HURFORD, T. *&* GEISSLER P.

submitted to *Icarus* (1996)

# The long term dynamics of Dactyl's orbit

J.M. Petit
C.N.R.S., Observatoire de Nice, France
&
L.P.L., University of Arizona, Arizona

D. D. Durda, Richard Greenberg, T. A. Hurford, P. E. Geissler
L.P.L., University of Arizona, Arizona

**Abstract**

Asteroid Ida's satellite Dactyl was observed over $5\frac{1}{2}$ hours by the Galileo spacecraft imaging system. The observed motion has been fit by a family of orbits parameterized by the mass of Ida (Byrnes and D'Amario 1994, Belton *et al.* 1996). We have tested the stability of these orbits and found that those with semi-major axis $a \leq 65$ km (corresponding to Ida's density $\geq 3.1$ g cm$^{-3}$) are unstable over time scales of a few days to a few months. Ida's density must therefore be $\leq 3.1$ g cm$^{-3}$. It is most likely that Dactyl was created at the same time as Ida (Durda 1996). Therefore, it is important to understand the longer term stability of candidate orbits. We have numerically integrated motion about a realistically shaped model for Ida. We find periodic orbits at the 5:1, 6:1, 7:1, etc., commensurabilities with Ida's rotation that are stable for at least 3000 years. These orbits have their pericenter locked at longitudes in the range $35° - 70°$ on Ida (not exactly at the short principal axis because of Ida's irregular shape). The 11:2, 13:2, etc. commensurabilities are also stable over those timescales. However, among the low-order mean-motion resonances only the 5:1 resonance allows Dactyl to have been at longitudes that fit the Galileo observations and is large enough to trap the satellite. If Dactyl's orbit is stabilized for the long term by commensurabilities, the mass (and hence density) of Ida may be constrained to one specific value, *e.g.* 2.93 g cm$^{-3}$, the value corresponding to the 5:1 spin-orbit resonance. A comparison between the stability of the 5:1 orbit and neighboring non-resonant orbits is underway, and preliminary results are presented.

# 1 Introduction

243 Ida's satellite Dactyl was discovered in imaging data returned by the Galileo spacecraft after its August 28, 1993 encounter with the asteroid (Belton and Carlson 1994). Because Galileo was targeted to fly by Ida at too great a distance for Doppler data to be useful in determining the asteroid's mass, the discovery provided a fortuitous opportunity to find the mass and mean density of Ida if an orbit for the satellite could be determined. Dactyl appeared in 47 of the returned images, defining an orbital arc some $5\frac{1}{2}$ hours long. Unfortunately, Dactyl's orbital plane about Ida was nearly coincident with Galileo's trajectory past the asteroid, so no single orbit could be unambiguously defined from the limited imaging data alone. The observed orbital motion fits a family of Keplerian orbits, each corresponding to a particular assumed mass of Ida (Byrnes and D'Amario 1994, Belton *et al.* 1995, 1996). These orbits range from nearly circular orbits to hyperbolic trajectories for the lowest density cases. The hyperbolic cases may be ruled out based on Hubble Space Telescope searches for Dactyl in the vicinity of Ida after its discovery (Belton *et al.* 1995) and from statistical arguments on the likelihood of observing a passing asteroid just during the short time of the encounter so close to Ida (Belton *et al.* 1996). Further constraints on the orbit of Dactyl, and thus on the mass and density of Ida, come from dynamical studies of orbital stability. In Section II we describe the dynamical constraints that can be placed on the range of allowable Dactyl orbits that have been found by short term integration using both a simple triaxial ellipsoidal shape model and a realistic irregular shape model for Ida (Thomas *et al.* 1996). We then describe in Section III our search for stable resonant orbits about a realistically shaped Ida which fit the observed locations of Dactyl. The long-term stability of resonant orbits is discussed in Section IV and our conclusions summarized in Section V.

# 2 Dynamical constraints on allowable orbits

To study the stability of orbits around a rapidly rotating, elongated asteroid, we first modeled the primary asteroid as a triaxial ellipsoid with parameters similar to Ida: principal axes 56 km, 24 km, and 21 km, assumed density of $3.5\,\mathrm{g\ cm^{-3}}$ and rotating with a period of 4.6 hours (Petit *et al.* 1994). With such a primary, at Dactyl's close distance orbits can be quite irregular. We found that the most critical parameter governing stability is the pericenter distance. For low enough pericenter distances, orbits impact the primary or escape within a few hundred periods. For retrograde orbits (motion opposite the sense of rotation), the critical pericenter distance is about 50 km. For prograde orbits (Dactyl's orbit is prograde), motion is unstable if the pericenter is lower than about 70 km. Such a restriction would rule out a large

range of orbital fits by Byrnes and D'Amario, and allow only those orbits that correspond to a mass density less than about 3 g cm$^{-3}$.

In order to test that result, we systematically examined the stability of elliptical osculating orbits fit to the Galileo images by Byrnes and D'Amario. This series of numerical experiments used the improved "best fit" triaxial ellipsoid reported by Thomas *et al.* (1996), with axes 59.8 km, 25.4 km, and 18.6 km. We integrate the equations of motion in a rotating frame where Ida is at rest:

$$\begin{aligned} \ddot{x} &= 2\omega\dot{y} + \omega^2 x - \frac{\partial U(x,y,z)}{\partial x}, \\ \ddot{y} &= -2\omega\dot{x} + \omega^2 y - \frac{\partial U(x,y,z)}{\partial y}, \\ \ddot{z} &= -\frac{\partial U(x,y,z)}{\partial z}, \end{aligned} \tag{1}$$

where $U(x,y,z)$ is the gravitational potential of the primary and $\omega$ the rotation speed, the rotation axis being parallel to the $z$ axis. The potential of the triaxial ellispoid is then given by elliptical integrals (Chauvineau *et al.* 1993). We used Bulirsch and Stoer integrator, with self adaptive time step, and required a relative precision of $10^{-10}$ at each time step. It turns out that the precision was about $10^{-12}$ to $10^{-13}$ for each time step and even after more than a million orbits, the Jacobi constant had changed by only $10^{-8}$ in relative value. The initial conditions are derived from the osculating elements for each mass of Ida, the density values being somewhat different from those listed by Byrnes and D'Amario due to the improved volume determination by Thomas *et al.* (1996). The shape model used here has more symmetry than the real Ida has, hence one might expect to have more stable orbits. But even so, prograde orbits with a pericenter distance less than 65 km exhibit strong short term instabilities. They either impact Ida or escape the asteroid in only a few hours to days. An orbit is said to escape if its binding energy is positive while its distance to Ida is large enough, namely 500 km. All significant changes in energy occur near pericenter. At large distances, the rotation of Ida averages out its effect on energy. Fig. 1 shows the results of a numerical integration of the osculating orbit from Byrnes and D'Amario with Ida's mass M $= 5.70 \times 10^{16}$ kg, $\rho = 3.54$ g cm$^{-3}$. We have plotted the orbit in the frame rotating with Ida. The central ellipse represents the projection of Ida on the $(x,y)$ plane. The semimajor axis of this orbit is 70.0 km with a pericenter distance of about 40.6 km (for the keplerian orbit). The satellite makes five revolutions before impacting Ida after a little under 37 hours. Fig. 2 illustrates a slightly longer-lived escaping case (GM $= 0.0034$ km$^3$s$^{-2}$) with a pericenter of about 56.2 km. Between each close approach, the energy oscillates around a constant value (Fig. 2b). At each passage at pericenter, the energy gets an impulse and changes by a potentially large amount. The satellite has four close

encounters with Ida and during the last one, after 29 days in orbit, receives enough energy to escape.

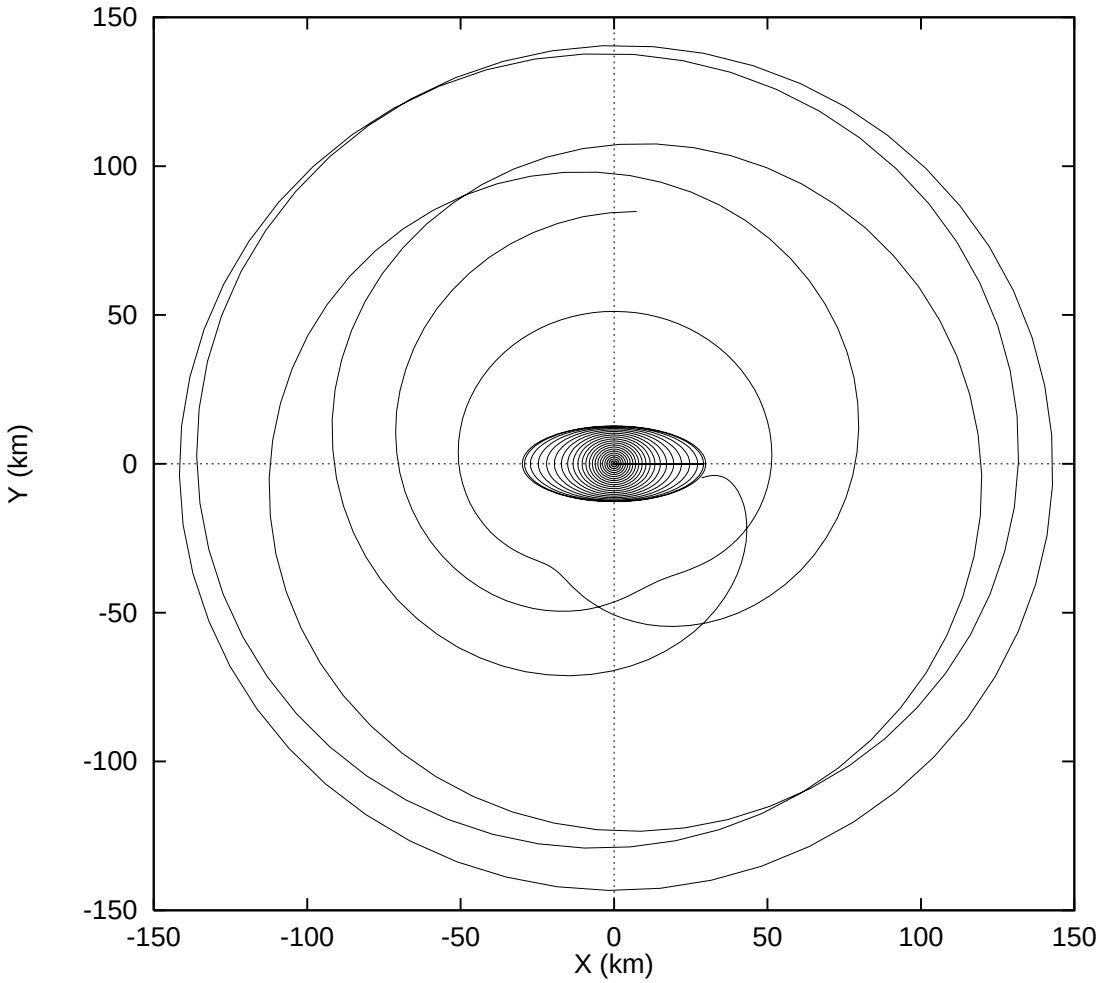



**Fig. 1:** Short term instability orbits for large mass of the primary. Orbit is plotted in the frame rotating with Ida. The central ellipse represents the projection of primary (a triaxial ellipsoid) on the $(x, y)$. $\mathrm{GM} = 0.0038\,\mathrm{km}^3\mathrm{s}^{-2}$.

