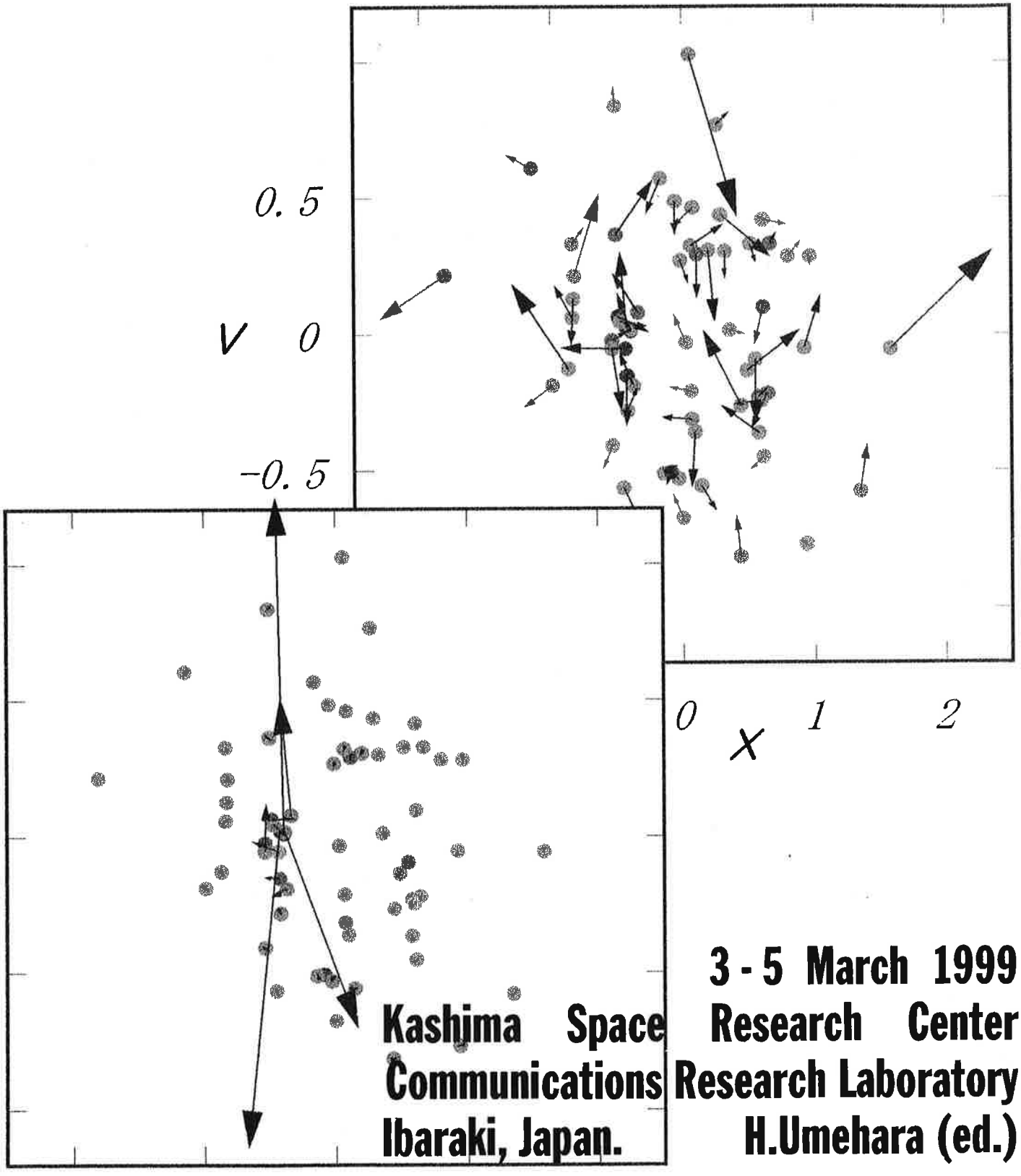
Proceedings of the 31st Symposium on Celestial Mechanics

天体力学・N体力学研究会集録



3 - 5 March 1999

Kashima Space Research Center

Communications Research Laboratory

Ibaraki, Japan.

H.Umehara (ed.)

Preface

『天体力学・N体力学研究会 (The 31st Symposium on Celestial Mechanics)』は、1999年3月3日より5日までの日程で、郵政省通信総合研究所関東支所鹿島宇宙通信センター (茨城県鹿嶋市) において開催されました。出席者総数は約70名、招待講演2、口頭発表22、ポスター発表23におよび、昼夜問わず議論の絶えない盛況な研究会となりました。

本研究会は、昨年まで国立天文台 (東京都三鷹市) 等で行なわれた『第n回天体力学研究会』($n \leq 30$) に引き続くものです。第30回 (1998年) からは研究会をより発展させるため、開催地・世話人などを流動化させる方針にしました。更に、今回は主要テーマを掲げ、N体が相互作用をした結果おこる集団運動の理論づくりを中心とした非線形力学系の問題や、この現象の工学への応用もとり上げました。したがって、研究会の日本語名称を従来のものから変えました。

このような主旨のもと、複数分野で活躍されている相澤洋二氏 (早稲田大学)・土屋俊夫氏 (京都大学) に招待講演をお願いしました。相澤氏には、N体ハミルトン系のクラスター形成現象が非双曲力学系などと関連していることなどについて御講演を頂きました。土屋氏には、多体運動で特徴的な運動をリヤプノフ・ベクトル等を用いて見出したことなどについてお話しして頂きました。その他、N体運動を取り扱う様々な分野の研究交流によって、問題解決や問題提起を生む研究会になったかと思います。

最後になりましたが、第30回で世話人をされた福島登志夫氏・伊藤孝士氏 (国立天文台)・布施哲治氏 (総合研究大学院大学) や、吉川真氏 (宇宙科学研究所) にはプログラム編成などSOC的な援助を頂きました。第29回まで世話人をされていた木下宙氏・中井宏氏 (国立天文台) にも運営のノウハウを御指導頂きました。また、開催にあたり、鹿島宇宙通信センターの藤田正晴氏・川瀬成一郎氏・栗原則幸氏・木村和宏氏、さらに、菅谷明彦氏・山口隆司氏・杉山明彦氏をはじめ同センター管理課の方々より多大な援助を頂きました。その他にも不慣れな世話人ゆえ、多くの方々にお世話になりました。深く御礼申し上げます。

1999年3月 Editor/ 世話人 梅原広明 (Umehara, Hiroaki)
通信総合研究所 関東支所鹿島宇宙通信センター 宇宙制御技術研究室
Kashima Space Research Center, Communications Research Laboratory
E-mail: ume@crl.go.jp, FAX: 0299-84-7160
ume@nict.go.jp (as of 2018)

表紙について

京都大学理学部宇宙物理学教室の土屋俊夫氏提供の1次元重力多体系におけるリヤプノフ・ベクトル図です。詳しくは、本集録 p.9 を参照して下さい。

Table of Contents

Invited Lectures

Many-Body Chaos in Physics

Chaotic Dynamical Systems

Friction by Many Bodies

Three-Body Problems

Planetary Formation and Resonances

Resonances in the Outer Solar System

Distributions around the Solar System

Many Bodies in the Universe

Numerical Techniques

Astrodynamics of N spacecrafts

Long Time Tails in N-body Hamiltonian Systems

Yoji Aizawa

Department of Applied Physics, Faculty of Science and Engineering,
Waseda University, 3-4-1 Ookubo, Shinjuku, Tokyo, Japan

One of the most striking phenomena in chaotic dynamics is the appearence of the long time tails such as the 1/f fluctuations [1,2]. In the nearly integrable hamiltonian systems, the long time tails are universally generated due to the stagnant motions near the invariant KAM tori;

$$H(I,\theta) = H_0(I) + \varepsilon H_1(I,\theta) \tag{1}$$

Here the parameter ε stands for the perturbation to the integrable hamiltonian H_0. An important theorem (Nekhoroshev,1977) explained that the residence time T in the stagnant layer obeys [3],

$$T \simeq \frac{1}{\varepsilon} exp[\varepsilon^{-b}] \ , \ (\varepsilon \ll 1) \tag{2}$$

where b is a positive constant determined by the unperturbed hamiltonian H_0 in Eq.(1).
The significant point in Eq.(2) is that the divergence of T does not obey the inverse power law but exhibits an essential singularity when ε goes to zero. In 1980's, the origin of such singularity was studied in terms of the scaling theory for the stagnant layers mentioned above, where the hierarchical structure of resonant tori (islands arround island) plays an essential role to induce the long time tails in dynamical quantities. The stagnant layer theory (Aizawa , 1989) demonstrated that the distribution of the residence time, say P(T), obeys a universal law [2],

$$P(T) \simeq \frac{1}{T[logT]^c} \ , \ (T \gg 1) \tag{3}$$

where c is a positive constant larger than unity. Equation(3) has been confirmed by simulations (Aizawa *et al*, 1989). The essential singularity in Eq.(2) reflects the onset of $1/T$ divergence in Eq.(3). The point is that the distribution is not normalizable,i.e., a typical infinite measure.
Firstly, my lecture will be directed to the review of the stagnant layer theory and some numerical evidence in many body systems. Secondly, the onset of a new type of long time tails will be discussed carrying out with the clustering motions in N-body systems with short range attractive forces, where the distribution of the trapping time T obeys another universal law,

$$P(T) \simeq T^{-\beta-1} exp[-aT^{-\beta}] \tag{4}$$

where a is a positive constant which depends on the size of cluster. The regularly varying part of Eq.(4) denotes the tail with $T^{-\beta-1}(T \gg 1)$, and the parameter β depends on the dimension of the cluster under consideration. The stability of the clustering motions will be explained based on the long time tail of Eq.(4).

References

[1] Aizawa Y 1989a *Prog.Theor.Phys.Suppl.* **99**149; Aizawa Y 1999 *Chaos, Soliton and Fractals* (in press); Tanaka K and Aizawa Y 1993 *Prog.Theor.Phys.* **90(3)**

[2] Aizawa Y 1989b *Prog.Theor.Phys.* **81(2)** 249; Aizawa Y *et al* 1989c *Prog.Theor.Phys.Suppl.* **98** 37; Aizawa Y 1995 *J.Korean Phys.Soc.* **28** 310; Aizawa Y 1991 *Dynamical Theory of $f^{-\nu}$ spectral chaos*, eds. Musha.T *et al*,pp483-487

[3] Nekhoroshev N N 1977 *Russ.Math.Surveys* **32** 1

Relaxation of One-Dimensional Self-Gravitating Many-Body Systems

Toshio Tsuchiya

Department of Astronomy, Kyoto University, Kyoto 606-8502, Japan

Abstract

The relation between relaxation, the time scale of Lyapunov instabilities, and the Kolmogorov-Sinai time in a one-dimensional gravitating sheet system is studied. Both the maximum Lyapunov exponent and the Kolmogorov-Sinai entropy decrease as proportional to $N^{-1/5}$. The time scales determined by these quantities evidently differ from any type of relaxation time found in the previous investigations. The relaxation time to quasiequilibria (microscopic relaxation) is found to coincide with the inverse of the minimum positive Lyapunov exponent. The relaxation time to the final thermal equilibrium differs to the inverse of the Lyapunov exponents and the Kolmogorov-Sinai time.

I. INTRODUCTION

Relaxation is the most fundamental process in evolution of many-body system. The classical statistical theory is based on ergodic property, which is considered to be established after relaxation. However, not all systems do not show such an idealistic relaxation. A historical example is FPU (Fermi-Pasta-Ulam) problem [1], which experiences the induction phenomenon (e.g., Ref [2,3]) and does not relax to the equipartition for very long time.

From nearly thirty years of investigation, one-dimensional self-gravitating sheet systems (OGS) have been known by their strange behavior in evolution. Hohl [4–6] first asserted that OGS relaxes to the thermodynamical equilibrium (the isothermal distribution) in a time scale of about $N^2\, t_c$, where N is the number of sheets, and t_c is

typical time for a sheet to cross the system. Later, more precise numerical experiments determined that the Hohl's result was not right, and then arguments for the relaxation time arose in 1980's. A Belgian group [7,8] claimed the OGS relaxed in shorter than Nt_c, whereas a Texas group [9,10] showed that the system showed long lived correlation and never relaxed even after $2N^2t_c$. Tsuchiya, Gouda and Konishi (1996) [11] (hereafter TGK) suggested that this contradiction can be resolved in the view of two different types of relaxations: the *microscopic* and the *macroscopic* relaxations.

At the time scale of Nt_c, cumulative effect of the mean field fluctuation makes the energies of the individual particles change noticeably. Figure 1 shows the nature of the energy fluctuation. If the evolution of the system is ergodic in the Γ-space, the long time average of the specific energy takes a unique value for all i, i.e.

$$\overline{\varepsilon_i} \equiv \lim_{T\to\infty} \frac{1}{T}\int_0^T \varepsilon_i(t)dt = \varepsilon_0 \equiv 5E/3. \tag{1}$$

The degree of the deviation from the equipartition is measured by the quantity,

$$\Delta(t) \equiv \varepsilon_0^{-1}\sqrt{\frac{1}{N}\sum_{i=1}^{N}(\overline{\varepsilon_i}(t)-\varepsilon_0)^2}, \tag{2}$$

where $\overline{\varepsilon_i}(t)$ is the averaged value until t. If the system is ergodic and has a finite correlation time (the relaxation time), it behaves like a random number from Markovian process and we can estimate the temporal evolution of $\Delta(t)$. In this case, a trajectory in the Γ-space visits almost every point in the ergodic region, thus the individual particle energy relaxes to equilibrium value and $\Delta(t)$ decreases as $t^{-1/2}$ for the time longer than the relaxation time, according to the central limit theorem. In Fig. 1, *Delta*(t) decreases as $t^{-1/2}$ after $t \sim 100$. This means that the equipartition among particles is established, and the fluctuation is just like thermal noise. Thus there is a relaxation at this time scale. This relaxation is a random walk diffusion in phase space, which is confirmed by the power spectrum density of the energy fluctuation (Fig. 2). In shorter time scale (the frequency $f \gtrsim 10^{-3}$), the spectrum has power of -2, which is typical for random walk diffusion.

FIGURES

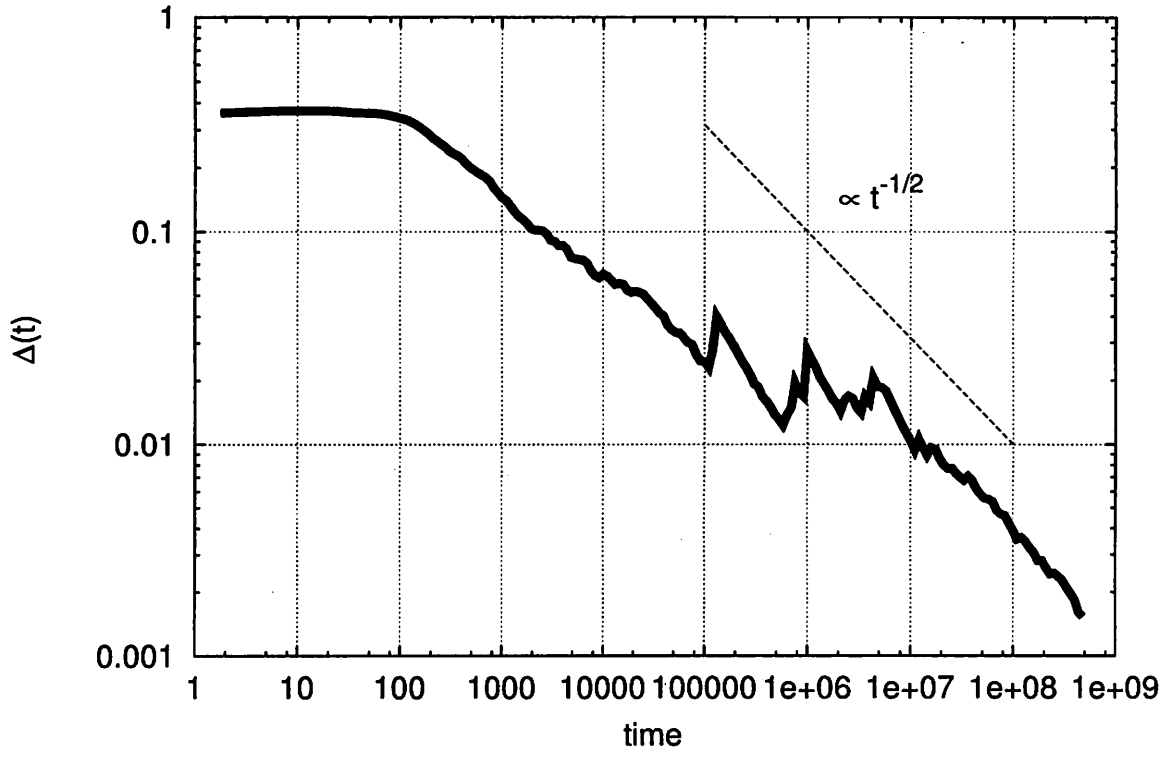


FIG. 1. Deviation from the equipartition, $\Delta(t) \equiv \varepsilon_0^{-1}\sqrt{\{\frac{1}{N}\sum_{i=1}^{N}[\overline{\varepsilon_i}(t) - \varepsilon_0]^2\}}$, where $\overline{\varepsilon_i}(t)$ is the energy of individual particle averaged over time t, and ε_0 is the energy of the equipartition.

By this relaxation the system is led not to the thermal equilibrium but only to a quasiequilibrium. The global shape of the one-body distribution remains different from that of the thermal equilibrium. This relaxation appears only in the microscopic dynamics, thus it is called the microscopic relaxation. The global shape of the one-body distribution transforms in much longer time scale. For example, a quasiequilibrium (the water-bag distribution, which has the longest life time) begins to transform at $4 \times 10^4 N t_c$ in average. This *slow relaxation* is confirmed also in the power spectrum density (Fig. 2). For longer time scale ($10^{-6} \lesssim f \lesssim 10^{-3}$), the power of the spectrum of the energy fluctuation is less than -2. This implies the long time correlation and slow diffusion in the phase space.

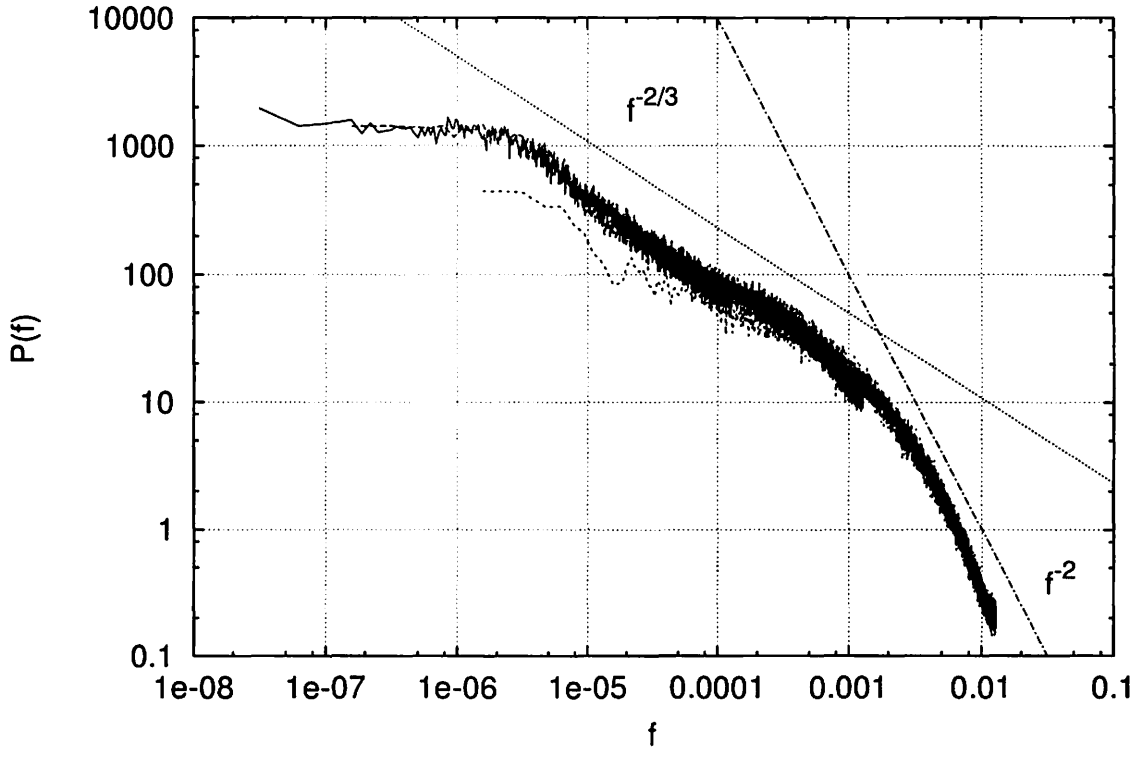


FIG. 2. Power spectrum density of the energy fluctuation of the individual particles.

TGK called this transformation the macroscopic relaxation, but later in Tsuchiya, Gouda and Konishi (1998) [12], it is shown that this transformation is onset of the *itinerant stage*. In this stage, the one-body distribution stays in a quasiequilibrium for some time and then changes to other quasiequilibrium. This transformation continues forever. This itinerancy can be seen as the several peaks in the regime ($10^5 \lesssim t \lesssim 10^7$) in Fig. 1. Probability density of the life time of the quasiequilibria has a power law distribution with a long time cut-off and the longest life time is $\sim 10^4 N t_c$. Power law distributions are often seen in chaotic dynamics, which do not posses typical time scale. Hence it is surmised that some chaotic dynamics is essential in the slow relaxation in our systems.

Only by averaging over a time longer than the longest life time of the quasiequilibria, the one-body distribution becomes that of the thermal equilibrium, which is defined as the maximum entropy state. Yawn and Miller [13,14] also showed that the ergodicity is established not in $10^4 N t_c$, but in several $10^5 N t_c$. Therefore the time $\sim 10^6 N t_c$ is necessary for relaxation to the thermal equilibrium, and called the *thermal relaxation*

time. Although there are some attempts to clarify the mechanisms of these relaxations [11,15,16,12], the reason why the system does not relax for such a long time is still unclear.

At the view of chaotic theory of dynamical systems, relaxation is understood as mixing in phase space, and its time scale is given by the Kolmogorov-Sinai time (KS time), $\tau_{\rm KS} = 1/h_{\rm KS}$, where $h_{\rm KS}$ is the Kolmogorov-Sinai entropy. However, it does not simply correspond to the relaxation of the one-body distribution function, which is of interest in many-body systems. Recently, Dellago and Posch [17] showed that in a hard sphere gas, the KS time equals the mixing time of neighboring orbits in the phase space, whereas the relaxation of the one-body distribution function corresponds to the collision time between particles. Now, it is fruitful to study relation between relaxation and some dynamical quantities, such as the KS entropy and the Lyapunov exponents, in the OGS. Milanović et al. [16] showed the Lyapunov spectrum and the Kolmogorov-Sinai entropy in the OGS for $10 \leq N \leq 24$. However, since it is known that the chaotic behavior changes for $N \sim 30$ for the OGS [18], it is considerably important to extend the analysis to the system larger than $N \sim 30$. In this paper, we extend the number of sheets to $N = 256$ and follow the evolution numerically up to $T \sim 10^6 N t_c$, which is long enough for the thermal relaxation [12].

II. NUMERICAL SIMULATIONS

The OGS comprises N identical plane-parallel mass sheets, each of which has uniform mass density and infinite in, say, the y and z direction. They move only in the x direction under their mutual gravity. When two of the sheets intersect, they pass through each other. The Hamiltonian of the system has the form

$$H = \frac{m}{2}\sum_{i=1}^{N} v_i^2 + (2\pi G m^2)\sum_{i<j}|x_j - x_i|, \tag{3}$$

where m, v_i, and x_i are the mass (surface density), velocity, and position of the ith sheet, respectively. Since the gravitational field is uniform, the individual particles moves parabolically, until they intersect with the neighbors. Thus the evolution of the system can be followed by solving quadratic equations. This property helps us to

calculate long time evolution with a high accuracy. Since length and velocity (thus also energy) can be scaled in the system, the number of the sheets N is the only free parameter. The crossing time is defined by

$$t_c = (1/4\pi GM)(4E/M)^{1/2}, \tag{4}$$

where M and E is the total mass and total energy of the system. Detailed descriptions of the evolution of the OGS can be found in our previous papers [19,11,12].

In order to investigate dynamical aspects of the system, we calculated the Lyapunov spectrum. The basic numerical algorithm follows Shimada and Nagashima [20], and detailed description of the procedure for the OGS can be found in ref [19,16]. We made numerical integration for $8 \leq N \leq 128$ up to $10^8 t_c$, which is enough time for the system to relax, and up to $1.8 \times 10^7 t_c$ for $N = 256$ for reference.

III. RESULTS

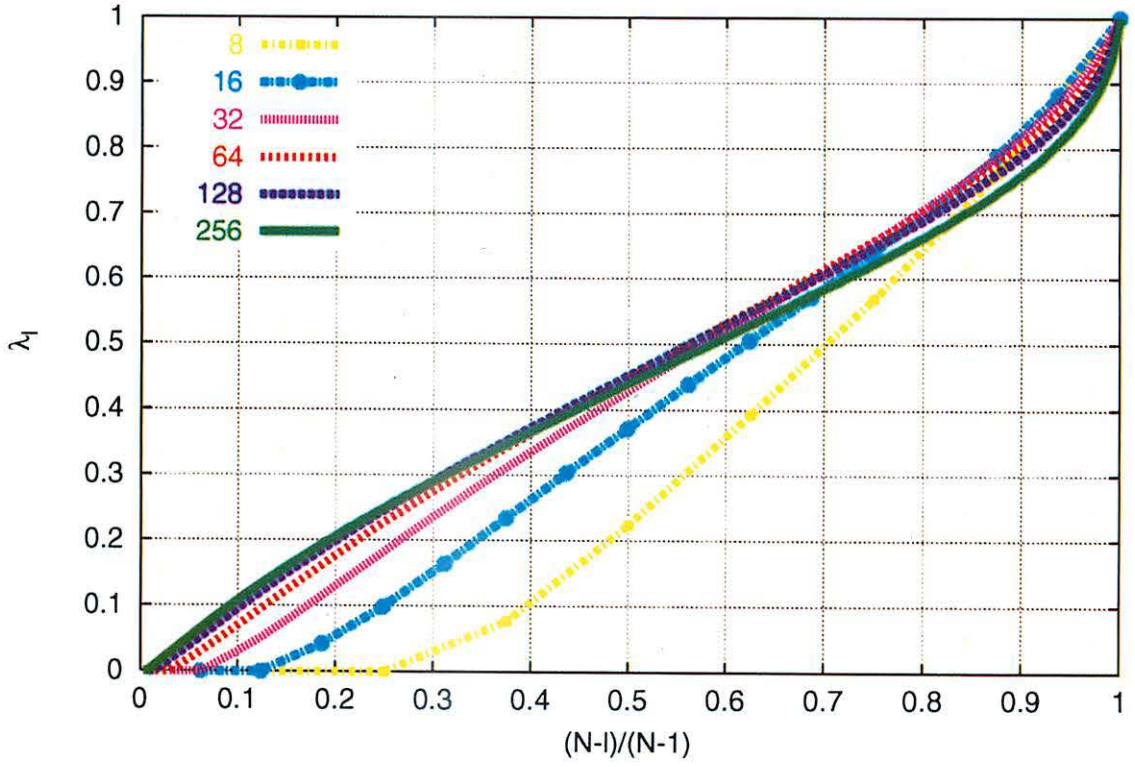


FIG. 3. Spectrum of the positive Lyapunov exponents for various N. The index of the Lyapunov exponents is scaled to 0 to 1.0. The vertical axis shows the Lyapunov exponents normalized by the value of the maximum Lyapunov exponent.

Figure 3 shows the spectrum of the Lyapunov exponents, $\{\lambda_i\}$, where their unit is $1/t_c$. This figure is the same diagram as Fig. 6 in Milanović et al [16], but the range of N is extended to $8 \leq N \leq 256$. In the horizontal axis, l is the index of the Lyapunov exponents, which is labeled in the order from the maximum to the minimum. Thus all the positive Lyapunov exponents ($l \leq N$) is scaled between 0 to 1 in the axis. The vertical axis shows the Lyapunov exponents normalized by the maximum Lyapunov exponents, λ_1. Milanović et al [16] stated that the shape of the spectrum approximately converges for large N.A closer look, however, shows bending of the spectrum, which is most clearly seen at $(N-l)/(N-1) \sim 0.9$. This bending seems increase with N for $N \geq 32$. A further investigation is needed to give a definite conclusion about the convergence of the shape of the spectrum.

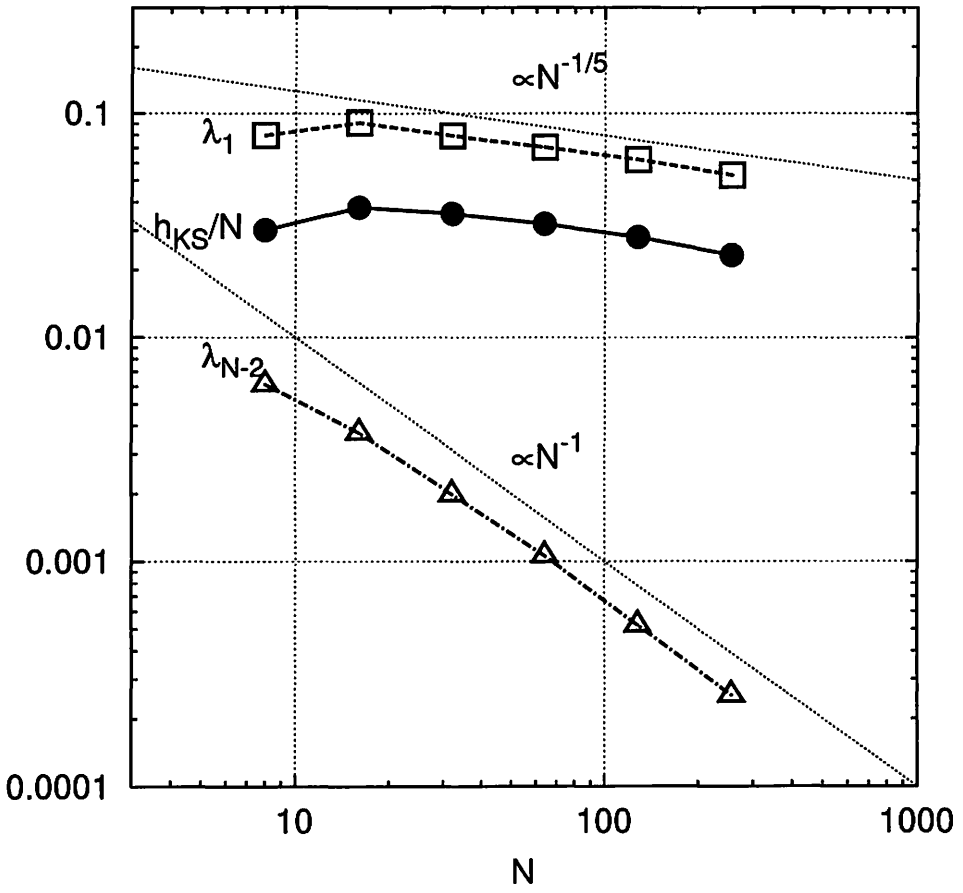


FIG. 4. Dependence of the KS entropy (solid line with the symbol •), the maximum Lyapunov exponent (long dashed curve with the symbol □), and the minimum positive Lyapunov exponent (dashed dotted curve with the symbol △).

Figure 4 shows N-dependence of the maximum (λ_1), the minimum positive Lyapunov exponent (λ_{N-2}), and the KS entropy h_{KS} per the number of freedom. λ_1 is already shown in Fig.13 in Tsuchiya et al. [19], and it is proportional to $N^{-1/5}$ for $N \geq 32$. Decreasing nature of the Lyapunov exponent may indicate that the OGS approaches closer to an integrable system for larger N. It is very interesting that the power of $-1/5$ is different from that observed in some other systems [22,23], which is $-1/3$. In those systems the power can be explained by means of a random matrix [24] approximation, where it is unclear in the OGS.

As expected from the spectrum the KS entropy divided by N is also proportional to $N^{-1/5}$. Therefore the conjecture by Benettin et al. [21] that h_{KS} increases linearly with N is not right. It is clear that the inverses of both the maximum Lyapunov exponents and the KS entropy do not give the time scale of any type of relaxation time.

The N-dependence of small positive Lyapunov exponents are quite different from larger ones. In Fig.4, the minimum positive Lyapunov exponent, λ_{N-2}, is shown by a dashed dotted line with the symbol $\triangle$. It decreases linearly for $N \geq 32$, and its time scale $1/\lambda_{N-2}$ is about the same as the microscopic relaxation time ($\sim Nt_c$).

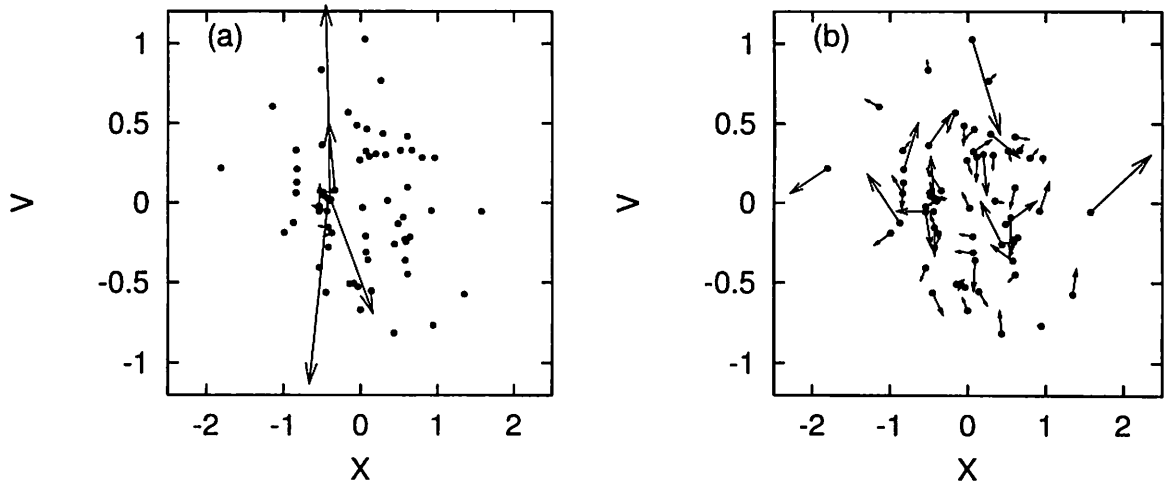


FIG. 5. Lyapunov vectors for $N = 64$. Filled circles indicate positions of N sheets and the arrows give direction of the Lyapunov vector: (a) the Lyapunov vector for λ_1, (b) that for λ_{N-2}.

The Lyapunov vectors also give useful information about instabilities associated with the Lyapunov exponents. The Lyapunov vector for λ_i is a unit vector in the phase space, and the instability grows with the ith fastest rate in that direction. Figure 5 shows projection of the Lyapunov vector for $N = 64$ on to the one-body phase space. Filled circles indicate positions of N sheets at a moment and the arrows give the direction of the Lyapunov vector at that time. The length of the vectors are scaled so as to see the direction clearly. Fig 5(a) is for the maximum Lyapunov exponent λ_1, and Fig.5(b) is for the minimum positive one, λ_{N-2}. The direction of the Lyapunov vectors change in time, but the characteristics of the instabilities are the same. For λ_1, the instability is carried only by a few particles, which are interacting in a very small region. The instability is thus not for global transformation. On the other hand, the instability with λ_{N-2} makes all particles mix in the phase space. This is the very effect of relaxation. These features are commonly seen for different N.

The results that the coincidence of the $1/\lambda_{N-2}$ and the microscopic relaxation time, and the direction of the Lyapunov vector, may be suggesting that the microscopic relaxation time is determined by the growing time of the weakest instability, which is determined by the minimum positive Lyapunov exponent; in other words, this time is necessary for the phase space orbit to mix in the phase space in the all directions of freedom. In our working model of the evolution of the OGS [11,12], the phase space is derived by some barriers which keep the phase orbit inside for a long time. The microscopic relaxation is considered to be a diffusion process in the barierred region [11,15], and in the time $\sim Nt_c$, restricted ergodicity is established within the barierred region. This time may correspond to the diffusion time in the slowest direction.

IV. CONCLUSIONS AND DISCUSSION

In the ergodic theory, the KS time represents the time scale of "Mixing" in the phase space. On the other hand, the relaxation of the one-body distribution is of the most interest in systems with large degrees of freedom. We have shown that the time scale of the relaxation of one-body distribution (both the microscopic and thermal relaxation) is certainly different from that of the KS time, and found that the growing time of the weakest Lyapunov instability is about the same as the microscopic relaxation time. In addition, taking into account the direction of the eigen vector of the weakest Lyapunov exponent, it is suggested that the microscopic relaxation is determined by the weakest Lyapunov instability.

The KS entropy is defined as a typical time for the system to increase "information". This definition does not depend on the number of degrees of freedom. In higher dimensions, however, even very small growth of instability can increase information quite rapidly. Therefore the KS time does not seem suitable to characterize the relaxation of the one-body distribution function.

The relaxation of the one-body distribution function implies ergodicity. To attain ergodicity, the phase space orbits should diffuse over all accessible phase space. For the microscopic relaxation, even though it is not true thermal relaxation, the system shows ergodicity which is restricted in a part of the phase space [11]. Therefore there is the slowest diffusion in a quasiequilibrium which corresponds to the microscopic relaxation. Since the microscopic relaxation time is universal for different quasiequilibria, a small Lyapunov exponent almost corresponds to the microscopic relaxation time in each restricted part of the phase space though the Lyapunov exponents are computed as average over the thermal relaxation time, which is much longer than the itinerancy among the quasiequilibrium. We found that not the thermal relaxation time but that of the microscopic relaxation is about the same as the minimum positive Lyapunov exponent. One may expect that there exists a smaller Lyapunov exponent which corresponds to the thermal relaxation time, but actually such a smaller Lyapunov exponent is absent.

A remaining problem is why the KS time and any of the Lyapunov times do not give the much long time scale of the thermal relaxation in the OGS. In our working model, the thermal relaxation is the successive transitions of the phase space orbit among the barierred regions, which corresponds to the quasiequilibria. Actual time of the thermal relaxation is the maximum time of transition among quasiequilibria. The fact that the Lyapunov exponents do not give the correct time of thermalization indicates that the transition is due to a different mechanism from local instabilities. There are some pieces of evidence that collective effects are responsible for keeping the system in a quasiequilibrium [12,25]. This may suggest that we need a new dynamical quantity which characterizes the slow diffusion.

Existing of the long-lived quasiequilibria is reported in various systems, such as one-dimensional systems with attractive pair potential $|x_i - x_j|^\nu$, where ν is a positive parameter [16], globally coupled spin models [26,22], and a two-dimensional system with long-range forces [27]. Therefore the slow relaxation seems universal property in systems with long-range forces. It is important to clarify the mechanism of the slow relaxation to construct a new statistical mechanics of the many-body systems with long-range forces.

REFERENCES

[1] E. Fermi, J. R. Pasta, and S. Ulam, *Collected Works of Enrico Fermi* (University of Chicago Press, Chicago, 1965), Vol. 2, p. 978.

[2] H. Hirooka and N. Saito, J. Phys. Soc. Jpn. **26**, 624 (1969).

[3] N. Saito, N. Ooyama, Y. Aizawa, and H. Hirooka, Prog. Theor. Phys. Suppl. **45**, 209 (1970).

[4] F. Hohl and D. T. Broaddus, Phys. Lett. A **25**, 713 (1967).

[5] F. Hohl and M. R. Feix, Astrophys. J. **147**, 1164 (1967).

[6] F. Hohl and J. W. Campbell, Astron. J. **73**, 611 (1968).

[7] M. Luwel, G. Severne, and P. J. Rousseeuw, Astrophys. Space Sci. **100**, 261 (1984).

[8] G. Severne, M. Luwel, and P. J. Rousseeuw, Astron. Astrophys. **138**, 365 (1984).

[9] H. L. Wright, B. N. Miller, and W. E. Stein, Astrophys. Space Sci. **84**, 421 (1982).

[10] B. N. Miller and C. J. Reidl, Jr., Astrophys. J. **348**, 203 (1990).

[11] T. Tsuchiya, N. Gouda, and T. Konishi, Phys. Rev. E **53**, 2210 (1996) (TGK).

[12] T. Tsuchiya, N. Gouda, and T. Konishi, Astrophys. Space Sci. **257**, 319 (1998).

[13] K. R. Yawn and B. N. Miller, Phys. Rev. E **56**, 2429 (1997).

[14] K. R. Yawn and B. N. Miller, Phys. Rev. Lett. **79**, 3561 (1997).

[15] B. N. Miller, Phys. Rev. E **53**, R4279 (1996).

[16] L. Milanović, H. A. Posch, and W. Thirring, Phys. Rev. E **57**, 2763 (1998).

[17] C. Dellago and H. A. Posch, Phys. Rev. E **55**, R9 (1997).

[18] C. J. Reidl, Jr. and B. N. Miller, Phys. Rev. E **51**, 884 (1995).

[19] T. Tsuchiya, T. Konishi, and N. Gouda, Phys. Rev. E **50**, 2607 (1994).

[20] I. Shimada and T. Nagashima, Prog. Theor. Phys. **61**, 1605 (1979).

[21] G. Benettin, C. Froeschle, and J. P. Scheidecker, Phys. Rev. A **19**, 2454 (1979).

[22] V. Latora, A. Rapisarda, and S. Ruffo, Phys. Rev. Lett. **80**, 692 (1998).

[23] A. Torcini and M. Antoni, Phys. Rev. E (1999) in press.

[24] G. Parisi and A. Vulpiani, J. Phys. A **19**, L425 (1986).

[25] J. L. Rouet and M. R. Feix, Phys. Rev. E **59**, 73 (1999).

[26] Y. Y. Yamaguchi, Progr. Theor. Phys. **95**, 717 (1996).

[27] M. Antoni and A. Torcini, Phys. Rev. E **57**, R6233 (1998).

Lyapunov Spectra and Structure of Phase Space

YAMAGUCHI Y. Yoshiyuki *
The general research organization of science and engineering,
Ritsumeikan University †

Abstract

To understand slow relaxation in Hamiltonian systems with many degrees of freedom, we introduce a new class of Moderately Chaotic systems and consider what structure of phase space is universal irrespective of details of systems by using Lyapunov spectra and time series of orbits. Some properties of moderately chaotic systems suggest that whiskered tori play an important role to the slow relaxation.

1 Introduction

Clusters, proteins, condensed matters, and galaxies are developing as Hamiltonian systems with many degrees of freedom, although they have different spatial scales. Hence dynamical properties of Hamiltonian systems with many degrees of freedom are an important topic.

We have two methods to approach the dynamical properties. One is to construct phenomenological model equations and to investigate properties of the model. This method gives details of the systems, but we must consider one by one. The other method is to research universal properties irrespective of details of systems , for instance, interactions. We cannot know details of the systems, but we can understand them uniformly. Hereafter, we adopt the latter method, and focus on structure of phase space and dynamical properties led from the structure.

Universal structures in Hamiltonian systems with N degrees of freedom are found in integrable, nearly integrable and fully developed chaotic systems. The features of structures of phase spaces are as follows.

- Integrable systems
  Phase space is foliated by N-dimensional tori and separatrices.
  Power spectra have some peaks.

- Nearly Integrable systems
  N-dimensional tori and chaotic sea coexist,
  and the former constructs self-similar hierarchical structure.
  Power spectra are power types as $1/f^{\nu}$ $(0 < \nu < 2)$ [Kar83, CS84].

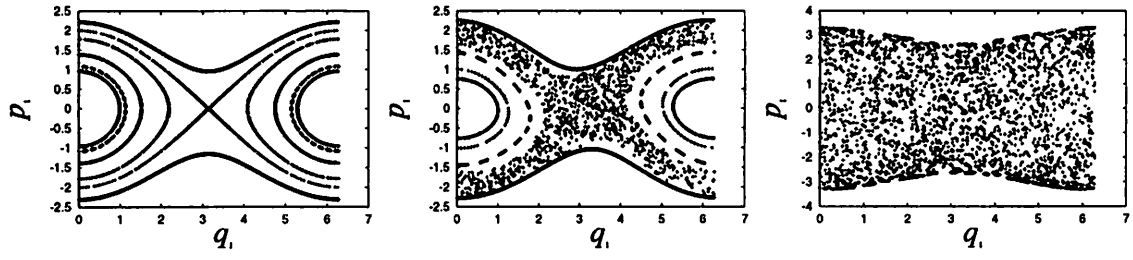
*e-mail: yyama@kuamp.kyoto-u.ac.jp
†Present Institute: Graduate School of Informatics, Kyoto University

- Fully Developed chaotic systems
  Only chaotic seas exist.
  Power spectra are white noise or Lorentzian $1/(a+f^2)$.

Some N-dimensional tori survive in nearly integrable systems, and the tori are called as Kolmogorov-Arnold-Moser (KAM) tori. For examples, we consider the following system

$$H = \frac{1}{2}(p_1^2 + p_2^2) + (1 - \cos q_1) + \frac{1}{2}q_2^2 + \epsilon(1 - \cos q_1)q_2. \tag{1}$$

This system is integrable when perturbation parameter ϵ is zero. Poincaré sections in the three classes are shown in Fig.1. An example of self-similar hierarchical structure in



(a) Integrable system$\epsilon = 0$ (b) Nearly Integrable system $\epsilon = 0.1$ (c) Fully Developed Chaotic system $\epsilon = 2$

Figure 1: Universal structures of phase spaces in Hamiltonian systems with 2 degrees of freedom. Figures show Poincaré sections (q_1, p_1). Tori are destroyed as perturbation parameter ϵ increases.

Hénon-Heiles system

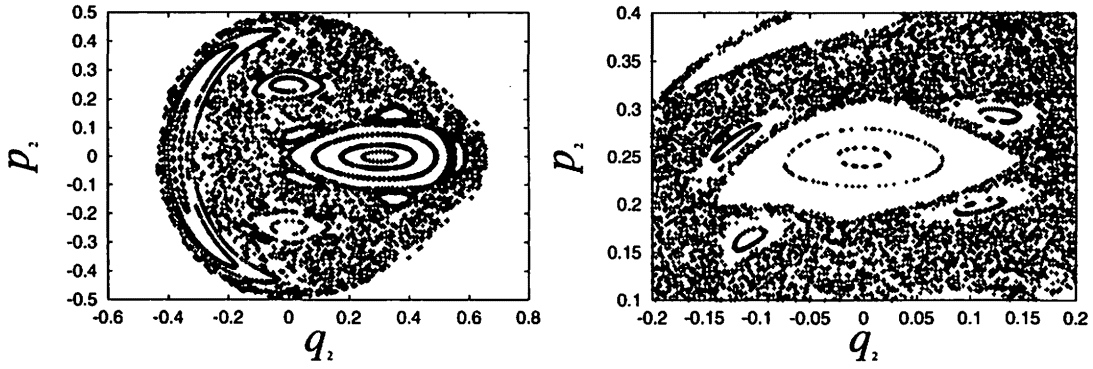
$$H = \frac{1}{2}(p_1^2 + p_2^2) + \frac{1}{2}(q_1^2 + q_2^2) + q_1^2 q_2 - \frac{1}{3}q_2^3 \tag{2}$$

is also shown in Fig.2.

The $1/f^\nu$ spectra are observed in various systems [Kar83, CS84, BHSO97, Yam97], and they are understood by using KAM tori in nearly integrable systems [Aiz84, AKH+89, Mei86, MO86]. However, it is believed that the region of ϵ rapidly becomes narrow where nearly integrable systems exist as N increases, and hence the understanding is not good for systems with many degrees of freedom. We therefore introduce a new class of moderately chaotic systems between nearly integrable systems and fully developed chaotic systems. The characteristic properties of this new class are N-dimensional tori are not observed although power spectra are power types $1/f^\nu$. The existence of this class is shown in Sec.sec:phase-transition by considering the critical point of phase transition. The purpose of this report is to understand universal structure of phase space and dynamical properties in moderately chaotic Hamiltonian systems with many degrees of freedom.

We use Lyapunov spectra $\{\lambda_i\}_{i=1,2,\ldots,2N}$ [LL92] as a tool of analyses. Lyapunov spectra indicate instability of orbits, and have information of the whole dimensional phase space. Here, Lyapunov spectra satisfy the following equation from symplectic properties:

$$\lambda_{2N-i+1} = \lambda_i, \tag{3}$$



(a) The whole space

(b) A Magnification of (a)

Figure 2: Poincaré sections (q_2, p_2) in Hénon-Heiles system. $E = 0.125$. (a) The whole space. (b) A magnification of (a). Some small tori exist around a big torus, and this structure hierarchically continues.

where

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{2N}. \tag{4}$$

We hence observe only the half of the spectra, $\{\lambda_i\}$ $(i = 1, 2, \ldots, N)$.

This report is constructed as follows. We clarify properties of moderately chaotic systems by showing that a system having phase transition is moderately chaotic at the critical point in Sec.2. In Sec.3 we show that Lyapunov spectra have an universality in moderately chaotic systems. We consider the structure of phase space which produces the universality in Sec.4. Section 5 is devoted to summary and discussions.

2 Second Order Phase Transition and Moderate Chaos

In this section we show that a system having second order phase transition is a moderately chaotic system at the critical point [Yam97]. The considered system is globally coupled XY spin system:

$$H = \frac{1}{2}\sum_{j=1}^{N} p_j^2 + \frac{1}{2N}\sum_{i,j}[1 - \cos(q_i - q_j)]\,. \tag{5}$$

Statistical mechanics states that the critical point is $E_c/N = 3/4$. We show the following two facts:

1. A power spectrum is one of power type
2. KAM tori are not observed

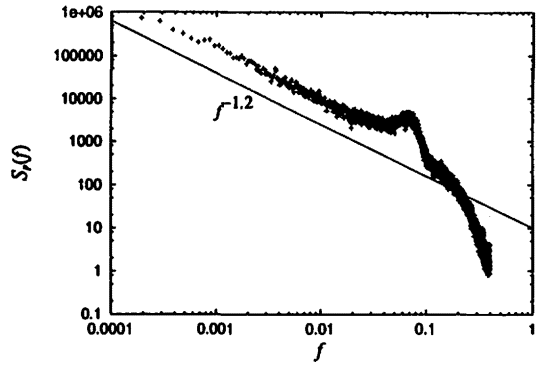


Figure 3: An average of power spectra of momenta. $N = 80, N_0 = 40, E = E_c$.

2.1 That a power spectrum is one of power type

We observe a power spectrum $S_p(f)$ which is an average of power spectra of momenta $S_{p_j}(f)$

$$S_p(f) = \frac{1}{N_0} \sum_{j=1}^{N_0} S_{p_j}(f), \tag{6}$$

The power spectrum $S_p(f)$ is shown in Fig.3, and it is one of power type with the exponent -1.2.

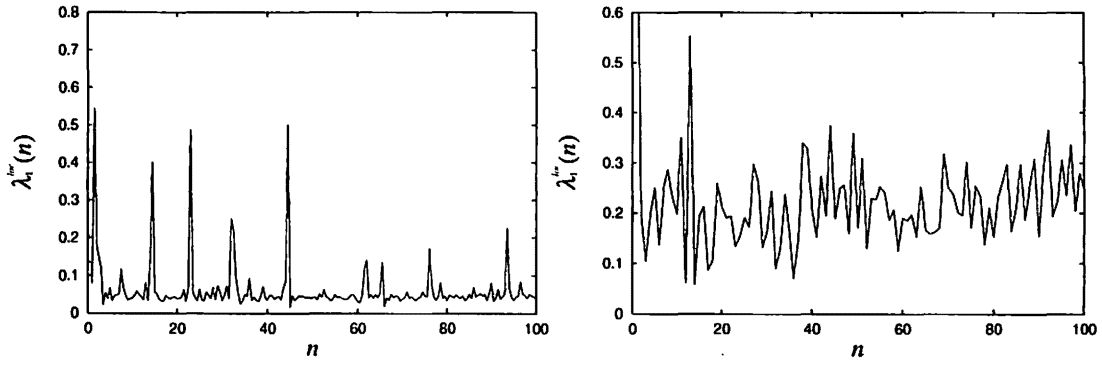
2.2 That KAM tori are not observed

Orbits are (quasi-)periodic on KAM tori, and hence orbits around KAM tori also behave (quasi-) periodically for a while, and they go to chaotic seas again. Values of local Lyapunov exponent is therefore intermittently suppressed when orbits are near KAM tori (see Fig.4(a)). Here, the definition of local Lyapunov exponent is

$$\begin{aligned} \lambda_1^{loc}(n) &= \frac{1}{\tau} \int_{n\tau}^{(n+1)\tau} \lambda_1(t) dt, \\ \lambda_1(t) &= \frac{d}{dt} \log |X(t)|, \end{aligned} \tag{7}$$

where $X(t)$ is a $2N$-dimensional tangent vector which governed by Jacobian of linearized equations of motion. On the other hand, in a moderately chaotic system, local Lyapunov exponent is uniformly apart from zero, and no KAM tori are observed.

Consequently, we show that the system (5) is a moderately chaotic system at the critical point from Secs.2.1 and 2.2.



(a) Nearly Integrable system (b) Moderately Chaotic system

Figure 4: Time series of local Lyapunov exponent in nearly integrable and moderately chaotic systems. (a) XY system with $E = 1$ defined in Sec.3. Local Lyapunov exponent is intermittently suppressed. (b) The system (5) at the critical point. Instability is uniform and no intermittency is found.

3 Universality of Lyapunov spectra

In this section we describe universality of Lyapunov spectra, which appear in moderately chaotic systems. We consider the following four systems to show the universality:

$$H = \frac{1}{2}\sum_{j=1}^{N} p_j^2 + U(q), \tag{8}$$

and

$$\begin{aligned} U_{XY} &= \sum_{<ij>} [1 - \cos(q_i - q_j)], \\ U_{DW} &= \frac{1}{2}\sum_{<ij>} (q_i - q_j)^2 + \sum_j \left(-\frac{1}{2}q_j^2 + \frac{1}{4}q_j^4\right), n \\ U_{SW} &= \frac{1}{2}\sum_{<ij>} (q_i - q_j)^2 + \sum_j \left(\frac{1}{2}q_j^2 + \frac{1}{4}q_j^4\right), \\ U_{LO} &= \frac{1}{2}\sum_{<ij>} (q_i - q_j)^2 + \sum_j \frac{q_j^2}{1 + q_j^2}. \end{aligned} \tag{9}$$

Here each particle is on a lattice point of simple cubic lattice, the sign $< ij >$ means to take sum between the nearest neighbourhood, and boundary condition is periodic. We confirmed that these systems with moderate energy are moderately chaotic systems by the same procedure described in the previous section. Lyapunov spectra in the energy region are shown in Fig.5. Now we focus on the forms of Lyapunov spectra, we permit to uniformly rescale the vertical axis:

$$\lambda_i \mapsto \gamma\lambda_i, \tag{10}$$

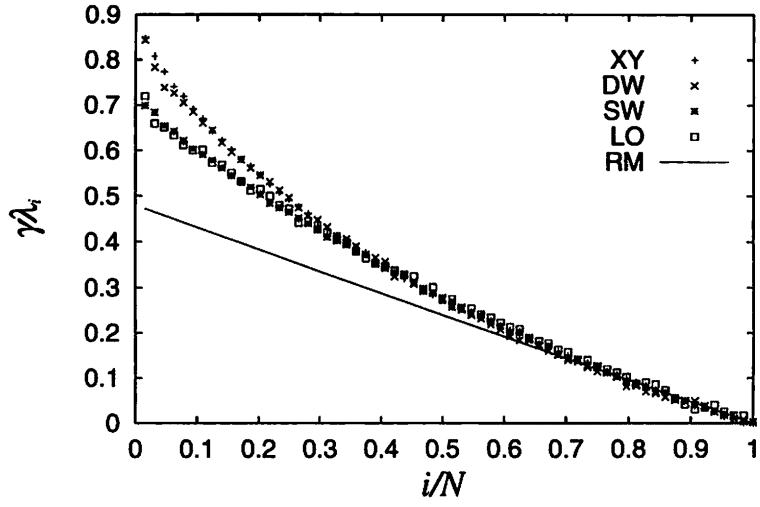


Figure 5: Universality of Lyapunov spectra in the four systems. $N = 64$. The horizontal axis is renormalized number of Lyapunov exponents i/N, and the vertical axis is scaled values of Lyapunov exponents $\gamma\lambda_i$. The four spectra are well agreement in the region $0.4 \leq i/N \leq 1$. It is known that Lyapunov spectra are straight in fully developed chaotic systems by using Random Matrices (RM), and the universal spectrum is not approximated by the straight line.

where suitable $\gamma > 0$ is picked up for each spectra.

We find the following two facts from Fig.5:

- Lyapunov spectra have universal form in the region of large i/N ($0.4 \leq i/N \leq 1$).
- The universal form is not approximated by the straight line, which is universal in fully developed chaotic systems [LPR86, EW88].

From these facts, the obtained universal form is characteristic for moderately chaotic systems, and we suppose that moderately chaotic systems have universal structure in subspace of phase space which corresponds to the region of large i/N [Yam98].

4 Structure of Phase Space

In the previous section universal spectrum in moderately chaotic systems is shown, and it is supposed that the universality reflects universal structure of phase space. In this section we therefore consider what is the universal structure by using time series of orbits in the configuration space. Let us consider coupled standard map

$$\begin{cases} p'_j &= p_j - \frac{K}{2\pi}\left[\cos 2\pi(x_{j-1} - x_j) + \cos 2\pi(x_j - x_{j+1})\right] \\ x'_j &= x_j + p'_j \end{cases}, \tag{11}$$

where boundary condition is periodic.

We show time series of x_3 and x_6 in Figs. 7(a) and (b) respectively, but these time series seem random. From interaction terms of the system (11), we guess that time series

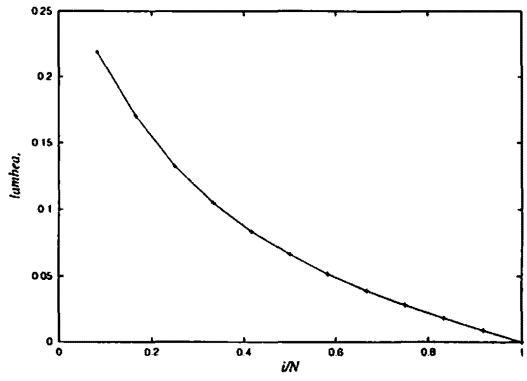


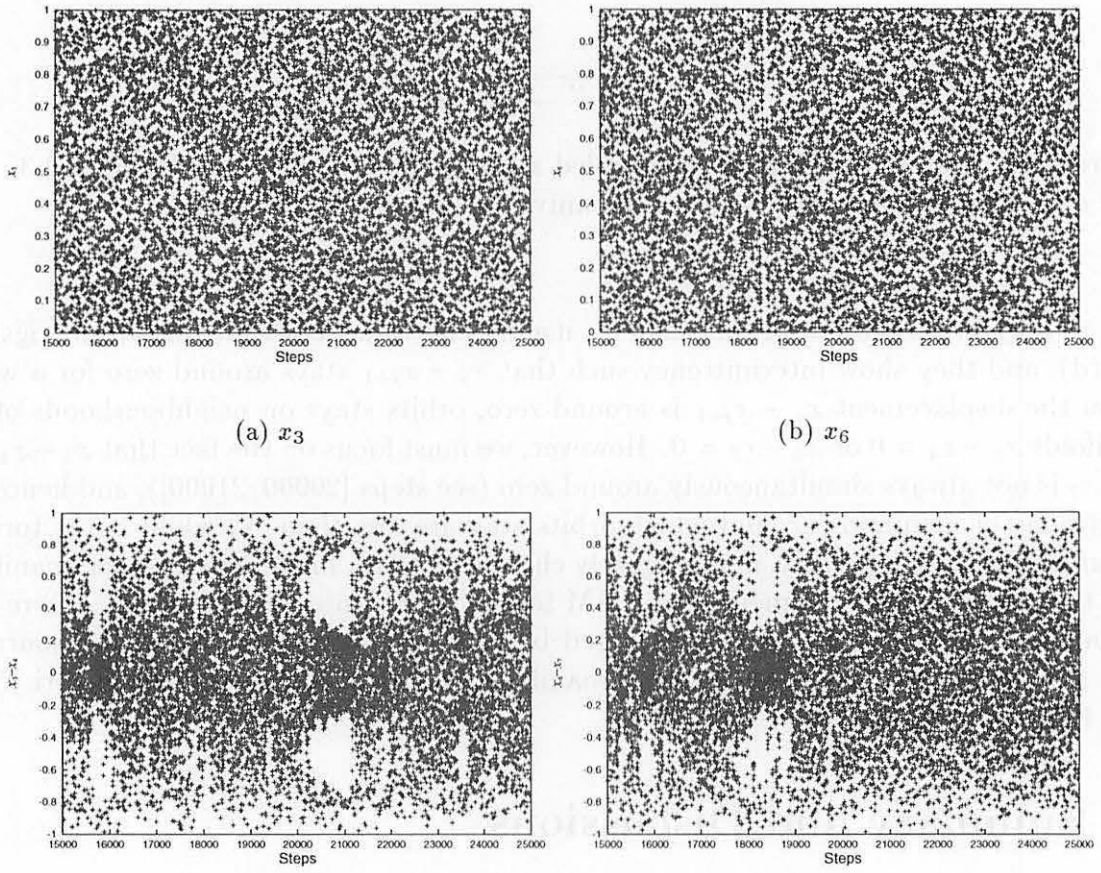
Figure 6: A Lyapunov spectrum in coupled standard map (11). $N = 12, K = 0.3$. The form of this spectrum is the same as the universal one observed in Fig.5.

of $x_j - x_{j\pm1}$ are more important than x_j itself. The time series are shown in Figs.7(c) and (d), and they show intermittency such that $x_j - x_{j+1}$ stays around zero for a while. When the displacement $x_j - x_{j+1}$ is around zero, orbits stays on neighbourhoods of the manifolds $x_3 - x_4 = 0$ or $x_6 - x_7 = 0$. However, we must focus on the fact that $x_3 - x_4$ and $x_6 - x_7$ is not always simultaneously around zero (see steps [20000, 21000]), and hence the dimensions of manifolds around which orbits stay are less than N, while KAM tori are N-dimension. Consequently, in moderately chaotic systems, higher dimensional manifolds play the same role as N-dimensional KAM tori in nearly integrable systems, where self-similar hierarchically structure constructed by KAM tori are supposed as the source of $1/f^{\nu}$ spectra. The higher dimensional manifolds are expected as whiskered tori which have hyperbolicity [YK98].

5 Summary and Discussions

In this report we introduced an important new class of moderately chaotic systems, and we consider universal structure of phase space in the class. First, we show that the system having second order phase transition is a moderately chaotic system at the critical point. Second, we clarify that Lyapunov spectra have an universal form in moderately chaotic systems in a region of large i/N. Finally, by observing motion in configuration space, we suggested that whiskered tori play an important role to realize slow relaxation of $1/f^{\nu}$ spectra. This suggestion is well agreement with the universality of Lyapunov spectra appears only in a region of large i/N, since it is supposed that phase space has similar properties as nearly integrable systems except for dimensions corresponding to the whiskers.

One of future works is to prove the above-mentioned suggestion qualitatively and quantitatively. Another one is to limit systems and degrees of freedom which have the universality since we consider systems with nearest neighbour interactions only and globally coupled systems are still not investigated. Finally, we must clarify the reason why Lyapunov spectra take the universal form. The understanding of the reason will give details of structure of phase space.



(a) x_3 (b) x_6

(c) $x_3 - x_4$ (d) $x_6 - x_7$

Figure 7: Time series in the system (6). (a) x_3 (b) x_6 (c) $x_3 - x_4$ (d) $x_6 - x_7$. The regions of the vertical axes are (a)(b) $[0, 1)$ and (c)(d) $[-1, 1)$. Time series in (c)(d) intermittently stay near the 0. Other particles show similar behaviours although we show only two samples for x_j and $x_j - x_{j+1}$.

References

[Aiz84] Y. Aizawa. "Symbolic dynamics approach to the two-dimensional chaos in area preserving maps: a fractal geometrical model". *Prog. Theor. Phys.*, 71:1419–21, 1984.

[AKH+89] Y. Aizawa, Y. Kikuchi, T. Harayama, K. Yamamoto, M. Ota, and K. Tanaka. *Prog. Theor. Phys. Suppl.*, 98:36, 1989.

[BHSO97] A. Baba, Y. Hirata, S. Saito, and I. Ohmine. "Fluctuation, relaxation and rearrangement dynamics of a model $(H_2O)_{20}$ cluster: non-statistical dynamical behavior". *J. Chem. Phys.*, 106:3329–37, 1997.

[CS84] B. V. Chirikov and D. L. Shepelyansky. "Correlation properties of dynamical chaos in Hamiltonian systems". *Physica D*, 13:395–400, 1984.

[EW88] J.-P. Eckmann and C. E. Wayne. "Liapunov spectra for infinite chains of nonlinear oscillators". *J. Stat. Phys.*, 50:853–78, 1988.

[Kar83] C. F. F. Karney. "Long-time correlations in the stochastic regime". *Physica D*, 8:360, 1983.

[LL92] A. J. Lichtenberg and M. A. Lieberman. *"Regular and Chaotic Dynamics. Second Edition"*. Springer, 1992.

[LPR86] R. Livi, A. Politi, and S. Ruffo. "Distribution of characteristic exponents in the thermodynamic limit". *J. Phys. A*, 19:2033–40, 1986.

[Mei86] J. D. Meiss. "Class renormalization: Islands around islands". *Phys. Rev. A*, 3:2375–83, 1986.

[MO86] J. D. Meiss and E. Ott. "Markov tree model of transport in area-preserving maps". *Physica D*, 20:387–402, 1986.

[Yam97] Y. Y. Yamaguchi. "Second order phase transition in a highly chaotic Hamiltonian system with many degrees of freedom". *International Journal of Bifurcation and Chaos*, 7(4), 1997.

[Yam98] Y. Y. Yamaguchi. "New universality of Lyapunov spectra in Hamiltonian systems". *J. Phys. A*, pages 195–207, 1998.

[YK98] Y. Y. Yamaguchi and T. Konishi. "A geometrical model for stagnant motions in Hamiltonian systems with many degrees of freedom". *Prog. Theor. Phys.*, pages 139–44, 1998.

Arnold Web in a Vibrationally Highly Excited Molecule

Mikito Toda

Department of Physics, Kyoto University

Kyoto 606-8502, Japan

When a molecule is put to be in a vibrationally excited state, a flow of vibrational energy is induced within the molecule, which leads to a breakup of a chemical bond or to a structural change in the molecular conformation. This process is called intramolecular vibrational redistribution of energy (IVR), and is one of the fundamental processes in chemical reactions[1,2]. On the other hand, this process is considered to be a typical example of Hamiltonian dynamics of many degrees of freedom. It is well known that generic Hamiltonian systems of many degrees of freedom exhibit chaos. Therefore, IVR is supposed to be closely related to chaotic motions of the molecules.

Arnold diffusion is among the characteristic effects of Hamiltonian chaos in systems of n degrees of freedom of ($n \geq 3$)[3]. In general, it takes place along those regions where nonlinear resonances occur and form a network called Arnold web. We think that IVR in highly excited molecules would offer an interesting example of Arnold web. We expect that comparison between the study of Arnold web and experiments on IVR would give dynamical understanding of how IVR proceeds.

Our example is the IVR in highly excited acetylene. Because of its relatively small size, detailed experiments on the IVR are possible for acetylene, and a large amount of experimental results have been already accumulated[2,4,5,6]. Therefore this is one of the best examples for examining our ideas.

In the electronic ground state, acetylene is a linear molecule. Using the harmonic approximation, vibrational energy levels in the electronic ground state can be labeled as $(v_1, v_2, v_3, v_4^{l_4}, v_5^{l_5})^l$ where v_1 is the quantum number of the symmetric CH stretch, v_2 the CC stretch, v_3 the antisymmetric CH stretch, v_4 the *trans*-bend with the vibrational angular momentum l_4, v_5 the *cis*-bend with the vibrational angular momentum l_5, and $l = l_4 + l_5$ is the total vibrational angular momentum. In the following, we put the total angular momentum $J = 0$ for simplicity. We also assume that the vibrational modes $v_i (i = 1, \cdots, 5)$ and the vibrational angular momenta $l_t (t = 4, 5)$ can be dealt with separately.

Then, the experimental results can be fitted to construct an integrable approximate Hamiltonian of a vibrationally excited acetylene. It is given by

$$H_0 = \sum_i \omega_i^0 v_i + \sum_{i \leq j} x_{ij}^0 v_i v_j + \sum_{r \leq s \leq t} y_{rst} v_r v_s v_t, \tag{1}$$

where i and j take the values from 1 to 5, and r, s and t from 4 to 5. The values of the constants for the linear and quadratic terms are taken from Table I of Ref.[4], and those for the third-order terms from Table II of Ref.[6].

The processes of IVR result from couplings among anharmonic modes of the Dunham expression. The effects of the couplings are most significant when resonances among the modes occur. A coupling term is a product of the annihilation and/or the creation operator of the modes v_i, and is denoted as $\mathbf{n} = (n_1, n_2, n_3, n_4, n_5)$ where $n_i (i = 1, \cdots, 5)$ are integers, and a positive integer means the product of the creation operator while a negative one the annihilation operator. Then, the location where the nonlinear resonance takes place for the coupling $\mathbf{n}$ is approximately given by

$$\mathbf{n} \cdot \frac{\partial H_0}{\partial \mathbf{v}} = 0 \tag{2}$$

where $\mathbf{v} = (v_1, v_2, v_3, v_4, v_5)$.

The Arnold webs of all the possible resonances up to 7th order are studied for those initial states $(0, v_2, 0, v_4, 0)$. Here the order of a resonance is defined as $\sum |n_i|$. One of the most important outcomes of this study is that the symmetries of the vibrational modes play a crucial role. Classically, there exists a lot of nonlinear resonances, implying that a fast IVR would proceed. However, the selection rules resulting from the symmetries of the modes prohibit most of these resonances. Only a small part of the nonlinear resonances are allowed to take place. This indicates that combined effects of symmetry and resonance would be indispensable in understanding the dynamical processes of IVR. Detailed analyses will be published elsewhere.

Reference

[1] See, e.g., the following volumes.

(a) P.M.Felker and A.H.Zewail, Adv.Chem.Phys.**70-I**(1988)265.

(b) L.R.Khundkar and A.H.Zewail, Annu.Rev.Phys.Chem.**41**(1990)15.

[2] P.Gaspard and I Burghhardt, ed., *Chemical Reactions and Their Control on the Femtosecond Time Scale*, Adv.Chem.Phys.**101**(1997).

[3] V.I.Arnold, Sov.Math.Dokl.,**5**(1964)581.

[4] D.M.Jonas et al., J.Chem.Phys.,**99**(1993)7350.

[5] M.A.Temsamani and M.Herman, J.Chem.Phys.,**102**(1995)6371.

[6] M.P.Jacobson et al., J.Chem.Phys.,**109**(1998)121.

Statistical behaviour in one-dimensional many-body system

Mitsusada M. Sano

Department of Fundamental Sciences
FIHS, Kyoto University
Sakyo-ku,Kyoto, 606-8501, Japan

e-mail: sano@phys.h.kyoto-u.ac.jp

Abstract

We investigate stationary heat conduction in Ding-Dong model coupled with heat reservoir. The Fourier law and Gallavotti-Cohen fluctuation theorem are checked numerically.

1 Introduction

In general, many body problems in one dimension are classified into two types of hamiltonian forms. (1) The first type is that all particles interact each other. $H = \sum_{i=1}^{N} \frac{p_i^2}{2m_i} + \sum_{i \neq j} V(|\mathbf{r}_i - \mathbf{r}_j|)$. Celestial mechanics and atom-molecular dynamics belong to this type. (2)The second type is that the particles are arranged on a lattice, $H = \sum_{i=1}^{N} \frac{p_i^2}{2m_i} + \sum_i V(|\mathbf{r}_i - \mathbf{r}_{i+1}|)$, or that the particle are arranged on a line with elastic constraint, $H = \sum_{i=1}^{N} \frac{p_i^2}{2m_i} + \sum_i V(\mathbf{r}_i)$. In the latter type, several models are known to exhibit chaotic behaviour. We introduce such two systems. (a) Balls in one-dimensional gravity[1, 2, 3]: $H = \sum_{i=1}^{N} \frac{p_i^2}{2m_i} + \sum_{i=1}^{N} m_i q_i$ with the constraint $0 \leq q_1 \leq q_2 \leq \cdots \leq q_N$. $q = 0$ is floor. The lowest particle can bounce elastically with floor. The other particles collide with the nearest neighbour particles elastically. For this model, some mathematicians give mathematical statements: If $m_1 = m_2 = \cdots = m_N$, the system is completely integrable. If $m_1 \geq m_2 \geq \cdots m_N$(but all mass are not equal), The Lyapunov exponent is positive and the system is hyperbolic. It is conjectured that the system is ergodic in the above condition. (b) Dawson model[5, 6, 7, 8, 9] or Ding-Dong model[4]: One-dimensional plasma sheet model(Dawson model) can be reduced to the system with the following hamiltonian (Ding-Dong model).

$$H = \sum_{i=0}^{N} \frac{p_i^2}{2} + \frac{q_i^2}{2}, \tag{1}$$

with the elastic constraint

$$q_{i+1} + 1 \geq q_i. \tag{2}$$

q_i represents the distance from the i-th lattice point. For our convenience, we set the lattice constant to $a = 1$. Thus Dawson model is equivalent to Ding-Dong model.

Particles move as harmonic oscillator around lattice point, but can collide with the nearest neighbour particles to exchange energy. Kitahara et al have investigated the properties of one-dimensional plasma [6, 7, 8, 9]. Prosen and Robnik have investigated heat conduction(Fourier law) by extensive numerical calculation[4]. Unfortunately, although many numerical works have been done, there is no mathematical investigation to this model.

Here we comment one point of numerical aspects on these models (a) and (b). Since particles are arranged in one-dimensional line and collide each other, these models have common numerical merit. Almost time of numerical time-evolution is spent for the calculation of the next collision time (For the formulation of time-evolution of Ding-Dong model, see the next section). In these case, the heap sort algorithm accelerates numerical calculation.

In this report, we investigate non-equilibrium (stationary) states in the Ding-Dong model coupled with heat reservoir, especially Fourier law and entropy production, motivated by recent development by chaos theory, especially Gallavotti and Cohen[11]. For Fermi-Pasta-Ulam problem(β-model) which has thermostat at the ends of a chain, the Gallavotti-Cohen fluctuation theorem was tested[10]. Furthermore, stimulating by the work by Gallavotti and Cohen, entropy production in stochastically driven non-linear lattice chain was rigorously proved[12]. Our aim in this report is to check the Gallavotti-Cohen fluctuation theorem in Ding-Dong model which is a different model from [10].

2 Ding-Dong model

In this section, we show set-up of time-evolution of the Ding-Dong model and explain the behaviour of the system, i.e. chaoticity, roughly.

First we remark the following point. We investigate the statistical behaviour of the Ding-Dong model. Thus we consider only two-body collisions, since three-body and higher-body collisions have measure zero. The "collision" means two-body collision hereafter.