Orbits with pericenter distances larger than 65 km were found to be stable for at least a year, corresponding to hundreds of orbital periods. Fig. 3 shows such a case for $\mathrm{GM} = 0.0032\,\mathrm{km}^3\mathrm{s}^{-2}$,

with pericenter distance of 68.3 km. The orbital energy remains constant and stable, except for oscillations due to the rotation of Ida, showing no signs of impending instability over the 100 day timescale of the integrations.

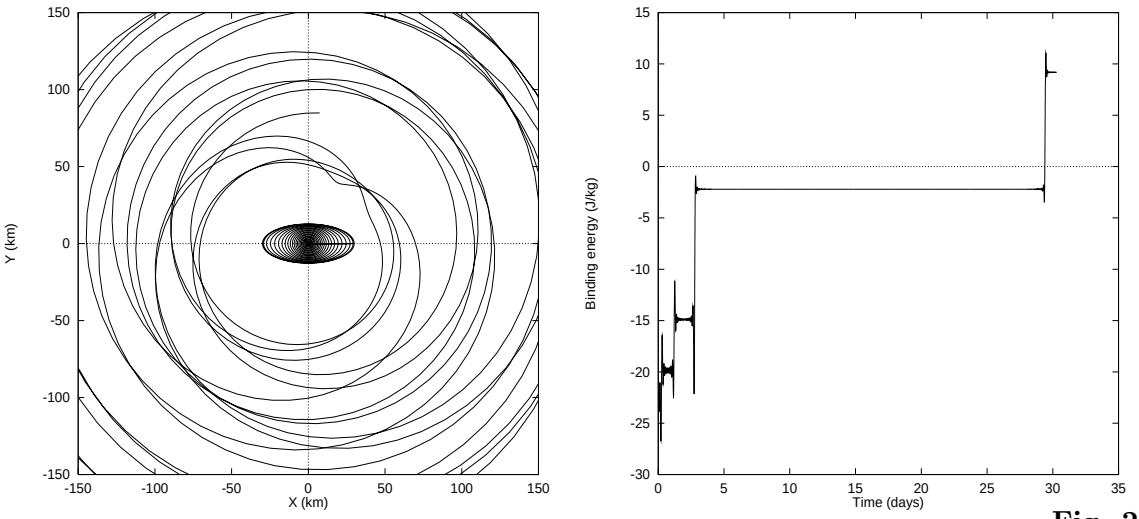



**Fig. 2:** (a) Same as Fig. 1 for GM = 0.0034 $\mathrm{km^3 s^{-2}}$. (b) Time evolution of energy per unit mass of the satellite. Negative energy corresponds to a bounded orbit, positive energy to an escaping orbit.

We next explored the influence of the irregular shape of Ida on the short-term instability limits. We used the shape model for Ida derived by Thomas *et al.* (1996) and filled it with spheres. The spheres have different masses and hence sizes: small spheres on the outer parts of Ida, to better represent the irregular mass distribution, and larger spheres inside the body, to save computing time. Masses of the packed spheres are chosen to approximate uniform density over the entire asteroid. The fate of individual orbits is different from what we found in the triaxial ellipsoid case, but the boundary between instability (impact or escape) and stability is not changed. The real shape model is slightly rounder than the triaxial ellipsoid, especially at the ends which may explain why we found fewer colliding orbits with the real shape. Fig. 4 shows, for the realistic shape model, the same initial orbits as Fig. 1. In both cases the satellite escapes after a few days to a little over one year in orbit. In order to better determine the limit between stable and unstable orbits, we interpolated osculating elements for parameter values GM in the range 0.00320 $\mathrm{km^3 s^{-2}}$ – 0.00340 $\mathrm{km^3 s^{-2}}$. For the triaxial ellipsoid, the limit of stability is GM = 0.00329 $\mathrm{km^3 s^{-2}}$, and for a realistic shape model including 52 spheres used in this study, GM = 0.00333 $\mathrm{km^3 s^{-2}}$. We used different shape models (*e.g.* different filling with spheres of the Thomas *et al.* shape model) and found that the limit did not change much.

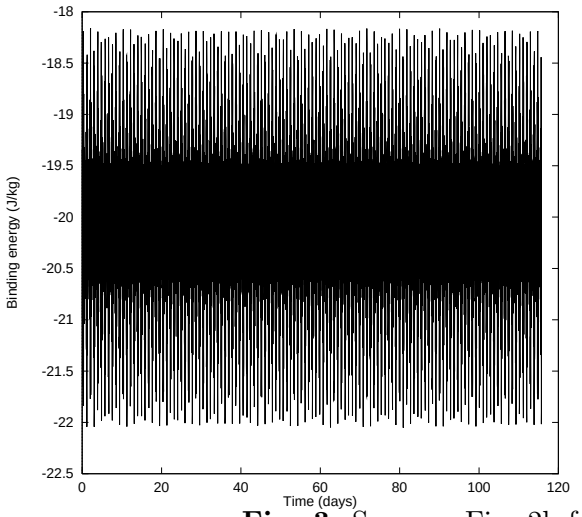



**Fig. 3:** Same as Fig. 2b for GM = $0.0032\,\mathrm{km}^3\mathrm{s}^{-2}$.

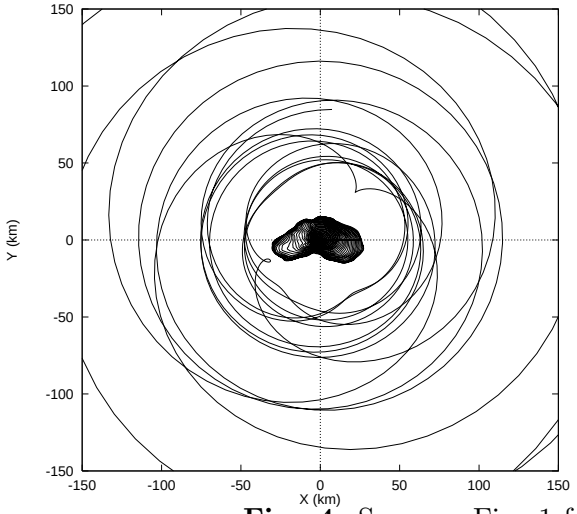



**Fig. 4:** Same as Fig. 1 for a realistic shape model for the primary.

On the "low-mass" end of the osculating orbit family, the orbits are hyperbolic. As discussed in the introduction, such an orbit has been ruled out for Dactyl (Belton *et al.* 1995, 1996). This means that values of $\mathrm{M} \leq 3.38 \times 10^{16}\,\mathrm{kg}$ can be disgarded. We therefore focussed on elliptical orbits and estimated the Lyapunov characteristic number (LCN, Benettin *et al.* 1980,

Froeschle C. 1984). This number characterizes the divergence rate of two neighbouring orbits. Let us consider a differential equation:

$$\frac{dX}{dt}(t) = F(X,t). \tag{2}$$

We want to know how the point $X(t)+\epsilon(t)$ diverges from $X(t)$, where $\epsilon(t)$ is a small quantity. We expand equation (2) to find the equation governing the evolution of $\epsilon(t)$:

$$\frac{dX}{dt}(t) + \frac{d\epsilon}{dt}(t) = F(X+\epsilon,t) = F(X,t) + L_{(X,t)} \cdot \epsilon + O(\epsilon^2), \tag{3}$$

where $L_{(X,t)}$ is the Jacobian matrix at point $X$ and time $t$. Since $\epsilon$ is an infinitely small vector, its evolution is governed by the well know variation equation:

$$\frac{d\epsilon}{dt}(t) = L_{(X,t)} \cdot \epsilon. \tag{4}$$

This is a linear equation, so we can normalize $\epsilon$ as we wish. It is well know that the quantity:

$$\Gamma(\epsilon(0),T) = \frac{1}{T}\frac{||\epsilon(T)||}{||\epsilon(0)||} \tag{5}$$

is independent of $\epsilon(0)$ and tends to a limit, which is the largest LCN, as $T$ goes to infinity. A strictly positive LCN means that two neighbouring points will diverge exponentially in time. This is what we call a chaotic motion. In practice, we define the Lyapunov characteristic indicators as the truncated values of the LCNs for a finite time. We found that the orbits with $M \leq 3.9 \times 10^{16}$ kg (*e.g.* large $e$) are chaotic. The Lyapunov characteristic indicator levels off after some $10^7$ to $10^9$ seconds. However a chaotic orbit for Dactyl can not be ruled out. While an orbit may be formally chaotic, its behavior may be well-bounded, regular, and with an indefinite lifetime. An example of such a case is the motion of terrestrial planets (Laskar 1990).

In our case, orbits with $M \leq 3.6 \times 10^{16}$ kg are chaotic leading to eventual satellite escape. Fig. 5a shows the time evolution of the Lyapunov characteristic indicator for $GM = 0.0024\,km^3s^{-2}$. Here again we stopped the integration when the satellite was detected to be on an escaping orbit (positive binding energy at a distance larger than 500 km). We computed the semi-major axis and eccentricity at apocenter. The osculating elements defined at apocenter are in good agreement with the values averaged over one orbit and match fairly well the two-body osculating elements given by Byrnes and D'Amario. The semi-major axis and eccentricity follow a kind of random walk with, however, a slight tendency to increase. This results in a decrease of

the "averaged pericenter distance" defined by $q = a(1 - e)$ where $a$ and $e$ are the values at apocenter (Fig. 5b). When this pericenter distance reaches 74 km, the process speeds up. This occurs when $e$ is about 0.9 and $a$ is 740 km. The semi-major axis starts to diverge and increases toward 1 by larger increments. After a few more revolutions, the satellite goes on a hyperbolic orbit and escapes.

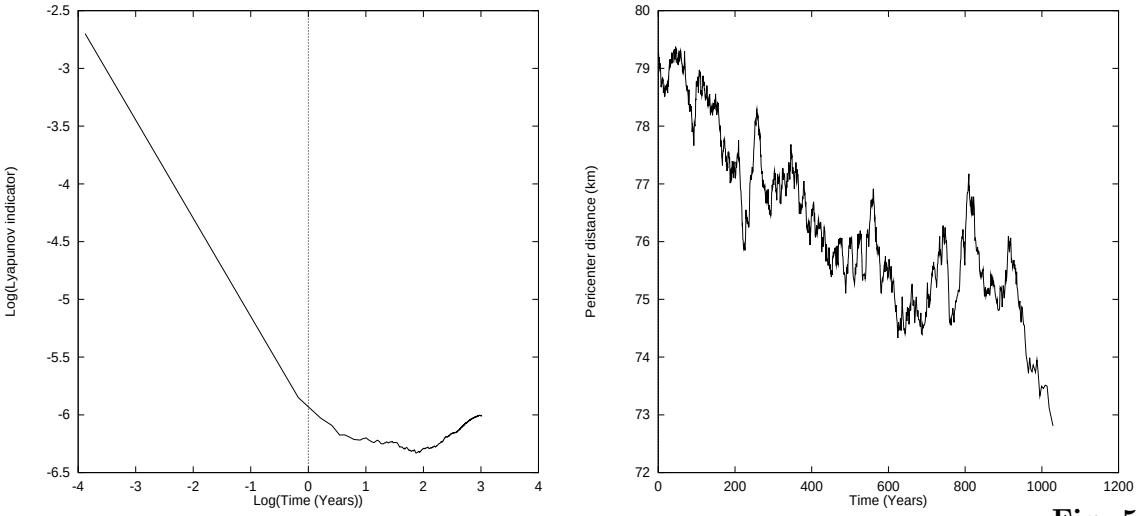



**Fig. 5:** (a) Time evolution of the Lyapunov characteristic indicator for GM $= 0.0024\,\mathrm{km^3 s^{-2}}$. The orbit is chaotic and the satellite eventually escapes. (b) Time evolution of the "averaged pericenter distance" $q = a(1 - e)$ where $a$ and $e$ are the semi-major axis and eccentricity defined at apocenter.

On the other hand, the orbit for GM $= 0.0025\,\mathrm{km^3 s^{-2}}$, even though chaotic, seems to be stable over the $10^{11}$ sec (about 3000 years) of integration. Fig. 6a shows the time evolution of the Lyapunov characteristic indicator which levels off at a value of the order of $10^{-8.1}$ after some $10^9$ sec. Despite this formal "chaos", the osculating elements seem to be trapped in a finite region not connected with the region of escaping orbits. Fig. 6b shows the time evolution of the semi-major axis. Although chaotic, its range of variation is limited to 350 m around a value of 195.32 km. Similarly, its eccentricity has a value of $0.5836 \pm 0.0006$.

None of the orbits given by Belton *et al.* (1996) with $4.2 \times 10^{16}\,\mathrm{kg} \leq \mathrm{M} \leq 4.8 \times 10^{16}\,\mathrm{kg}$ showed a sign of instability over the $10^{11}$ sec of integration. Nor are they chaotic: their Lyapunov characteristic indicator decreases regularly with time, linearly in a log-log plot (*e.g.* Fig. 7, GM $= 0.0031\,\mathrm{km^3 s^{-2}}$).

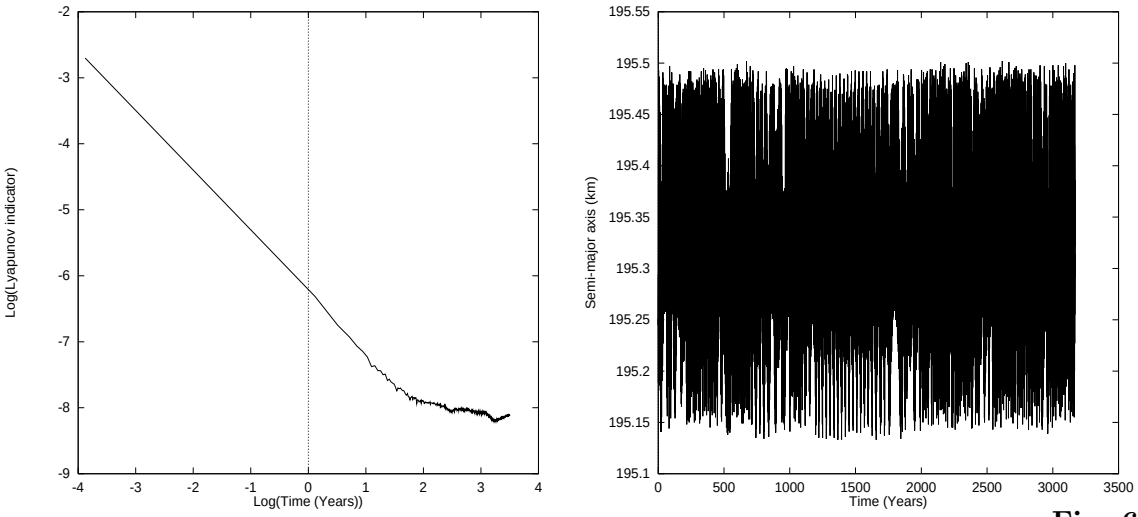



**Fig. 6:** (a) Same as Fig. 5a for GM = 0.0025 $\mathrm{km}^3\mathrm{s}^{-2}$. The orbit is chaotic but the satellite does not escape. (b) Time evolution of the semi-major axis at apocenter.

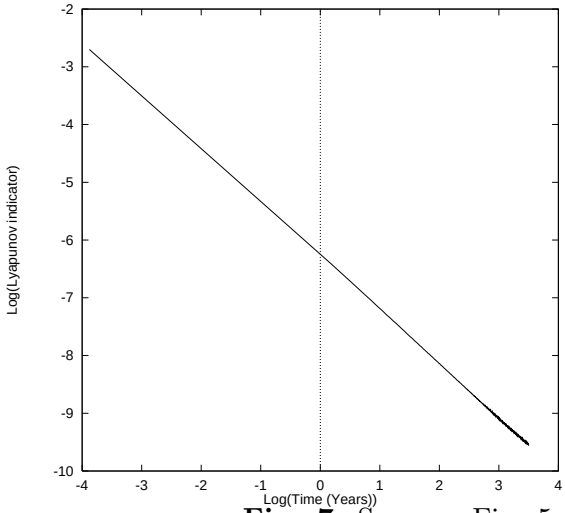



**Fig. 7:** Same as Fig. 5a for GM = 0.0031 $\mathrm{km}^3\mathrm{s}^{-2}$. The orbit is stable for over 3000 years.

The boundary for stable prograde orbits found from these integrations is indicated in Fig. 8, where the family of candidate orbits obtained by Belton *et al.* are also shown. The range of

possible orbits is $0.0025\,\mathrm{km^3 s^{-2}} \leq \mathrm{GM} \leq 0.0033\,\mathrm{km^3 s^{-2}}$ or $3.7 \times 10^{16}\,\mathrm{kg} \leq \mathrm{M} \leq 4.95 \times 10^{16}\,\mathrm{kg}$. Assuming a volume of 16,100 $\mathrm{km^3}$ (Thomas *et al.* 1996) this results in a mean density $\rho$ in the range $2.3\,\mathrm{g\ cm^{-3}} - 3.07\,\mathrm{g\ cm^{-3}}$.

To further restrict the range of possible masses we must consider longer term stability.

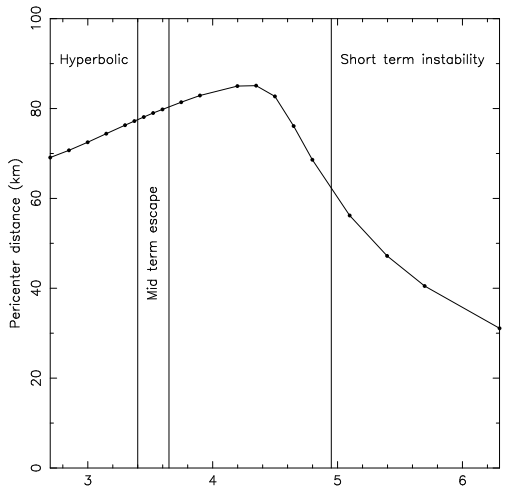



**Fig. 8:** Pericenter distance of Galileo fit orbits versus mass of Ida. Only the central part with Ida's mass in the range $3.65 \times 10^{16}\,\mathrm{kg} - 4.95 \times 10^{16}\,\mathrm{kg}$ allows for long lived orbits.