The hamiltonian of Ding-Dong model is Eq.(1) with the constraint Eq.(2). We denote $t = t_k$ the time that k-th collision occurs. and $\{t_i\}_{i=-\infty}^{\infty}$ a set of collision time. Further we set the inter-collision time τ_k

$$\tau_k = t_k - t_{k-1}. \tag{3}$$

Now the collision is elastic. Then if i-th particle collides with $(i+1)$-th particle, after collision, their momentum are exchanged.

$$p_i^+ = p_{i+1}^-, \tag{4}$$
$$p_{i+1}^+ = p_i^-. \tag{5}$$

Between the collision and the next collision, the system obeys the following hamil-

tonian equation of motion.

$$\frac{d}{dt}q_i = p_i, \qquad \frac{d}{dt}p_i = -q_i, \qquad , \tag{6}$$

The map from the collision to the next collision can be constructed as

$$\Phi = \Phi_{\text{collision}} \circ \Phi^{\tau(\mathbf{x})}_{\text{oscillator}}, \tag{7}$$

where $\mathbf{x} = (q_1, q_2, \cdots, q_N, p_1, p_2, \cdots, p_N)$. $\tau(\mathbf{x})$ is the time that the next collision occurs started from the point $\mathbf{x}$. $\tau(\mathbf{x})$ is called the ceiling function. $\Phi_{\text{collision}}$ for the collision between i-th and $i+1$-th particles is given as

$$\begin{pmatrix} q_1^+ \\ q_2^+ \\ \vdots \\ q_N^+ \\ p_1^+ \\ p_2^+ \\ \vdots \\ p_i^+ \\ p_{i+1}^+ \\ \vdots \\ p_N^+ \end{pmatrix} = \begin{pmatrix} 1 & 0 & & 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 1 & & 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & & \ddots & & \vdots & & \ddots & 0 & 0 & \cdots & \vdots \\ 0 & 0 & & 1 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & 1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & & \ddots & \vdots & \vdots & & & & & & \vdots \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 1 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 1 & 0 & \cdots & 0 \\ \vdots & & \ddots & \vdots & \vdots & & & & & & \vdots \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 1 \end{pmatrix} \begin{pmatrix} q_1^- \\ q_2^- \\ \vdots \\ q_N^- \\ p_1^- \\ p_2^- \\ \vdots \\ p_i^- \\ p_{i+1}^- \\ \vdots \\ p_N^- \end{pmatrix}. \tag{8}$$

$\Phi^{\tau(\mathbf{x})}_{\text{oscillator}}$ is given as

$$\begin{pmatrix} q_1(t_i + \tau(\mathbf{x}_i)) \\ q_2(t_i + \tau(\mathbf{x}_i)) \\ \vdots \\ q_N(t_i + \tau(\mathbf{x}_i)) \\ p_1(t_i + \tau(\mathbf{x}_i)) \\ p_2(t_i + \tau(\mathbf{x}_i)) \\ \vdots \\ p_N(t_i + \tau(\mathbf{x}_i)) \end{pmatrix} = \begin{pmatrix} p_1(t_i)\sin(\tau(\mathbf{x}_i)) + q_1(t_i)\cos(\tau(\mathbf{x}_i)) \\ p_2(t_i)\sin(\tau(\mathbf{x}_i)) + q_2(t_i)\cos(\tau(\mathbf{x}_i)) \\ \vdots \\ p_N(t_i)\sin(\tau(\mathbf{x}_i)) + q_N(t_i)\cos(\tau(\mathbf{x}_i)) \\ p_1(t_i)\cos(\tau(\mathbf{x}_i)) - q_1(t_i)\sin(\tau(\mathbf{x}_i)) \\ p_2(t_i)\cos(\tau(\mathbf{x}_i)) - q_2(t_i)\sin(\tau(\mathbf{x}_i)) \\ \vdots \\ p_N(t_i)\cos(\tau(\mathbf{x}_i)) - q_N(t_i)\sin(\tau(\mathbf{x}_i)) \end{pmatrix}, \tag{9}$$

where the i-th collision point is $\mathbf{x}_i$.

With the condition that the center of mass is zero, if we carry out some canonical transformation.

$$\mathbf{x} = (\mathbf{q}, \mathbf{p}) \mapsto \mathbf{X} = (\mathbf{Q}, \mathbf{P}), \tag{10}$$

where $\mathbf{Q} = (Q_1, \cdots, Q_{N-1}), \mathbf{P} = (P_1, \cdots, P_{N-1})$, then after canonical transformation and diagonalization, we obtain the hamiltonian finally.

$$H = \frac{1}{2}\sum_{i=1}^{N-1}(\lambda_i Q_i^2 + \mu_i P_i^2) = E. \tag{11}$$

The elastic constraint Eq.(2) now becomes

$$\mathcal{F}^{(i)}(\mathbf{Q}) \equiv \sum_{k=1}^{N-1} (U_{ik} - U_{i+1\,k}) Q_k \leq 1, \tag{12}$$

where $i = 1, 2, \cdots, N-1$ and U_{ik} is the matrix element for diagonalizatoin. The hamiltonian is for $(N-1)$-dimensional anisotropic harmonic oscillator with the hard walls satisfying Eq.(12). Hence if the total energy is sufficiently low, the system is integrable. Because the trajectories go around the stable equilibrium point. There is a critical value of the total energy for chaos. But when the total energy is high enough, the system behaves like an integrable system. In this case, the particle moves ballistically and collide with the walls. Therefore, in the intermediate total energy range, the system exhibits chaos. From numerical calculation, in this range, the system seems to be non-hyperbolic. But we expect that chaotic sea dominates the phase space in this energy range of the system.

3 Non-equilibrium stationary states, thermal conduction, and entropy production

In this section, we investigate the non-equilibrium stationary states of the Ding-Dong model, especially entropy production.

3.1 Phenomenology and thermodynamics

We know the following two important facts from thermodynamics. The system which is coupled to the heat reservoir at the boundary obeys the Fourier law.

$$\mathbf{j} = \kappa \nabla T, \tag{13}$$

where $\mathbf{j}$ is heat current vector and κ is the heat conductivity. From the thermodynamics, the entropy production $\sigma(\mathbf{x})$ is defined as

$$\sigma(\mathbf{x}) = \mathbf{j} \cdot \nabla \left(\frac{1}{T} \right). \tag{14}$$

Thus the total entropy production Σ of the domain with linear temperature gradient for one-dimensional system is

$$\Sigma = \int_L^R dx\, j \frac{d}{dx} \left(\frac{1}{T(x)} \right) = j \left(\frac{1}{T_R} - \frac{1}{T_L} \right), \tag{15}$$

where T_L and T_R are the temperature of left and right heat reservoir, respectively.

3.2 Gallavotti-Cohen fluctuation theorem

We here briefly summarize the Gallavotti-Cohen fluctuation theorem[11]. Original derivation is for two-dimensional shearing flow with thermostat.

$$\frac{d\mathbf{q}_j}{dt} = \mathbf{p}_j/m + \mathbf{i}\gamma y_j, \tag{16}$$

$$\frac{d\mathbf{p}_j}{dt} = \mathbf{F}_j - \mathbf{i}\gamma y_j p_{y_j} - \alpha \mathbf{p}_j. \tag{17}$$

$j = 1, \cdots, N$. α controls the property of thermostat. The entropy production is defined as the contraction rate of the phase space volume. Dynamical characterization of entropy production starts from the idea of Sinai-Ruelle-Bowen, i.e., the construction of SRB measure $\overline{\mu}$.

$$\int_C \overline{\mu}_{\tau,T}(d\mathbf{x})G(\mathbf{x}) = \frac{\sum_j \overline{\Lambda}_{u,\tau}^{-1}(\mathbf{x}_j)G(\mathbf{x}_j)}{\sum_j \overline{\Lambda}_{u,\tau}^{-1}(\mathbf{x}_j)}, \tag{18}$$

where Λ is defined from the Jacobian of time evolution operator and $G(\mathbf{x})$ is an observable. Here consider the finite time average of the entropy production at $\mathbf{x}$.

$$\sigma_\tau(\mathbf{x}) = \frac{1}{\tau}\sum_{j=-\tau/2}^{\tau/2-1} \sigma(S^j\mathbf{x}) = \langle\sigma\rangle\, a_\tau(\mathbf{x}), \tag{19}$$

where S is the Poincaré map. Fluctuation of the entropy production from the mean value is characterized by the $a_\tau(\mathbf{x})$. Consider the probability measure of a_τ. We denote $\pi_\tau(p)dp = P(a_\tau \in (p, p+dp))$ that. The Gallavotti-Cohen fluctuation theorem is the following relation.

$$\frac{\pi_\tau(p)}{\pi_\tau(-p)} = e^{\tau\langle\sigma\rangle p} \tag{20}$$

Using the construction of SRB measure, this relation can be rewritten in terms of dynamical quantities (i.e., Jacobian)[11]. For the detailed derivation, see [11].

For non-linear lattice system with thermostat, we define the following mean heat current[10]. [1].

$$\langle J\rangle_\tau = \frac{1}{\tau}\int_t^{t+\tau} J(t')dt'. \tag{21}$$

We define the distribution of $P_\tau(z)$ of the following quantity z.

$$z = \frac{\langle J\rangle_\tau}{\langle J\rangle_\infty}. \tag{22}$$

In this case, the Gallavotti-Cohen fluctuation theorem becomes

$$\ln\frac{P_\tau(z)}{P_\tau(-z)} = \tau z\Sigma = \tau z j\left(\frac{1}{T_R} - \frac{1}{T_L}\right). \tag{23}$$

In the next subsection, we consider the system does not possess thermostat, but has stochastic heat reservoir. Thus this relation Eq.(23) is not so trivial for our system. the numerical test of Eq.(23).

3.3 Numerical result

Let us define heat reservoir. We set the walls at the ends of the system. The other side of the wall is heat reservoir. The momentum of particles in heat reservoir is distributed due to distribution

$$P(p)dp = \frac{|p|}{T}\exp\left(-\frac{p^2}{2T}\right)dp. \tag{24}$$

[1] In the first stage of this investigation(Jan.1999), I did not know this work. I thank K.Saito for informing me this work.

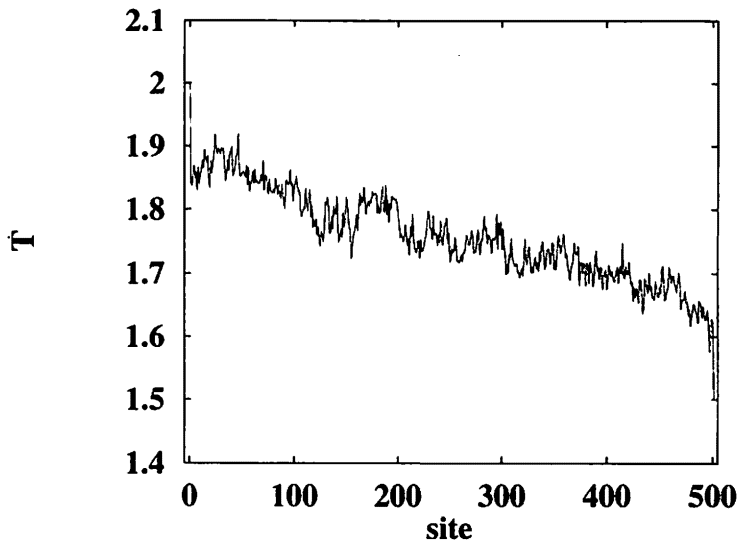


Figure 1: *Temperature distribution* : $N = 500, T_L = 2.0, T_R = 1.5$. *The plot is obtained starting some initial condition, and after* 5×10^6 *collisions.*

For the left(right) reservoir, we take plus(minus) sign of momentum. This setting of heat reservoir is very important[13]. We set temperature of the heat reservoir at the ends T_L(left) and T_R(right), respectively. When the particle at the ends ($i = 1, N$) collides with the wall, the particle instantaneously collides with the particle inside the heat reservoir. In other words, the particle($i = 1, N$) and particle inside the heat reservoir exchanges the momentum. Local temperature T_k is defined as

$$T_k = \left\langle p_k^2 \right\rangle . \tag{25}$$

$\langle \cdots \rangle$ means the time average or phase space average. For numerical calculation, we used the time average.

First we check Fourier law. Fig.1 is temperature profile($N = 500, T_L = 2.0, T_R = 1.5$). There is an edge effect. The temperature of the end in the left and right is not T_L or T_R, respectively. Relaxation to the non-equilibrium stationary state is very slow. The temperature profile is still zig-zag shape, not smooth, but we can see the linear dependence of the Fourier law.

In Fig.2, the distribution of inter-collision time is depicted($N = 250, T_L = 2.0, T_R = 1.5$). After 10^5 collisions and 3×10^6 collisions, 10^5 collisions are sampled. The mean inter-collision time is order of 10^{-2} for this case. The obtained distribution is well-fitted to exponential distribution. This manifests that the collisions occur randomly.

In Fig.3, the velocity distribution is shown. As expected, the distribution is well-fitted to the Gaussian distribution (Maxwell distribution). But tending to the stationary state, gradually the tail of distribution becomes asymmetric. This asymmetry suggests the energy transfer.

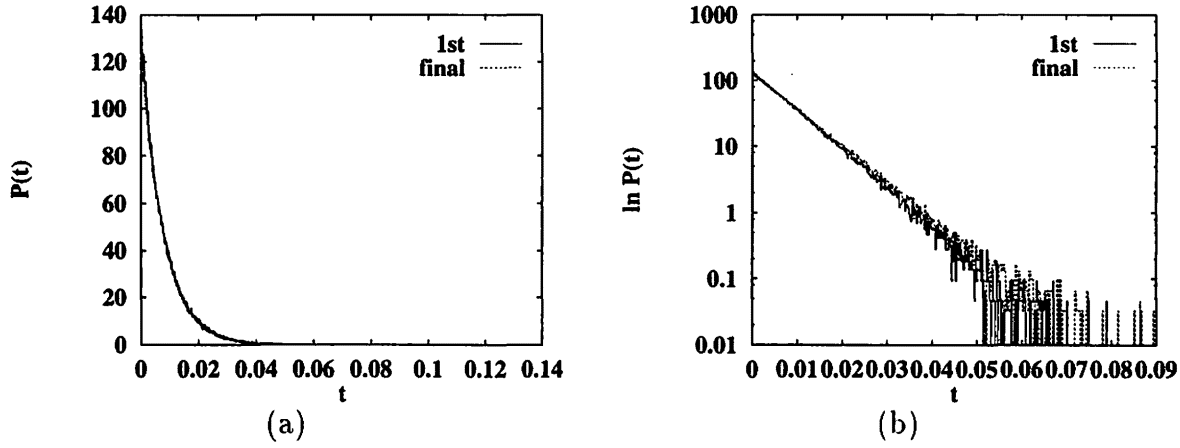


Figure 2: *Inter-collision time distribution* : $N = 250, T_L = 2.0, T_R = 1.5$. *(a)The plot is obtained after* 10^5 *collisions and* 3×10^6 *collisions. (b)log scale of (a).*

Time variation of heat current distribution is depicted in Fig.4. Remark that the distribution is not Gaussian. After long run, the tail of the distribution becomes asymmetric, which seems to be an evidence of stationary energy transfer.

In Fig.5(a), spatial distribution of time-averaged heat current $\overline{j_k} = 1/t \int_0^t j_k(t')dt'$ is depicted which was obtained after 3×10^6 collisions starting from some initial condition. Although there is still fluctuation, the heat current is almost constant over all sites. In Fig.5(b) the relaxation of the mean heat current $J = 1/N \sum_{k=1}^{N} \overline{j_k}$ is depicted corresponding to (a). The relaxation is very slow.

The entropy production can be checked by using Eq.(23). In Fig.6. we show the result($N = 100, T_L = 2.0, T_R = 1.5$). In Fig.6(a), the time evolution of the mean heat current distribution is depicted. Fig.6(a) is obtained from time-evolution of 50000 initial configurations. It seems that the distribution tends to the limiting form obeying the central limit theorem. In Fig.6(b), we check the Gallavotti-Cohen fluctuation theorem. We can see the linearity clearly which suggests that the Gallavotti-Cohen fluctuation theorem holds and the entropy production is non-zero. Fig.6 is the same form as the result of [10] for FPU β-chain with thermostat.

4 Summary

We have investigated non-equilibrium properties of the Ding-Dong model coupled with heat reservoir motivated by recent work of [11] and [10]. We have checked the Fourier law and entropy production in this model. Numerical calculation has shown that the Gallavotti-Cohen fluctuation theorem holds and the entropy production is surely non-zero. It can be stressed that original Gallavotti-Cohen theorem for thermostated dynamical systems extends to systems coupled with heat reservoir. In this sense, the Gallavotti-Cohen theorem is quite general. A stochastic extension of

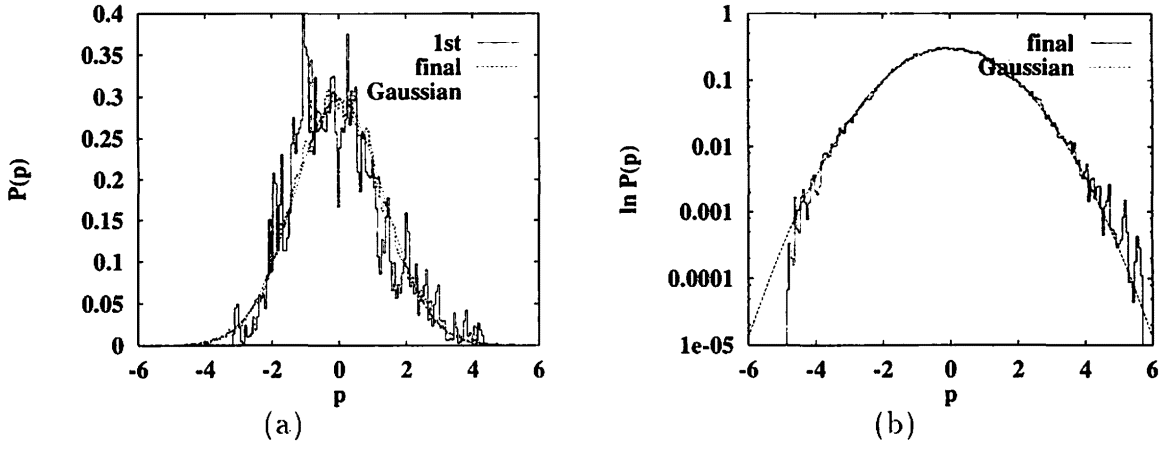


Figure 3: *Velocity distribution* : $N = 250, T_L = 2.0, T_R = 1.5$. *The plot is obtained starting from some initial condition, up to* 10^5 *collisions and* 3×10^6 *collisions. (a) initial(*10^5 *collisions) and final distribution(*3×10^6 *collisions). (b) log scale. The distribution tends to the Gaussian distribution. But the tail of the distribution is asymmetric.*

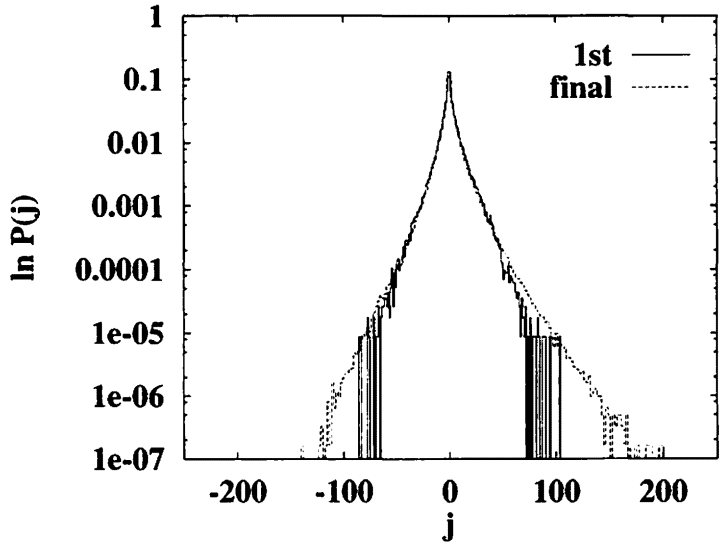


Figure 4: *Heat current distribution* : $N = 250, T_L = 2.0, T_R = 1.5$. *Starting some initial condition,* 10^5 *collisions and* 3×10^6 *collisions After sufficient time-evolution, the tail of the distribution becomes asymmetric.*

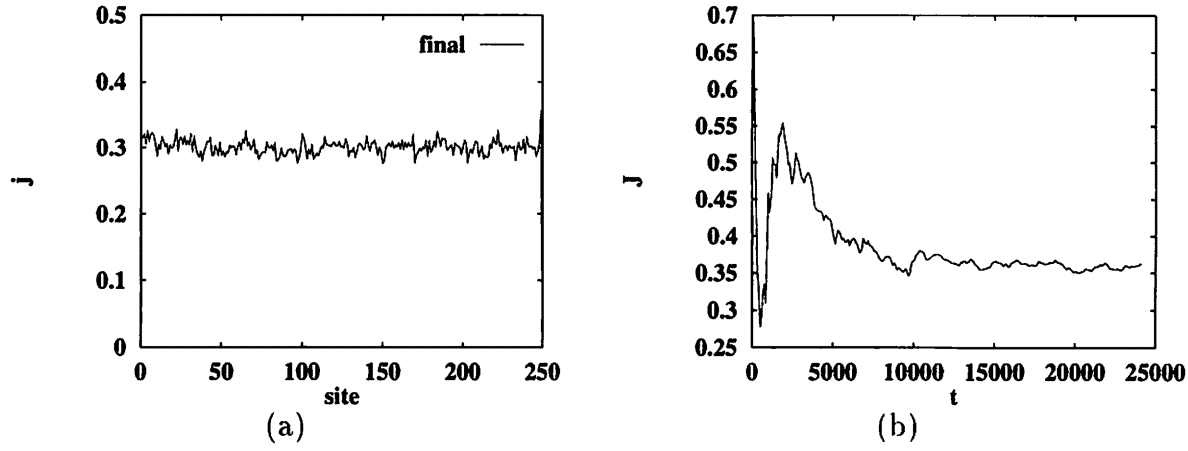


Figure 5: *Spatial distribution of heat current and relaxation of mean heat current* : $N = 250, T_L = 2.0, T_R = 1.5$. *(a) Spatial distribution of heat current which occurred up to* 3×10^6 *times collision starting some initial configuration. (b) Relaxation of mean heat current.*

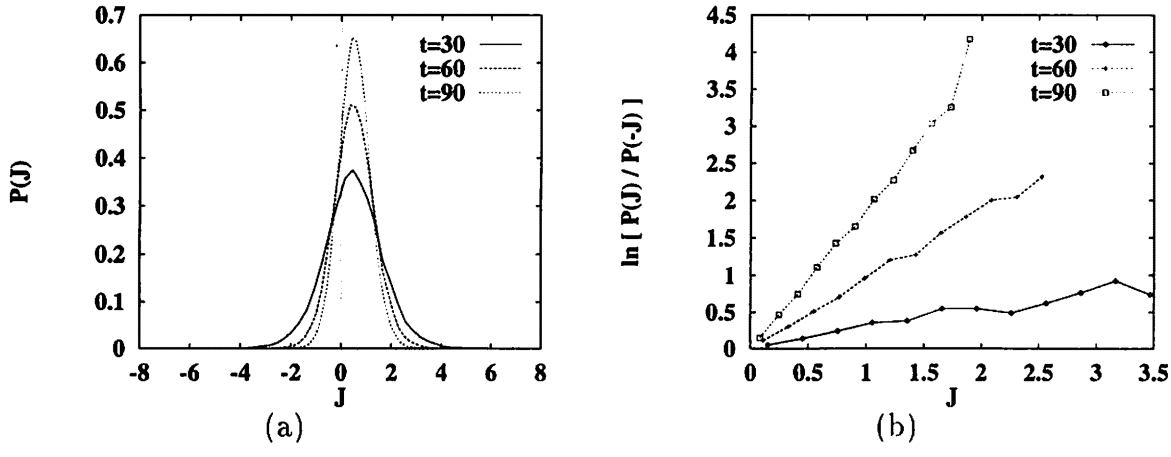


Figure 6: *Check of the Gallavotti-Cohen fluctuation theorem: (a) heat current distribution* $P_\tau(J)$, $\tau = 30, 60, 90$. *(b)* Jvs. $\ln(P_\tau(J)/P_\tau(-J))$, $\tau = 30, 60, 90$. $N = 100, T_L = 2.0, T_R = 1.5$. *The number of initial conditions is* 50000.

the Gallavotti-Cohen theorem can be expected.

References

[1] M.P.Wojtkowski *Comm.Math.Phys.* **126** (1990) 507; **127** (1990) 425; *preprint*, "Hamiltonian systems with linear potential and elastic constraints".

[2] N.I.Chernov *Physica* **D53** (1991) 233.

[3] N.Simányi *preprint*, "The characteristic exponents of the falling ball model".

[4] T.Prosen and M.Robnik *J.Phys.* **A25** (1992) 3449.

[5] J.M.Dawson *Phys.Fluids*, **5** (1962) 445; *Meth.Comp.Phys.*, **9** (1970) 1.

[6] K.Kitahara, N.Ishibashi and K.Tanno *Some problems on the Theory of Dynamical Systems in Applied Sciences*, ed. H.Kawakami, World Scientific, 1991, p.65.

[7] Y.Mitsumori *Bachelor Thesis*. Tokyo Institute of Technology, 1993.

[8] A.Kubota *Master Thesis*, Tokyo Institute of Technology, 1993.

[9] T.Yanagisawa *Master Thesis*, Tokyo Institute of Technology, 1995.

[10] S.Lepri, R.Livi, A.Politi *preprint*, "Energy transport in anharmonic lattices close and far from equilibrium", cond-mat/9709156.

[11] G.Gallavotti, E.G.D.Cohen *Phys.Rev.Lett.*, **74** (1995) 2694; *J.Stat.Phys.*, **80** (1995) 931; *Doc.Math.J.DMV Extra Volume ICM I* (1998) 65.

[12] J.-P. Eckmann, C.-A. Pillet, L.Rey-Bellet *preprint*, "Entropy production in non-linear, thermally driven hamiltonian systems", chao-dyn/9811001.

[13] R.Techver, F.Toigo, J.Koplik, J.R.Banavar *Phys.Rev.E*, **57** (1998) R17.

An Example of Nonlinear Phenomena in High Energy Accelerators

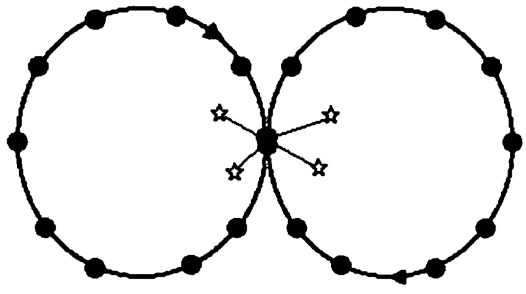
Kohji HIRATA[1]
So-Ken-Dai, The Graduate University for Advanced Studies
and
KEK, High Energy Accelerator Research Organization

abstract

The modern high energy accelerators are introduced which show many nonlinear multiparticle phenomena. One example, the beam-beam interaction, is discussed more in detail. A simple model is presented for a so-called strong-strong scheme. The model reproduces a genuine nonlinear and coherent phenomenon for the accelerator beams.

1 Introduction

Here is a typical layout of high energy accelerators. In a ring, many (typically 10^{11}) electrons form a cluster (called bunch) and running in the same direction. They meet a positron bunch at a point (Interaction Point, IP) producing high energy collisions. A detector sitting at the IP observes the products of the collision.



The most important parameters are 1) the collision energy and 2) the luminosity. The collision energy is

$$E_{col} = E_+ + E_-,$$

the sum of the energies of two colliding particles in the rest frame of their center of gravity. The luminosity L is the measure of the efficiency of the accelerator to produce high energy events, defined by

$$\text{event rate} = \sigma \times L.$$

Here σ is a cross section of the event. When two colliding beams are very short and are Gaussian in transverse distribution, L is given by

$$L = f\frac{N_+N_-}{4\pi\sigma_x\sigma_y}.$$

Here, f:is the collision frequency, $N_\pm$ is the number of particles in $e^\pm$ bunches, $\sigma_{x,y}$ is the transverse rms beam size of the bunches.

[1]hirata@soken.ac.jp
http://www-acc-theory.kek.jp/members/HIRATA.html

When two bunches collide, only a few particles make the real collision and most of the particles pass through the opposite bunch. These particles feel a macroscopic electromagnetic field produced by the opposite bunch. Thus, each particle in both bunches feel a nonlinear force. This is the beam-beam interaction and is one of the most important effects which limit the luminosity.

2 Beam Dynamics in Accelerators

Let us simplify the whole dynamics to the 1D problem. Between the collision at the IP to the next collision after one turn, an electron (positron) obeys the following mapping[1].

$$\begin{pmatrix} X \\ P \end{pmatrix} \to \lambda \begin{pmatrix} \cos\mu & \sin\mu \\ -\sin\mu & \cos\mu \end{pmatrix} \begin{pmatrix} X \\ P \end{pmatrix} + \sqrt{1-\lambda^2} \begin{pmatrix} \hat{r}_1 \\ \hat{r}_2 \end{pmatrix}. \tag{1}$$

Here, X and P are canonical variables for the transverse coordinate at the IP. They are properly normalized. $\hat{r}$'s are the standard Gaussian random variables and λ is a constant slightly smaller than unity. They represent the effects of the synchrotron radiation[3]. (The case $\lambda = 1$ corresponds to proton and hadron mchines where the synchrotron radiation can be ignored).

If the beam-beam interaction is absent at the IP, each particle repeats Eq.(1) many times. Eventually , they will be distributed in the Gaussian form.

$$\psi(X,P) = \frac{1}{2\pi} \exp - \left[\frac{X^2}{2} + \frac{P^2}{2} \right]. \tag{2}$$

3 Beam-Beam Interaction

With the beam-beam interaction, each particle feels the macroscopic electromagnetic field produced by the opposite bunch. Here, we will use a model consistent with the 1D approximation. If the space is 1D, the elctric or magnetic force field produced by an electron sitting at $X = X_0$ is represented by

$$F(X) = \text{constant} \times H(X - X_0), \tag{3}$$

where H is the step function

$$H(X) = \begin{cases} 1 & \text{if } X > 0 \\ -1 & \text{if } X < 0 \end{cases}.$$

There are two approaches to the beam-beam interaction:

- The weak-strong approach: it is an approximation that one of the beam consists only of a single electron and the positron bunch is not affected by the beam-beam force. This approximation is quite unnatural but because of its simplicity it is used frequently.
- the strong-strong approach: it is more realistic but is difficult to traet even numerically.

3.1 Weak-Strong Approach

Let us assume that $N_+ \gg 1$ and $N_- = 1$. The e^+ beam is in the Gaussian distribution, Eq. (2). Then the force (kick) felt by an e^- particle is given by

$$F(X) = \text{const.} \times \int dX_* \rho_*(X_*) H(X - X_*) = -2\pi^{3/2}\eta \ \text{erf}\left[(X - \bar{X}_*)/(\sqrt{2}\sigma_*)\right],$$

where ρ is the 1D Gaussian distribution and η is a constant, called the beam-beam parameter, depending on N_+, beam energy etc. An e^- feels a kick as

$$P \longrightarrow P + F(X), \quad X \longrightarrow X \tag{4}$$

at the IP. Thus, an e^- follows the maps Eq.(1) and Eq.(4) alternatively. After numbers of turns, depending on λ, the e^- particles fall into an equilibrium distribution. In Fig.1, the left figures are the Poincaré surface plots in (X, P) space when $\lambda = 1$; the right ones are the equilibrium distribution with $\lambda \lesssim 1$.

It can be seen that the particles are distributed equally to the resonances. How the particles are distributed between all different types of resonances is one of the interesting topics now under investigation.

3.2 Strong-Strong Approach

From Eq.(3), the macroscopic kick felt by each particle at the IP is

$$\Delta P = -C \int_{-\infty}^{\infty} dX, \ \rho_*(\bar{X}) \ H_1(X - \bar{X}). \tag{5}$$

Here, ρ_* is the distribution function of the opposite bunch. Equation (5) can be rewritten as

$$P \longrightarrow P - 2\pi^{3/2}\eta(N_*^a - N_*^b)/(N_*^a + N_*^b),$$

where N_*^a and N_*^b is the number of particles in the opposite bunch above and below the particle, respectively.

Starting from various configurations, we observe the steady states. See Fig.2. Contrary to the weak-strong cases, there are states where all the particles in a bunch get together forming a cluster. That is, somehow, particles prefer to behave together, instead of being distributed into different resonances. Also, the symmetry between two beams is broken spontaneously. Further, the steady state is not unique for a given parameter set. All these are the coherent phenomena and can not be imagined from the weak-strong approach.

When we change the parameters continuously and slowly, the state changes from one to another at certain point. Here, we can observe the hysteresis. In Fig.3, the phase space distribution are shown for the case with $\mu = 2\pi \times 0.48$. When $\eta_+ \gg \eta_-$, the e^- bunch is split into two clusters as shown in the weak-strong case. If, for example, we decrease η_+ on the line $\eta_+ + \eta_-$ =const., the e^- particles suddenly start to form a cluster at certain point, entering into the strong-strong regime. If we go down further, eventually the e^+ particles start splitting into two clusters also suddenly at certain point. The points for these sudden changes are different if we go up on the same line.

We have shown briefly how beams behave under the beam-beam interaction. This is one of the typical nonlinear behavior of the accelerator beams. The real accelerator beam is under a more complicated situation but the simple model presented here helps us to understand the complicated behaviour more intuitively. To understand the real beam, we need more realistic simulation, which is sometimes too difficult even with the present-day computers[5].

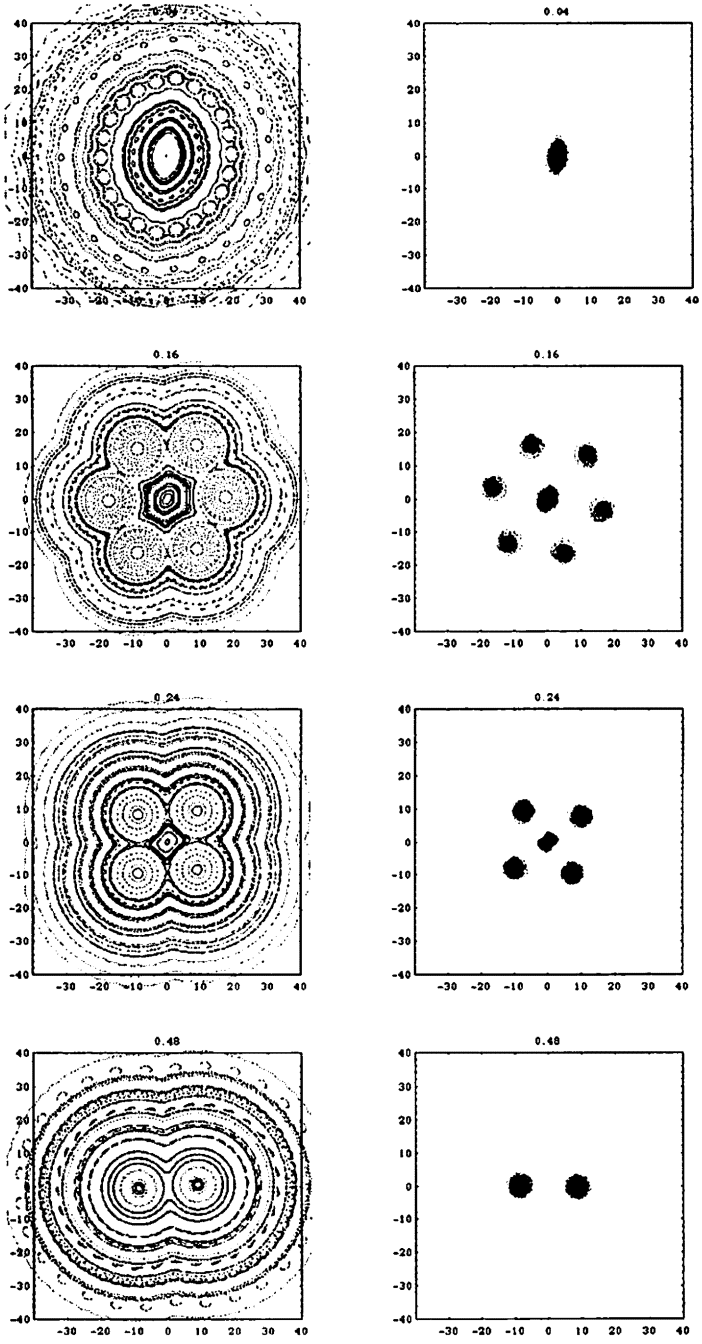


Figure 1: the phase space for the weak-strong approach. The Poincaré surface plot (left) and the equilibrium with radiation effect. The parameter μ's are $2\pi \times 0.04$, $2\pi \times 0.16$, $2\pi \times 0.24$, and $2\pi \times 0.48$, from the top to the bottom.

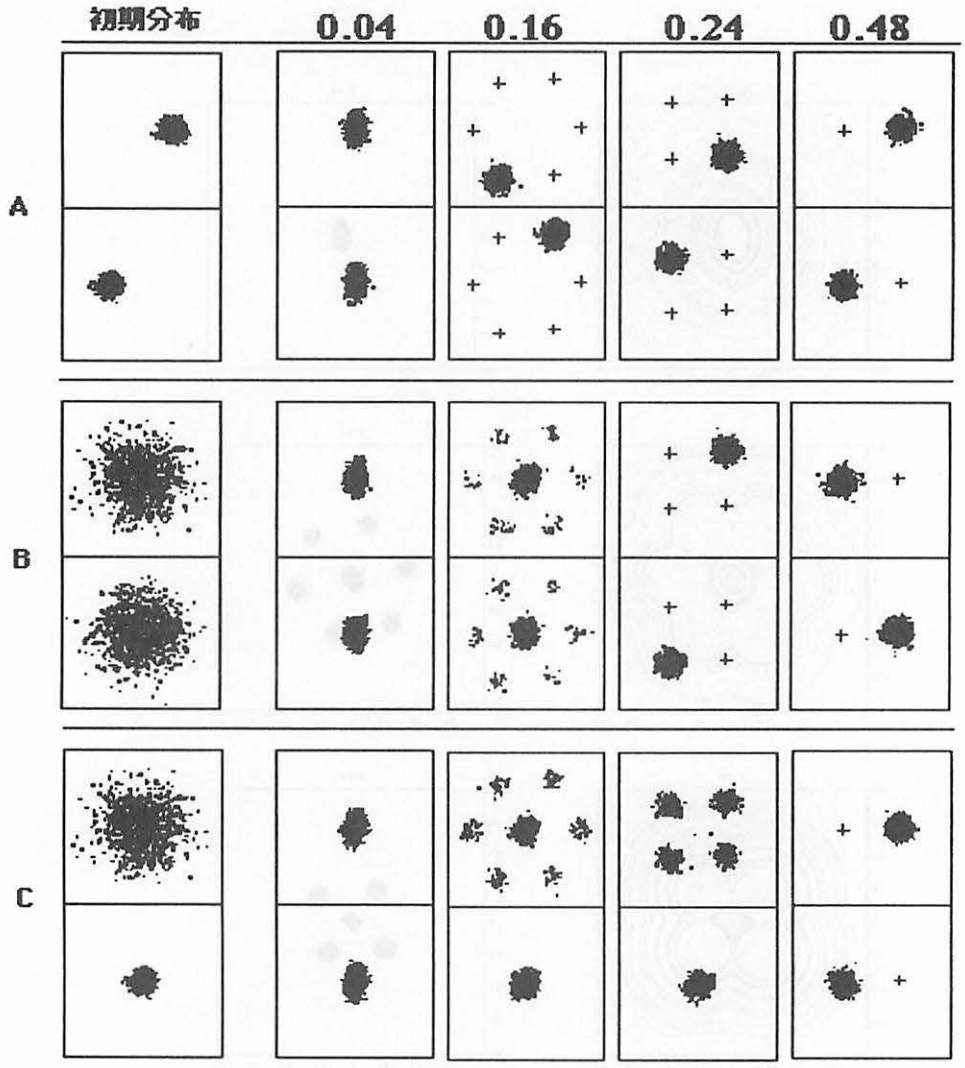


Figure 2: Phase space distribution of the both beams.The left is the initial distributions. The others are the steady states starting from them. Parameters: $\nu = 0.04$, 0.16, 0.24, 0.48, $\eta = 0.06$, where $\nu = \mu/(2\pi)$. The symbol + indicates the position of barycenters in other turns. For example, when we start from the states A with $\nu = 0.16$, the steady state is that all the e^+ and e^- particles form a cluster sitting in the opposite position in the phase space. They change their position every turn as indicated by + and come back to the same position after 6 turns.