# 3 Stable resonant orbits

## 3.1 Orbital interactions with elongated primaries

Dactyl has almost certainly been orbiting Ida over a geologic timescale. It is extremely unlikely that we would have encountered Ida at the one rare moment that a small, dynamically unstable satellite was formed. It is most likely that Dactyl was created with Ida (Durda 1996), altough it is possible that it formed later, *e.g.* as ejecta from a large cratering event on the asteroid. In any case, it is unlikely that a short-lived satellite would be in orbit just as Galileo encountered Ida. In Ida's complex dynamical environment irregular motion may likely be the rule. Resonant orbits may offer a means to stabilize orbits for time scales comparable to the age of Ida.

In our studies with triaxial primaries we explored special classes of stable, resonant, periodic orbits that may extend the range of possible conditions that can lead to a long-lived satellite.

Here, in order to reduce the volume of phase space to explore, we restrict ourselves to zero inclination orbits, *e.g.* in Ida's equatorial plane. One class of orbit, prograde and synchronous at a longitude 90° from the long axis, was thought *a priori* to be stable. Such an orbit is in effect trapped in the gravitational harmonic "topography". We found a set of stable synchronous at that longitude about a primary with dimensions 36 km:29 km:27 km. However, we discovered that such orbits proved to be unstable for a triaxial primary as elongated as Ida. Similar results were reported by Scheeres (1994).

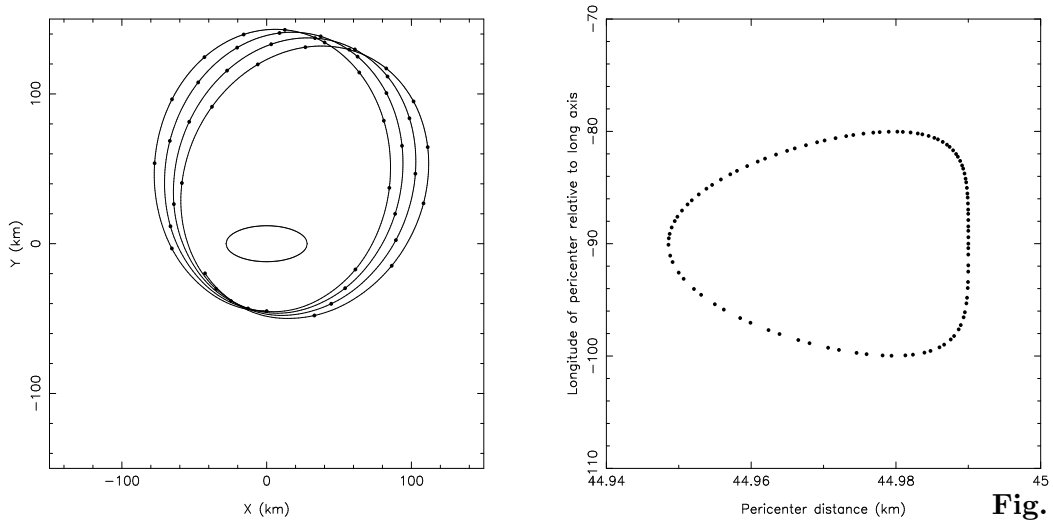



**Fig. 9:** (a) $p$:1 commensurable retrograde orbit plotted in a fixed frame. Dots indicate the position of the satellite when the primary is orientated as shown (central ellipse). This orbit is highly eccentric: $e \sim 0.47$. (b) Pericenter longitude relative to primary's long axis versus pericenter distance.

We did find a class of stable, highly eccentric, retrograde orbits that are commensurate with Ida's rotation period. This discovery came about by noting that, for motion near (but not exactly at) a $p$:1 commensurability, pericenter longitude precesses through 360°, but pericenter distance $q = a(1 - e)$ is correlated with the rotational phase of Ida: when pericenter is located near the long axis of Ida, the pericenter distance is greater than average. Further investigation revealed that orbits could be found which were trapped with pericenter 90° from the long axis. In the case of prograde orbits, the stability zone does not extend as close to the primary as for retrograde orbits. However, we found the same kind of trapping phenomenon. This type of orbit is very strongly stabilized. Consider the case shown in Fig. 9a. This figure shows four revolutions of a retrograde orbit with pericenter near 48 km. The dots indicate the position of the satellite at times when the primary is oriented as shown (central ellipse). Fig. 9b shows

that the longitude of pericenter is locked to $-90°$. The dots in Fig. 9b correspond to the passage at pericenter. This motion is stabilized by crossing pericenter only near the shorter axis of Ida. Similar orbits can be found for the 5:1, 6:1, 7:1, etc., commensurabilities.

Next consider commensurabilities of the form 11:2, 13:2, 15:2, etc. In such cases, the location of pericenter alternates, first on one side of Ida and then on the other. Numerical integration shows that pericenter longitudes still avoid the long axis, just as for the $p$:1 cases. Such a retrograde orbit is shown in Fig. 10a. Fig. 10b (analogous to Fig. 9b) shows the alternating position of pericenter. Even though pericenter can come very close to Ida, the orbit is stable and pericenter is confined far from the long axis so as to protect the satellite from impact.

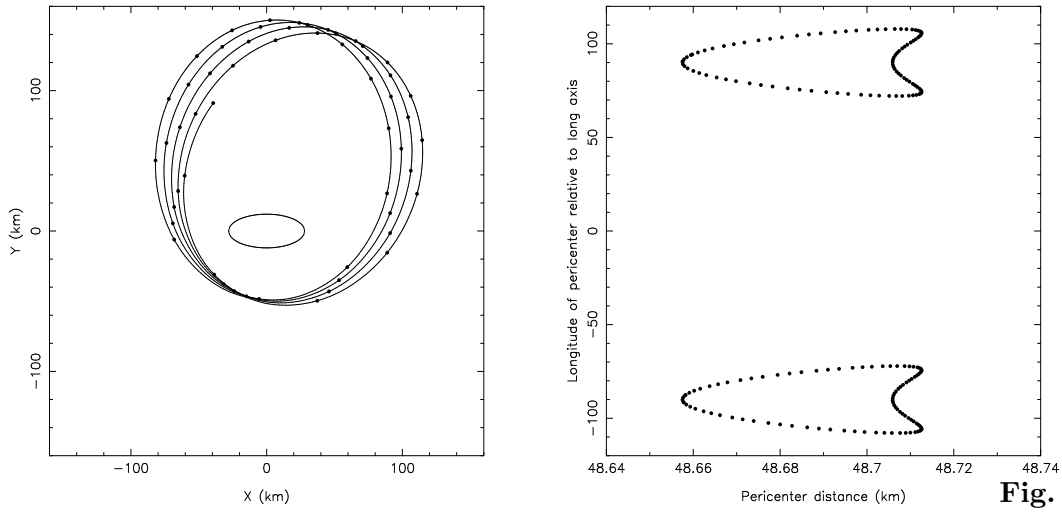



**Fig. 10:** Same as Fig. 9 for a $(2p+1)$:2 commensurability.

These studies indicate that exchanges of energy and angular momentum must occur when the satellite is close to the primary, especially close to the bulge. For the stable orbits, the energy oscillates, but on average does not change. This kind of orbit stabilization may be able to counteract long-term perturbations, such as solar tides or tidal torques, which might otherwise destabilize orbits on timescales comparable to, or less than, the age of Ida.

These stability studies with a triaxial primary identify a locking mechanism between the rotational phase of the asteroid and the sub-satellite longitude and distance of pericenter, maintained by the force due to the elongated bulge at close approach. By holding pericenter away from the bulge, this effect helps stabilize the orbit. To examine the underlying physics of this stabilizing interaction between the primary's bulge and an orbiting satellite, we can approximate the primary's shape as a set of three spheres: a larger central sphere whose

diameter was chosen to match the shorter principal axes of Ida, and two smaller spheres placed

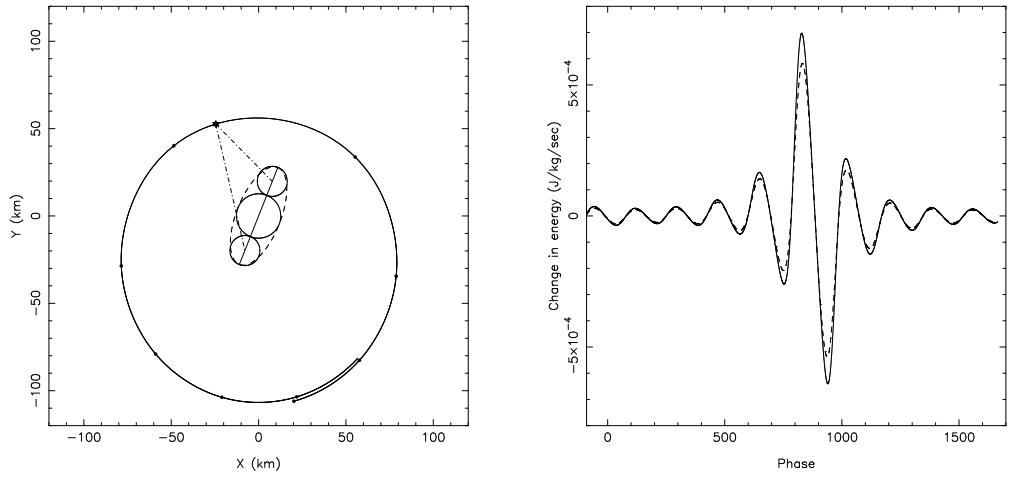



**Fig. 11:** (a) 5:1 commensurable prograde orbit plotted in a fixed frame. Dots indicate the position of the satellite when it is aligned with the primary's long axis. The open star indicates the satellite position when it is at 45° off short axis before passage at pericenter. The solid straight line through the focus of the orbit is the long axis of the primary at that time. Dash-dotted lines indicate the orientation of the perturbing forces on the satellite. The dashed ellipse is the projection of the primary and the solid circles the three sphere model. (b) Change in energy per unit mass per unit time for the triaxial primary (dashed line) and the three sphere model (solid line). (c) Energy per unit mass of the satellite versus phase relative to the primary's long axis for one orbit.

opposite each other roughly filling the remainder of Ida's long dimension (central body in Fig. 11a).