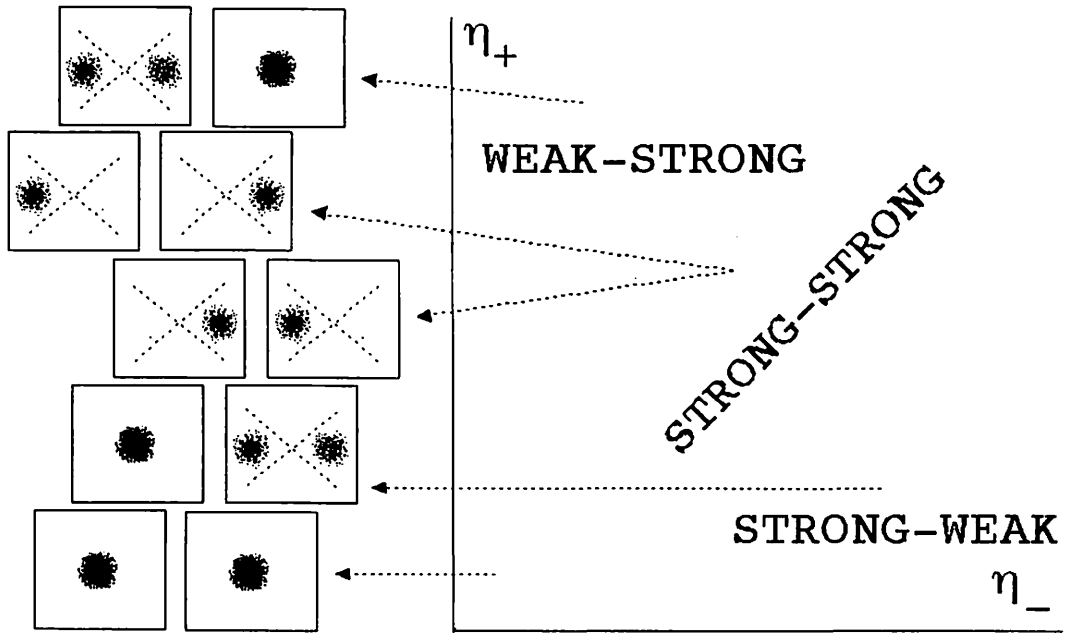


Figure 3: The hysteresis of the states under the smooth change of η_+ and η_-.

4 Beam Physics for High Energy Accelerators

For long time, the dynamics study of the accelerator beams has been done in more engineering attitude. Only recently, it is realized that there are many subjects there to be studied more deeply in physics sense. The reason may be as follows:

- because of the need to have more advanced accelerators, the empirical approach is no longer effective,
- some acdemic approaches turned out to be quite succesful,
- physicists found it quite interesting.

This tendency joins with other stream in other disciplines and is forming a new discipline called the beam physics, which treats accalerator beams, plasmas, lasers, and their mutual interactions[6].

References

[1] An easily readable introduction to the beam dynamics for electron/positrons is Ref.[2]. It can be recommended for non-specialist. For more professional beam dynamics theorists, see Refs.[3, 4].

[2] M. Sands, *The Physics of Electron Storage Rings*, SLAC-121 (1970).

[3] K. Hirata and F. Ruggiero, Part. Accel.**28**, 137 (1990).

[4] K. Ohmi, K. Hirata and K. Oide, *From the Beam Envelope Matrix to Synchrotron Radiation Integrals*, Phys. Rev. **E49**, 751 (1994).

[5] For more discussions, see K. Hirata, in *Handbook of Accelerator Physics and Engineering*, A. W. Chao and M. Tigner eds., World Scientific (1999), and references quoted there.

[6] See our www site (in Japanese): http://www-acc-theory.kek.jp/BP/BPclub.html

Effects of Interaction between Heat Reservoir and Cluster on Spontaneuos Alloying Phenomenon

T.R. Kobayashi,* K.S. Ikeda †, S. Sawada‡ Y. Shimizu‡

Abstract

"Spontaneuos alloying" for nanometer size metal clusters observed by Yasuda, Mori et al. and simulated with classical Hamiltonian dynamics by Ikeda, Sawada, Shimizu is simulated under isothermal conditions using classical molecular dynamics with Langevin dynamics and velocity scaling method.(Section 2) Very rapid allowing was comfirmed similar classical Hamiltonian dynamics simulations, i.e. "Peeling dynamics" (Subsection 4.1) and "2-step diving" (Subsection 4.2) occur under isothermal conditions. However, typical time scale of those phenomena seem to vary order of $10^{\pm 1}$. (Section 4) It is not considered that these results are caused by either increment of the temperature or dynamical perturbation yet.

1 Introduction

"Spontaneuos alloying" is a phenomenon that nanometer size metal clusters (host) which is covered with other appropriate metallic atoms (guest) alloy at anomalously rapid rate even at room temperature in comparison with bulk system [2].

Remarkable feartures of spontaneuos alloying are following.

- **The heat of solution is the most dominante factor.**
  The rapid alloying occurs for a wide class of conbinations of metals with negative heat of solution.
- **The spontaneuose alloying proceeds in the solid phase.**
  Through the alloying process, the global crystalline structure is not changed.
- **There is a critical size of clusters.**
  The spontaneuose alloying does not occure when the size of the cluster is over a critical size.

The alloying speed is anomalously rapid, and the diffusion rate is $10^{9\sim10}$ times larger than that of the corresponding bluk system. (table 1)

*Department of Physics, Faculty of Science and Engineering, Ritsumeikan University, Nojihigashi 1-1-1, Kusatsu 525-8577, Japan

†Department of Physics,Kwansei Gakuin University, 1-1-155, Uegahara, Nishinomiya 662-0891, Japan

‡Institute for Fundamental Chemistry, Takano-Nishihiraki-cho 34-4, Sakyou-ku, Kyoto 606-8103, Japan

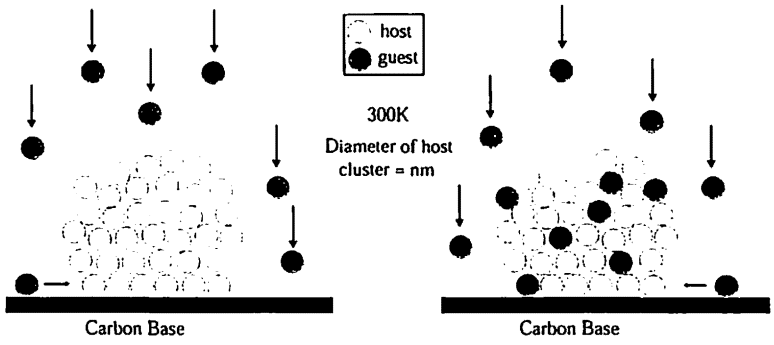


Figure 1: Conceptual illustration of spontaneuse alloying. Copper atoms (guest) are vapor-deposited onto nm-size gold cluster (host) on base of carbon graphite film (left). Then a rapid dissolution of guest atoms into host cluster takes place (right).

Table 1: A comparison of diffusion coefficient between bulk system and cluster. (Temperature = 300K)

$(Cu \rightarrow Au)$	diffusion coefficient
Bulk	$2.4 \times 10^{-28}[m^2/s]$
Cluster	$2.0 \times 10^{-19}[m^2/s]$

2 Modeling

2.1 Model potential

According to results of experiments, spontaneous alloying is a quite universal phenomenon which is supposed to be controlled dominantly by the heat of solution. Therefor Morse potential which is typical two-body potential is adopted [1]. Morse potential for i numbered atom can be written as,

$$V_i = \sum_{i \neq j} \epsilon_{i,j}[\exp(-2\beta(r_{ij} - \sigma)) - 2\exp(-\beta(r_{ij} - \sigma))] \tag{1}$$

where σ specifies the average of bond length, β means the decay rate of potential, ϵ is depth of potential.

The differences of interaction between same and different species are represented by the heat of solution. Thus

$$\epsilon_{i,j} = \begin{cases} \alpha\epsilon & i,j \text{ different species}(\alpha \neq 1) \\ \epsilon & i,j \text{ same species} \end{cases} \tag{2}$$

2.2 Langevin dynamics method

Langevin eq. is following.

$$\dot{v}(t) = -\gamma_L v + \frac{R(t)}{m} \tag{3}$$

Its discretized form is adopeted.

$$v(t+dt) - v(t) = -\gamma_L v(t) + \sqrt{\frac{2\gamma_L k_B T}{mdt}} R_{Gauss}(t) \tag{4}$$

where $R_{Gauss}(t)$ is random number which obeys Gauss distribution, $v(t)$ is velocity of an atom at the time of t. γ_L is a inverse factor of decay rate of time.

Langevin dynamics method reproduces effects of Brownian motion.

2.3 Velocity scaling method

Velocity scaling metod is one of the correction method, such as kinetic energy is converged at setting point.

$$v'(t) = (1 + \gamma_v \frac{T_{setting} - \int_{t-\delta t}^{t} \frac{T(t)}{\delta t} dt}{T_{setting}}) v(t) \tag{5}$$

Where γ_v specifies a inverse factor of decay rate of time.

Velocity scaling method provides pure temperature control without dynamical disturbance.

3 Simulation conditions

3.1 Setting of temperature

Alloying proceeds in the solid phase. Thus whole process is required under melting temperature conditions. We adopt two methods in order to decide the melting temperature. One of the method is Lindemann index which indicate the degree of mobility of atoms.

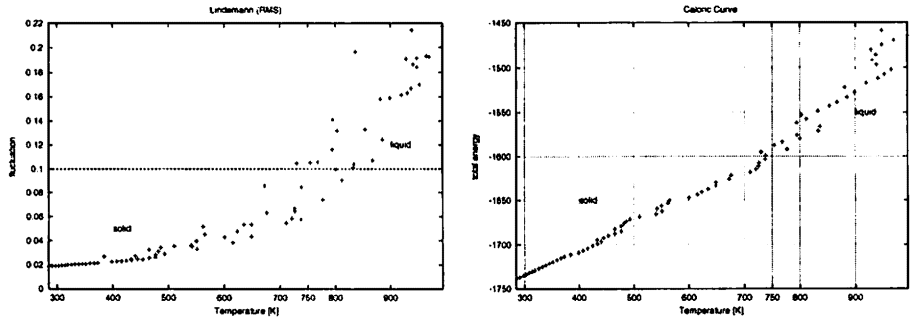
$$\delta = \frac{2}{N(N-1)} \sum_{i<j} \frac{\sqrt{\langle r_{ij}^2 \rangle_t - \langle r_{ij} \rangle_t^2}}{\langle r_{ij} \rangle_t} \tag{6}$$

The melting temperature is point of $\delta = 0.1$.

Two of the method is decided by latent heat from caloric curve (Fig. ??).

3.2 Setting of decay time

Estimation from experiments suggests that the time scale, which heat of cohesion escape from cluster to carbon base and system acheive equilibrium, is about 10^{-3} sec. Rescaleing this time scale to the system of simulations, we set decay time about $10^{-9 \sim 10}$ sec.



(a) Lindemann (b) Caloric curve

Figure 2: These results indicate melting temperature at 750 K. System parameters are Cu55(host)+12(guest) $\alpha = 1.2$. Each point is average of 10^5 steps.

4 Results

4.1 Peeling dynamics

Peeling dynamics was named after behavior of deviation which alloying process proceeds in the solid phase states. Inner part of host cluster can not move about such as liquid motion. Thus movable atoms can exist surface of cluster and they are diffusing there. On the process of surface atoms' diffusion surface atoms sometimes gather one side of cluster. These motions cause what inner part of cluster move opposite side by reaction, then a part of them expose. Therefor a mechanism of alloying proceeding in the solid phase is generated by reiteration of these replaceing processes. (Fig. 3)

4.2 2-step diving

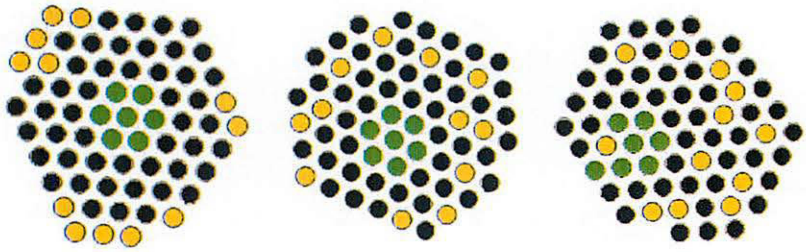
We observed that alloying behavior is classified two steps. First stage, guest atoms rapid dive into second layer from the surface. Second stage, then guest atoms diffuse into more inner part of the cluster. (Fig. 4)

We introduce following physical quantity for analysis.

1. **Bond number between differnt species**
   The number of host atoms which is nearest neighbor of guest atoms, and convert per atom.

$$b(t) = \sum_{i=1}^{N_{guest}} \frac{N_{host}^{(i)}}{N_{guest}} \tag{7}$$

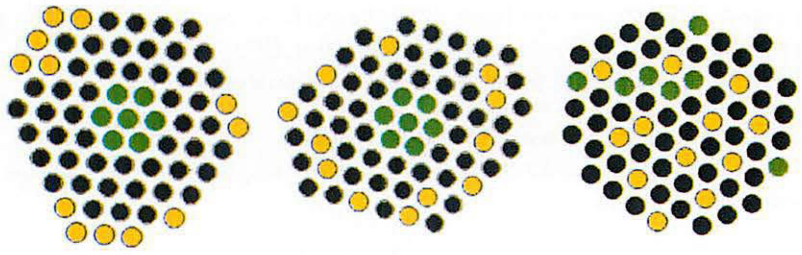
2. **averaged distance from center to guest atoms**



(a) initial configuration (b) 787 ns (c) 979 ns

Figure 3: figures of Peeling dynamics. 7 atoms which exist center of cluster at first time is transported furface of cluster after 979 ns.

(a) averaged distance from center to guest atoms



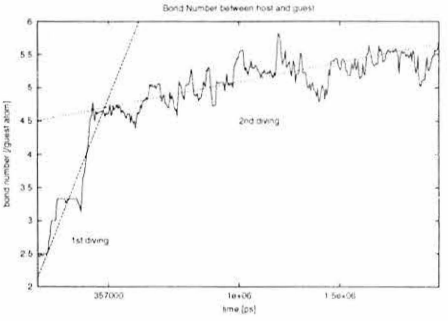
(b) initial configuration (c) 1st stage completed (d) ideal mixing state

Figure 4: figures of 2-step diving 12 guest atoms which exist surface of the cluster at initial configuration have dived one layer until time at "1st satge completed", then change motion from diving to diffusion.

$$dive(t) = \sum_{N_{guest}} \frac{r_{N_{guest}}}{N_{guest}} \tag{8}$$

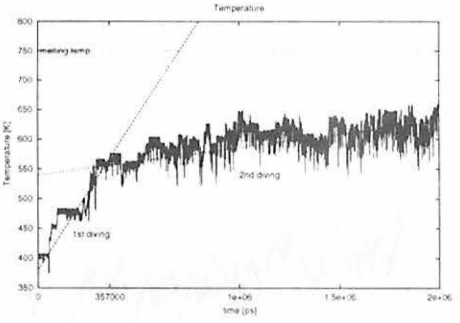
4.3 Isolated system

For the reference we elect isolated sysytem which initial temperature is 400 K. 1st stage is completed at 550 K and ideal mixing is completed at 600 K. Thus temperature of Langevin dynamics and velocity scaling simulations are setting to 550,600 K.

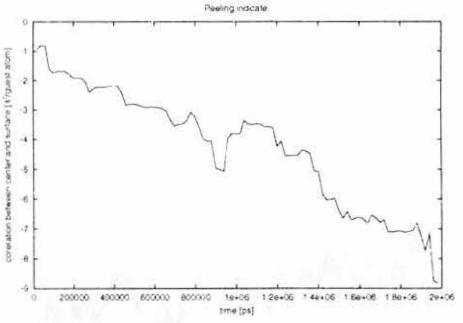


(a) Bond number between differnt species

(b) Average of distance from center to guest atoms



(c) Temperature

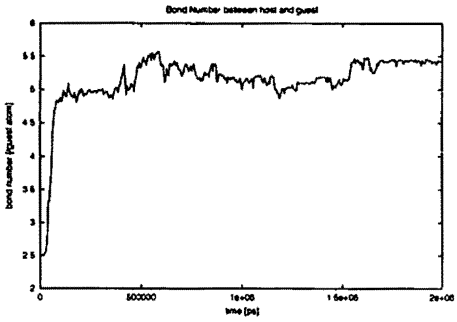


(d) Correlation of motion between center atom and average of surface atoms

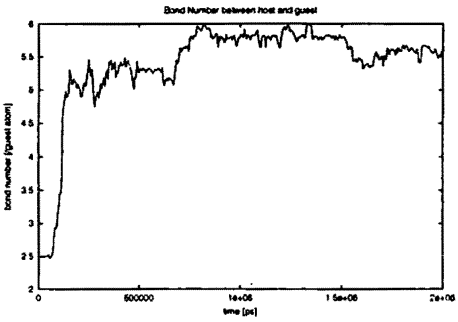
Figure 5: At 357 ns 1st diving stage is completed. In this condition it is about 2 μs until ideal mixing.

4.4 Langevin dynamics

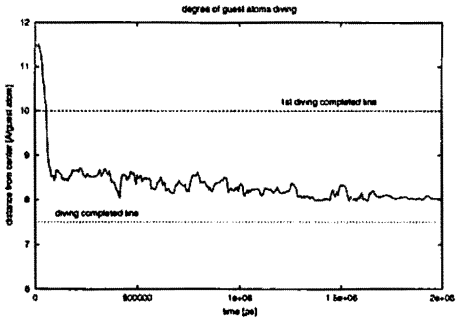
Temperature is 550 K and 600K.



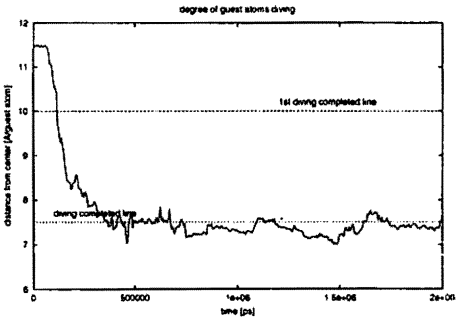
(a) Bond number between differnt species 550K



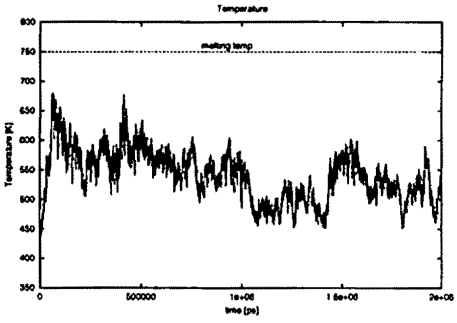
(b) Bond number between differnt species 600K



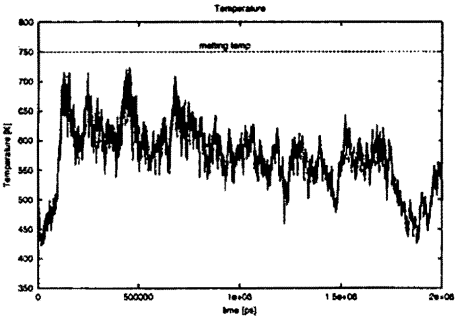
(c) average of distance from center to guest atoms 550K



(d) average of distance from center to guest atoms 600K



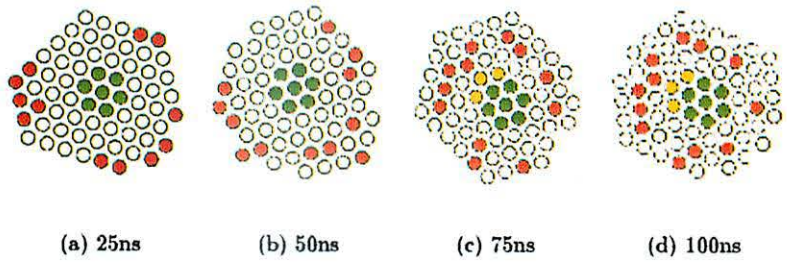
(e) Temperature 550K



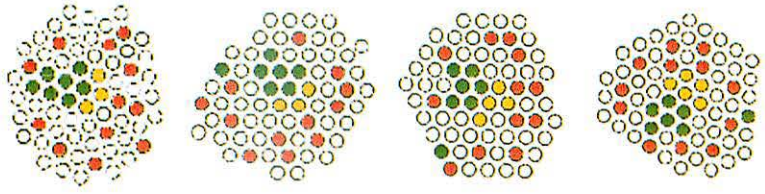
(f) Temperature 600K

Figure 6: Langevin time serieses.

Snap shots of cluster. Red balls are geust atoms. Green balls are host atoms which existed center at initial condition. Yellow balls are host atoms moved to center and became to interacte with heat reservoir.

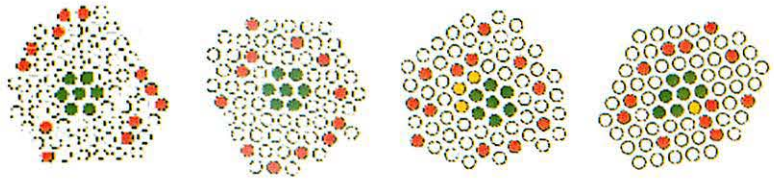


(a) 25ns (b) 50ns (c) 75ns (d) 100ns

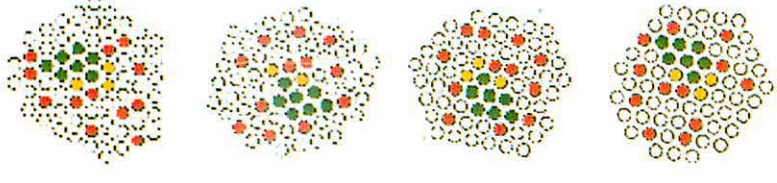


(e) 500ns (f) 1000ns (g) 1500ns (h) 2000ns

Figure 7: Snap shots of Langevin 550K.



(a) 100ns (b) 125ns (c) 150ns (d) 180ns



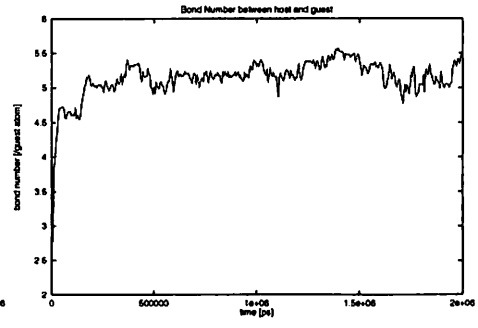
(e) 500ns (f) 1000ns (g) 1500ns (h) 2000ns

Figure 8: Snap shots of Langevin 600K.

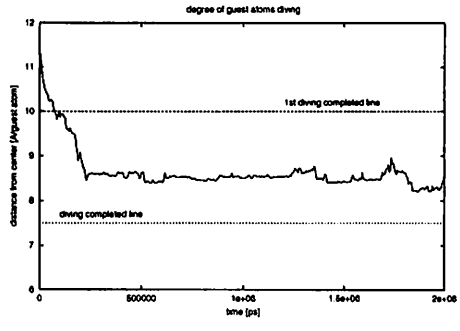
4.5 Velocity scaling

Temperature is 550 K and 600K.

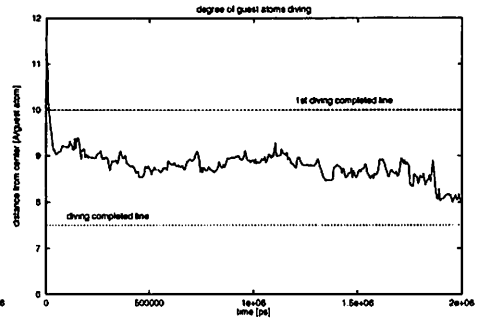
(a) Bond number between differnt species 550K



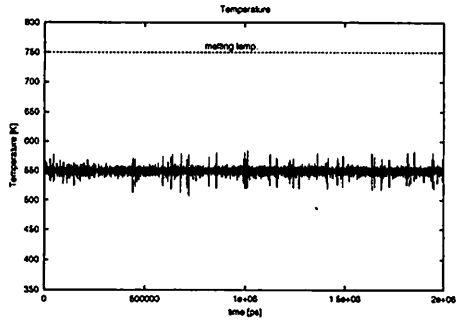
(b) Bond number between differnt species 600K



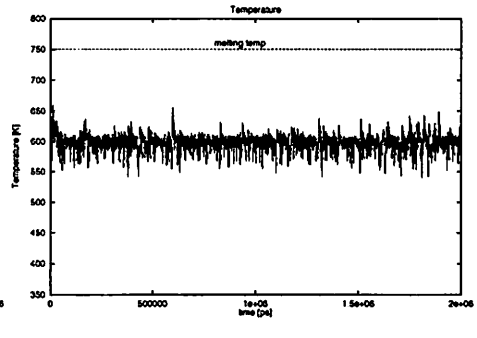
(c) average of distance from center to guest atoms 550K



(d) average of distance from center to guest atoms 600K



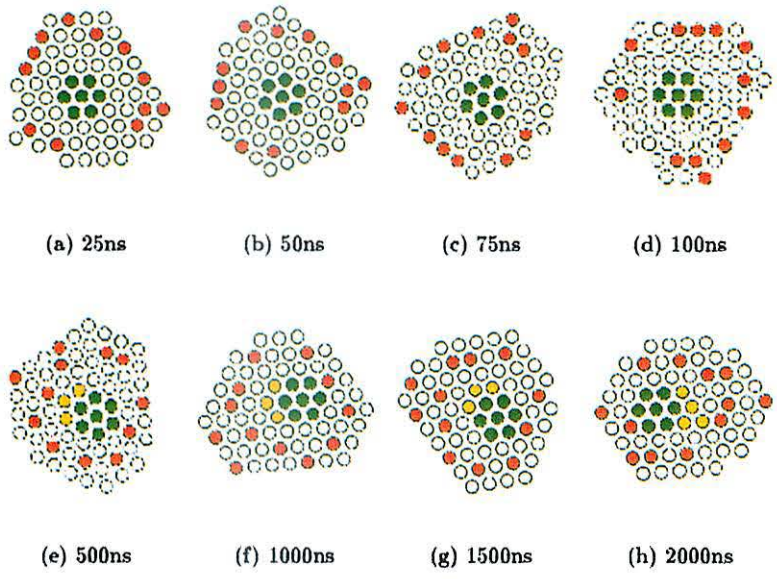
(e) Temperature 550K



(f) Temperature 600K

Figure 9: Velocity scaling time serieses.

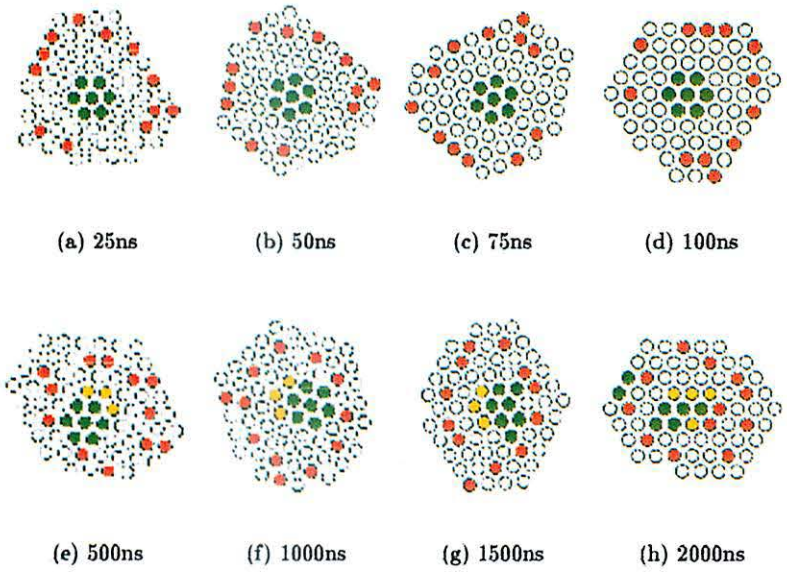
Snap shots of cluster. Red balls are geust atoms. Green balls are host atoms which existed center at initial condition. Yellow balls are host atoms moved to center and became to interacte with heat reservoir.



(a) 25ns (b) 50ns (c) 75ns (d) 100ns

(e) 500ns (f) 1000ns (g) 1500ns (h) 2000ns

Figure 10: Snap shots of velocity scaling 550K.



(a) 25ns (b) 50ns (c) 75ns (d) 100ns

(e) 500ns (f) 1000ns (g) 1500ns (h) 2000ns

Figure 11: Snap shots of velocity scaling 600K.

References

[1] L. A. Girifalco and V. G. Weizer. Application of the morse potential function to cubic metals. *Phys. Rev.*, 114(3):687–690, 1959.

[2] H. Yasuda, H. Mori, M. Komatsu, K. Takeda, and H. Fujita. In situ observation of spontaneous alloying in atom clusters. *J. Electron Microsc.*, 41, 1992.

The intersection angles between N-dimensional stable and unstable manifolds in $2N$-dimensional symplectic mappings

Yoshihiro Hirata* , Kazuhiro Nozaki and Tetsuro Konishi
Department of Physics, Nagoya University, Nagoya, 464-8602, Japan

Abstract

We asymptotically compute the intersection angles between N-dimensional stable and unstable manifolds in $2N$-dimensional symplectic mappings. In general, there exist the particular 1-dimensional stable and unstable sub-manifolds which experience exponentially small splitting of separatrix. We show that the angle along the sub-manifolds is exponentially small with respect to the perturbation parameter ϵ, and the other angles are power of ϵ.

1 Introduction

The existence of a transversal intersecting point between stable and unstable manifolds leads to non-integrability of a Hamiltonian system. Besides, in 2-dimensional symplectic mappings (area-preserving mappings), which correspond to Poincaré mappings of Hamiltonian systems with 2 degrees of freedom, the area of the lobes enclosed by stable and unstable manifolds stands for phase space flux from inside (resp. outside) of separatrix to outside (resp. inside). Because the intersection angle at the principal intersecting point is related with the area of the lobes, the intersection angle is also related with phase space transport. Therefore the intersection angle is essentially important to understand the global structure of phase space.

Melnikov's method (or "Melnikov's integral") is a powerful tool to detect transversal homoclinic points and to compute the area of lobes[1]. The method is, however, based on the particular feature of 2-dimensional mappings. Hence one cannot straightforwardly apply this method to many-dimensional systems, especially the fixed points which possess two or more hyperbolic modes.

On the other hand, for some symplectic maps, e.g. the standard map, Hénon map, and the double-well map, one can analytically construct asymptotic expansions of stable and unstable manifolds[2–4]. Hence one can also compute the asymptotic form of the intersection angles between stable and unstable manifolds in such maps. This asymptotic method is called "asymptotic expansion beyond all orders." We have been extending the method to 4- or more-dimensional symplectic maps.

In [5] we constructed the asymptotic expansions of particular 1-dimensional stable and unstable sub-manifolds which experience exponentially small splitting of separatrix, in a 4-dimensional double-well symplectic mapping. Furthermore, in [6], the neighborhood of the sub-manifolds was asymptotically studied and exponentially small oscillating terms were successfully obtained. In this paper we treat $2N$-dimensional double-well symplectic mappings with nearest-neighbor two-body couplings and asymptotically compute the intersection angles between N-dimensional stable and unstable manifolds.

The outline of this paper is as follows. In Sec. 2 we introduce the $2N$-dimensional symplectic map we treat in this paper. Sec. 3 is spared explanation of the approximation and the coordinate system we use. Sec. 4 is the main part of this paper and we show the intersection angles. In Sec. 5 we give conclusion and discussion on the results.

* E-mail address: yhirata@allegro.phys.nagoya-u.ac.jp

2 Model

The canonical coordinates we use are $\vec{q} = (q_1, \cdots, q_N) \in \mathbf{R}^N$, $\vec{p} = (p_1, \cdots, p_N) \in \mathbf{R}^N$ and $\vec{\gamma} = (\vec{q}, \vec{p})$. We start with $2N$-dimensional symplectic mappings $\vec{\gamma} \mapsto \vec{\gamma}'$,

$$\begin{cases} p_j' &= p_j - \epsilon \dfrac{\partial V_0(q_j)}{\partial q_j} - \epsilon^3 \dfrac{\partial V_1(\vec{q})}{\partial q_j}, \\ q_j' &= q_j + \epsilon p_j', \end{cases} \tag{1}$$

where $j = 1, 2, \cdots, N$[†] and $|\epsilon| \ll 1$. In this paper we take account of the two different mapping: One is the double-well mapping (DW), the other is the standard mapping (SM). The potential function $V_0(q_j)$ is given as

$$V_0(q) = \begin{cases} \dfrac{1}{4}q^4 - \dfrac{1}{2}q^2, & \text{for DW} \\ \cos q, & \text{for SM} \end{cases} \tag{2}$$

respectively. For both of the potential functions, the origin $(0, \cdots, 0)$ is a $2N$-dimensional hyperbolic fixed point (a direct product of N 2-dimensional hyperbolic fixed points), and there exist N-dimensional stable and unstable manifolds near the origin. The coupling potential $V_1(\vec{q})$ is given as a polynomial of the distance between nearest-neighbors of order 4, i.e.,

$$V_1(\vec{q}) = \sum_{j=1}^{N} \frac{a_j}{2}(q_j - q_{j+1})^2 + \sum_{j=1}^{N} \frac{b_j}{4}(q_j - q_{j+1})^4, \tag{3}$$

where a_j's and b_j's are real constants and $q_{N+1} = q_1$. Now we restrict $a_1 = a_2 = \cdots = a_N = a$ and $b_1 = b_2 = \cdots = b_N = b$ to simplify the following computation.

3 The outer approximation and the unperturbed solution

One can rewrite the map (1) into the second order difference equation

$$\Delta_\epsilon^2 q_j(t) = f_0(q_j) + \epsilon^2 f_{j,1}(\vec{q}) \tag{4}$$

where $\Delta_\epsilon^2 q(t) = \{q(t+\epsilon) - 2q(t) + q(t-\epsilon)\}/\epsilon^2$, $f_0(q_j) = -\dfrac{\partial V_0(q_j)}{\partial q_j}$ and $f_{j,1}(\vec{q}) = -\dfrac{\partial V_1(\vec{q})}{\partial q_j}$. The stable and unstable manifolds $q_j^{\pm}(t;\epsilon)$ is asymptotic to the origin as $t \to \pm\infty$, respectively. Hence the boundary condition with which we solve the difference equation (4) is given as $q_j^{\pm}(t;\epsilon) \searrow 0\,(t \to \pm\infty)$. Expanding $q_j(t \pm \epsilon) = \sum_{l=0}^{\infty} \dfrac{(\pm\epsilon)^l}{l!} \dfrac{d^l q_j}{dt^l}$, one obtains

$$\frac{d^2 q_j}{dt^2} = f_0(q_j) + \epsilon^2 f_{j,1}(\vec{q}) - 2\sum_{l=2}^{\infty} \frac{\epsilon^{2l-2}}{(2l)!} \frac{d^{2l} q_j}{dt^{2l}}, \tag{5}$$

which we call the outer equation. In general, although the infinite summation in the r.h.s. of the equation (5) does not converge, the equation gives a good approximation of the map (1).

As the outer approximation, we construct the solution to the outer equation by expanding $q_j(t;\epsilon)$ into the power series of ϵ^2, i.e., $q_j(t;\epsilon) = q_{j,0}(t) + \epsilon^2 q_{j,1}(t) + \cdots$. In the same way, the stable and unstable manifolds $\vec{\gamma}$ is expanded as $\vec{\gamma} = \vec{\gamma}_0 + \epsilon^2 \vec{\gamma}_1 + \cdots$.

[†] We shall use the subscript j as meaning $j = 1, 2, \cdots, N$ without further comment.

By substituting the expansion into the outer equation (5), one gets the unperturbed equation

$$\frac{d^2 q_{j,0}}{dt^2} = f_0(q_{j,0}), \tag{6}$$

and the perturbed equation (the linearized equation)

$$\frac{d^2 q_{j,l}}{dt^2} = f_0'(q_{j,0}) q_{j,l} + \tilde{f}_{j,l}(t), \tag{7}$$

for $l = 1, 2, \cdots$, where the inhomogeneous term $\tilde{f}_{j,l}$ is a polynomial of $q_{j,0}(t)$'s and their derivatives, e.g., $\tilde{f}_{j,1} = f_{j,1}(q_{j,0}) - \frac{1}{12}\frac{d^4 q_{j,0}}{dt^4}$. In the same way the boundary condition is also expanded as $q_{j,l}^{\pm}(t) \searrow 0\,(t \to \pm\infty)$, $l = 0, 1, \cdots$. Note that the ordinary differential equations for $q_{j,l}$, $l = 0, 1, \cdots$ form a Hamiltonian system with N degrees of freedom in order-by-order.

The unperturbed solution $q_{j,0}(t)$ is obtained as $q_{j,0} = s(t + t_j)$ where $s(t)$ is written as

$$s(t) = \begin{cases} \mathrm{sech}(t), & \text{for DW} \\ 4\tan^{-1}(e^t), & \text{for SM} \end{cases} \tag{8}$$

and t_j's are constants of integration. Let us put $s(t + t_j) = s_j(t)$ to simplify notations. Now we have N free parameters, t_j's. Because the map (1) is autonomous, it is translational invariant with respect to t. Hence one of the parameters is not essential and can be vanished. To simplify the further computation we transform t_j's into new variables α_j's as

$$t_j = \alpha_{j-1} - \alpha_j, \tag{9}$$

where $t_0 = t_N$, $t_{N+1} = t_1$, and $\alpha_0 = \alpha_N = 0$. These $N - 1$ variables, α_j's, and t span the N-dimensional stable and unstable manifolds. We put $\vec{\gamma}_{0,k} = \frac{\partial \vec{\gamma}_0}{\partial \alpha_k}(t, 0)$, $k = 1, \cdots, N-1$ and $\vec{\gamma}_{0,N} = \frac{\partial \vec{\gamma}_0}{\partial t}(t, 0)$. Owing to the relation (9), $\vec{\gamma}_{0,k}$'s are orthogonal to $\vec{\gamma}_{0,N}$ and $\vec{\gamma}_{0,k+m}$, $m \geq 2$. Note that $\vec{\gamma}_{0,k}$ $(k \neq N-1)$ are not orthogonal to $\vec{\gamma}_{0,k+1}$.