When the satellite is about 45° away from the long axis, the perturbation of the two small spheres is maximum. Obviously, any effect is stronger close to pericenter than at any other location along the orbit. In Fig. 11a, we represent an orbit at the 5:1 commensurability, in the inertial frame, the dots show the location of the satellite when it is aligned with the long axis of the primary. The open star denotes its position when it is at 45° off the short axis, just before passage at pericenter, and the solid straight line through the focus of the orbit is the long axis of the primary at that time. The two dash-dotted lines indicate the orientation of the perturbing forces on the satellite, the dashed ellipse is the projection of the ellipsoid and the three solid circles are the projections of the three spheres. If pericenter is just aligned with the small axis of the primary, then the locations when the satellite is 45° from the short-axis, both before and after passage at pericenter, are symetrical and the two perturbations cancel each

other. On the other hand, if passage at pericenter occurs before (or after) passage along the short axis, then the perturbation felt for the preceeding passage at 45° is stronger (or weaker) than the one felt for the following passage at 45°. We estimated the change in orbital energy of the satellite due to the gravitational interactions with the two long axis bulges (simulated by the two smaller spheres) along the orbit. For this, we computed the difference in acceleration due to the three spheres and the keplerian acceleration due to a point mass with the same mass as the primary, and multiplied it (dot product) by the orbital velocity. This energy change was compared with the energy change found directly from the numerical integration with the triaxial primary (Fig. 11b, solid is for the three sphere model, dashed line for the triaxial ellispoid model). The magnitude and sign of the energy changes found in both cases compared quite favorably, verifying the perturbing influence of the bulges in changing the orbital energy of the satellite. The net effect is a change of the mean energy of the orbit, resulting in a change in the orbital period. Fig. 11c shows the energy per unit of mass of the satellite during the orbit shown in Fig. 11a. One can easily see the assymetry around passage at pericenter (phase between 800° and 1000°). For example, the average energy per unit mass before pericenter is about -19.1 J/kg, while after, it is about -18.9 J/kg.

This change in orbital energy (and, hence, semimajor axis) results in a slight change in the orbital period, causing the time of the next pericenter passage to slightly advanced or retarded, depending on the phase of pericenter relative to the long axis. This causes a precession in the longitude of pericenter which always acts to return the longitude of pericenter toward its location at 90° longitude. Thus, a pericenter location aligned with the short axis of a triaxial primary exists in a kind of potential minimum: any precession of pericenter off the short axis results in a "torque" (due to interactions with the long axis bulges) returning pericenter toward 90° longitude. The orbit of such a hypothetical satellite is thus stabilized, motivating us to search for similar stable, commensurate orbits about the real Ida which would be consistent with the observed orbit and locations of Dactyl.

## 3.2 Numerical Integrations With a Realistic Shape Model

Using a realistic shape model for Ida (Thomas *et al.* 1996) and numerical integrations of Dactyl's motion, we searched for resonant orbits which were consistent with the fit Keplerian orbits (Hurford *et al.* 1995). Here again, we consider only orbits in Ida's equatorial plane. However, this can be inforced only for the initial conditions. Then, due to the irregular shape of Ida, the orbit goes off-plane, but the inclination never exceeds a few hundredth of a degree. The 5:1, 6:1, and 7:1 commensurabilities were low-order cases within the range of likely stable orbits found from our previous simulations, and within the range of candidate osculating orbital periods from the Galileo data (Byrnes and D'Amario 1994). For a candidate commensurable

period, we interpolate the Byrnes and D'Amario fit orbits with a step of 0.01 g $cm^{-3}$ in density over the range bracketing the correct orbital period. We use the density corresponding to the interpolated orbital period closest to the one under consideration, and get inital conditions of velocity and position for Dactyl. After numerically integrating the motion of the satellite for one complete orbit, the initial conditions are adjusted to reach the correct period. Once the commensurable period have been found, the orbital eccentricity is compared with the interpolated Byrnes and D'Amario value.

As we have already seen in the triaxial simulations, the rotation of elongated Ida under the orbiting satellite causes regular, periodic variations in the orbital osculating elements. The short period variation, synodic with the revolution of Dactyl, is caused by rotation of the long axis of Ida under the satellite. This variation is modulated by the longer orbital period of Dactyl, with the larger variation in orbital elements corresponding with the times of pericenter passage. In order to compare our values of the orbital parameters with the Galileo fit osculating elements, we time averaged the elements obtained from our numerical integrations over one complete orbital period.

If the value of $e$ does not fit the osculating $e$ for the specified density, the initial conditions are adjusted, along the "commensurable" orbit family, until an $e$ matching the Keplerian fit has been found. Given these density and initial conditions, we perform an integration for many orbit periods to verify that the orbit is stable (pericenter librating about a fixed Ida longitude). If the orbit is indeed stable and librating, we compare the averaged semi-major axis $a$ to the interpolated one. If it is larger, then we change the density to the next lower, if it is smaller, we change to the next higher density and we iterate the procedure. This gives a family of orbits commensurable with Ida's rotational period which actually intersect with the observational fit orbits in only one point, as seen in Fig. 12.

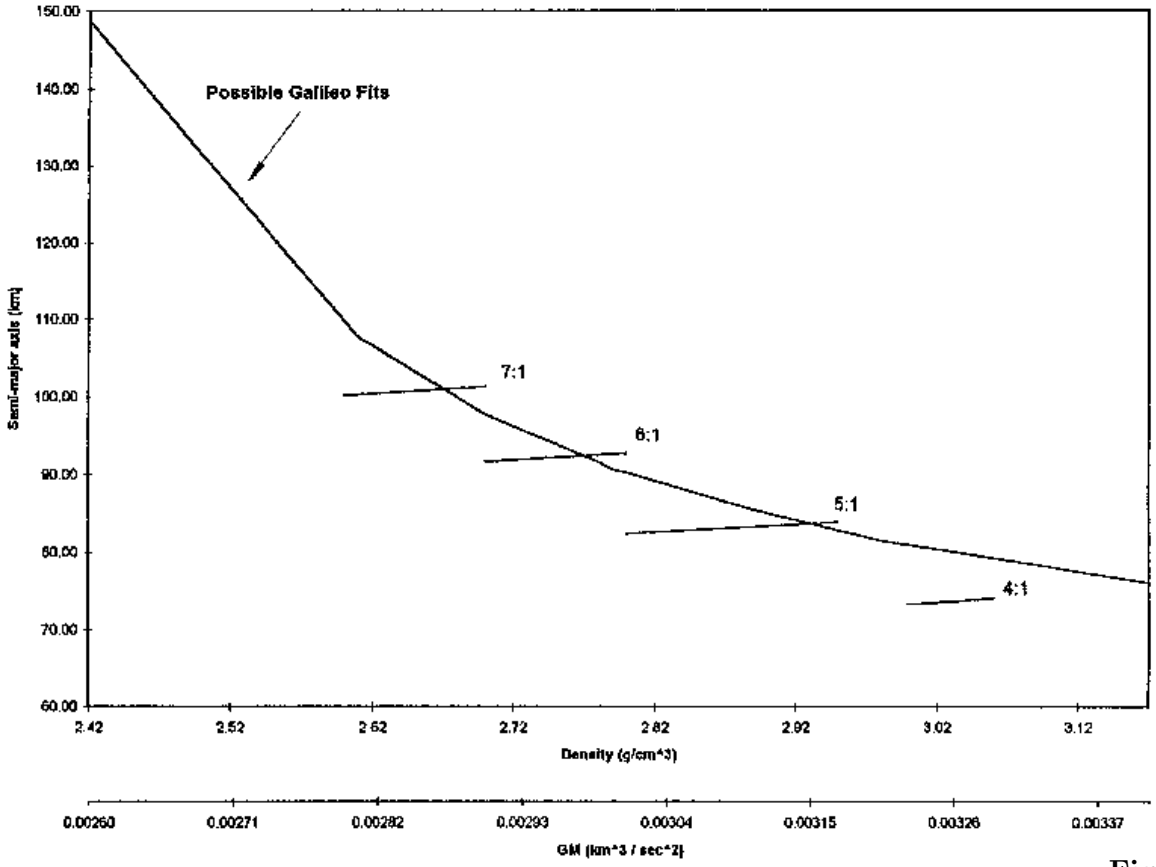



**Fig. 12:** Comparison between Galileo fit orbits and $p$:1 commensurable orbits. The commensurable orbits are forced to have the eccentricity of the fit orbit with the same mass of Ida.

As an example, Fig. 13a shows the oscillation of the longitude of pericenter for the 5:1 resonant orbit having values of $a$ and $e$ matching the osculating orbits found by Byrnes and D'Amario. Fig. 13b shows the orbit with Ida in its orientation when Dactyl is at pericenter (filled circle). As seen in our triaxial models, pericenter is locked away from the longer axis of Ida, although centered near a longitude of 45°, only on one side, instead of 90°, not at all surprising considering Ida's complex, irregular shape. For the 6:1 and 7:1 commensurabilities also the pericenter librates about a longitude which is not aligned with the short axis. The pericenter longitude of the zero libration resonant orbit for the 6:1 and the 7:1 commensurability is around 40°. The libration amplitude may be reduced from the one illustrated by adjusting the initial conditions of the integration while maintaining the values of $a$ and $e$ which allow the

orbits to match those observed by Galileo. The pericenter of the zero libration resonant orbit is located near the centroid of the plotted libration paths.