4 The linearized equation and the intersection angles

Next we construct the solutions to the linearized equation (7). The fundamental system of solutions to the homogeneous equation is given as $\dot{s}_j(t) = \dot{s}(t + t_j)$ and $g_j(t) = g(t + t_j)$ where $g(t) = \dot{s} \int \frac{dt}{\dot{s}^2}$. Note that the Wronskian determinant $W(t) = \begin{vmatrix} \dot{s}(t) & g(t) \\ \ddot{s}(t) & \dot{g}(t) \end{vmatrix} \equiv 1$. The stable and unstable solutions for the equation (7) can be given as

$$q_{j,1}^{\pm}(t, \vec{\alpha}) = g_j(t) \int_{\pm\infty}^{t} \dot{s}_j(t) \tilde{f}_{j,1}(t) dt - \dot{s}_j(t) \int_0^t g_j(t) \tilde{f}_{j,1}(t) dt. \tag{10}$$

When $\vec{\alpha} = 0$, the r.h.s. of the equation (10) is easily integrated. The result does not depend on j and the sign $\pm$, hence putting $q_1(t) = q_{j,1}^{\pm}(t, 0)$,

$$q_1(t) = \begin{cases} \frac{1}{3}\mathrm{sech}^3(t) - \frac{7}{24}\mathrm{sech}(t) + \frac{t}{24}\sinh(t)\mathrm{sech}^2(t), & \text{for DW} \\ \frac{1}{4}\sinh(t)\mathrm{sech}^2(t) - \frac{t}{12}\mathrm{sech}(t). & \text{for SM} \end{cases} \tag{11}$$

The solution $q_j^{\pm}(t,0;\epsilon) = s(t) + \epsilon^2 q_1(t) + O(\epsilon^4)$ stands for the 1-dimensional stable and unstable sub-manifolds.

To compute the intersection angles at $\vec{\gamma}(t,0)$, we shall analyze the neighborhood of the sub-manifold as a perturbation problem with respect to α_j's. We expand $q_{j,1}^{\pm}(t,\vec{\alpha})$ with respect to α_j's as

$$q_{j,1}^{\pm}(t,\vec{\alpha}) = q_1(t) + \sum_{k=0}^{N-1} \alpha_k \tilde{q}_{j,k}^{\pm}(t) + O(\alpha_j^2), \tag{12}$$

where $\tilde{q}_{j,k}^{\pm}(t) = \left.\dfrac{\partial q_{j,1}^{\pm}(t,\vec{\alpha})}{\alpha_k}\right|_{\vec{\alpha}=0}$. By using (10), straightforwardly computation gives $\tilde{q}_{j,k}^{\pm}(t)$ as

$$\tilde{q}_{j,k}^{\pm}(t) = t_{j,k} F_1^{\pm}(t) + (t_{j+1,k} + t_{j-1,k}) F_2^{\pm}(t), \tag{13}$$

where $t_{j,k} = \dfrac{\partial t_j}{\partial \alpha_k}$. The functions $F_1^{\pm}(t)$ and $F_2^{\pm}(t)$ is explicitly written by using the functions $\dot{s}(t)$ and $g(t)$ as

$$\begin{aligned} F_1^{\pm}(t) &= \dot{q}_1(t) \pm a\beta g(t) - a\delta \dot{s}(t)\sinh^2(t), \\ F_2^{\pm}(t) &= \mp\frac{1}{2} a\beta g(t), \end{aligned} \tag{14}$$

where $\beta = \displaystyle\int_{-\infty}^{\infty} \dot{s}(t)^2 dt = \frac{2}{3}$ (DW), 8 (SM), and $\gamma = \dfrac{1}{3}$ (DW), 1 (SM). Note that the expression (14) contains the parameter a, which determines the strength of the linear coupling.

By using the perturbed solution (13) we asymptotically compute the intersection angles between the stable and unstable manifolds. For this purpose, let us define $\Delta\vec{\gamma}_l(t,\vec{\alpha}) = \vec{\gamma}_l(t,\vec{\alpha}) - \vec{\gamma}_l(t,0)$, $l = 0,1$, and $\Delta\vec{\gamma}^{\pm}(t,\vec{\alpha}) = \Delta\vec{\gamma}_0(t,\vec{\alpha}) + \epsilon^2 \Delta\vec{\gamma}_1^{\pm}(t,\vec{\alpha})$ (see Fig. 1). We expand $\Delta\vec{\gamma}^{\pm}(t,\vec{\alpha})$ into the power series of α_k as

$$\Delta\vec{\gamma}^{\pm}(t,\vec{\alpha}) = \sum_{k=1}^{N-1} \alpha_k (\vec{\gamma}_{0,k} + \epsilon^2 \tilde{\gamma}_k^{\pm}), \tag{15}$$

up to $O(\alpha_k^2)$. We put

$$\vec{\alpha}_k = (0,\cdots,0,\overset{k}{\check{\alpha}_k},0,\cdots,0), \tag{16}$$

for $k = 1,\cdots,N-1$. Let us suppose that the stable and unstable manifolds intersect at $(t_0,0)$. We can define the intersection angles at the point $(t_0,0)$ as follows. For $m,n \in \{1,\cdots,N-1\}$, the angle $\theta_{m,n}$ is defined as

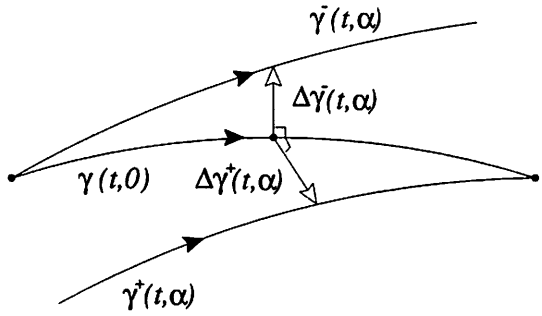


Figure 1: The stable and unstable manifolds $\gamma^{\pm}(t,\alpha)$ and the submanifold $\gamma(t,0)$. $\Delta\gamma^{\pm}(t,\alpha)$ stands for the vectors from $\gamma(t,0)$ to $\gamma^{\pm}(t,\alpha)$.

$$\theta_{m,n} = \cos^{-1} \frac{\Delta\vec{\gamma}^{+}(t_0,\vec{\alpha}_m)\cdot\Delta\vec{\gamma}^{-}(t_0,\vec{\alpha}_n)}{|\Delta\vec{\gamma}^{+}(t_0,\vec{\alpha}_m)||\Delta\vec{\gamma}^{-}(t_0,\vec{\alpha}_n)|}, \tag{17}$$

and

$$\theta_{N,N} = \cos^{-1} \frac{\dot{\vec{\gamma}}^{+}(t_0,0)\cdot\dot{\vec{\gamma}}^{-}(t_0,0)}{|\dot{\vec{\gamma}}^{+}(t_0,0)||\dot{\vec{\gamma}}^{-}(t_0,0)|}, \tag{18}$$

where $\cdot$ stands for the inner product. Straightforward computation gives

$$\begin{aligned}
\theta_{k,k}(t_0) &= \frac{\epsilon^2|\beta|}{\dot{s}(t_0)^2+\ddot{s}(t_0)^2}\left[9+(g(t_0)^2+\dot{g}(t_0)^2)(\dot{s}(t_0)^2+\ddot{s}(t_0)^2)\right]^{\frac{1}{2}}, \\
\theta_{k,k+1} &= \frac{2\pi}{3}+O(\epsilon^2), \\
\theta_{k,k+m} &= \frac{\pi}{2}+O(\epsilon^2)\ (m\geq 2),
\end{aligned} \tag{19}$$

for $k=1,\cdots,N-1$. $\theta_{N,N}$ is exponentially small with respect to ϵ[2–4, 7], i.e., there exist $C_1, C_2>0$ such that

$$|\theta_{N,N}| \leq \begin{cases} \dfrac{C_1}{\epsilon^5}\exp\left(-\dfrac{\pi^2}{\epsilon}\right), & \text{for DW} \\ \dfrac{C_2}{\epsilon^3}\exp\left(-\dfrac{\pi^2}{\epsilon}\right). & \text{for SM} \end{cases} \tag{20}$$

5 Summary and discussion

We successfully compute the intersection angles between the N-dimensional stable and unstable manifolds. In $2N$-dimensional phase space, there exist N independent intersection angles between two N-dimensional manifolds. The results (19) and (20) imply that in the set of the intersection angles between the N-dimensional stable and unstable manifolds, one is exponentially small and the others are power of ϵ.

The reason why one angle is exponentially small is that the outer equation is a Hamiltonian system, i.e., there exists an energy integral in order-by-order, and therefore the 1-dimensional homoclinic trajectory can exist under the outer approximation. This homoclinic trajectory in reality splits into two different manifolds, the stable and unstable sub-manifolds, and the splitting angle is exponentially small with respect to the perturbation parameter.

The intersection angles are related to phase space flux. We suppose that escape rate of trajectories gets large as N increases because the number of the angles in the power of ϵ dominates overwhelmingly. We would like to confirm this supposition with numerical experiments in the future work.

References

[1] J. Guckenheimer and P. Holmes, *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*, Springer-Verlag (1983).

[2] V. Hakim and K. Mallick, Nonlinearity **6** (1993) 57.

[3] A. Tovbis, M. Tsuchiya and C. Jaffé, Chaos, **8** (1998) 665.

[4] K. Nakamura and M. Hamada, J. Phys **A29** (1996) 7315.

[5] Y. Hirata and T. Konishi, *http://xxx.lanl.gov/abs/chao-dyn/9802009.*

[6] Y. Hirata, K. Nozaki and T. Konishi, Prog. Theo. Phys. **101**(5) (1999) to appear.

[7] V. G. Gelfreich, V. F. Lazutkin and N. V. Svanidze, Physica **71D** (1994), 82.

2重の対称性をもつ近可積分写像系における周期点分布

丸尾 剛 (名古屋大学)

The distribution of periodic point in nearly integrable maps with dual symmetries

Tsuyoshi Maruo

Department of Physics, Nagoya University
Nagoya 464-8602, Japan

abstract

一般の可逆写像系について、その系の写像から包合写像 (Involution) というものを構成できるとき、系の位相空間における運動の対称性や周期点の分布などを見出せることが知られている。さらに、いくつかの2次元シンプレクティック写像、特に標準写像においては、2種類の包合写像が構成できることが知られている。今回、2種類の包合写像を構成できる2次元近可積分写像系において、周期 n の共鳴領域にある n 周期点が 4n 個以上存在しなければならない場合があることが分かったので、その報告を行う。また、現在知られているタイプの包合写像とは異なる、未知のタイプの包合写像が存在する可能性についても述べる。

In general, if a invertible map has *involution*, we can find out a *symmetry* of the map and periodic point. Further, it is known that some two-dimensional symplectic maps, such as standard map, have two kinds of involution. In this paper, it is shown that in two-dimension nearly integrable maps the number of points of period n in a resonance must be not less than $4n$. And we also offer possibility that unknown type of involution exists.

1 Introduction

ここではまず、本研究における keyword となる位相空間の対称性、包合写像および対称線などの説明をしておく。

ある可逆写像 M に対して、その逆写像が

$$M^{-1} = RMR, \qquad R^2 = 1 \tag{1}$$

となるような鏡映写像Rが存在するとする。このことは、写像M^{-1}が写像Mに対する運動の時間反転であることから、鏡映写像Rに対する不動点の集合Γを鏡映線とする、運動の時間反転対称性が存在することを意味する[1]。

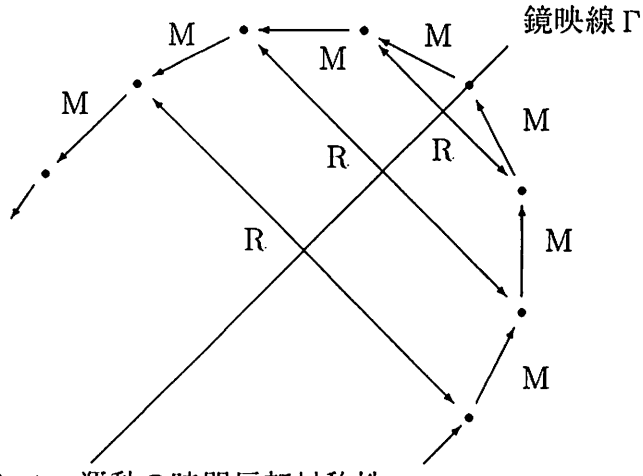


Fig.1　運動の時間反転対称性
symmetry in motion for the time inversion

ここで、次のように定義される包合写像(Involution)

$$I_n \equiv M^n R \qquad (n:\text{整数}) \tag{2}$$

を考える[2]。このとき、(1)から $M = RM^{-1}R$ であるので、I_nは

$$\begin{aligned} I_n &= M^n R \\ &= (RM^{-1}R)^n R \\ &= (RM^{-n}R)R \quad = \quad RM^{-n} \end{aligned} \tag{3}$$

と書くこともできる。ここで、M_nとI_nには、次のような関係があることが示される。

I1)	$M^j I_i$	$=$	$M^j(M^i R)$	$=$	$M^{(i+j)}R$	$=$	I_{i+j}
I2)	$I_i M^j$	$=$	$(RM^{-i})M^j$	$=$	$RM^{-(i-j)}$	$=$	I_{i-j}
I3)	$I_i I_j$	$=$	$(M^i R)(RM^{-j})$			$=$	M^{i-j}

またI3)で、$i=j$のときから

I3')	$I_i^2 = E$　　(E:恒等写像)

である。このようにして、生成された包合写像I_nに対して

$$I_n M I_n = I_n I_{n+1} = M^{-1} \tag{4}$$

であることから、(1)より、あらゆるnに対してどのI_nも、もとの鏡映写像Rにとることができ、添字の違いに対して本質的な違いはない。

また、I1)とI3')より

$$M = I_{i+1} I_i \tag{5}$$

となることが分かるが、これは、写像 M を 2 つの包合写像 $(I_i^2 = E, I_{i+1}{}^2 = E)$ の積の形に分解する問題でもあることが分かる。この問題はバーコフの研究 [3] にみることができる。

さて、ここで次のような包合写像 I_n の不動点の集合

$$\Gamma_n = \{x \mid I_n x = x\} \tag{6}$$

を考える。これは、一般に "対称線 (symmetry line)" と呼ばれている。この対称線に関する性質を、次に示す。

まず、対称線Γ_i上の点 x $(I_i x = x)$ に対して、$M^n x = y$ である点 y について

$$\begin{aligned} I_{i+2n} y &= I_{i+2n} M^n x \\ &= I_{i+n} x \quad [\leftarrow I2)\,] \\ &= M^n I_i x \quad [\leftarrow I1)\,] \\ &= M^n x \qquad = \; y \end{aligned}$$

であり、点 y は対称線Γ_{i+2n}上にあることになる。逆に、同様にして、対称線Γ_{i+2n}上の点は、逆写像 M^{-n}によって対称線Γ_i上に写されることも示せる。よって、まず次の性質があることが分かる。

S1)	任意の $i,\ n$ に対して、対称線Γ_iは、写像 M^nによって対称線Γ_{i+2n}に写される。

また、対称線Γ_i上の点 x に対して、$Rx = y$ である点 y について

$$\begin{aligned} y &= Rx \\ &= M^{-i} I_i x \quad [\leftarrow (2)\,] \\ &= M^{-i} x \end{aligned}$$

より、点 y は点 x を写像 M^{-i}で写した点だが、これは S1) の性質により対称線Γ_{-i}上にあることが分かる。よって、次のことが言える。

S2)	任意の i に対して、対称線Γ_iは、写像 R によって、対称線Γ_{-i}に写される。

次に、ある 2 つの対称線Γ_iとΓ_j $(i-j=n)$ の交点 x $(I_i x = x,\ I_j x = x)$ について、

$$\begin{aligned} M^n x &= M^{i-j} x \\ &= I_i I_j x \quad [\leftarrow I3)\,] \\ &= x \end{aligned}$$

となり、点 x は n 回写像 M^nに対して不動であり、次の性質があることが分かる。

S3)	任意の $i,\ j$ に対して、2 つの対称線Γ_iと$\Gamma_j (i \neq j,\ i-j=n)$ の交点は、写像 M の n 回写像 (M^n) の不動点である。

さらに、ある対称線Γ_i上の点xがn周期点であるとする。このとき、任意の整数lに対して、

$$\begin{aligned} I_{i-ln}x &= I_i M^{ln} x \quad [\leftarrow I2)\,] \\ &= I_i (M^n)^l x \\ &= I_i x \\ &= x \end{aligned}$$

であるので、点xは対称線Γ_{i-ln}上の点でもあり、次のことが言える。

S4)	ある対称線Γ_i上の点xがn周期点であるとき、点xは任意の整数lに対して対称線Γ_{i-ln}上の点でもある。

特に、S3) の性質は有用であり、周期点を求める際、2次元写像系であれば1次元の問題に帰着することができる。対称線と周期点の関係については、DeVogelaereによって研究されており [2]、またS3) の性質は、実際に具体的な写像についての周期点の解析に利用されている [4][5][6][7]。

2　シンプレクティック写像の2重対称性

まず、次のような一般的なシンプレクティック写像$M : (q,p) \mapsto (q',p')$を考える。

$$M \begin{cases} q' &= q + \epsilon p' &= q + \epsilon(p + \epsilon F(q)) \\ p' &= p + \epsilon F(q) \end{cases}$$

これは、母関数が

$$W = qp' + \epsilon H(q,p')$$

$$H = \frac{1}{2}p'^2 + V(q), \qquad -\frac{dV(q)}{dq} \equiv F(q)$$

である正準変換であり、

$$q' = \frac{\partial W}{\partial p}, \qquad p = \frac{\partial W}{\partial q'}$$

から、写像Mが構成される。また、逆写像$M^{-1} : (q',p') \mapsto (q,p)$は、

$$M^{-1} \begin{cases} q &= q' - \epsilon p' \\ p &= p' - \epsilon F(q) &= p' - \epsilon F(q' - \epsilon p') \end{cases}$$

である。パラメーターϵが無限小であるとき、この無限小変換は上のHをハミルトニアンとする時間推進変換となる。パラメーターϵを有限の値にとったこの写像系は、シンプレクティック写像の一般的なタイプとして用いられている。

さて、このタイプの写像は、まず次のような鏡映写像R

$$R\begin{pmatrix} q \\ p \end{pmatrix} = \begin{pmatrix} q-\epsilon p \\ -p \end{pmatrix} \qquad \text{鏡映線 } \Gamma:\ p=0$$

をもつことが知られており、写像 M は、鏡映写像 R による運動の時間反転対称性を有している (Fig.3)。

さらに、関数 $F(q)$ が奇関数 ($F(-q)=-F(q)$) であると、次のような鏡映写像 $\tilde{R}$

$$\tilde{R}\begin{pmatrix} q \\ p \end{pmatrix} = \begin{pmatrix} -q+\epsilon p \\ p \end{pmatrix} \qquad \text{鏡映線 } \Lambda:\ q-\tfrac{\epsilon}{2}p=0$$

をもち、これによる対称性も存在する (Fig.3)。これら 2 つの写像 R, $\tilde{R}$が、関数 $F(q)$ に依らないというのは、注目すべき点である。Fig.2 は、$\epsilon=1$, $F(q)=K\sin(q)$ の場合 (標準写像といわれる) の 2 つの対称性の例である。

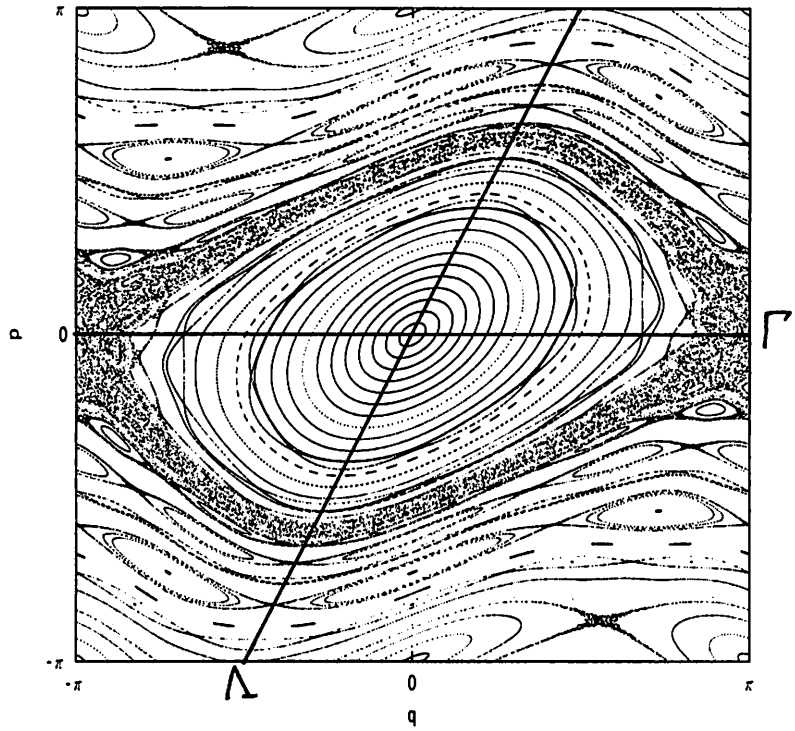


Fig.2 $F(q)=K\sin(q)$, $(K=0.9)$ の場合の 2 つの対称性
dual symmetry in case of $F(q)=K\sin(q)$, $(K=0.9)$

このように、1 つの写像系がもつ異なる 2 つの対称性を、その写像系における "2 重対称性" と呼び、2 重対称性を有する系を "2 重対称系" と呼ぶことにする。ただし、このとき 2 つの対称性を互いに『独立な』ものとするため、一般に、2 つの鏡映写像 R, $\tilde{R}$が『独立である』とは、次の条件を満たしている場合であるとする。

$$\forall n \quad \tilde{R} \neq M^n R \qquad (n: \text{整数})$$

これは、1 つの鏡映写像から生成される包合写像 I_nに対する対称性を、同質のものとみなすためのものである。もちろん、シンプレクティック写像 M における R, $\tilde{R}$ は、この条件を満たしている。

さて、ここでシンプレクティック写像MのRと$\tilde{R}$には、次の関係があることに注意した。すなわち、次のように定義される写像S

$$S \equiv R\tilde{R} \tag{7}$$

に対して、

$$S^2 = E \qquad (E:\text{恒等写像}) \tag{8}$$

であるということである(Fig.3)。このとき、

$$R\tilde{R} = \tilde{R}R = S \tag{9}$$

である。シンプレクティック写像Mにおける写像Sは、

$$S\begin{pmatrix} q \\ p \end{pmatrix} = R\tilde{R}\begin{pmatrix} q \\ p \end{pmatrix} = \begin{pmatrix} -q \\ -p \end{pmatrix}$$

であり、条件(8)を満たしている。また、写像Sの不動点は、$(q,p)=(0,0)$であり、もとの写像Mの不動点と一致している。今回は、この点も写像Sに対する条件として採用する。

次節からは、(7)で定義される写像Sが、上の条件を満たすときの一般論について論じる。

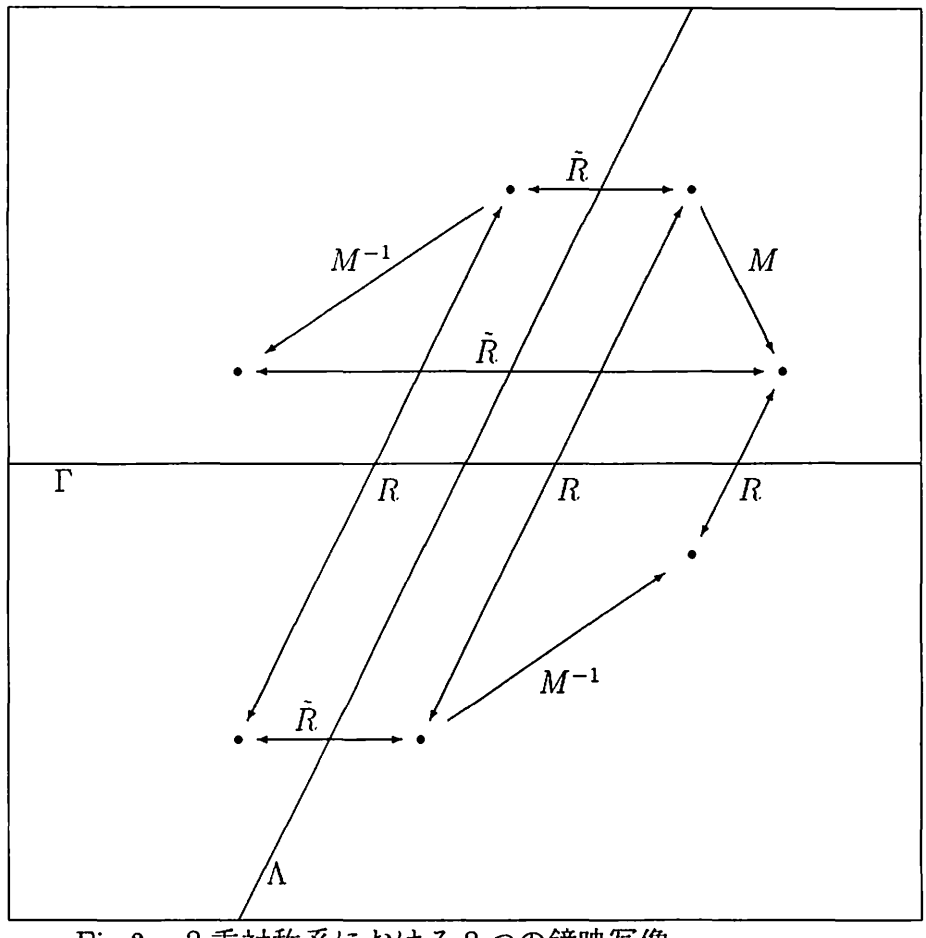


Fig.3　2重対称系における2つの鏡映写像
two kinds of involution in dual symmetric system

3　2重対称系における対称線と周期点

3-1　対称線の性質

ここでは、(7) で定義された写像 S と対称線の一般的な性質について以下に示す。

まず、2つの包合写像 $I_n \equiv M^n R,\ J_n \equiv M^n \tilde{R}$ に対するそれぞれの対称線を

$$\Gamma_i \ = \ \{x \mid I_i x = x\}, \qquad \Lambda_j \ = \ \{x \mid J_j x = x\} \tag{10}$$

とする。ここで、対称線Γ_i上の点 x $(I_i x = x)$ に対して、$Sx = y$である点 y について

$$\begin{aligned} I_i y &= I_i S x \\ &= (M^i R)(R\tilde{R})x \quad [\leftarrow (2),\ (7)\ ] \\ &= M^i \tilde{R} x \\ &= \tilde{R} M^{-i} x \quad [\leftarrow (2),\ (3)\ ] \\ &= (SR) M^{-i} x \quad [\leftarrow (9)\ ] \\ &= S I_i x \quad [\leftarrow (3)\ ] \\ &= Sx \qquad = \ y \end{aligned}$$

より、点 y は対称線Γ_i上の点であり、また対称線Λ_jについても同様のことが示せる。よって、次のことが言える。

SS1)　すべての $i,\ j$ に対して、対称線$\Gamma_i,\ \Lambda_j$は写像 S の不変集合である。

さらに、対称線Γ_i上の点 x に対して、点 y を $Rx = y$である点とし、これは S2) より対称線Γ_{-i}上の点である。このとき、$\tilde{R}x = y'$である点 y' について

$$\begin{aligned} y' &= \tilde{R} x \\ &= SRx \quad [\leftarrow (9)\ ] \\ &= Sy \end{aligned}$$

であるが、SS1) より点 y が写像 S によって写された点は対称線Γ_{-i}上の点であるので、点 y'は対称線Γ_{-i}上の点であることが分かる。また、対称線Λ_jと写像 R についても同様である。よって、次のことが言える。

SS2)　任意の i に対して、対称線Γ_iは写像 $\tilde{R}$ によって対称線Γ_{-i}に、また 任意の j に対して、対称線Λ_jは写像 R によって対称線Λ_{-j}に写される。

SS2) において、特に $i = j = 0$ のときから、次のことが言える。

SS3)　対称線Γ_0は写像 $\tilde{R}$ の、対称線Λ_0は写像 R の不変集合である。

また、さらに包合写像 I_n, J_n は、

$$\begin{aligned} I_n J_n &= R M^{-n} M^n \tilde{R} \quad [\leftarrow (2), (3)] \\ &= R\tilde{R} = S \end{aligned}$$

であることから、R, $\tilde{R}$ と同様の性質をもつ。よって、SS3) より次のことが言える。

SS4)	任意の i, j に対して、対称線Γ_iは写像 J_iの、対称線Λ_jは写像 I_jの不変集合である。

3-2 対称線と周期点の関係

さて、次に 2 重対称系における対称線と周期点の関係について調べる。

まず、点 x を、対称線Γ_iと対称線Λ_jの交点とする ($I_i J_j x = x$) 。そして、点 y を、$J_i x = y$ である点とし、これは SS4) より対称線Γ_i上の点である ($I_i y = y$)。このとき、点 x について

$$\begin{aligned} x &= I_i J_j x \\ &= M^{i-j} I_j M^{j-i} J_i x \quad [\leftarrow I1)] \\ &= M^{i-j} I_i y \quad [\leftarrow I2)] \\ &= M^{i-j} y \\ &= M^{i-j} J_i x \\ &= M^{i-j} M^{i-j} J_j x \quad [\leftarrow I1)] \\ &= M^{2n} x \qquad (n = i - j) \end{aligned}$$

となり、点 x は $2n$ 回写像 M^{2n} に対して不動であり、次の性質があることが分かる。

SS5)	任意の i, j に対して、2 つの対称線Γ_iとΛ_j($i \neq j$, $i - j = n$) の交点は、写像 M の $2n$ 回写像 (M^{2n}) の不動点である。

これは、$2n$ 周期点を見つける際に有用な性質である (Fig.4)。

また、$i = j$ のとき、すなわち 2 つの対称線Γ_iとΛ_iの交点 ($I_i J_i x = x$) に関しては、

$$\begin{aligned} x &= I_i J_i x \\ &= (R M^{-i})(M^i \tilde{R}) x \\ &= R\tilde{R}x = Sx \end{aligned}$$

より、写像 S の不動点であり、これは写像 M の不動点である。

性質 SS5) に関して、2 つの対称線Γ_iとΛ_jからは、一見、偶数周期点しか“拾えない”かのように見えるが、実は、M^{2n}の不動点というのは、奇数周期点である可能

性もある。すなわち、nが、その約数に奇数をもつ場合である。この状況を、以下のように考える。

対称線Γ_iとΛ_jの交点xが$2n$周期点であるとき、さらに

$$n = kl \qquad (l:\ \text{奇数})$$

であるとする。ここで、点xがl周期点であると仮定する。このとき、S4) (**1 Introduction**) より、点xは任意の整数mに対して、対称線Γ_{i-ml}上の点でもあることになる。

ここで、$m = k$であるとき

$$\begin{aligned} i - ml &= i - kl \\ &= i - n \quad = \quad j \end{aligned}$$

より、点xは対称線Γ_j上の点でもあることになるが、点xは対称線Λ_j上の点でもあるため、対称線Γ_jとΛ_jの交点となり、それは、不動点となってしまう。よって、点xはl(奇数)周期点ではあり得ず、改めて、次のように言うことができる。

SS6)	2つの対称線Γ_iとΛ_jの交点は、常に偶数周期点である。

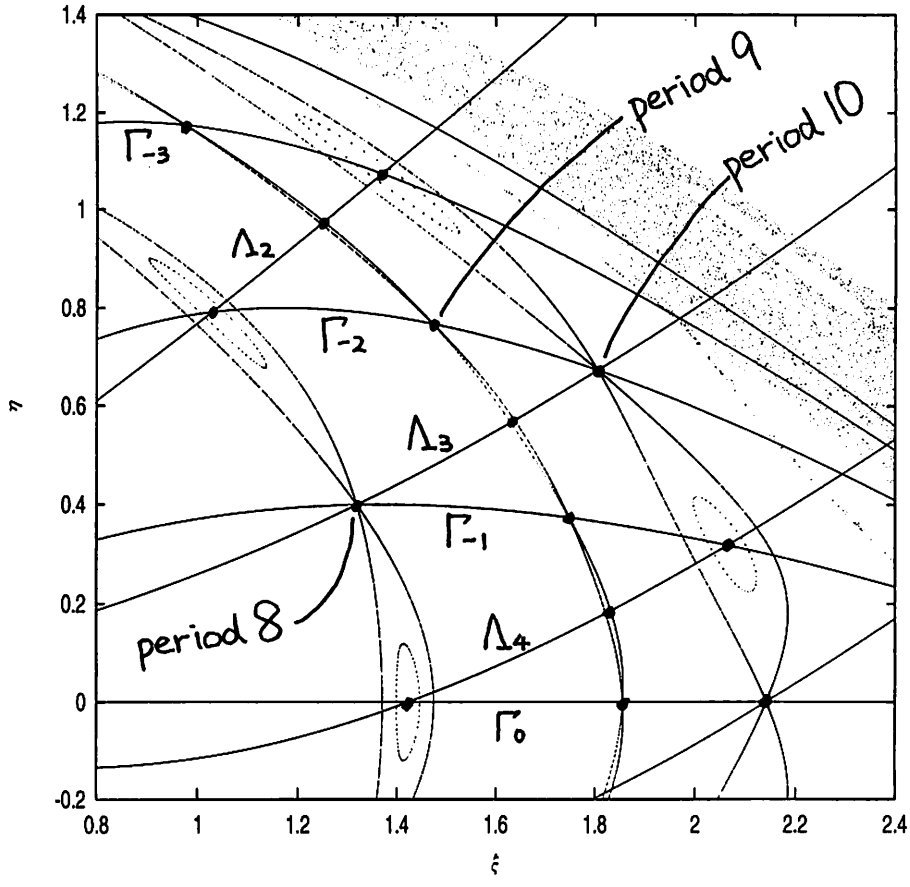


Fig.4　2重対称系における対称線と周期点
symmetry lines and periodic points in dual symmetric system

3-3　近可積分2重対称系における周期点分布

2次元可積分写像系における周期点(トーラス状に稠密に分布)は、それに摂動を加えた近可積分系においては、楕円型、双曲型の有理数周期点および後者から延びるセパラトリクスなどから成る領域に分裂する。この領域は"共鳴領域"といわれる[8]。ここでは、この近可積分系における共鳴領域に関する考察を行う。

近可積分系において、ある周期nの共鳴領域にあるn周期点は、$2n$個(楕円型と双曲型がそれぞれn個づつ)以上存在することが、ポアンカレ・バーコフの定理として知られている。この共鳴領域を、ある対称線が横切っている場合、この対称線による議論からも上の定理を理解することができる。

まず、ある対称線Γ_iがある周期nの共鳴領域にあるn周期点x上を通っているとする。このとき、S4) (**1 Introduction**)より点xは対称線Γ_{i+n}上の点でもあり、対称線Γ_iと対称線Γ_{i+n}の交点として、点xを$[\Gamma_i, \Gamma_{i+n}]$と書くことにする。また、S1) (**1 Introduction**)より、対称線Γ_iは、n回写像後、対称線Γ_{i+2n}に写されるので、点xは$[\Gamma_{i+2n}, \Gamma_{i+3n}]$でもある。これらのことから、上の共鳴領域にある周期点は、点xのほかに次の異なる周期点

$$[\Gamma_{i+1}, \Gamma_{i+1+n}],\ [\Gamma_{i+2}, \Gamma_{i+2+n}],\ \cdots,\ [\Gamma_{i+(2n-1)}, \Gamma_{i+(2n-1)+n}]$$

があることになる。よって、ある周期nの共鳴領域には、少なくとも$2n$個の周期点が存在することが分かる。

さて、ここで2重対称系における、対称性と周期点の関係について考える。まず、2重対称性を形成する2つの対称線Γ_iとΛ_jが、両方とも、ある周期nの共鳴領域を横切っているとする。ここでさらに、もし周期nが奇数であるとすると、対称線Γ_iがこの共鳴領域のあるn周期点x上を通っているとき、対称線Λ_jは、SS6) (3-2節)より点x上を通ることはできない。よって、対称線Λ_jは、対称線Γ_iとは別の周期点列

$$[\Lambda_j, \Lambda_{j+n}],\ [\Lambda_{j+1}, \Lambda_{j+1+n}],\ \cdots,\ [\Lambda_{j+(2n-1)}, \Lambda_{j+(2n-1)+n}]$$

を形成することになる。このとき、この共鳴領域内には、対称線Γ_iとΛ_jからそれぞれ異なる$2n$個の周期点が見出せることになる。よって、2重対称性をもつ近可積分系に対しては、次のことが言える。

SS7)	あるn(奇数)周期の共鳴領域内のn(奇数)周期点は、 $\underline{4n}$個以上存在する。

ここで、nが偶数であるとき、上の点x上で2つの対称線Γ_iとΛ_jは交わることが許されるので、周期点の個数の可能性は、やはり$2n$個以上である。

4　別種の2重対称系における周期点分布

4-1　別種の2重対称系

前章では、ある写像 M が2つの鏡映写像 $R,\ \tilde{R}$をもつとき、それらが、

$$(R\tilde{R})^2 = E \qquad (E:\text{恒等写像}) \tag{11}$$

であるという仮定の下に一般論を展開したが、ここで次のような場合を考える。すなわち、ある写像 M が2つの鏡映写像 $R,\ \tilde{R}$をもつとき、それらが、

$$(R\tilde{R})^2 = M \tag{12}$$

であると仮定する。このとき、実は、前章と同様な議論を行うことにより、このときの対称線と周期点との関係は次のようになることが導ける。

> 2つの対称線Γ_iと$\Lambda_j (i-j=n)$の交点は、写像 M の $(2n-1)$ 回写像 $(M^{(2n-1)})$ の不動点であり、また偶数周期点ではない。

そして、この種の2重対称性をもつ近可積分系について、次のことが言える。

> ある n(偶数) 周期の共鳴領域内の n(偶数) 周期点は、$4n$ 個以上存在する。

これらの結果は、前章で導いたものに対して、ちょうど偶数と奇数が入れ替わったものになっており、前章のものと対称的な結果となっている。一般にこのような種の2重対称系が存在するかどうかは現在のところ不明であるが、この種の2重対称系である可能性のある写像系が存在することを見出したので、次節で紹介しておく。

4-2　摂動McMillan mapにおける周期点分布

次のようなシンプレクティック写像 $M_m\ :\ (q,p) \mapsto (q',p')$

$$M_m \begin{cases} q' &= -p + \frac{2\mu q}{1+q^2} + \epsilon g(q), \qquad \mu > 1, \quad |\epsilon| \ll 1 \\ p' &= q \end{cases}$$

を考える。この写像は、摂動パラメータ $\epsilon = 0$ で可積分写像であり McMillan map といわれるものである。

この写像は、$\epsilon = 0$ で次のような2つの鏡映写像をもつ。

$$R\begin{pmatrix} q \\ p \end{pmatrix} = \begin{pmatrix} p \\ q \end{pmatrix} \qquad \text{鏡映線}\Gamma : p - q = 0$$
$$\tilde{R}\begin{pmatrix} q \\ p \end{pmatrix} = \begin{pmatrix} -p \\ -q \end{pmatrix} \qquad \text{鏡映線}\Lambda : p + q = 0$$

これらは、

$$(R\tilde{R})^2 = E \qquad (E : \text{恒等写像}) \tag{13}$$

を満たしており、3章の議論を適用できる。$\epsilon \neq 0$のとき、関数$g(q)$の偶奇性によって、系の対称性は次の2つの場合に分かれる。

$g(-q) = -g(q)$ ： 2重対称系 (鏡映写像：I, J)
$g(-q) \neq -g(q)$ ： 単一対称系 (鏡映写像：Iのみ)

ここで、この2つの場合の簡単なものとして、$g(q) = q^3$のとき (2重対称系) と、$g(q) = q^4$のとき (単一対称系) を考える。このときの数値計算結果を比較したものが、Fig.5, 6である。

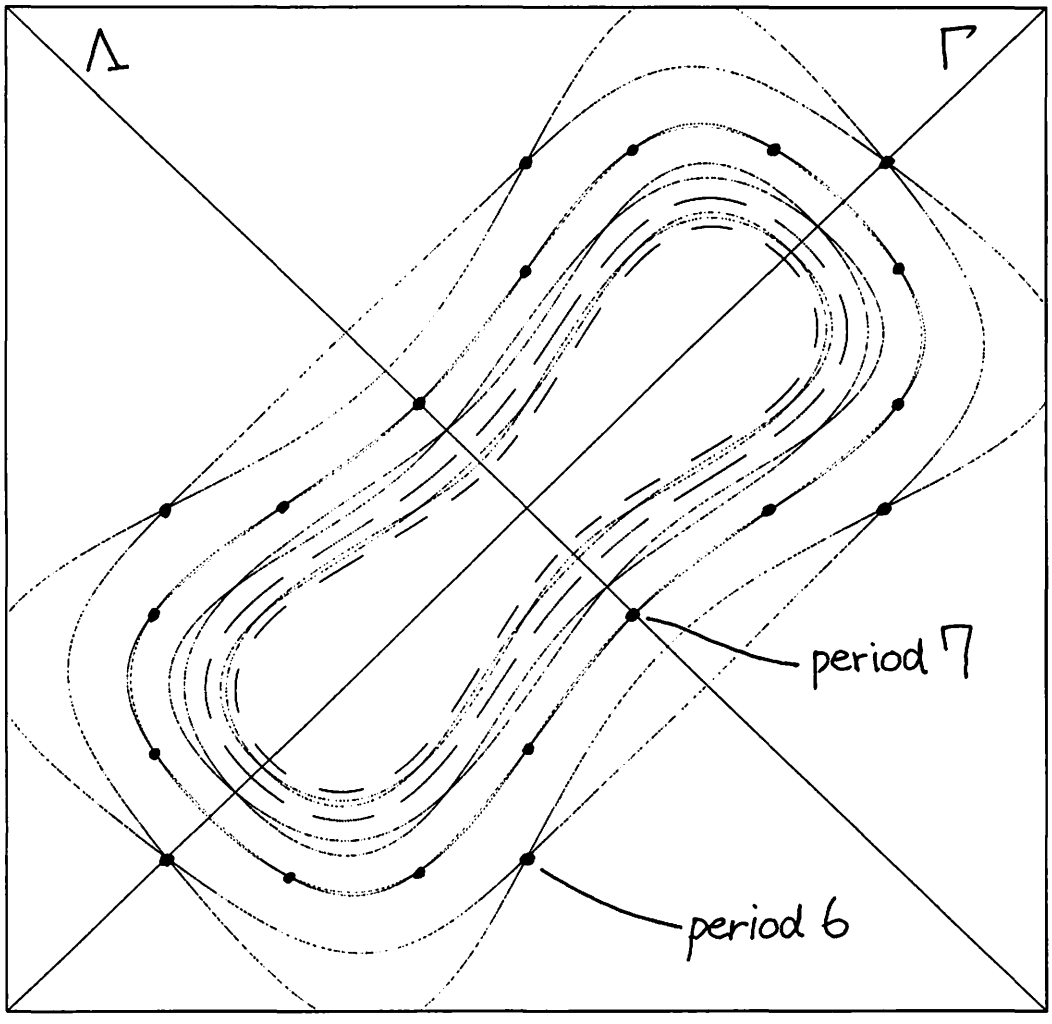


Fig.5 $g(q) = q^3$の場合の写像M_mの位相空間
phase space of the map M_m : $g(q) = q^3$

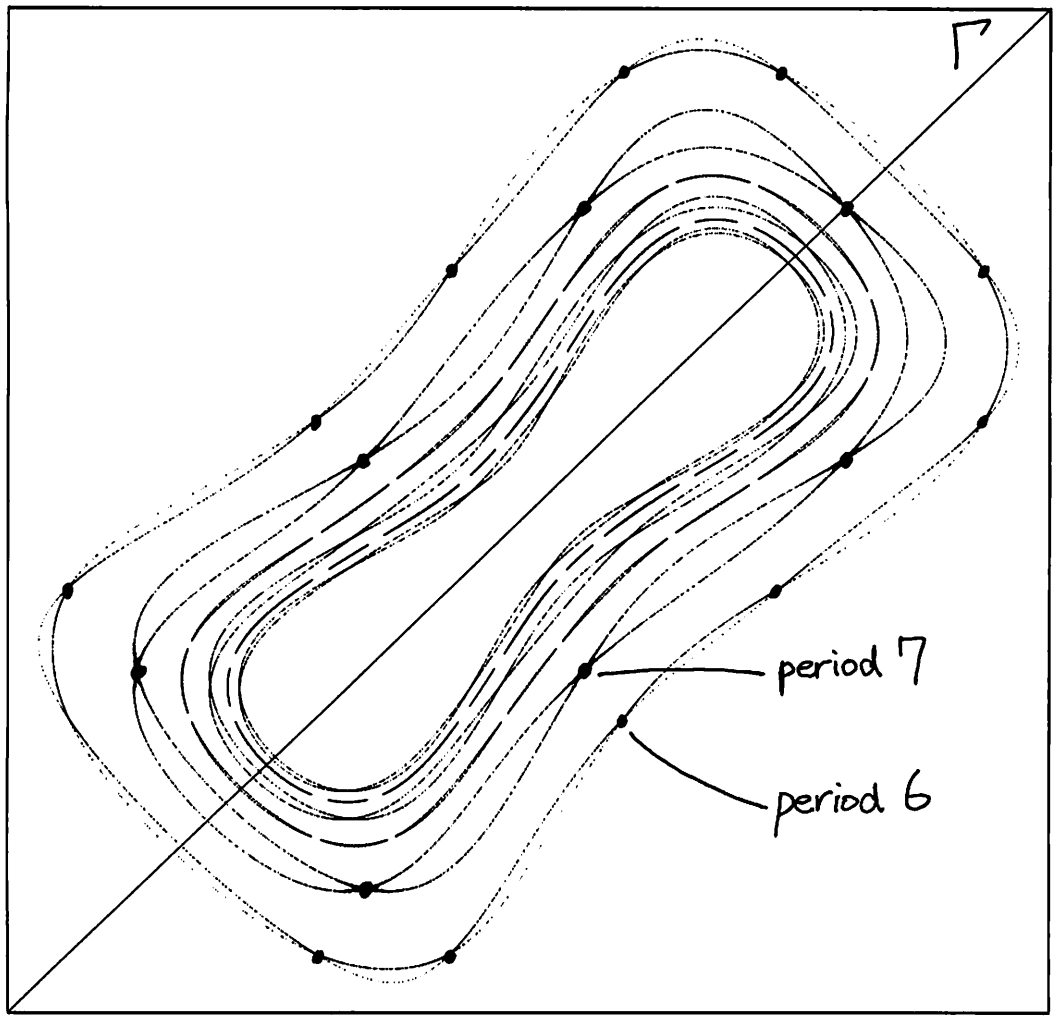


Fig.6　$g(q) = q^4$の場合の写像 M_mの位相空間
phase space of the map M_m : $g(q) = q^4$

図 には、双方とも外側から６、７、８、９、１０、１１周期の共鳴領域を描いている。

２重対称系である $g(q) = q^3$のときには n(奇数) 周期点は $4n$ 個であったが、一方の対称性が破れた $g(q) = q^4$のときには n(奇数) 周期点は $2n$ 個となっており、２重対称系における定理 SS7) の制約が無くなったことを反映している。このことについては予想内の結果となっているが、ここで注目すべきことは、逆に n(偶数) 周期点が $4n$ 個現われていることである。これは、もちろん３章の議論からは予想されるものではなく、それぞれの共鳴領域において、奇数周期点と偶数周期点はその個数の $2n$ 個と $4n$ 個という立場が入れ替わっており、位相空間の構造は大きな違いを見せている。

ここで、$4n$ 個現われた偶数周期点について、これが予想されない結果であるとはいえ、もともと、ボアンカレ・バーコフの定理 (3-3 節) によれば、近可積分系において n 周期点は $2n$ 個以上存在するとなっており、上の計算結果はもちろんこれに反するものではない。しかしながら、このような系統的な周期点分布、すなわち n(偶

数) 周期点が $4n$ 個、n(奇数) 周期点が $2n$ 個存在するというのは、なにか意味のあることのように思える。

さて、では実際になにか意味があるのだとしたら、どのような可能性がありうるのだろうか。ここで、再び対称性によって考えてみることにする。

$g(q) = q^4$のときの位相空間をその対称性によって調べようとしたとき、実は、$4n$ 個ある n(偶数) 周期点のうち対称線によって調べることができるのは楕円型不動点のみであり、双曲型不動点は対称線から洩れている。現在のところ、系のもつ対称線がすべての周期点を拾うのかどうかは不明であるが、ここで、対称線から洩れている不動点は、新たな別の対称性の存在を示唆していると考えることができるだろう。では、どのような対称性が考えられるだろうか。もし、別の対称性が存在するとしたら、この系は新たな 2 重対称系となる。ただし、奇数周期点は $2n$ 個しかないことから、第 2 章で考えたような 2 重対称系ではあり得ない。ここで 4-1 節で考えた 2 重対称系が候補として挙がることになる。

このように、4-1 節で考えた 2 重対称系は、$g(q) = q^4$のときの位相空間の状況に説明を与える有力な候補であるが、今のところ実際に (12) を満たすような鏡映写像が存在するかどうかは不明であり、これは今後の課題である。

5 議論と展望

3 章で得られた最終的な帰結は、次のようなものであった。

> 近可積分 2 重対称系において、ある n(奇数) 周期の共鳴領域内の n(奇数) 周期点は、$\underline{4n}$個以上存在する。

一方、ポアンカレ·バーコフの定理は、近可積分系において、ある n(奇数) 周期の共鳴領域にある n 周期点は、$2\,n$ 個以上存在するというものであったが、いくつかの写像における実際の位相空間は、特に本論文でみたような 2 重対称性をもたない系においては、ほとんどがその周期点の数は単に $2\,n$ 個である。このような状況において、$2\,n$ 個ではなく $4\,n$ 個以上存在しなくてはならない場合があることが、その系のもつ対称性によって系統的に見出せたことは、意味のあることである。

今後は、4 章でみた対称性のような未知の対称性が実際に存在するのか、また、今回の結果を拡張し、多次元系における多重対称性といったものを考えることで多次元位相空間における周期点分布を議論することは可能か、といったところが課題となるところである。

参考文献

[1] Hénon, H. : Quart. Appl. Math. 27 (1969), 291

[2] R. de Vogelaere : *Contributions to the Theory of Nonlinear Oscillations*, ed. S. Lefschetz (Prinston Univ. Press, Princeton, New Jersey, 1958) Vol. IV , p.53

[3] G. D. Birkoff : *Dynamical Systems* (Amer. Math. Soc. , 1927) p.186

[4] J. M. Green : J. Math. Phys. 20 (1979), 1183

[5] J. M. Green, R. S. Mackay,F. Vivaldi and M. J. Feigenbaum : Physica 3D (1981), 468

[6] Yoshi H. Ichikawa, Tetsuo Kamimura, Tadatsugu Hatori and S. Y. Kim : Prog. Theor. Phys. Suppl. No.98 (1989), 1

[7] D. del-Castillo-Negrete, J. M. Green, P. J. Morrison : Physica D91 (1996), 1

[8] R. S. Mackay, J. D. Meiss and I. C. Percival : Physica 27D (1987), 1

ねじれ写像におけるKAM曲線の崩壊

Breakup of KAM curves in Twist Mappings

山口喜博 (帝京平成大学)　　谷川清隆 (国立天文台)

Y. Yamaguchi
Teikyo Heisei University
yy-chaos@jb3.so-net.ne.jp

K. Tanikawa
National Astronomical Observatory
tanikawa@cc.nao.ac.jp

Abstract

The structure of KAM(Kolmogorov-Arnold-Moser) curve at the critical situation(a_c) that the KAM curve is destroyed is investigated in the one-parameter(a) analytical twist mappings. We give the numerical result showing that the KAM curve at the critical situation is C^1 but its derivative is not the function of bounded varitaion. The fractal distribution function of mapped points on such KAM curve is caused by the unbounded variation property. At $a > a_c$, we can not construct the reversible twist mapping including the wandering intervals constructed by Denjoy's procedure.

1　はじめに

保存力学系において可積分系に近い状態ではKAM曲線は非常に滑らかであるが, 可積分系より離れるにつれて滑らかさは徐々に失われてくると考えられている[1]. ねじれ写像においてKAM曲線が崩壊する臨界点で,KAM曲線は $C^{1+\gamma}(0 \leq \gamma < 1)$ であることが分かっている[2]-[5]. 我々はいくつかのねじれ写像 Tを利用して数値的で次の結果を得たので報告する. 詳細は[10] を見ていただきたい.

「臨界点 a_c においてKAM曲線は C^1 級であるが, その1階微分は非有界変動である.」

次にこの結果より臨界点ではKAM曲線上の写像点の分布がフラクタル構造を持つ事を示す. 左右の対称性を有する系ではDenjoy[6] の手法によって構成された遊走区間を含む円写像を構成できないことを示す.

2　円筒面上のねじれ写像

1つのパラメータ a を含む円筒面上のねじれ写像 $T(x,y)$ $[-\infty<y<\infty, 0 \leq x<2\pi]$ は次のように定義される.

$$y_{n+1} = y_n + f_a(x_n), \tag{1}$$

$$x_{n+1} = x_n + g(y_{n+1})(\mathrm{Mod}\ 2\pi) \tag{2}$$

ここで $f_a(x)$ は 2π-周期関数であり,$dg/dy > 0$ (ねじれ条件) である. $f_a(x)$ または $g(y)$ が奇関数である場合, T は reversibility[7] を持つ. $f_a(x)$ が奇関数の場合, T は左右の対称性を持つ. また g が奇関数の場合, T は上下の対称性を持つ. 更に $f_a(x)$ と $g(y)$ が共に奇関数の場合, T は上下左右の 2 重の対称性を持つ [8]. 以下の議論では次の 4 つのねじれ写像を扱う.

[TM1] 2 重の対称性を持つ Standard Mapping: $f_a(x) = -a\sin x,\ g(y) = y.$
[TM2] 左右の対称性のみをもつ Separatrix Mapping: $f_a(x) = -a\sin x,\ g(y) = 2\pi \ln \frac{1}{|y|}.$
[TM3] 上下の対称性のみを持つ写像: $f_a(x) = -a\sin x + b\cos 2x,\ g(y) = y.$
[TM4] 対称性を持たない写像: $f_a(x) = -a\sin x + b\cos 2x,\ g(y) = 2\pi \ln \frac{1}{|y|}.$

3 円写像 $x_{n+1} = G(x_n)$ の導入

相平面における不変曲線を $y = F(x)$ と記す. また, 不変曲線上の x_n と x_{n+1} の関係を $x_{n+1} = G(x_n)$ (円写像と呼ぶ) と記す.$F(x)$ と $G(x)$ の満たす関数方程式と, 両者の関係式は次のように得られる.

$$F(x + F(x) + f_a(x)) = F(x) + f_a(x), \tag{3}$$
$$G(x) + G^{-1}(x) = 2x + f_a(x), \tag{4}$$
$$G(x) = x + F(x) + f_a(x). \tag{5}$$

左右の対称性をもつねじれ写像(TM[1],TM[2]) では$G(x)$ を偶関数$G_{even}(x)$ と奇関数$G_{odd}(x)$ に分けられ, 奇関数部分は次のように得られる.

$$G_{odd}(x) = x + \frac{1}{2}f_a(x). \tag{6}$$

4 臨界点における KAM 曲線の性質

パラメタ a を a_c に近づけたとき KAM 曲線の滑らかさの変化を調べるために,$G(x)$ 並びに n 階微分 $G^{(n)}(x)$ のグラフの長さを計算する. 場合によっては $G^{(n)}_{even}(x)(n \geq 0)$ のグラフの長さを計算してもよい. ここでは Standard Mapping[TM1] の場合の結果 (Fig.1) を紹介する. 結果は次のようにまとめられる.
[1] KAM 曲線はパラメタの増大につれて滑らかさが失われるのではなく,$a = a_c$ において突然滑らかさが C^1級となる.
[2] KAM 曲線は $a = a_c$ において 1 階微分可能であるが, 1 階微分のグラフの長さは発散する. 高階微分のグラフの長さも同様に発散する.

上記の結果は他の [TM2]-[TM4] の写像においても成立していることを確かめた.[1],[2] は前に示した結果を意味している. 1 階微分が有界変動性を持たないことは, 幾何学的には関数 $G'(x)$ が無限に折りたたまれた構造をしていることを意味している. つまり $G'(x)$ がもつ構造は, Weierstrass 関数とか Takagi 関数に似ていると考えられる.

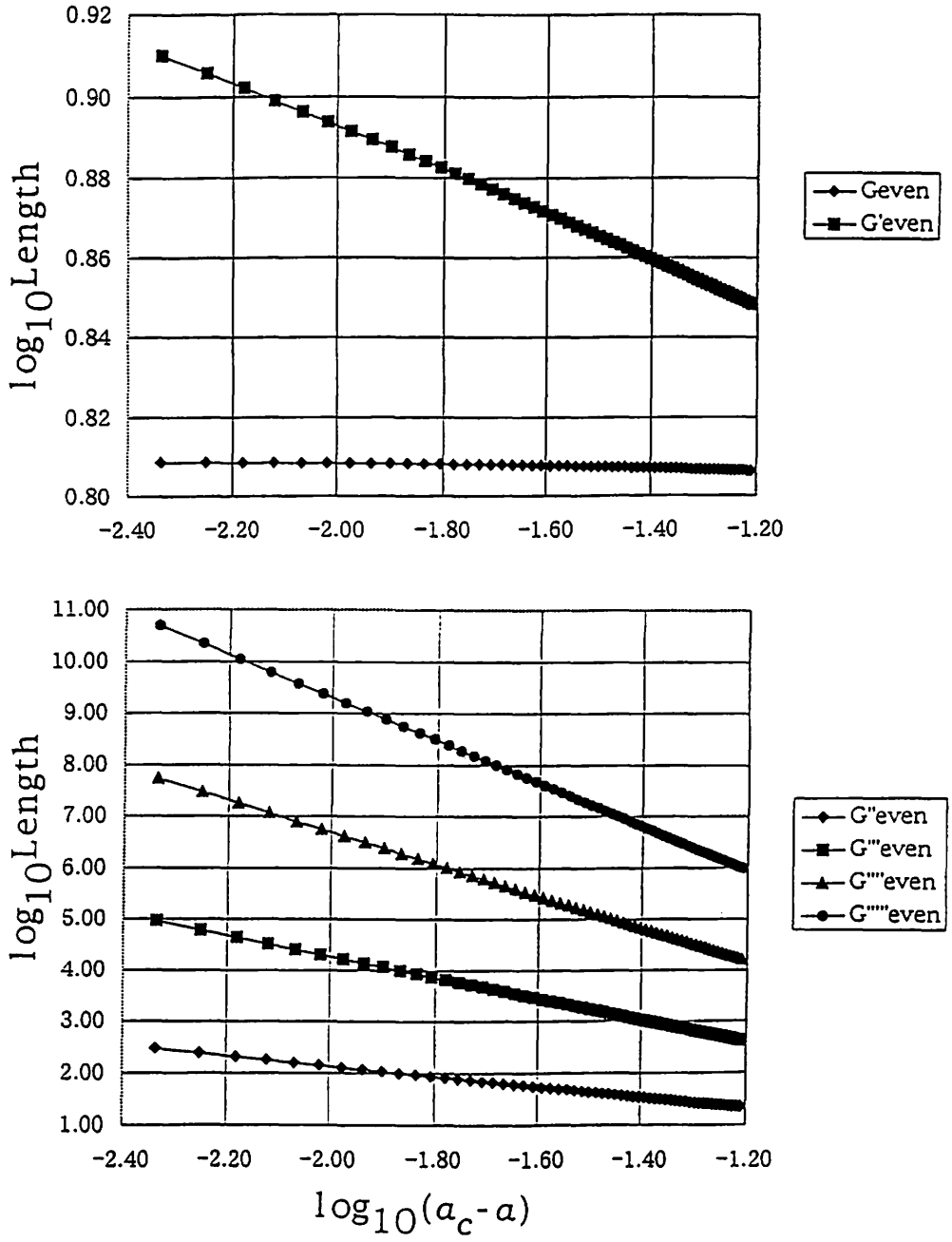


Fig.1 The lengths $L(G^{(n)}_{even}(x))(0 \leq n \leq 5)$ of graph of $G^{(n)}_{even}(x)$ for the golden mean are shown.

次に臨界点における KAM 曲線上の写像点の度数分布について議論する.KAM 曲線を記述する円写像 $G(x)$ と写像点の度数分布 $\rho(x)$ の関係は, 次の Frobenius-Perron 関数方程式で記述される.

$$\rho(x_{n+1}) = \frac{\rho(x_n)}{G'(x_n)}. \tag{7}$$

この関数方程式より $G'(x)$ の有界変動性が破れると度数分布が定義できなることが分かる.

Standard Mapping における, いくつかの回転数に対する KAM 曲線とその周りの周期点の様子と,KAM 曲線上の度数分布を Fig.2 に示してある. パラメタの値は, それぞれの KAM 曲線の壊れる臨界値に非常に近い値を設定してある. これらの図より次の事が分かる.
[1] 対称線 $x = 0$ と KAM 曲線との交点 u では度数分布が減少する. しかし他の対称線との交点では度数分布が増大する.
[2] 交点 u には楕円点が集積している.

無理数の回転数を有理数で近似したとき, 近似のしやすい回転数をもっている KAM 曲線は崩壊しやすい. また $x-y$ 平面における KAM 曲線とサドル, 楕円点との距離が近い KAM 曲線は崩壊しやすい. 数値計算の結果を見る限り, 前者の影響より後者の影響の方が強いのではないかと考えられる. これは今後の検討課題である.

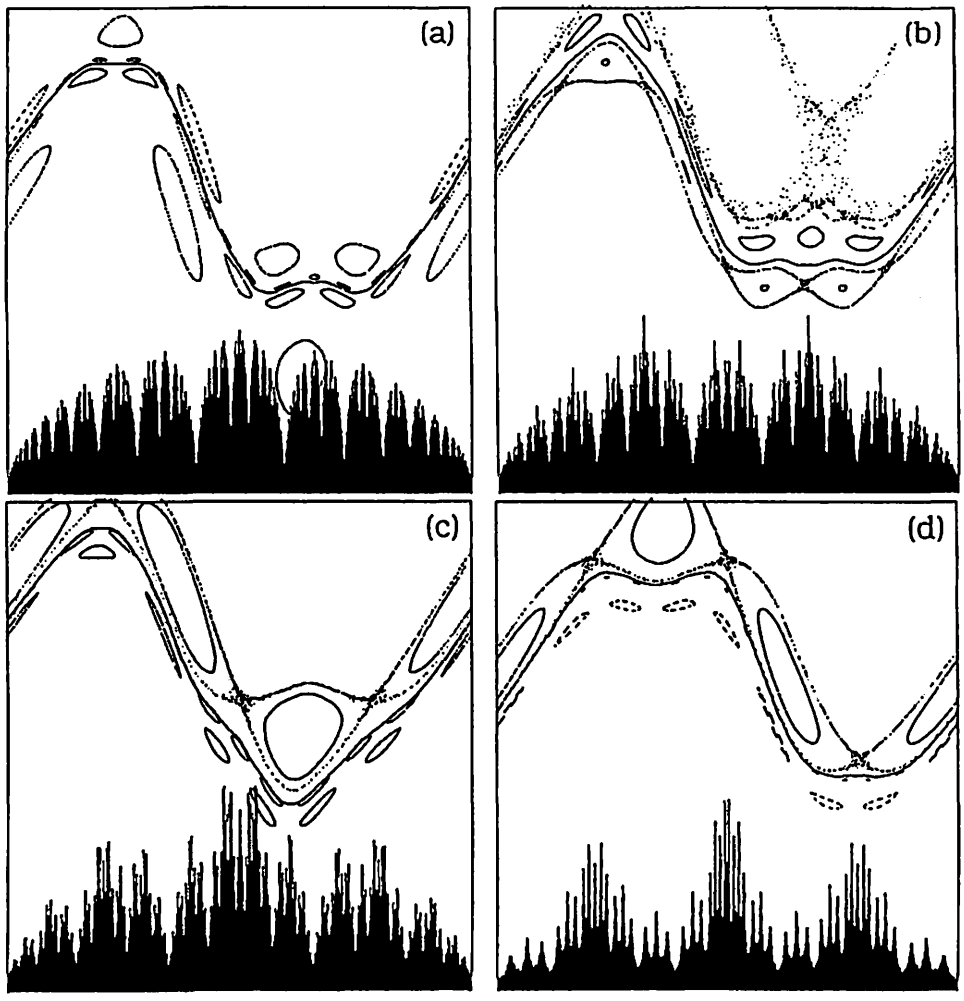


Fig.2 The distribution functions $\rho(x)$ for several rotation numbers. The phase plot around the KAM curve is shown in the upper part of each figure. The horizontal axis shows x axis($0 \leq x < 2\pi$). (a) golden mean($a = 0.97$),(b) silver mean($a = 0.888$), (c) $(\sqrt{13} - 3)/3(a = 0.888)$, and (d) $\sqrt{7} - 2(a = 0.875)$.

5 円写像における遊走区間の非存在

ここでは左右の対称性をもつねじれ写像を考える.

補題：円写像 $G(x)$ のグラフで次の性質を持つ 2 点 $P=(x_-,y_-), Q=(x_+,y_+)$ を選ぶ.
(1) P と Q は $y=-x(\mathbf{Mod}\ 2\pi)$ について対称な位置にある.

(2) $x_+=G(x_-)$.

このとき, 円写像面上の P,Q を通過する軌道は, $x-y$ 平面では $x=0$ と KAM の交点 u を通過する軌道であるか, または $x=\pi$ と KAM の交点 v を通過する軌道である. また逆に $x-y$平面で $x=0$ と KAM の交点 u を通過する軌道と,$x=\pi$ と KAM の交点 v を通過する軌道は円写像面上で $y=-x(\mathbf{Mod}\ 2\pi)$ について対称な軌道である.

証明：$y=-x$ に対称な場合は,$x_-=-y_+, y_-=-x_+$ である.$y_-=G(x_-)$ であるから, $y_-=x_+$. つまり $x_+=0$. これは $x-y$ 平面において $x=0$ と KAM の交点 u を通過する軌道であることを意味している.$y=-x+2\pi$に対称な場合は, $x=\pi, y=\pi$を新しい原点とする座標 (x',y') で表示を書き直すと前の場合と同じ式:$x'_+=0$ を得る. つまり $x_+=\pi$ を得る. これは $x-y$ 平面において $x=\pi$ と KAM の交点 v を通過する軌道であることを意味している. 逆は, Reversibility から自明.(QED)

定理：$y=-x(\mathbf{Mod}\ 2\pi)$ に対称な円写像 $G(x)$ に対して, 次の 2 つの性質を持つ新しい円写像を構成できない.
(1) Denjoy の手法によって構成された遊走区間を持つ.
(2) $y=-x(\mathbf{Mod}\ 2\pi)$ に対称.

証明：補題により $y=-x(\mathbf{Mod}\ 2\pi)$ に対称な軌道として,$x-y$平面では u を初期点とするものを選ぶ. u に対応する円写像の軌道は $x=0$ を初期点 x_0 とするものである. 円写像 G のグラフを $G^n(0)(n=\pm1,\pm2,\ldots)$ の位置で切れ目を入れて分割する. 分割された弧を $L_k(k=0,\pm1,\ldots)$ と記す. $L_k(k\geq 0)$ と L_{k+1} の間に遊走区間 l_k を挿入し, $L_{-k}(k\geq 0)$ と $L_{-(k+1)}$ の間に遊走区間 l_{-k} を挿入する. ただし L_k と L_{-k} が, かつ l_k と l_{-k} が $y=-x$ に対称になるように配置する. また遊走区間 l_k の長さは例えば $(\lim_{|k|\to\infty} l_k/l_{k+1}=1, \sum_{k=-\infty}^{\infty} l_k<\infty)$

$$l_{\pm k}=\frac{1}{(|k|+2)(|k|+3)}$$

とする. また端点では 1 回微分が連続であるように接続を行う. 以上の手続き行った結果が Fig.3 である. L_0 を未来へ写像したものが $L_k(k\geq 1)$ に, L_0 を過去へ写像したものが $L_{-k}(k\geq 1)$ になっている. l_1を未来へ写像したものが $l_k(k\geq 2)$ に, l_{-1} を過去へ写像したものが $l_{-k}(k\geq 2)$ になっている. しかし L_0 を未来へ写像した像が l_1 を含み, l_{-1} を未来へ写像した像が L_0 に含まれる. つまりこの構成によって得られた Fig.3 には矛盾がある.(QED)

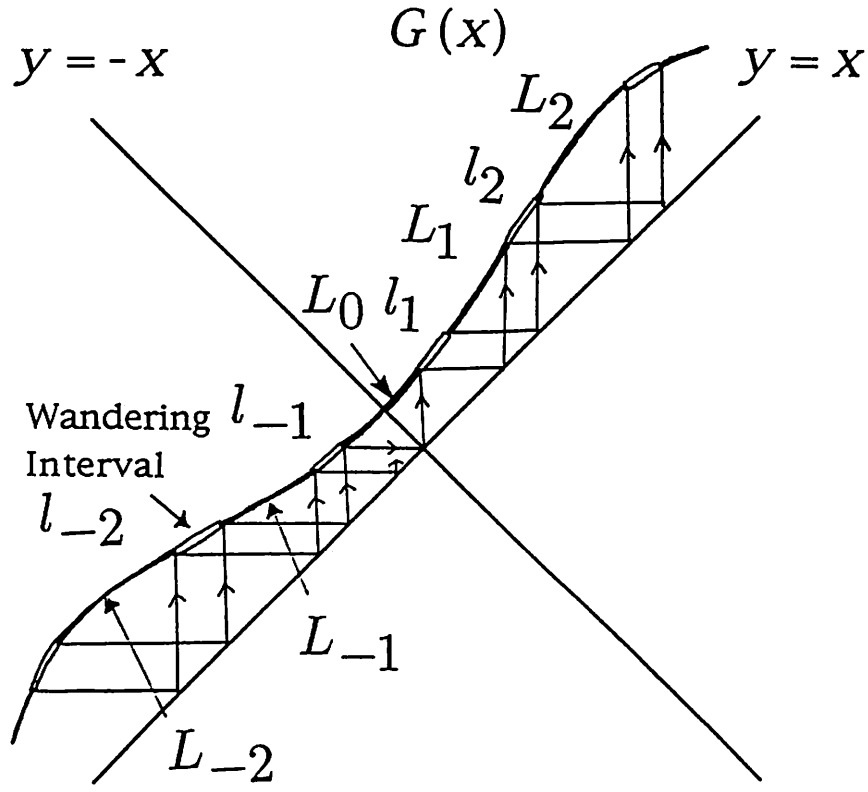


Fig.3 The construction of circle mapping including the wandering intervals $l_i(i = \pm 1, \pm 2, \cdots)$.

$a = a_c$において生じた「穴」が, $a > a_c$ では「gap」になると考えられる. 一方 Aubry-Mather 集合は Lipschitz 曲線上に存在する. 定理より,gap はこの Lipschitz 曲線上に拘束されないことが分かる. 臨界状態では Birkhoff-サドルの安定多様体, 不安定多様体が KAM 曲線に漸近している. 漸近していたこれらの多様体は gap を通過することになる. このとき生じる多様体の構造変化も今後の課題である.

参考文献

[1] J. Moser, *SIAM REVIEW* **28** (1986),459.

[2] Y.Katznelson and D.Ornstein, *J.Analyse* **60** (1989),157.

[3] S.Stirnemann, *Commun. Math. Phys.***152** (1993),369.

[4] A.H.Osbaldestin and M.Y.Sarkis, *J.Phys.A:Math.Gen.* **20** (1987), L953.

[5] B.R.Hunt, K.M.Khanin, Y.G.Sinai, and J.A.Yorke,*J.Stat.Phys.* **85** (1996),261.

[6] A.Denjoy, *J.Math.Pures Appl.*(9) **11** (1932),333.

[7] R.DeVogelaere, in *Contribution to the theory of nonlinear oscillation*, Vol.IV ed.by S.Lefschetz. pp.53-84(1958).

[8] K.Tanikawa and Y.Yamaguchi, *J.Math.Phys.* **28** (1987),921.**30** (1989),608.

[9] S.Aubry, *Physica*,**7D** (1983),240. J.N.Mather, *Publ.Math.I.H.E.S.*,**63** (1986),153.

[10] Y.Yamaguchi and K.Tanikawa, *Prog. Theor. Phys.*,**101** (1999),1.

力学系の再構成におけるカオスとノイズ

湯浅　学, 孫野　成子

Effects of Chaos and Noise in the Reproduction of Dynamical Systems

Manabu Yuasa
Research Institute for Science & Technology,
Kinki University, Higashi-Osaka 577,Japan
Shigeko Magono
Department of Information Science, Faculty of Science,
Nara Women's University, Nara 630, Japan

Abstract

Based on the principal component analysis, reproduction of Hénon-Heiles' dynamical system is performed using the various data sets which are prepared by the numerical integration of the original differential equations with different initial conditions. The accuracy of the reproduction seems to be strongly correlative with the chaotic motion. The influence of the white noise introduced into the data upon the accuracy of the reproduction is also investigated.