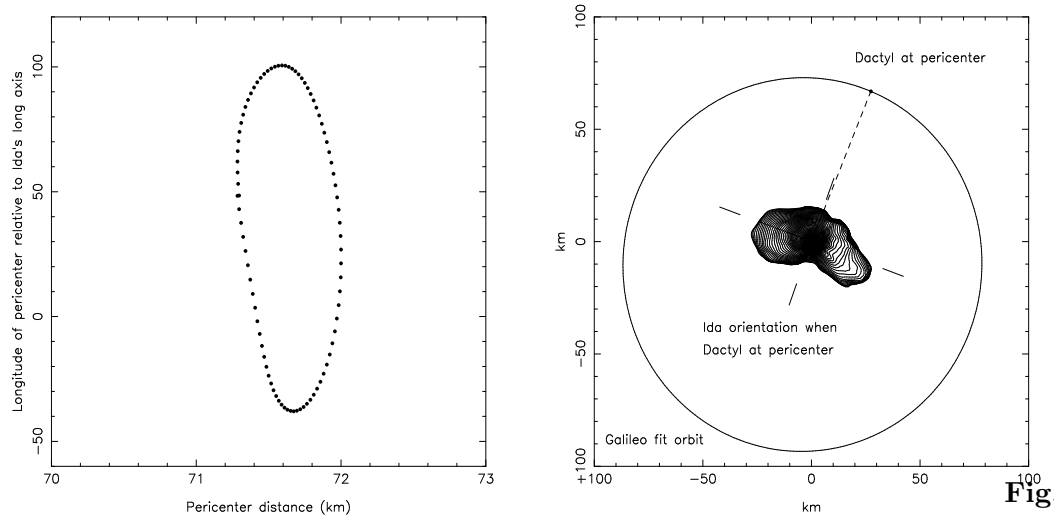



**Fig. 13:** (a) Longitude of pericenter relative to Ida's long axis versus pericenter distance for a possible fit orbit in the 5:1 commensurability. (b) Fit orbit with Dactyl at pericenter and Ida oriented as should be from Galileo images and time necessary for Dactyl to reach pericenter.

## 3.3 Global view of $p$:1 commensurabilities

We want to have a more global view of the $p$:1 commensurabilities, in terms of their location and width in the ($a$, $e$) plane, for each of the allowed fit orbits of Byrnes and D'Amario. Therefore, we consider the hamiltonian of the system in the rotating frame, written in term of "actions" and "angles". For a resonant orbit, one of the angles and its conjugate action do not evolve. Hence, the hamiltonian has an extremum value with respect to these variables and the resonance. Let us write the hamiltonian in cartisian coordinates:

$$\mathcal{H} = \frac{\mu}{2}(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) - \omega\mu(x\dot{y} - y\dot{x}) + U(x, y, z), \tag{6}$$

where the dots designate time derivatives, $\mu$ is the product of the mass of the primary by the gravitational constant, and $U(x, y, z)$ is the gravitational potential generated by the primary. We introduce the Delaunay variables:

$$l \;=\; \text{mean anomaly}$$

$$
\begin{aligned}
L &= \sqrt{\mu a} \\
g &= \text{argument of pericenter} \\
G &= \sqrt{\mu a(1-e^2)} \\
h &= \text{longitude of node} \\
H &= G\cos(i)
\end{aligned}
\tag{7}
$$

where $i$ is the inclination of the orbit. Here and in the following, the capital letters represent action variables, while lowercase letters stand for the the conjugate angle variable. We then rewrite the hamiltonian as:

$$
\mathcal{H} = -\frac{\mu^2}{2L^2} - \omega H + U_{pert}(x,y,z), \tag{8}
$$

with

$$
\begin{aligned}
\mu x &= \left[\cos(h)\cos(g) - \frac{H}{G}\sin(h)\sin(g)\right] L^2(\cos(E)-e) \\
&\quad + \left[-\cos(h)\sin(g) - \frac{H}{G}\sin(h)\cos(g)\right] GL\sin(E) \\
\mu y &= \left[\sin(h)\cos(g) + \frac{H}{G}\cos(h)\sin(g)\right] L^2(\cos(E)-e) \\
&\quad + \left[-\sin(h)\sin(g) + \frac{H}{G}\cos(h)\cos(g)\right] GL\sin(E) \\
\mu z &= \sqrt{1-\frac{H^2}{G^2}}\sin(g)L^2(\cos(E)-e) \\
E - e\sin(E) &= l
\end{aligned}
\tag{9}
$$

$E$ being the eccentric anomaly, and $U_{pert}(x,y,z)$ the perturbation potential, *e.g.* the difference between the true potential and the potential generated by a point with the same mass as the primary and located in its center of mass. In order to reduce the number of dimensions and make the problem computationally tractable, we restrict ourselves to orbits lying in the equatorial plane, *e.g.* $i = 0$. To be sure that an orbit starting in the equatorial plane will remain in this plane, we use a shape model symmetric with respect to the $(x,y)$ plane. This leads to a canonical change of variables:

$$
\begin{aligned}
\gamma &= g + h \\
\Gamma &= H \\
m &= g \\
M &= G - H
\end{aligned}
\tag{10}
$$

$\gamma$ represents the longitude of pericenter. Since Ida is rotating fast, the pericenter longitude is moving fast in retrograde direction with respect to Ida's long axis. In the equatorial plane $M = 0$ and the angle $m = g$ is undefined, so $\mathcal{H}$ can be considered as a function of 4 variables only. We are interested in commensurability between the orbital period and the rotational period of the primary. Hence, we expect the angle $-pl - \gamma$ to be almost constant, or slowly varying. So we introduce another canonical change of variables:

$$\begin{aligned} \sigma &= -pl - \gamma \\ S &= -\Gamma \\ n &= l \\ N &= L - p\Gamma \end{aligned} \tag{11}$$

When we are close to a $p$:1 commensurability, $n = l$ varies rapidly compared to $\sigma$. So we look at the averaged hamiltonian

$$\bar{\mathcal{H}} = \frac{1}{2\Pi} \int_0^{2\Pi} \mathcal{H} \, \mathrm{d}n. \tag{12}$$

The averaged hamiltonian $\bar{\mathcal{H}}$ no longer depends on $n$, the mean anomaly. Remember that $n$ and $N$ are conjugate variables and $\dot{n} = \frac{\partial \bar{\mathcal{H}}}{\partial N}$ and $\dot{N} = -\frac{\partial \bar{\mathcal{H}}}{\partial n}$. Hence $N$ is a constant of motion and can be considered as a parameter. There are only 2 variables left, and we look for extrema of the surface $\bar{\mathcal{H}}(S, \sigma)$. A contour plot of such a surface is given in Fig. 14 for $\mu = 0.00315\,\mathrm{km^3 s^{-2}}$, $N = -2020250\,\mathrm{m^2 s^{-1}}$ corresponding to $a = 82.64$ km and $e = 0.127$ for the value of $S = -506085\,\mathrm{m^2 s^{-1}}$ at resonance point A. One can see that there exists only one allowed orientation of the pericenter for a resonant orbit, on one side of the primary, and not 2 as in the case of a symmetrical primary.

For a given value of Ida's mass, we derive the order of commensurability $p$ as being the integer closest to the fit orbit's orbital period divided by the rotational period of Ida. Then we consider a range of values for $N$, and for each of these values, we look for the extremum of $\bar{\mathcal{H}}(S, \sigma)$. Once we have obtained $N$ and $S$, we compute $a$ and $e$ which give us the location of the exactly resonant orbit for that particular mass of Ida. For the same value of $\sigma$ as for this extremum, we search for the two values of $S$ yielding to the same value of $\bar{\mathcal{H}}(S, \sigma)$ as at the saddle point (Fig. 14, point B). This gives the limits of the resonance zone. The solid lines in Fig. 15 represent the resonant orbits, while the dashed lines limit the resonance zones. Fit orbits are on the dot-dashed line, and the symbols represent the interpolated fit orbit for the for the same mass of Ida as for the resonant orbits. Actually, the values of $a$ and $e$ are averaged values over one orbit. But they are good approximations of $a$ and $e$ values of commensurable orbits at apocenter. Hence we use these values as osculating elements at apocenter, derive position and velocity of the satellite at apocenter assuming a keplerian orbit, and then integrate the orbit

for at least one libration around the resonant point. Then we average the $a$, $e$ and direction of apocenter and iterate the process. After two or three iterations, one gets an orbit which is

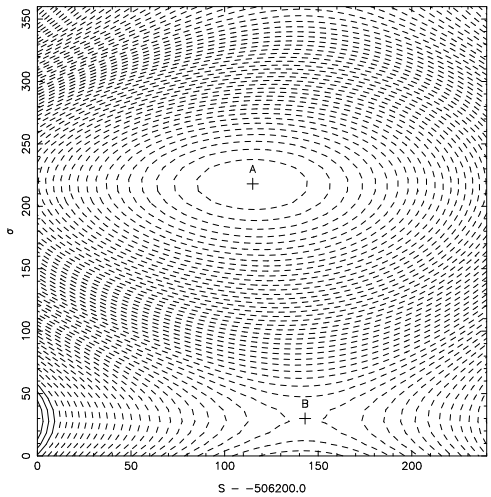



**Fig. 14:** Contour plot of the averaged hamiltonian as a function of $S$ and $\sigma$ for $N = -2020250\,\mathrm{m^2 s^{-1}}$ and $p = 5$. Point A denote the resonance point, and point B the saddle point.

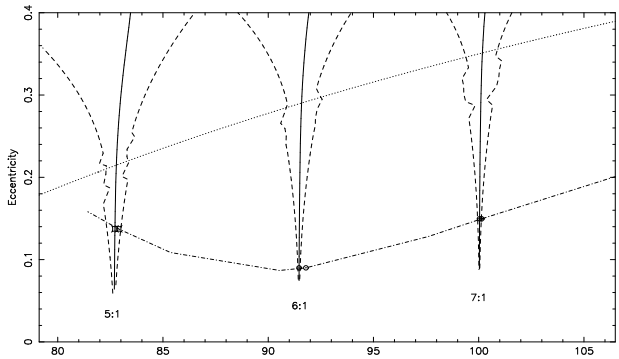



**Fig. 15:** Location of resonant orbits in the $(a, e)$ plane. Solid lines represent the commensurabilities, dashed lines limit the regions of librating orbits, the dotted line corresponds to a pericenter distance of 65 km and the dash-dotted line shows the fit orbit familly. The pairs of symbol show the location of resonant orbits together with the corresponding fit orbit (same $e$, same mass of Ida). The roughness of the resonance zones is due to the appearance of a second resonance on the other side of Ida for large orbits.

fairly close to the resonant one (small libration).