Key Words: Hénon-Heiles' dynamical system, Principal component analysis, Chaos, White noise

1 Introduction

いくつかの物理量についての観測データが多数の異なる時間(time)又は場所(site)あるいはその両方について得られているとする。それらの個々のデータがなぜそのような値を取るか、物理量の間の依存関係はどうか、あるいは、データの変化をもたらしている主たる原因は何であるか等を明らかにする為には、これらのデータから、物理量の間に成り立つ微分方程式系を導くことが出来れば非常に見通しが明るくなる。Unno(1995)はこの目的のために主成分解析法を繰り返し用いることによりデータ群から微分方程式系を構成する新しい力学系構成法を提示した。その後、この新しい方法は、数理経済学(Unno et al. 1996;Yuasa,Unno 1996;Yuasa et al. 1997)や、生物科学(Unno et al. 1997)、そして天文学(Yuasa et al. 1998)の問題に応用され、いくつかの新しい重要な結果が得られている。又、この方法の正当性を裏付けるために、微分方程式系で与えられた力学系を数値積分することによってデータを作成し、作成したデータから元の力学系を再現することも試みられている(Yuasa, Magono 1997)。その結果、我々はHénon-Heilesの力学系について、1つの初期値について元の力学系を記述する微分方程式系の係数を3桁の精度で再現することに成功

した。Hénon-Heilesの力学系は、全エネルギーを保存し、軌道のふるまいがchaoticとなるエネルギーのcritical valueが存在する。この論文では、種々のエネルギー値について同じ方法で力学系の再構成を行い構成した微分方程式系の係数の決まり方の精度がエネルギーと共にどのように変化するかを調べた。その結果、エネルギーの増加と共に係数の決まり方の精度は悪くなり軌道がchaoticなふるまいをする傾向と、係数の決まり方の精度の悪くなる傾向に、大きな類似性のあることを見出した。また、元の微分方程式系を数値積分することによって作成したデータに、white noiseを加えた場合に、noiseの大きさと共に係数の決まり方がどのように乱されるかも調べた。

2 Hénon-Heiles系

HénonとHeiles(Hénon,Heiles 1964)は、銀河系における恒星の運動を扱う恒星系力学を研究していた。そして恒星系の運動における、エネルギー積分、角運動量積分以外の第3積分を捜していたが、問題を単純化して次の微分方程式系で与えられる自由度2の4次元の力学系を考えた。

$$\frac{dq_1}{dt} = \frac{\partial H}{\partial p_1} = p_1, \tag{1}$$

$$\frac{dp_1}{dt} = -\frac{\partial H}{\partial q_1} = -q_1 - 2q_1q_2, \tag{2}$$

$$\frac{dq_2}{dt} = \frac{\partial H}{\partial p_2} = p_2, \tag{3}$$

$$\frac{dp_2}{dt} = -\frac{\partial H}{\partial q_2} = -q_2 - q_1^2 + q_2^2, \tag{4}$$

ただし、

$$H = \frac{1}{2}(p_1^2 + p_2^2) + \frac{1}{2}(q_1^2 + q_2^2 + 2q_1^2q_2 - \frac{2}{3}q_2^3). \tag{5}$$

である。

ハミルトニアンHは、陽に時間tを含まないので、時間の変化に対してその値Eを保存する(エネルギー積分の存在)。また、この系は、$E < \frac{1}{6}$の時、座標q_1とq_2は$(q_1, q_2) = (0, 1), (\pm\frac{\sqrt{3}}{2}, -\frac{1}{2})$で表される3点のうちの2点を通る3直線で囲まれる三角形領域の内部に限られる。この力学系は、比較的簡単な微分方程式系で書けているが、その軌道のふるまいはエネルギーの増加と共に複雑になりchaotic motionとなる。HénonとHeilesはこの系の軌道をPoincaréのsurface of section法によって解析した。即ち$q_1 = 0$で$p_1 > 0$である点が$q_2 - p_2$平面上で時間と共に次々とどのように写像されていくか(Poincaré mapping)を調べたのである。

その結果、全エネルギーEが$E \leq \frac{1}{9}$の場合と$\frac{1}{9} < E < \frac{1}{6}$の場合でPoincaré写像が定性的に大きく異なる事を発見した。$E \leq \frac{1}{9}$の場合には図1 (a)のようにPoincaré写像の点列は，種々の初期条件についてすべてある1つの滑らかな曲線上に乗る。一方、$\frac{1}{9} < E < \frac{1}{6}$の場合には、図1(b)のように初期条件によっては滑らかな曲線上に乗るものもあるが、滑らかな曲線上に整然と乗るのではなく規則性の見られないランダムな点として分布する初期条件のものが出現する。

図1 (b)で散らばった点は1つの初期条件から出発したPoincaré写像である。この散らばった点となる領域は、エネルギーEの増加と共に増加する。HénonとHeiles

は、エネルギーEを横軸にとり、縦軸に滑らかな曲線で満たされる領域の全領域($p_2^2+q_2^2-\frac{2}{3}q_2^3<2E$)を満たす領域) に対する相対領域をとり図2を得た。当時(1964年)Hénon と Heiles は、chaos という概念にまでは至らなかったが、散らばった点を与える軌道はchaos 軌道である。エネルギーEの値が$\frac{1}{9}$を境として、chaos 軌道が出現し、エネルギーの増加と共に、chaotic motion となる初期値が増加しているのである。

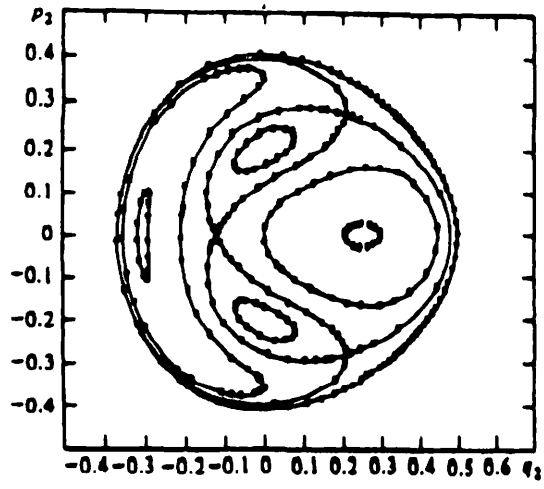


図1(a) Poincaré mapping (E=0.083), Hénon & Heiles 1964

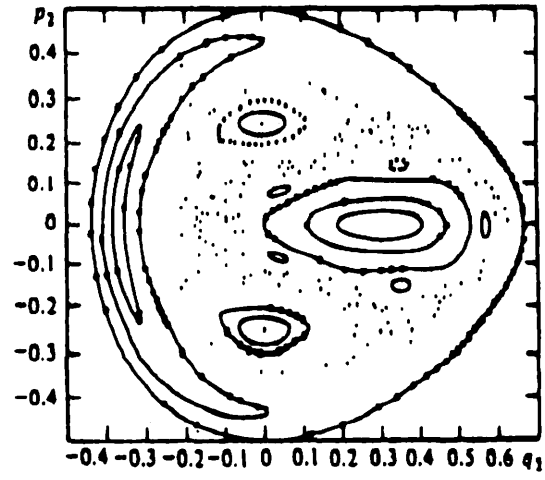


図1(b) Poincaré mapping (E=0.125), Hénon & Heiles 1964

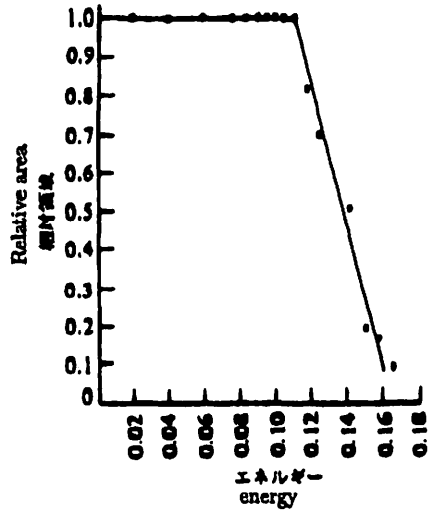


図2 滑らかな曲線で満たされた領域の割合のエネルギーによる変化
Relative area covered by the smooth curves as a function of energy, Hénon & Heiles 1964

3　力学系の再構成とChaos

我々は既に昨年Hénon-Heilesの力学系を、初期条件が $q_1(0) = 0.1$, $p_1(0) = -0.3$, $q_2(0) = 0.3$, $p_2(0) = 0.1$, E $= 0.0940$ の場合について、主成分解析法を用いた新しい力学系構築法により係数を3桁の精度で再現することに成功している(Yuasa, Magono 1997)。今回は同じ解析法で，$q_1, p_1, q_2, p_2, q_1^2, q_2^2, q_1q_2$ を7次元空間に埋め込みデータのアドレス数を３３２に固定して、エネルギー値だけをパラメーターにして変化させ、係数の決定精度を調べた。８個のエネルギー値について数値積分によりデータを作成し解析を行ったが、各エネルギー値に対して採用した初期条件を表1に示す。

表1 作成データの初期値とエネルギー値

データ番号	$q_1(0)$	$p_1(0)$	$q_2(0)$	$p_2(0)$	E
1	0.1	−0.3	0.3	0.1	0.094
2	0.1	−0.35	0.3	0.1	0.110
3	0.1	−0.36	0.3	0.2	0.128
4	0.1	−0.4	0.3	0.2	0.144
5	0.1	−0.42	0.3	0.2	0.152
6	0.1	−0.435	0.3	0.2	0.159
7	0.1	−0.45	0.3	0.2	0.165
8	0.1	−0.453	0.3	0.2	0.167

それぞれの初期条件から作成したデータ群を解析して再構成された方程式(1) ~ (4)に対応する微分方程式系の右辺は、$c_0 \sim c_7$を定数として、

$$c_0 + c_1q_1 + c_2p_1 + c_3q_2 + c_4p_2 + c_5q_1^2 + c_6q_2^2 + c_7q_1q_2 \tag{6}$$

の形で得られる。一方、元の微分方程式(1) ~ (4)の右辺を、$a_0 \sim a_7$を定数として(ただし$a_0 = 0$)

$$a_0 + a_1q_1 + a_2p_1 + a_3q_2 + a_4p_4 + a_5q_1^2 + a_6q_2^2 + a_7q_1q_2 \tag{7}$$

とする。

ここで、微分方程式系の再構成の精度の目安を与える量として、係数誤差の和

$$c = \sqrt{\sum_{i=0}^{7} \Delta c_i^2} = \sqrt{\sum_{i=0}^{7} (c_i - a_i)^2} \tag{8}$$

を８種の各エネルギー値について作成したデータから再構成した４個のそれぞれの微分方程式について計算した。この結果を、エネルギー値を横軸にとり、cを縦軸にとって図3 (a)に示した。図3 (a)を見ると、dp_2/dtの右端の２つのケースを除いて、係数誤差の和cはエネルギーと共に増加し、その増加は、カオスの出現するcritical energy　$E = 1/9 = 0.11$付近から始まっている事がわかる。

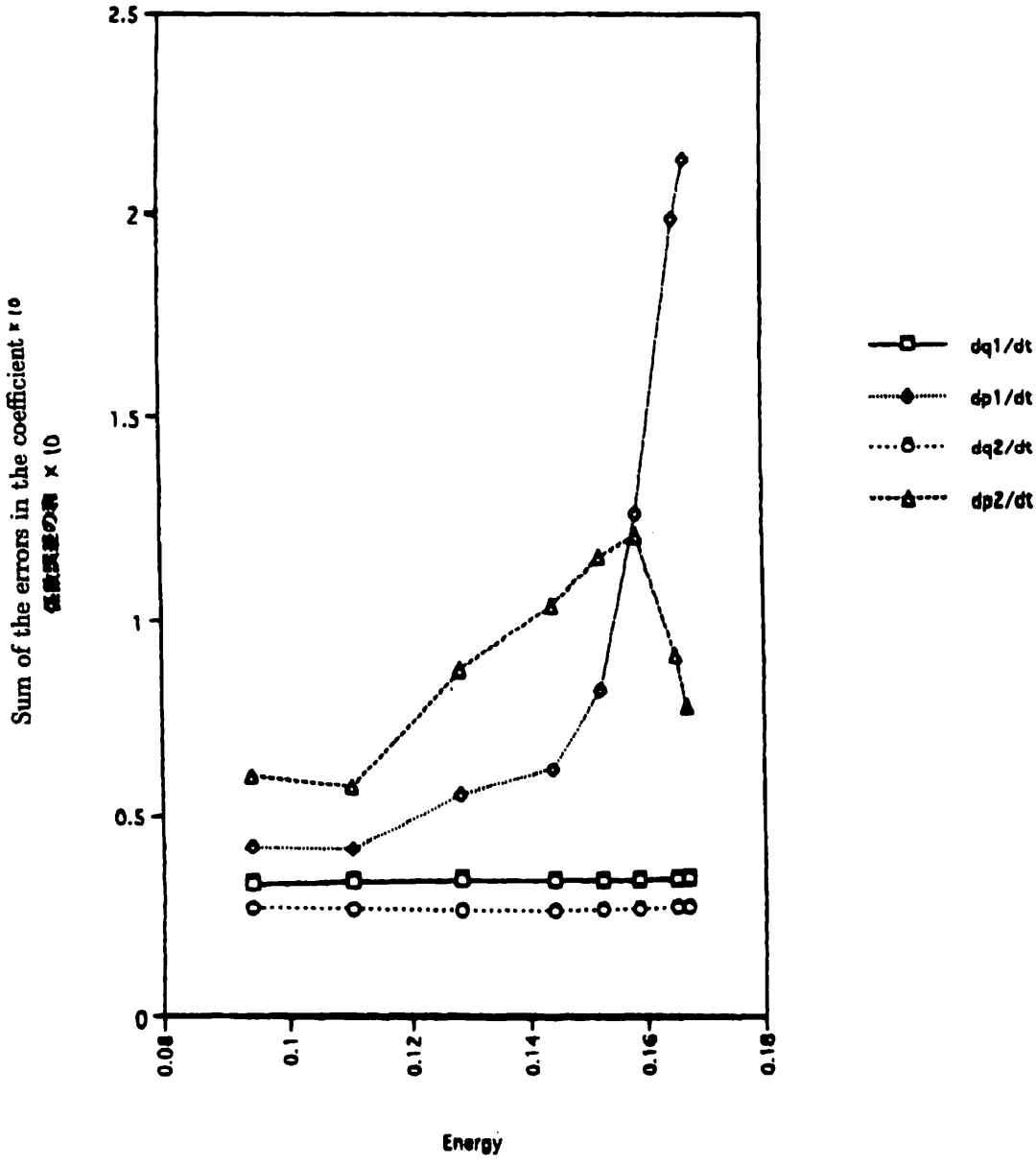


図3(a) (1) ～ (4) の各微分方程式の係数誤差の和のエネルギーによる変化

Sum of the errors in the coefficient of each reproduced equation, as a function of energy

さらに、同一データから再構成した$dq_1/dt, dp_1/dt, dq_2/dt, dp_2/dt$の４個の微分方程式は連立系なので同一初期条件から再構成した４個の微分方程式のｃをすべて加えたものが系全体としての再構成の精度を与える。そこで、４個のｃの和を縦軸にとったものが図３(b)である。図３(b)は、Hénon と Heiles によって、Poincaré 写像が散らばった点列とならないで滑らかな曲線上に乗る相対面積とエネルギーの関係、として描かれた図(図２)と裏返しの関係でよく一致していることがわかる。図３(b)においても、係数誤差の和の合計$\sum c$も$E = 1/9 = 0.11$付近から増加しはじめていることがわかる。従って、我々の採用している主成分解析法を用いた新しい力学系構成法においても、chaos の出現と共に係数の決定精度が低下し、再構成した係数の誤差と chaotic motion との間には明らかな相関関係が存在することがわかる。

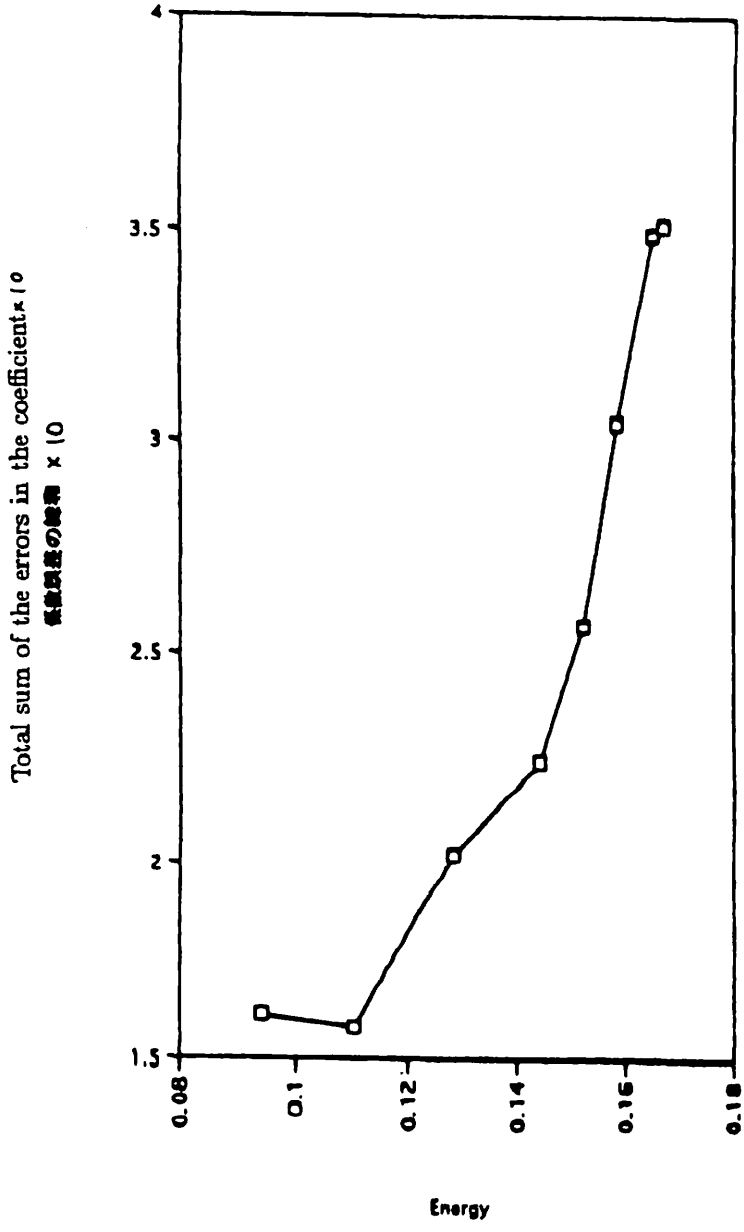


図3(b) 系全体の係数誤差の総和のエネルギーによる変化
Total sum of the errors in the coefficient of the reproduced 4 equations, as a function of energy

4　力学系の再構成とNoise

観測から得られたデータは、必然的に観測誤差を含む。一方ここで扱っているデータは、元の微分方程式から作成したもので、全エネルギーが１０桁の精度で保存されている。従って使用したデータ $q_1(t), p_1(t), q_2(t), p_2(t)$ もほぼ１０桁近くまで正しいと考えられる。このデータにwhite noiseを加えた場合に、再構成した微分方程式系の係数誤差がどのように変化するかを調べた。データ $q_1(t), p_1(t), q_2(t), p_2(t)$ は、10^{-1} 程度の量であるが、これに $10^{-5}, 10^{-4}, 10^{-3}$ のwhite noiseをかぶせたデータからそれぞれ力学系の再構成を行った。図４(a),(b)は、$E = 0.1103$ と $E = 0.144$ についての dp_1/dt と dp_2/dt についての解析結果である。10^{-3} のnoiseを加えると、急激に精度が落ちることが明らかになった。

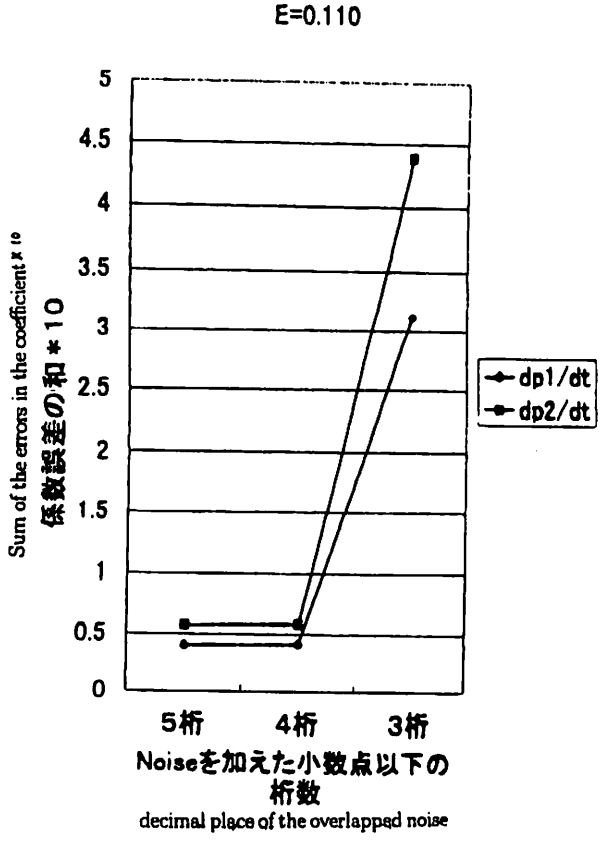


図4(a) dp_1/dt と dp_2/dt の係数誤差の和とNoiseを加えた小数点以下の桁数 (E=0.110)
Sum of the errors in the coefficient of the reproduced equation, as a function of decimal place of the overlapped noise (E=0.110).

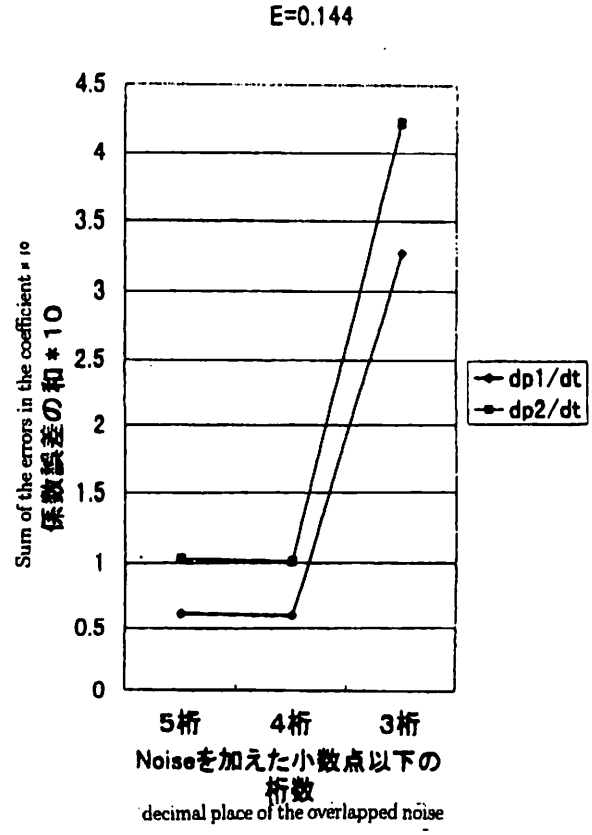


図4(b) dp_1/dt と dp_2/dt の係数誤差の和とNoiseを加えた小数点以下の桁数 (E=0.144)
Sum of the errors in the coefficient of the reproduced equation, as a function of decimal place of the overlapped noise (E=0.144).

5 Discussion

力学系の再構成と軌道のchaotic motionの間に明白な相関関係が存在すること、及び主成分解析法に基づく我々の方法において元の微分方程式系を正確に再現するためには、観測データは、少なくとも有効数字3桁の精度を持つことが望ましいことが明らかになった。chaotic motionとの関連については、更にLorenzの力学系のような他のchaos系について同様な傾向が見られるかどうか調べることが今後の課題である。

Noiseを加えた場合の精度については、有効数字3桁というのは、今回扱ったように、微分方程式系をrigorousに再現する場合に観測データに要求される精度であり、力学系(データ群)の定性的な性質をさぐる解析においては、もう少し低い精度でも十分データ群の特徴を抽出することが出来るものと考えられる。

References

(1) Hénon,M.,Heiles,C. 1964, A.J 69,73
(2) Unno,W. 1995,Science & Technology, Kinki Univ. 7,1
(3) Unno,W.,Yuasa,M.,Onishi,T. 1996,Science & Technology, Kinki Univ. 8,61
(4) Yuasa,M.,Unno,W. 1996,Science & Technology, Kinki Univ.8,67
(5) Yuasa,M.,Unno,W., Onishi,T. 1997,Science & Technology, Kinki Univ. 9,19
(6) Unno W., Onishi,T., Yuasa M., 1997,Viva Origino 25,139
(7) Yuasa,M.,Magono,S. 1998,Science & Technology, Kinki Univ. 10,21
(8) Yuasa,M.,Unno,W.,Magono,S. 1998,"Distance Determination of Mass-losing Stars", accepted for publication in PASJ 51

Method of constructing multi-dimensional exactly solvable chaos

Ken Umeno[1]
Communications Research Laboratory
Ministry of Posts and Telecommunications
4-2-1 Nukui-Kitamachi, Koganei, Tokyo 184-8795, Japan

Abstract

Multi-dimensional ergodic dynamical systems with absolutely continuous invariant probability measures are explicitly constructed by successive use of skew-product transformations of parametric families of one-dimensional ergodic transformations.

1 Introduction

It was over 50 years ago that S. Ulam and J. von Neumann [1] discovered that the one-dimensional logistic map $X_{n+1} = 4X_n(1 - X_n)$ is ergodic with respect to the invariant measure

$$\mu(dx) = \frac{dx}{\pi\sqrt{x(1-x)}} \tag{1}$$

over the unit interval $[0, 1]$. Since such ergodic dynamical systems with explicit invariant measures can offer us good examples of constructively explaining the microscopic origin of statistical mechanics, it is of great interest to investigate the following problem: To what a degree is the extension of the analytic and explicit result by Ulam and von Neumann made to higher-dimensional dynamical systems? As for analytical investigation of high-dimensional chaotic systems, so-called *the curse of dimensionality* seems to exist. The purpose of the present paper is to show that we can construct high-dimensional chaotic dynamical systems whose statistical characteristics (absolutely continuous invariant measures) are analytically obtained. Such construction of high-dimensional maps with explicit invariant measures which are absolutely continuous with respect to the Lebesgue measure is the first one as far as the author knows. The key for generating such systems is the use of *skew product* transformations of ergodic maps, which will be explained in the subsequent sections.

2 Skew Product Transformations

Let us consider a dynamical system $x_{n+1} = f(x_n)$ on the phase space X and a parametric family of dynamical systems $\{y_{n+1} = g_\lambda(y_n)\}_{\lambda \in X}$ on the phase space Y. We assume that f and g_λ are ergodic with respect to the invariant measures $\sigma(x)dx$ and

[1]E-mail:umeno@crl.go.jp

$\nu_\lambda(y)dy$ respectively. A skew product of f and $\{g_\lambda\}_{\lambda\in X}$ is given by a transformation $h : X \times Y \to X \times Y$ defined by

$$h(x,y) = (f(x), g_{f(x)}(y)) \quad \text{for} \quad x \in X, y \in Y. \tag{2}$$

This dynamical system $z_{n+1} = h(z_n), z_n = (x_n, y_n) \in X \times Y (\equiv \Omega)$ has the following remarkable property:

Theorem 1 *The absolutely continuous measure on* $X \times Y$

$$\rho(x,y)dxdy \equiv \nu_x(y)\sigma(x)dxdy \tag{3}$$

is invariant under the transformation $h(x,y)$.

(Proof of Theorem 1)

From the assumption that $\sigma(x)dx$ is the invariant measure of f and $\nu_\lambda(y)dy$ is the invariant measure of $g_\lambda(y)$, we have the Perron-Frobenius(probability preservation) relations:

$$\sigma(x) = \sum_{f(x')=x} \sigma(x')/|\frac{df(u)}{du}|_{x=x'} \tag{4}$$

and

$$\nu_x(y) = \sum_{g_x(y')=y} \nu_x(y')/|\frac{dg_x(u)}{du}|_{u=y'}. \tag{5}$$

What we will prove here is the following Perron-Frobenius equation for skew-product transformation h:

$$\rho(y_1,y_2) = \sum_{h(x_1,x_2)=(y_1,y_2)} \rho(x_1,x_2)/|\det Dh(x,y)|_{(x,y)=(x_1,x_2)}. \tag{6}$$

We note that

$$(x',y') \equiv h^{-1}(x,y) = (f^{-1}(x), g_x^{-1}(y)). \tag{7}$$

By using Eq.(4) and Eq. (5, we can obtain the following identities

$$\begin{aligned}
\sum_{h(x',y')=(x,y)} \frac{\rho(x,y)}{|\det Dh(x_1,y_1)|_{(x_1,y_1)=(x',y')}} &= \sum_{(x',y')=h^{-1}(x,y)} \frac{\sigma(x')\nu_x(y')}{|f'(x')g'_x(y')|} \\
&= \sum_{x'=f^{-1}(x)} \{\frac{\sigma(x')}{|f'(x')|} \sum_{y'=g^{-1}_{f(x')}(y)} \frac{\nu_{f(x')}(y')}{|g'_{f(x')}(y')|}\} \\
&= \sum_{x'=f^{-1}(x)} \{\frac{\sigma(x')}{|f'(x')|}\nu_{f(x')}(y)\} \\
&= \{\sum_{x'=f^{-1}(x)} \frac{\sigma(x')}{|f'(x')|}\}\nu_x(y) \\
&= \sigma(x)\nu_x(y),
\end{aligned} \tag{8}$$

which proves the Perron-Frobenius equation (6).
Since we can obtain one-dimensional ergodic transformations with explicit invariant measures by the addition theorems of trigonometric functions and elliptic functions [3, 4], we can apply them as basic components and construct skew-product transformations in this scheme.

3 Examples

Two-Dimensional Examples
Two-dimensional mappings

$$\begin{aligned} X_{n+1} &= \frac{4X_n(1-X_n)(1-lX_n)}{(1-lX_n^2)^2} \\ Y_{n+1} &= \frac{4Y_n(1-Y_n)(1-Y_nX_{n+1})}{(1-X_{n+1}Y_n^2)^2} \end{aligned} \tag{9}$$

are given by the skew-product transformations of Katsura-Fukuda maps [2] (a generalized model of Ulam-von Neumann map [1]). They have an absolutely continuous invariant measure

$$\frac{dxdy}{4K(l)K(x)\sqrt{xy(1-x)(1-y)(1-lx)(1-xy)}} \tag{10}$$

over the two-dimensional domain $\Omega_2 = [0,1] \times [0,1]$, where $K(a)$ is the elliptic integral of the first kind given by $K(a) = \int_0^1 du/\sqrt{(1-u^2)(1-au^2)}$.

Three-Dimensional Examples
We can construct higher-dimensional maps with explicit invariant measures by successive use of skew-product transformations. Three-dimensional mappings(a)

$$\begin{aligned} X_{n+1} &= \frac{4X_n(1-X_n)(1-lX_n)}{(1-lX_n^2)^2} \\ Y_{n+1} &= \frac{4Y_n(1-Y_n)(1-Y_nX_{n+1})}{(1-X_{n+1}Y_n^2)^2} \\ Z_{n+1} &= \frac{4Z_n(1-Z_n)(1-Z_nY_{n+1})}{(1-Y_{n+1}Z_n^2)^2} \end{aligned} \tag{11}$$

constructed in this way are shown to have an invariant measure

$$\frac{dxdydz}{8K(l)K(x)K(y)\sqrt{xyz(1-x)(1-y)(1-z)(1-xy)(1-yz)}} \tag{12}$$

over the three-dimensional domain $\Omega_3 = [0,1] \times [0,1] \times [0,1]$.

Exactly Solvable Chaos in Arbitrary Dimension
In a similar way, we consider an N-dimensional mapping with explicit invariant measure by successive use of skew-product transformations.

$$X_{n+1}(j) = \frac{4X_n(j)[1-X_n(j)][1-X_n(j)X_{n+1}(j-1)]}{[1-X_{n+1}(j-1)X_n(j)^2]^2} \quad \text{for} \quad j = 1, \cdots N, \tag{13}$$

where the boundary condition is fixed as $X_n(0) = 0$ for $n = 1, \cdots$. This model can be viewed as a kind of *open flow systems* which were initially studied by Keneko [5].

Clearly, this model is an N-dimensional generalization of the above mappings (9) and (11). The above dynamical systems have explicit invariant measures

$$\frac{\prod_{j=1}^{N} dx_j}{\pi \prod_{j=1}^{N} \sqrt{x_j(1-x_j)} \prod_{j=2}^{N} \{2K(x_j)\sqrt{1-x_{j-1}x_j}\}} \tag{14}$$

which are absolutely continuous with respect to the Lebesgue measure over the N-dimensional cubic $\Omega_N = \overbrace{[0,1] \times \cdots \times [0,1]}^{N}$.

Kolmogorov-Sinai entropy We can also calculate Kolmogorov-Sinai entropy $h(\mu)$ which is the asymptotic rate of information creation by an iteration of F by the Pesin identity [7]:

$$h(\mu) = \sum_{j=1} \lambda_j = \int_\Omega \ln|\det[DF(x)]|\mu(dx) \tag{15}$$

$$= \int_\Omega \ln|\det[DF(x)]|\rho(x_1, \cdots, x_N)dx_1 \cdots dx_N, \tag{16}$$

where $\lambda_1, \lambda_2, \cdots, \lambda_N$ are N positive Lyapunov exponents of an N-dimensional map F. Although the existence of the unique invariant measures (*the Gibbs measures*) which are absolutely continuous with resepect to the Lebesgue measure such as in this example are proven by Bunimovich and Sinai[6] for generic multi-dimensional coupled map lattices with loose-coupling, the present method for showing multi-dimensional absolutely continuous invariant measures is a constructive one. Thus, we can implement these multi-dimensional chaotic maps in computers and then such chaotic maps with explicit invariant measures can be seen as a kind of multi-dimensional random-number generators.

References

[1] S. M. Ulam and J. von Neumann, On combination of stochastic and deterministic processes, *Bull. Am. Math. Soc.* **53** (1947) 1120.

[2] Katsura, S. and Fukuda, W.:Exactly solvable models showing chaotic behavior. *Physica*__130__ **A**,597-605 (1985).

[3] K. Umeno, Method of constructing exactly solvable chaos, *Phys. Rev.* **E 55** (1997) 5280-5284 (eprint chao-dyn/961009).

[4] K. Umeno, *Exactly solvable chaos and addition theorems of elliptic functions*,(eprint chao-dyn/9704007).

[5] K. Kaneko, Spatial period-doubling in open flow, *Phys. Lett.* **111A** (1985) 321-325.

[6] L. A. Bunimovich and Y. G. Sinai, Spacetime chaos in coupled map lattices, *Nonlinearity* **1** (1988) 491-516.

[7] Y. B. Pesin, Lyapunov characteristic exponents and smooth ergodic theory. *Russ. Math. Surv.* **32(4)**(1977) 55-114.

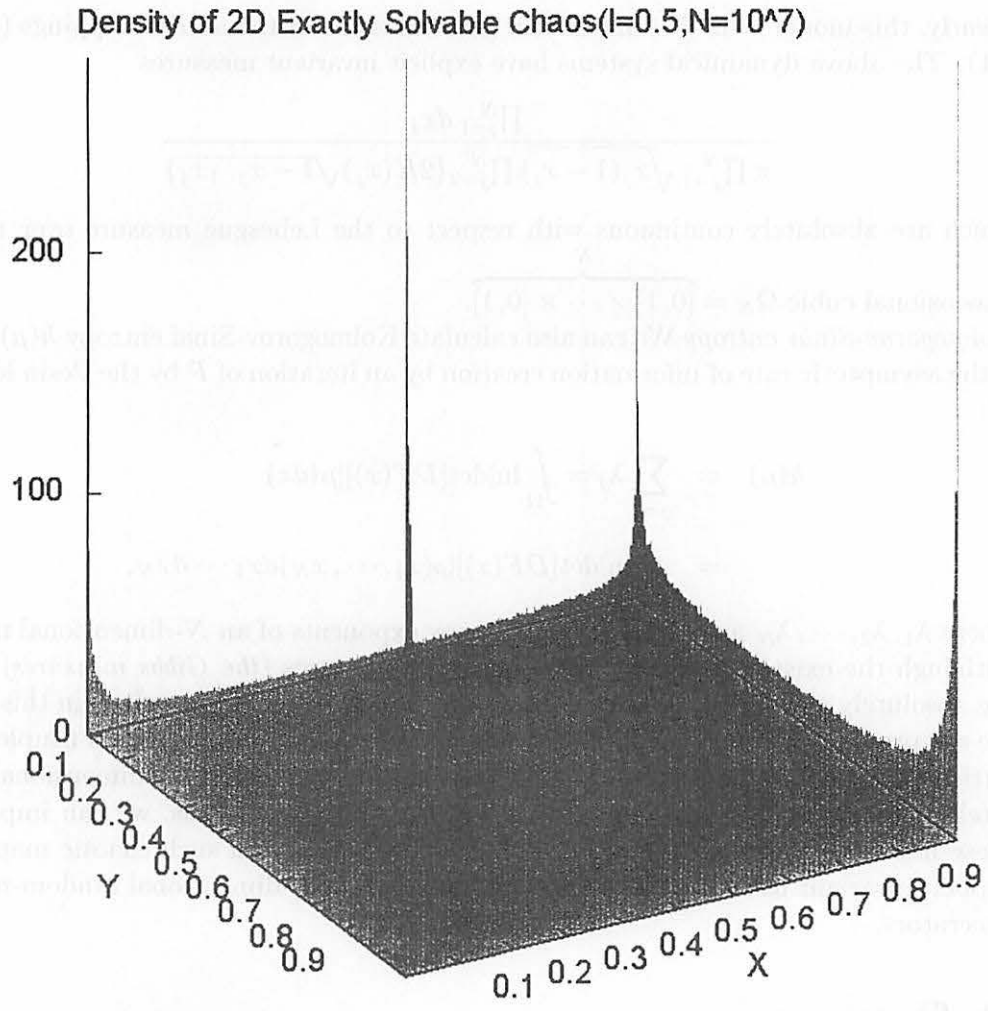


Fig. 1:
Computer simulation of density of invariant measure of 2D exactly solvable mapping (9) for $l = 0.5$ and $N = 10^7$(=iteration number). This is consistent with the analytic formula of the invariant measure in Eq. (10).

N体力学モデルによる摩擦検討
A Study of the Friction Like Phenomena by N body Molecular Dynamics Method

Kazuo HARAOKA, K.TSUTSUI[A], K.HAYASHI[B], K.SHIDA[C], & T.KAWAI[D]
Study Group for Friction[1]
[A] Keio University, Japan
[B] Kanazawa Institute of Technology, Japan
[C] Musashi Institute of Technology, Japan
[D] Chitose Institute of Science and Technology, Japan

Abstract

It is widely recongized that friction is associated with the energy transfromation from kinetic energy to thermal one. But the microscopic process of this transformation is not clear.

In this paper,to understand the energy transformation mechanism which is based on the frictional resistance, we are analyzing the simple model which is named *Friction Like Phenomena* by Molecular Dynamics Method with second ordered symplectic time integration. And we get the two typical results, one is dependency of number of particles, and the other is the dependency of spatial dimensionality.