The map of resonances gives a hint of the reason for the chaotic behaviour of orbits with $M \leq 3.6 \times 10^{16}$ kg. In Fig. 16, we present the $p$:1 commensurabilities in the vicinity of the fit orbit for GM $= 0.0024\,\mathrm{km^3 s^{-2}}$. The filled circle represents the fit orbit, its size being approximately the size of Dactyl. Remember however that the coordinates are only averaged values of $a$ and $e$. Hence, during its wandering around Ida, Dactyl will enter the resonances which overlap a lot for eccentricities not much larger than Dactyl's one. Overlapping resonances lead to chaos. And this chaotic region is easily connected to infinity, hence the possible escape of the satellite.

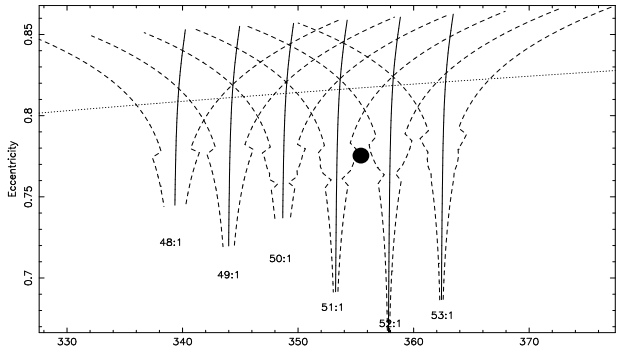



**Fig. 16:** Same as Fig. 15 for the fit orbit with GM $= 0.0024\,\mathrm{km^3 s^{-2}}$.

## 3.4 Which Resonant Orbits Fit The Galileo Data?

If Dactyl indeed exists on one of these stable, commensurate orbits, not only must the orbital elements of the satellite be consistent with the observations, the rotational orientation of Ida relative to Dactyl's position must match as well. At the epoch 28 August, 1993 16:52:05 UT all the osculating orbits pass through a reference point $\sim$ 90 km from Ida at longitude 85°, indicated in Fig. 17. Several of the osculating Byrnes and D'Amario orbits are shown for a range of discrete values of GM for Ida. For comparison the 5:1, 6:1, and 7:1 resonant orbits found from our numerical integrations have GM $= 0.00315$, 0.00298, and $0.00287\,\mathrm{km^3 s^{-2}}$, respectively (remember however that these values depend on the shape model used to perform the integrations). Knowing the period and orientation in space of the osculating orbits allows us to advance Dactyl along a given orbit from the reference position to the pericenter of that orbit. During the time of travel to pericenter Ida rotates from its indicated orientation, resulting in a pericenter passage for the 5:1, 6:1, and 7:1 resonant orbits at Ida longitude in the range $80° - 95°$, $250° - 260°$, and $90° - 100°$, respectively (Fig. 13b for the 5:1 case). In Fig. 18, we plot the Ida longitude of Dactyl at pericenter passage versus semi-major axis for one of the

shape models. Here again, the solid lines represent the resonant orbits and the dashed lines the limit of the resonance zone. Interpolated fit orbit are marked by symbols on the dot-dashed line which shows the family of fit orbits. Corresponding symbols on the solid lines show the resonant orbits for the same mass of Ida. From our numerical integrations we know that for each of the resonant orbits Dactyl's pericenter occurs at an Ida longitude in the range $\sim 30^\circ - 110^\circ$.

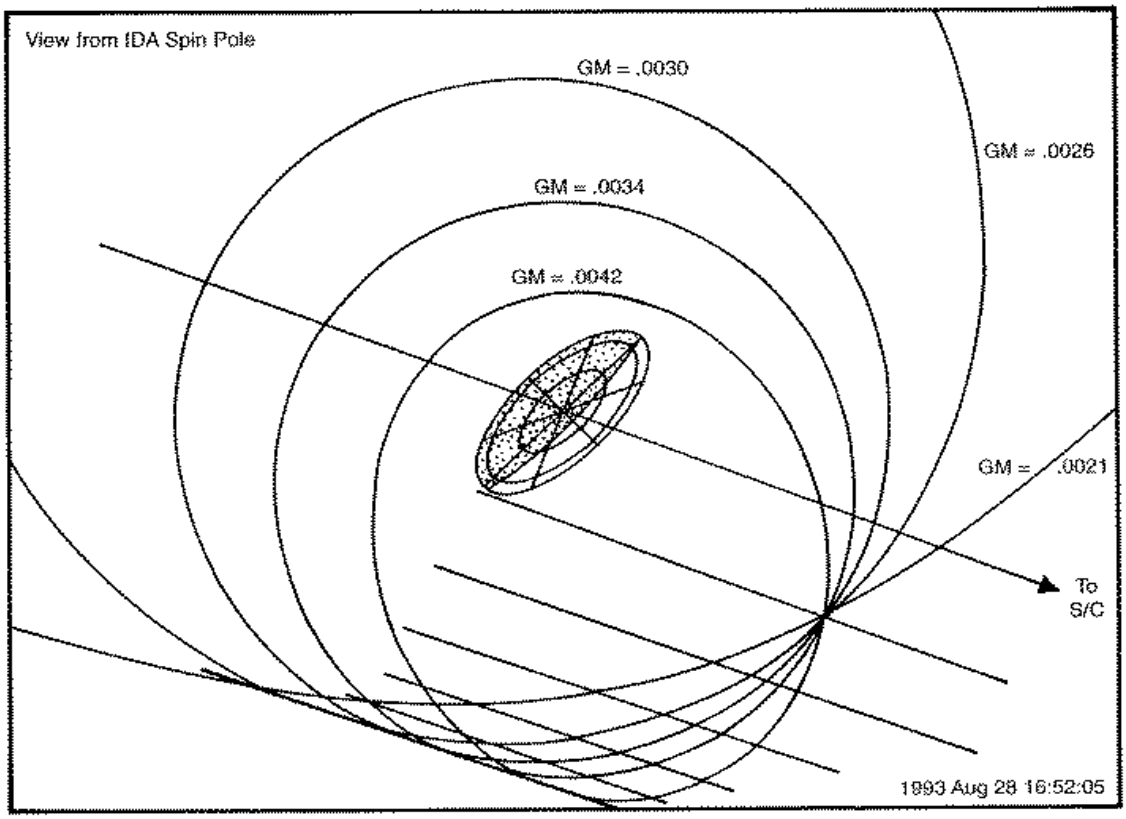



**Fig. 17:** Possible fit orbits of Dactyl for different values of parameter GM. The central ellipse shows the orientation of Ida at eqoch 28 August, 1993 16:52:05 UT, Dactyl being located at the intersection of all orbits. The arrows point toward the spacecraft.

The observed position of Dactyl and orientation of Ida at the Galileo fly-by epoch is consistent with pericenter in the stable range for the 5:1 and the 7:1 resonant orbits. For the nominal values of the true anomaly at the epoch of Galileo encounter, we need large amplitude librating orbits. The position of Dactyl is known with a rather high accuracy, but the uncertainty on the

velocity is of the order of 10%. Hence, the determination of the true anomaly, or the orientation of the pericenter, is quite uncertain. The error on that angle can be larger than 10°, which in terms of Ida's orientation translates to a 50° or 70° possible variation. On the 6:1 orbit Ida would be oriented almost 180° from the required phase at pericenter and can therefore be ruled out as a match to the observations. The 11:2 and 13:2 commensurabilities were examined as well and may also be ruled out due to inconsistent Ida orientations at Dactyl pericenter passage.

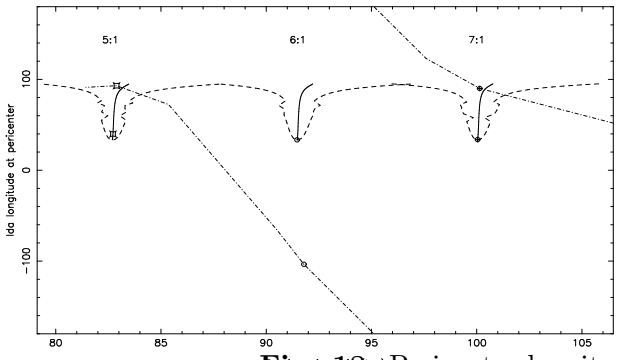



**Fig. 18:** Pericenter longitude relative to Ida's long axis versus semi-major axis. The meaning of each line is the same as in Fig. 15.

Looking at the width of the resonance zones (Fig. 15), one can see that it is very unlikely that a body the size of Dactyl can be trapped in the 7:1 resonance. For the nominal eccentricity the width is of the order of 200 m, and it reaches 1.4 km (size of Dactyl) only for $e > .3$. In the case of the 5:1 resonance, the width for the fit eccentricity is of the order of 700 m and one needs to get only to $e = .24$ to reach the size of Dactyl. Therefore, the 5:1 commensurability seems to be the one in which Dactyl is most likely to be trapped if it is indeed in a resonance.

The value of $M = 4.72 \times 10^{16}$ kg (Byrnes and D'Amario 1994) for the 5:1 orbit would correspond to a density of Ida of $\rho = 2.93$ g cm$^{-3}$ for a volume of 16,100 km$^3$ (Thomas *et al.* 1996).

# 4 Long-term stability of resonant orbits

We have shown that of the low order resonant orbits within the zone of stable orbits about Ida, it is likely that only the 5:1 commensurability allows Dactyl to have been at Ida longitudes consistent with the Galileo observations while being in a large isolated resonance. Next we consider whether this resonant orbit is any more stable over long timescales than nearby, non-commensurate orbits. To address this question we have performed a series of very long timescale

(at least thousands of years) numerical integrations of the 5:1 orbit and neighboring non-resonant orbits. Owing to our determination of commensurable orbits, we considered only orbits with zero inclination (*e.g.* in the equatorial plane of Ida) but with the same other osculating elements as the fit orbits of Byrnes and D'Amario. This is not greatly different from the actual Dactl's inclination of $\sim 8^\circ$ and this does not change much the results. Our preliminary results indicate that the resonant orbit offers no significant advantage over other orbits in stabilizing Dactyl for the long term. We have studied the time evolution of the Lyapunov characteristic indicator for the 5:1 orbit The result is very similar to the one shown in Fig. 7. The logarithm of the Lyapunov indicator is decreasing linearly with logarithm of time for at least $10^{11}$ sec, indicating a stable, non-chaotic orbit for Dactyl with zero Lyapunov exponent (the limit at infinite time of the Lyapunov indicator). We also looked at two nearby, non-resonant, initialy in equatorial plane orbits and found a time evolution of the Lyapunov characteristic indicator which is undistinguishable from the 5:1 commensurability one, showing that these to orbits are just as stable. Although this study is still in progress, our preliminary conclusion is that although consistent with the observed location of Dactyl, from the standpoint of relative stability the 5:1 orbit offers no special benefit in the long-term dynamical stability of Dactyl, in the absence of dissipation.

In another approach to understanding resonant orbits as a stabilizing influence, we are examining the role resonances play in stabilizing the decay of orbits due to dissipative forces. As an example we study the case of collisional drag. As Dactyl orbits Ida it also passes through a variable cloud of impact generated orbital debris, lofted into temporary trajectories about Ida after large cratering impacts (Geissler *et al.* 1996). Passage through this debris creates a small drag on the satellite, causing its orbit about Ida to slowly decay. Regardless of Dactyl's provenance, it may have since been captured into a stable resonant orbit after evolving inward toward Ida due to debris-induced drag. Fig. 19a shows the evolution of the semimajor axis of an Ida satellite near the 5:1 resonance under the influence of a gas-drag-like force. The orbit is seen to decay at a steady rate until captured into the 5:1 resonance. As is typical in a resonance with an induced drag force, the eccentricity of the orbit slowly increases (Fig. 19b) until the satellite enters a secondary resonance overlapping the primary one (Henrard and Moons 1992). The orbit then evolves along this secondary resonance on a chaotic path. Eventually, it crosses the separatrix of the primary resonance and continues to evolve inward. Of course the drag force studied in this case is very large and is simply used to illustrate the nature of the capture and stability properties of a resonance. Actual debris drag forces for Dactyl will be much smaller in magnitude and may not be strong enough to remove the satellite from a resonance once it has been captured. Therefore, a satellite evolving under dissipative forces will be trapped in the 5:1 resonance and may stay there for geological times. Further work on this problem is underway.

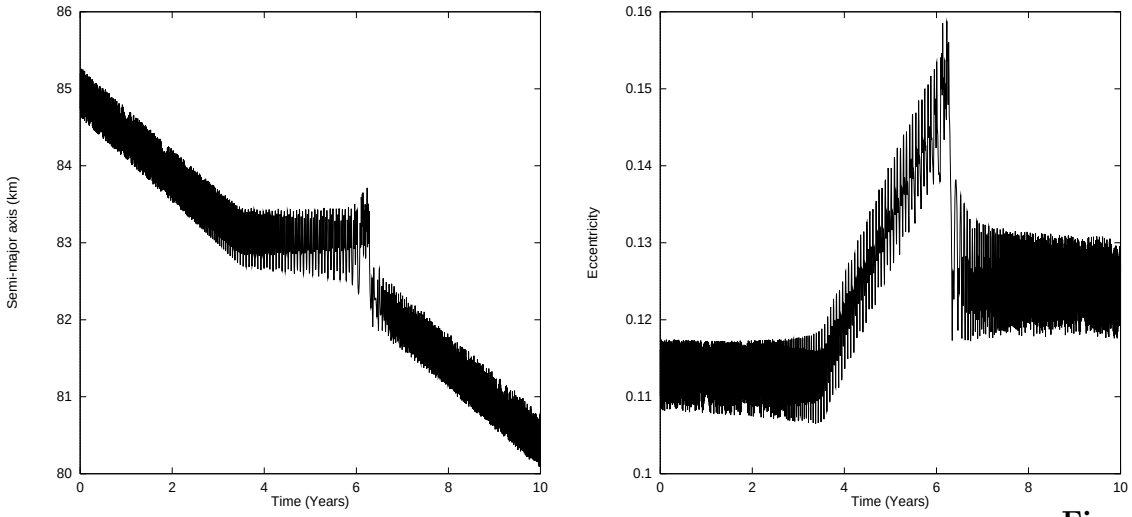



**Fig. 19:** (a) Time evolution of the semi-major axis of the fit orbit for GM $= 0.0031\,\mathrm{km}^3\mathrm{s}^{-2}$, when subject to a gaz drag like dissipative force. The dissipative force per unit mass is of the form $-\lambda\vec{V}$ where $\vec{V}$ is the velocity in the inertial frame, and $\lambda = 10^{-10}\,\mathrm{s}^{-1}$. Between $10^8$ and $2 \times 10^8$ seconds, the orbit is trapped in the 5:1 resonance. (b) Time evolution of the eccentricity. While the satellite is trapped in the 5:1 resonance, its eccentricity increases until it reaches a secondary resonances and eventually exits from the primary resonance.

# 5 Conclusions

We have tested the stability of orbits around Ida with numerical integrations and find that Dactyl is restricted to orbits with pericenter distances greater than 65 km. Galileo observations of Dactyl's positions are fit by a family of orbits parameterized by the mass of Ida. Our stability studies thus allow us to constrain the density of Ida to $\rho \leq 3.1\,\mathrm{g\ cm}^{-3}$. On the low mass end of the family, Belton *et al.* (1995, 1996) ruled out the hyperbolic orbits. Using numerical integrations with a realistic shape model for Ida, we were able to constrain the low end limit of allowed orbits to a higher density. For $\rho \leq 2.3\,\mathrm{g\ cm}^{-3}$ the elliptical fit orbits are actually chaotic and go on escaping orbits after some 1000 years. We numerically integrated the fit orbits for $\rho$ in the range $2.3\,\mathrm{g\ cm}^{-3} - 3.07\,\mathrm{g\ cm}^{-3}$ over more than 3000 years and found no hint of instability on this time scale. We argue that in order to have discovered Dactyl at all, it must exist in an orbit that is dynamically long-lived and we suggest orbits commensurate with the rotation of Ida as good candidates. We demonstrate that such resonant orbits are stabilized by

interactions with the long-axis bulges of an elongated primary. We find that of the candidate resonant orbits consistent with the observed family of orbits, only the 5:1 resonance allows Dactyl at Ida longitudes consistent with the Galileo observations of Dactyl's position and Ida's orientation. If Dactyl indeed exists on this resonant orbit, Ida's density is constrained to one specific value: $\rho = 2.93\,\mathrm{g\ cm^{-3}}$.

# Acknowledgments

This work was supported by NASA grant NAGW-1029.

# References

[1] Belton, M. J. S. & Carlson, R., 1994. *1993 (243).* IAU Circ. p. 5948.

[2] Belton, M., Chapman, C., Thomas, P., Davies, M., Greenberg, R., Klaasen, K., Byrnes, D., D'Amario, L., Synnott, S., Merline, W., Petit, J-M., Storrs, A., Zellner, B. & the Galileo Imaging Team, 1995. *The bulk density of asteroid 243 Ida from Dactyl's orbit.* Nature **374**, pp. 785–788.

[3] Belton, M., Mueller, B., D'Amario, L., Byrnes, D., Klaasen, K., Synnott, S., Breneman, H., Johnson, T., Thomas, P., Veverka, J., Harch, A., Davies, M., Merline, W., Chapman, C., Davis, D., Denk, T., Petit, J-M., Greenberg, R., McEwen, A., Storrs, A. & Zellner, B., 1996. *The discovery and orbit of 1993 (243) 1 Dactyl.*. Icarus **120**, pp. 185–199.

[4] Benettin, G.L., Giorgilli, A. & Strelcyn, J.M., 1980. Meccanica **15**, p. 10.

[5] Byrnes, D. & L. D'Amario, 1994. Galileo Imaging Team Report, unpublished.

[6] Chauvineau, B., Farinella, P. & Mignard, F., 1993. *Planar orbits about a triaxial body: Application to asteroidal satellites.* Icarus **105**, pp. 370–384.

[7] Davis, D., Chapman, C., Durda, D., Farinella, P. & Marzari, F., 1996. *On the formation and collisional/dynamical evolution of the Ida/Dactyl system as part of the Koronis family.* Icarus **120**, pp. 220–230.

[8] Durda, D., 1996. *The formation of asteroidal satellites in catastrophic collisions.* Icarus **120**, pp. 212–219.

[9] Froeschlé, C., 1984. Numéro spécial du journal de Mécanique théorique et appliquée p. 101.

[10] Geissler, P., Petit, J-M., Durda, D., Greenberg, R., Bottke, W., Nolan, M. & Moore, J., 1996. *Erosion and Ejecta Redistribution on 243 Ida and its Moon.* Icarus **120**, pp. 140–157.

[11] Henrard, J. & Moons, M, 1992. *Capture probabilities for secondary resonances.* Icarus **95**, pp. 244–252.

[12] Hurford. T., Durda, D., Greenberg, R., Petit, J-M., Geissler, P. & Bottke, W., 1995. *Stable periodic orbits around Ida: Several fit Dactyl's motion.* Bull. Am. Astron. Soc. **27**, pp. 1070–1071.

[13] Laskar, J., 1990. *The chaotic motion of the solar system: a numerical estimate of the size of the chaotic zones.* Icarus **88**, pp. 266–291.

[14] Petit, J.M., Greenberg, R. & Geissler, P., 1994. *Orbits around a small, highly elongated asteroid: Contraints on Ida's Moon.* Bull. Am. Astron. Soc. **26**, pp. 1157–1158.

[15] Scheeres, D., 1994. *Dynamics about uniformly rotating triaxial ellipsoids: application to asteroids.* Icarus **110**, pp. 225–238.

[16] Thomas, P., Belton, M., Carcich, B., Chapman, C., Davies, M., Sullivan, R. & Veverka, J., 1996. *The shape of Ida.* Icarus **120,**, pp. 20–30.