KEYWORD Frcition Like Phenomena, Energy Transformation,MD ,Number dependency

1 始めに

「摩擦」という現象は、我々の最も身近な物理現象の一つである。この振る舞いについては古より多くの研究がなされている。摩擦、もう少し正確に言うなら「すべり摩擦 (Sliding Friction)」を古典的に扱った法則として、アモントン-クーロンの法則がある。これは摩擦に関する経験の集大成、ともいえる。それらは、（1）摩擦力は接触面積には関係しない、（2）摩擦力は垂直抗力に比例する、（3）動摩擦係数は静止摩擦係数よりも小さい、である。静止摩擦の充分マクロな系は、これらは凝着説を用いておおよそ説明されている [1, 2]。しかし、サブミクロン程度の大きさになり、接触原子の「粒」としての描像が顕著に表れる系や、非常に整った面になると説明ができなくなる。つまり、凝着説における「真接触面積」と「見かけの接触面積」が十分異なる、という仮定の限界を超えてしまう。Krim らは、銀の上に乗せた一原子層若しくは２原子層の固体キセノンフィルムを用いて実験を行い面白い結果を得ている [3]。これは空間次元自由度に対する興味を起こす。また、同様の研究は、Sokoloff [4],Liebsch [5] でもなされている。

またこの様な小さな系としては例えば粉体摩擦を挙げることができよう [6]。粉体は散逸的であるから、温度が無視できる。このことに着目した解析も行われつつある [7]。

「摩擦」と切っても切り離せない現象が、「発熱」である。「手と手をこすりあわせると暖かくなる」、「マッチを擦ると火が点く」、「嵐の時に木がこすれて山火事が起きる」、これらはすべて摩擦による発熱が起こす、もしくは「摩擦」と呼ばれる現象によって「運動エネルギー」が「熱エネルギー」に転化された結果である。上述のアモントン-クーロンの法則は、これらについては言及していない。最近のマイクロマシーンの例を持ち出すまでもなく、摩擦やそれによるによる発熱現象の制御は、発熱による信頼性低下、劣化、摩耗等々を抑制する目的からも工業的にも重要な課題である。上述の粉体系ほどの大きさで、「熱」の発現機構にも注意をした解析が必要となる。最近、摩擦の非平衡熱力学系として幾つかの顕著な研究がある [8]。また、並進運動のエネルギーから内部運動への過程を論じた物には、Rapp モデルの解析がある [9]。

以上を鑑み、このレポートは、[摩擦現象に似た現象] を表すモデル、FLP モデルを提唱し、その系での全粒子の運動を動力学的に扱い、運動エネルギーが如何に熱に変わっていくか？についての解析を試みると共に、系の粒子数が系の挙動に与える影響等を論じる。

[1]Address : 1-24-8-201 Nishi-ga-Hara, Kita-ku, Tokyo 114-0024, Japan, E-mail : kharaoka@ca2.so-net.ne.jp

2 モデル (FLP) の説明

解析を行うモデル (FLP モデル) について説明する[2]。系は内部自由度を持たない 2 種類の粒子から成る。即ち有限個の入射粒子●と有限個のアレイ粒子○である[3](Figure1)。ここでは簡単のため、入射粒子は一つであるとする。入射粒子の位置と運動量はそれぞれ、Q, P、アレイ粒子はそ

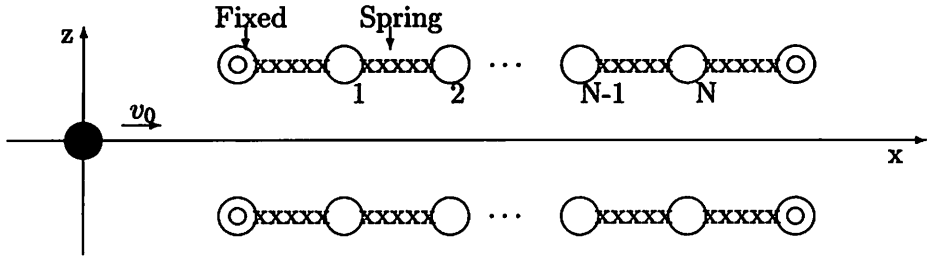


Figure 1: Rough Sketch of the FLP model

れぞれ q_i, p_i で表される。P, Q, p, q はそれぞれ本来 3 次元のベクトルである。アレイ粒子についている添え字は、粒子を識別するためのものである。アレイ粒子は、x 軸に対して対称に動くものとしている。これは、初期状態が x 軸に対称であれば、対称入射粒子が x 軸上を常に動くようにするための要請である。アレイ粒子の間は、相互作用ポテンシャル U のバネで繋がれている[4]。アレイ粒子の両端は、固定されている (図中◎)。アレイ粒子と入射粒子は、相互作用ポテンシャル V の相互作用を行う。

以上をまとめて、入射粒子が一つの場合の系全体のハミルトニアン $\mathcal{H}$ は

$$\mathcal{H} = \frac{P^2}{2} + 2\sum_i \frac{p_i^2}{2} + 2\sum_i V(Q, q_i) + 2\frac{1}{2}\sum_{i,j} U(q_i, q_j), \tag{1}$$

となる。p, V, U の和の係数の 2 は、アレイが 2 本ある事に由来する。粒子の質量は、全て 1 とした。今回の解析においては、アレイ粒子間のバネは、全て近接 2 粒子間の線形バネとし、ばね定数を k_1, また自然長 l_0 を $l_0 = 1$ とした ;

$$U(q_i,\ q_{i-1}) \ = \ \frac{1}{2}\, k_1\, (|q_i - q_{i_1}| - l_0)^2. \tag{2}$$

アレイ粒子と入射粒子の間は、下記の様な相互作用である ;

$$V(Q, q_i) = \begin{cases} \mathcal{A}\{\frac{1}{3}d^3 - \frac{1}{2}1.9l_0d^2 + 0.9l_0^2d - V_0\} & (0.9l_0<d<1.0l_0), \\ 0 & (1.0l_0<d). \end{cases} \tag{3}$$

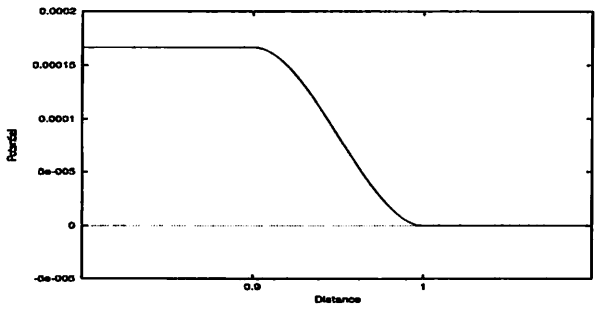


Figure 2: Interaction Potential between Injected particle and Array particle

[2]基本的には Frenkel − Kontrova モデルの拡張である
[3]このレポートではアレイというよりは「鎖」というべきである。
[4]このレポートでは、線型バネしか扱わない。

ここで $d = |Q - q_i|$、$\mathcal{A}$ は相互作用の強さを表すパラメタ、V_0 は $d = 1$ でポテンシャルを 0 にするためのオフセットである[5]。ポテンシャルの形状は Figure2 に示されている。この図に描いては、$\mathcal{A} = 1$ である。このポテンシャルは、アレイ粒子と入射粒子が、アレイ粒子の自然長と同じ距離に来た時に相互作用が滑らかに開始されることを示す。$0.9l_0$ よりは近づく事はない事を確認しながら計算は行われる。また、2 列のアレイ粒子の間隔は $1.9l_0$ であり、これらは相互作用を行う程接近する事がない様に計算される。

今回の解析においては、$t = 0$ においてアレイ粒子は静止している。入射粒子は、$t = 0$ において、アレイの外から (図では左から) 速度 v_0 で投げ込まれる。即ち、系の全エネルギーは、入射粒子の持っている運動エネルギーだけであり、保存される。

このモデルでは、入射粒子には初期速度を与える他は、アレイとの相互作用だけである。そのため、アレイにエネルギーが移り、アレイ内で静止するか、若しくは減速して通り抜けるか、またはアレイにエネルギーを与えた後に跳ね返されるかの 3 つの終状態が考えられる。入射粒子がエネルギーを失うことから、このモデルを「摩擦に似た現象」(**FLP**;Friction Like Phenomena) [6] と呼ぶ。

3 計算結果

2 章で述べた系の時間発展を、2 次のシンプレクティック法で時間積分した。また前述したが、入射粒子 − アレイ粒子の相互作用は Figure2 に示されている通りで、上下のアレイの間は $1.9l_0 (z = \pm 0.95l_0)$ であり、$l_0 = 1$, $k_1 = 1$ とする。粒子質量はすべて 1 である。

3.1 一次元での計算

解析の概要をイメージしやすくする為、まず FLP モデルでの一次元計算の結果を紹介する。この場合の一次元とは、アレイ粒子の運動が x 軸方向だけに制限されている事を指す。

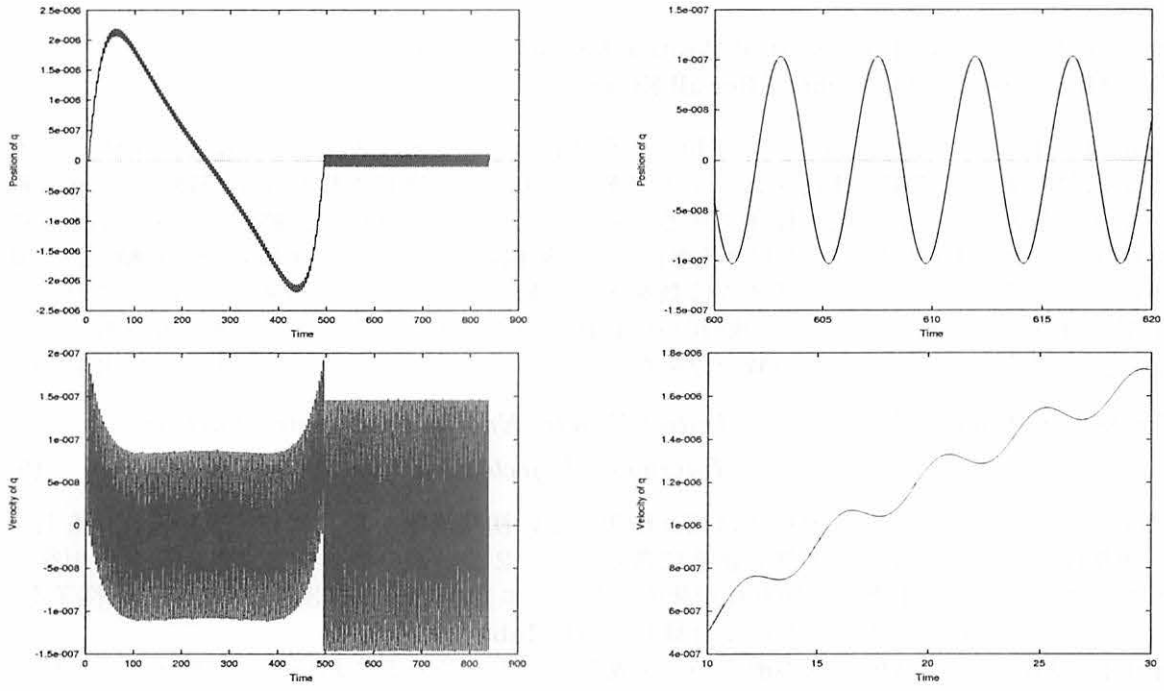


Figure 3: One Array Particle Model result. Up-Left: Distance from Neutral Position, Up-Right: detail of Up-Left, Down-Left: Verocity, Down-Right: detail of Down-Left

[5] この形状自体にも色々なモデルがあるが、滑らかに相互作用が始まる事、あまり非線形を強くしたくない事、からこの様な形を本レポートでは用いた。

[6] 「現象」と言っても、数値実験上で表れる現象である。また、モデルとして相互作用の形や、対称性の仮定など、まだモデルに対する議論も必要である。

アレイ粒子が各1つの場合、式(1)は、次の様に書き換える事ができる：

$$\mathcal{H} = \frac{P^2}{2} + 2\frac{p^2}{2} + 2V(Q,q) + 2k_1(q-l_0)^2. \tag{4}$$

アレイの粒子は一つなので添え字は省略した。また、P,Q,p,q は全てx軸上の値を表すスカラとして扱う。$\mathcal{A} = 0.01$、$v_0 = 0.002$ とした時の計算結果をFigure3に示す。

Figure3中、左上は、アレイ粒子の位置の時間発展を見ている。右上は、入射粒子が過ぎ去ったあとのアレイ粒子の位置を見る為に、時間を拡大したものである。左下はアレイ粒子の速度、右下は、立ち上がりの部分の拡大である。この計算では、まず、アレイの周期は $2^{1/2}\pi$ のみである。Figure3の右上を見ると、相互作用が無くなった場合の系は平衡系となり、周期運動を再現している。アレイ粒子は、入射粒子と相互作用を行う間(時間＝500近辺まで)は、基本周波数を載せた形の強制振動を行っている。この際、運動エネルギーを得ているが、平衡に落ち着いた際のアレイの内部エネルギーである。力は対称な形であるが、入射粒子が初速を持っているから、入射粒子がアレイ粒子の右にある場合と左にある場合で時間的に対称にはならず、力積は0ではない。速度に関してもほぼ同様の知見を得る事が出来る。

では、アレイ粒子を2個にしてみる(Figure4)。Figure3と異なるのは、入射粒子との相互作用の回数が2回に成る事と、アレイの基本周期が 2π と $2\pi/6^{1/2}$ に成ること、である。

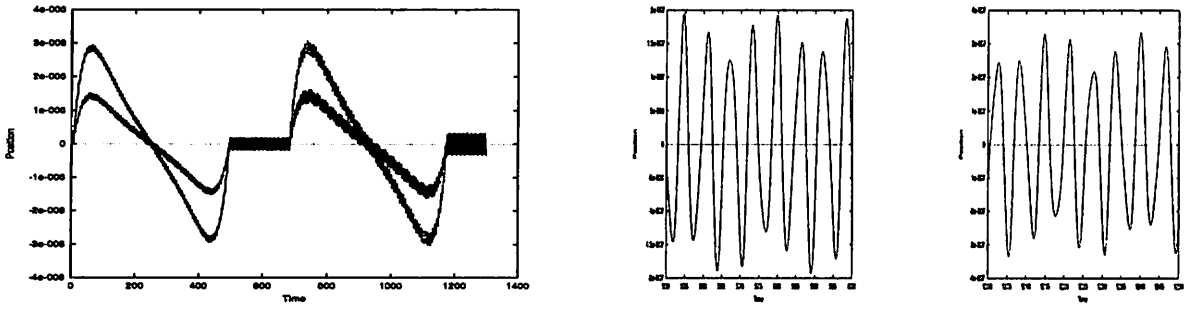


Figure 4: Left: Distance from Neutral Position for each particle,
Middle: After First Event, Right: After all Event

Figure4において、左の図は、アレイ粒子の自然長からのずれである。始めに大きなピークを持つ方が入射粒子と先に相互作用を起こす方である。中は、一度目の相互作用が終わった後、擬似的に平衡状態となっている時の様子である。右端、二度目の相互作用が終わった後である。中と右を見比べると、相互作用が2度起きる事により、振幅は大きくなる. 図においてy軸の値が異なる事に注意して欲しい。しかし基本周波数を大きく動かすほどではない事が分かる。これは入射粒子が充分遅いので高次モードへの励起が行われない為である。系のエネルギーが保存している事から、入射粒子のエネルギーの減少分をもってアレイが得たエネルギーとする事が出来る;

$$\begin{aligned} TransferedEnergy(E_T) \;=\; & Initial\ Kinetic\ Energy\ of\ Injected\ Partilce \\ & - Energy\ of\ Injected\ Partilce\ After\ Correlation. \end{aligned} \tag{5}$$

特に今回の計算においては、入射粒子は相互作用のない初期状態から、アレイと相互作用を行った後、通り抜けて再び相互作用の無い状態に落ち着く。2粒子系の場合、2度目の相互作用後のアレイのエネルギーは、1度目の相互作用後のアレイのエネルギーの約3倍となる。この様子を、1，2、4、8、16粒子の系において計算したのがTable1である。

Table1において、横は何回目の衝突か、を表している。この表においては、粒子アレイ粒子一つ当たりのエネルギーに換算してある。また一粒子の場合の遷移エネルギーを用いて規格化した。つまりTable1の値 $\mathcal{E}$ は下記の計算によって与えられる；

$$\mathcal{E} = \frac{E_T}{\{(Number\ of\ Array\ Particle)(Transfer\ Energy\ of\ One\ Particle\ Array\ Model)\}} \tag{6}$$

この様子から遷移エネルギーの割合は、粒子数8の場合に得意な値を示す事が分かる。これは、非共鳴適な場合、そして16個の場合は共鳴的な場合であると考えるが、詳細は今後である。

Table1. Transfered Energy(scaled by 1 particle E_T)

	1	2	3	4	5	6	7	8
1Particles	1.00	-	-	-	-	-	-	-
2Partilces	1.50	4.50	-	-	-	-	-	-
4Particles	1.75	9.25	17.50	18.25	-	-	-	-
8Particles	1.25	3.75	7.62	10.12	8.25	5.25	2.88	3.45
16Particles	2.00	15.62	55.87	140.94	288.68	507.36	793.85	1135.86
	9	10	11	12	13	14	15	16
16Particle	1512.60	1895.42	2253.98	2563.42	2803.87	2952.59	3047.44	3087.31

3.2 二次元での計算

2次元での計算結果を示す。ここでの2次元とは、アレイ粒子の運動を、Figure1 におけるx軸及びz軸とする事である。入射粒子は、対称性からx軸上のみの運動である。

z軸方向への運動を許すと、系は様相を変える (Figure5)。

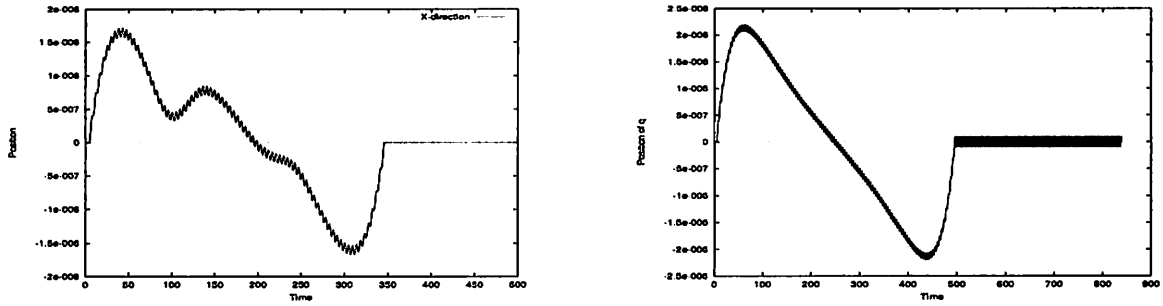


Figure 5: Comparison between One dimension and Two dimension for X direction
Left: two dimension model, Right: one dimension model(same as Figure3 Up-Left)

Figure5 は、一粒子の場合のx方向の動きを、二次元 (左) と一次元 (右) で比較している。相互作用等、他の条件はすべて同じであるにもかかわらず、様相が異なる事が分かる。Figure6 ではその理由を調べる為にx－z面上で動きを追いかけた[7]。

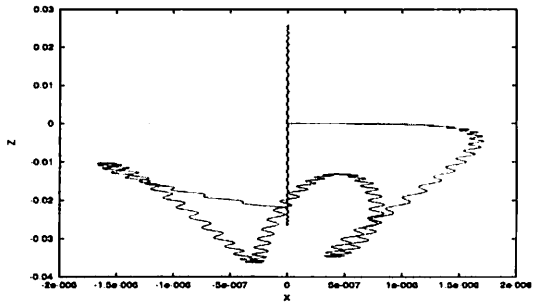


Figure 6: Trajectory of the Particle. At first, the particle moving toward x direction from origin.

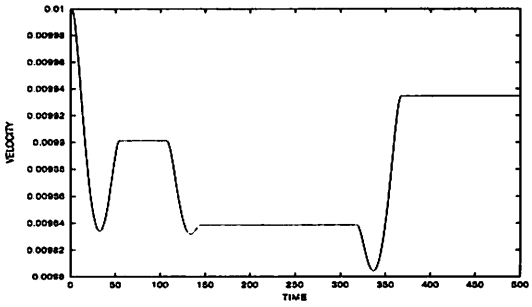


Figure 7: Velocity of the Injected Particle. Each Array consists of four particles.

まず、Figure6 は、軸の大きさが4桁異なる事に注意したい。即ち、動きのほとんどはz軸方向である。粒子は、まず自然長の点 (図中原点) からx方向に延び、一旦右回りをしてから左 (x<0) にいき、相互作用が無くなるとz方向への大きな単振動をする[8]。この計算の場合、入射粒子はその運動量の 12% 以上をアレイ粒子に渡しており、そのほとんどはz方向の振動へ移る。

Figure7 は、アレイ粒子が4つの場合の入射粒子の速度の時間変化である。アレイ粒子が4つあるから相互作用は4回あるはずだが、3つめの粒子とは、相互作用を行っていない。これは、z

[7]今回のモデルではアレイはx軸に対称に動くので、ここでは-z 領域にあるアレイの粒子の動きを示す。
[8] z方向は見かけ上ばね定数が小さい。

方向への変位が大きく、入射粒子が通過する際に、相互作用する領域に居なかった為である。また、４つめの粒子との相互作用においては、アレイから進行方向のエネルギーを取得した形になる。これはアレイ粒子間の共鳴で説明する事ができる。

２次元モデルの最後に、入射粒子のエネルギーと遷移エネルギーのアレイ粒子数依存を見る(Figure8)。この計算と上の Figure7 においてのみ、入射粒子の初速度は 0.01 である。横軸は 2 を

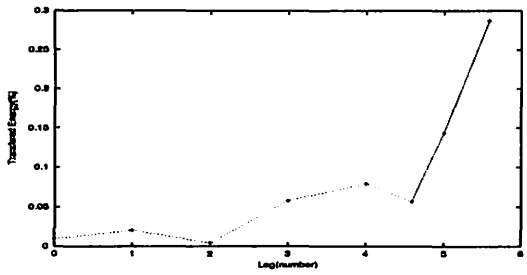


Figure 8: Dependency of the Particle number. The energy is transfered from Injected particle to Array.

底とした個数、縦軸は遷移したエネルギーの割合である。粒子数４の時は、上述の共鳴効果が顕著であり、特異な傾向を示しているが、全体としては右上がりの傾向が認めらる。

4 議論とまとめ

本レポートにおいては、摩擦現象における大きさの依存性とエネルギー移行の解析を目的に FLP モデルを提唱して、その上での解析を行った。まず、大きさに関する議論は、ここではアレイの粒子数依存性とおきかえる。１，２次元ともに粒子数に対する顕著な依存性が見られた。しかし、初期のアレイの状態をエネルギー 0 にする、入射粒子が充分遅い、等の初期状態依存性に付いての議論を行っていないから、それはこれからの課題である。また次元に関しては、FLP モデルのもつ特殊性が大きい。その為、z 軸方向へのエネルギー移行が大きい。しかし、現実系として薄膜等を考える場合には、z 軸方向のエネルギーを如何に扱うべきであるか、という事を示唆する。

今回解析したのは、粒子アレイを入射粒子にとっての環境とみなせば、入射粒子は環境との「摩擦」によって運動エネルギーを損失する過程であると考えることが可能である．初期状態において、入射粒子系と粒子アレイ系は互いに独立していて相関はないが、入射粒子が粒子アレイの上にやってくると、Figure2 で表される相互作用ハミルトニアンによって、両者はエネルギーをやり取りする．この相互作用によって、入射粒子系と粒子アレイ系は平衡に向かって時間発展するはずである．よって、入射粒子がエネルギーを損失する過程は、非可逆過程であり、粒子鎖で振動が励起される過程は、摩擦に伴う「発熱」であると考えられよう．今回調べた過程の統計熱力学的解釈は今後の課題である．

References

[1] 松川 宏 , 日本物理学会誌,Vol.51,pp584,1996

[2] I.L.Singer, J.Vac.Sci.Tech.,A12(5),pp2605,1994

[3] J.Krim, Physical Review Letters, Vol.76-5,pp803,1996

[4] J.B.Sokoloff, Cond-mat/9904164v2

[5] A.Liebsch, Cond-mat/9904202

[6] 那須野悟、早川尚男,「粉体の物理」(物理学最前線), 共立,1999

[7] J.M.Carlson, Physical ReviewE57,pp4986,1998

[8] C.Jarzynski, Physical Review Letters,Vol.78-14,pp2690,1997 and its references

[9] D.Rapp, J.of Chem.Phys.,Vol.32 pp735, 1960

また、本冊子に含まれている林の論文を参考して下さい。

Law of Wearless Friction in Sub-Micrometer Size Actuators Revealed by Molecular Dynamics Simulation

Keiji Hayashi [1], Akifumi Maeda [1], Satoru Abe [1], Noriyuki Sakudo [1], Kazuo Haraoka [2], Kiyoshi Tsutsui [3], and Toshio Kawai [4]

[1] Advanced Materials Science R&D Center, Kanazawa Institute of Technology, 3-1, Yatsukaho, Ishikawa 924-0838, Japan
[2] Study Group for Friction, 1-24-8-201, Nishigahara, Kita-ku, Tokyo 114-0024, Japan
[3] Faculty of Science & Technology, Keio University, 3-14-1, Hiyoshi, Kouhoku-ku, Yokohama 223-8522, Japan
[4] Chitose Institute of Science and Technology, 758-65, Bibi, Chitose, Hokkaido 066-8655, Japan

Law of wearless friction in sub-micrometer size machines such as motor is studied by means of molecular dynamics simulation based on an atomistic simplified two-dimensional model. A large number of simulation data obtained under various experimental conditions were systematically analyzed to examine frictional force as a function of load, temperature, and sliding velocity. Exploration of universal features lying in the function revealed unique dependence of the frictional force on sliding velocity and load. We discuss microscopic origin of the law observed under weakly coupled condition in terms of energy dissipation via forced vibration and subsequent energy transfer between normal phonon modes due to anharmonicity of inter-particle potential, laying the emphasis on effects of resonance.
Key words: Wearless friction, Molecular dynamics method, Nano-machine

1. はじめに

近年，高密度記録媒体やマイクロアクチュエーターに代表される機構部品の微細化，高性能化が長足の進歩を遂げた．機構部品において，動摩擦現象は，エネルギー効率を左右する重要な因子であると同時に，動摩擦に伴う発熱が劣化を促進するため信頼性を左右する要因でもあることはよく知られている．またブレーキに用いる機構部品には，動摩擦力が大きいことが要求される．機構部品の微細化が進むに連れて，部品へのエネルギーの供給が難しくなるためエネルギー効率の向上はますます重要な課題になり，また，部品の熱容量が小さくなるため熱発生の抑制が部品の信頼性向上の観点からますます重要な課題になる．さて，マクロな系においては，現象論的な摩擦法則として高等学校でも習うCoulomb-Amontonの三法則がよく知られており，その発現機構として，例外は多々あるものの，凝着説が広く受け入れられている．しかし，近い将来，機構部品のサイズがサブマイクロメートルスケールになると，動摩擦現象の起こる状況がマクロな動摩擦系の場合と根本的に異なってくる．サブマイクロメートルスケールの機構部品では，まず，二つのパーツ（例えば結晶格子）を原子スケールで平坦な界面で接触させることになり，真実接触面積は見かけ上の接触面積と同じオーダーになる．また，界面に垂直な方向のパーツの高さもサブマイクロメートルスケールになるため，界面にかかる荷重がマクロな摩擦系に比べて一般に桁違いに小さくなる．サブマイクロメートルスケールの系で，動摩擦現象の起こる状況がマクロな動摩擦系の場合と異なるとなれば，動摩擦現象の発現機構も凝着説とは異なり，したがって，現れる動摩擦法則もCoulomb-Amontonの法則とは異なって然るべきである．我々は，サブマイクロメートルスケールの機構部品に用いる所望の動摩擦特性をもつ材料の開発への応用を目指した基礎研究として，弾性接触条件下での動摩擦現象 [1] すなわち 磨耗を伴わない滑り摩擦現象 にみられる普遍的な法則性を，分子動力学(MD)法を用いてシミュレーション実験により調べてきた [2-4]．本研究では，弾性接触摩擦系の原子論的な単純化した二次元MDモデルに基づいて，様々な実験条件のもとで，２つの結晶格子を定常的な相対速度で滑らせるシミュレーションを行い，弾性接触条件下で発生する動摩擦力を調べた．その結果から，滑り速度をパラメタとし荷重（垂直抗力）と系の平均温度（または，界面近傍における局所疑似温度 [2]）を独立変数にとって動摩擦力をその関数として系統的に整理し，法則性を検討した．なお，三次元系では粒子の運動の仕方としてatomistic rockingが重要であることが指摘されており [5]，この点では二次元系と異なるが，二次元系では三次元系より明快な法則性が現れると期待され，二次元系と三次元系で動摩擦現象の微視的な発現機構は本質的には同じであると考えられるので，本研究では二次元のモデルを採用した．

2. 弾性接触摩擦系の 単純化したMDモデル

本研究でシミュレーションに用いた弾性接触摩

擦系の原子論的な単純化した二次元 MD モデルを図 1 に示す．二つの二次元最密格子 A および C を原子スケールで平坦な界面で接触させる．以下に示す実験では，粒子間に Lennard-Jones ポテンシャルを仮定した．A-A 粒子間および C-C 粒子間に仮定したポテンシャルの深さに比べ，A-C 粒子間に仮定したポテンシャルの深さを 1/20～3/10 の範囲で浅く採ることにより，界面の存在を考慮した．格子 A と C を接触させた後，動摩擦現象のシミュレーションに先立って，simulated anneal 法により粒子を平衡核配置に緩和させ，つぎにシミュレーションのランごとに Maxwell 分布に従う乱数を発生させて各粒子に初期速度として与え，さらに熱平衡系の分子動力学法により熱平衡状態にした．

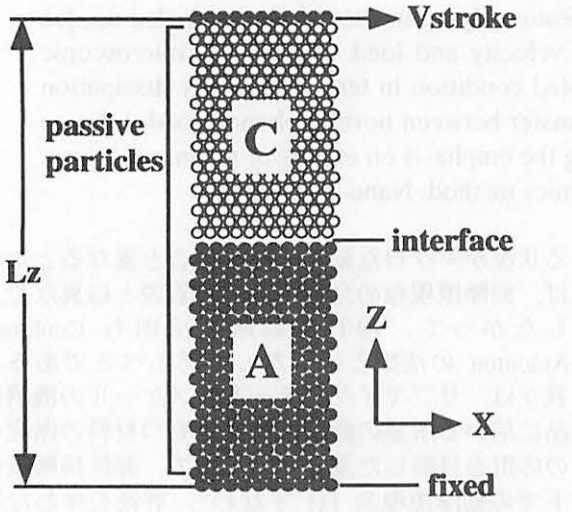


Fig. 1 An example of atomistic two-dimensional model.

その後，結晶格子 A，C を定常的な相対速度で滑らせるシミュレーションを行った．格子 A の最下端粒子層を固定し，格子 C の最上端粒子層を 固定層からの距離 L_Z を保って 界面と平行（x 方向）に シミュレーションの各ランを通し一定の滑り速度 V_{stroke} で 移動させた．格子 C の最上端粒子層を移動させ始める時刻を時刻 t = 0 とした．現実の機構部品では滑り速度が粒子の熱速度や結晶格子中の音速に比べ十分遅い場合が殆どであることを踏まえ，シミュレーションにおいても滑り速度は格子中の熱速度に比べて充分に遅く設定した．受動粒子すなわち格子 A の最下端粒子層と格子 C の最上端粒子層を除く全粒子の時間発展を，運動方程式を数値的に解くことにより追跡した．格子 A，C の左右の両端には周期境界条件を課した．なお，シミュレーションに際して，運動方程式に現象論的なエネルギー散逸項を付け加えたり，能勢の方法 [6] を用いるなど，温度一定でシミュレーションを行う手法をもちいることは，敢えて避けた．動摩擦現象はエネルギー散逸機構の典型例であり，エネルギー散逸のミクロな発現機構を論ずるモデルに現象論的にエネルギー散逸を仮定したのでは本末転倒になる．なお，結果的には，摩擦熱による温度上昇は系の初期温度に比べて充分小さく，系の平均温度がほぼ一定の条件下でのシミュレーションになった．粒子 C と A の質量比 m_C/m_A，粒子間に仮定したポテンシャルの深さの比 $\varepsilon_{CC}/\varepsilon_{AA}$，$\varepsilon_{AC}/\varepsilon_{AA}$，および平衡粒子間距離の比 σ_{CC}/σ_{AA}，σ_{AC}/σ_{AA}，格子 A および C を構成する粒子数 N_A，N_C，格子 A 最上端粒子層から格子 C 最下端粒子層までの距離 L_Z，系の平均温度 T_set，滑り速度 V_{stroke}，の設定値を それぞれ広い範囲にわたって振り，様々な実験条件で数多くのシミュレーションを行った．本稿では，基本単位として，粒子 A の質量 m_A を質量の単位に採り，A-A 粒子間に仮定した Lennard-Jones ポテンシャルの平衡粒子間距離 σ_{AA} および深さ ε_{AA} をそれぞれ距離およびエネルギーの単位に採る．また，温度の単位としても ε_{AA} を用いる．

時々刻々の各受動粒子の速度と位置が分かると，その時刻における受動粒子系の全運動エネルギーから格子 C の巨視的な並進の運動エネルギーを除いた部分 E_k および全位置エネルギー E_p がそれぞれ求まり，その和として系の全内部エネルギー E_{tot} が分かる．二次元系では，E_k を受動粒子の総数で割った値が系の平均温度 T になる．他方，荷重および動摩擦力は以下の手順で決めた．時々刻々の各受動粒子の位置が分かると，格子 A 最下端粒子層を構成する各粒子に働く力の合力の x 成分（水平成分）F_{fx} と z 成分（鉛直成分）F_{fz}，および格子 C 最上端粒子層を構成する各粒子に働く力の合力の x 成分 F_{mx} と z 成分 F_{mz} が求まる．$(F_{mz} - F_{fz})/2$ をその時刻における荷重 L，$(F_{fx} - F_{mx})/2$ をその時刻における動摩擦力 F_{fr} とみなした．動摩擦力 F_{fr} に滑り速度 V_{stroke} を掛け，時刻 0 から時間で積分することにより，受動粒子系に対して外からした摩擦仕事 W_{fr} が求まる．なお，観測された生の F_{fr} は揺らぎが大きいため，実際には，以上の手順で摩擦仕事 W_{fr} を求めた後，W_{fr} を時間の関数として冪級数で最小二乗近似し，その時間微分として動摩擦力 F_{fr} を決めている．シミュレーションのラン毎に，時刻 0 から任意の時刻までの E_{tot} の増分がその時刻の W_{fr} と一致していることを確かめてリャープノフ不安定などによる誤差が生じていないことを確認するとともに，界面近傍での局所疑似温度の上昇に伴って磨耗が起こっていないかを確かめた．以下の弾性接触動摩擦現象の解析には磨耗の起こっていない時間範囲でのデータのみを使用した．

3. 結果と考察

3.1 二次元系における 弾性接触 動摩擦法則

シミュレーション結果から，滑り速度をパラメ

ターとし荷重と系の平均温度を独立変数に採って動摩擦力をその関数として系統的に整理し，法則性を検討した．平均温度が或る値になる時刻を求め，その時刻における動摩擦力と荷重の関係を調べた．この結果，系の平均温度が異なっても荷重および他の実験条件が同じならば動摩擦力の値に有意の差が見られなかったことから，動摩擦力は系の平均温度に依存しないことが まず 分かった．つぎに，滑り速度をパラメターとして 動摩擦力の荷重依存性を調べ，さらに，荷重をパラメターとして 動摩擦力の滑り速度依存性に整理し直した．$m_C/m_A = 1.0$，$\sigma_{CC}/\sigma_{AA} = 10/11$，$\varepsilon_{CC}/\varepsilon_{AA} = 1.0$，$\varepsilon_{AC}/\varepsilon_{AA} = 0.1$，$N_A = 400$，$N_C = 440$ としてシミュレーションを行った例から得られた弾性接触動摩擦力と滑り速度の関係を図2に示す．

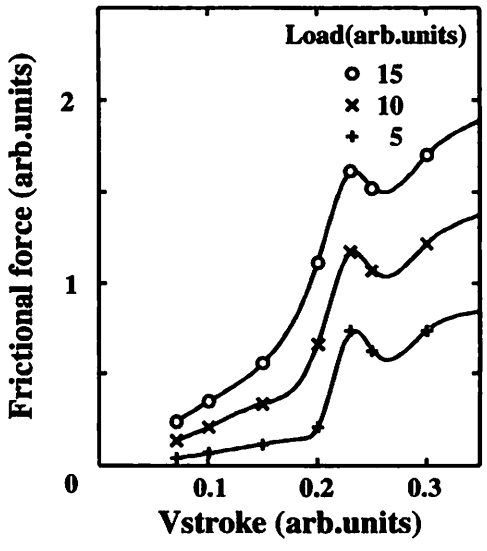


Fig. 2 Dependence of frictional force on sliding velocity.

図2から，滑り速度に閾値が存在することが分かる．滑り速度を閾値より遅く設定した場合には，設定値が遅いほど動摩擦力は単調に減少し，この関係を直線で近似し 滑り速度ゼロに外挿して良いと仮定すれば，滑り速度ゼロの極限で動摩擦力はゼロになる．一方，滑り速度を閾値より速く設定した場合には，遅く設定した場合に比べて，殊に低荷重のときに 動摩擦力が飛躍的に増大し，また，動摩擦力の滑り速度依存性は構造を持つ．図3には，動摩擦力と荷重の関係として，滑り速度を閾値より遅く設定した場合の典型例（$V_{stroke} = 0.15$）および 速く設定した場合の典型例（$V_{stroke} = 0.3$）を示す．滑り速度を閾値より速く設定した場合には，ゼロ荷重でも動摩擦力が正の値をとり，荷重の増加に伴って動摩擦力が直線的に増加する．ゼロ荷重でも動摩擦力が正の値をとることは，粒子間ポテンシャルの斥力項のみならず引力項によっても動摩擦現象が起こることを意味している．

一方，滑り速度を閾値より遅く設定した場合には，荷重にも閾値が存在する．閾値より大きい荷重領域では，荷重の増加に伴って 動摩擦力は直線的に増加する．一方，荷重が 閾値より小さいときは 大きいときに比べて 動摩擦力が極めて小さい値をとる．

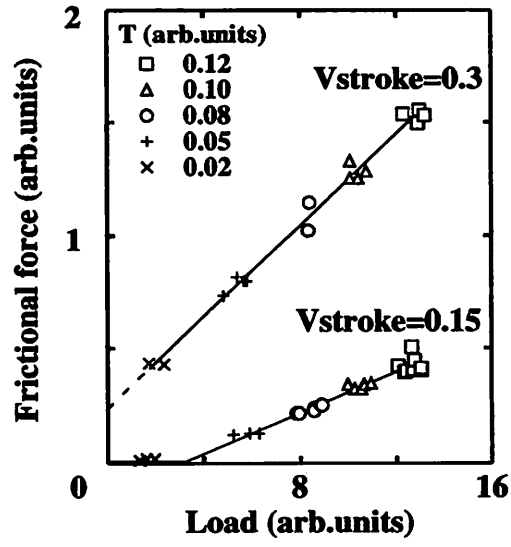


Fig. 3 Dependence of frictional force on load.

動摩擦現象は，滑りとしての巨視的な相対並進運動のエネルギーが固体中の様々なフォノンモードに分配されるエネルギー散逸過程に由来している．A-C 粒子間の相互作用 ε_{AC} が A-A，C-C 粒子間の相互作用 ε_{AA}，ε_{CC} に比べて充分小さく 結晶格子 A と C のカップリングが小さい場合には，状況の本質を単純化して 以下のように捉えることができる．滑りに伴って，結晶格子 A，C のそれぞれが，まず，滑り速度に比例した振動数をもつ弱い外力で強制振動される．さらに，結晶格子 A，C のそれぞれについて，強制振動のエネルギーが粒子間ポテンシャルの非線形性を介して 様々なノーマルフォノンモードに分配される．このエネルギー散逸に呼応して動摩擦仕事が現れる．ここで，粒子間ポテンシャルの非線形性は，粒子間距離が接近するほど，すなわち，平衡粒子間距離に対する粒子変位の大きさの比が大きいほど，顕著になる．この解釈に立つと，上述のシミュレーション結果は次のように説明される．滑り速度を閾値より速く設定した場合には，強制振動の振動数が 結晶格子 A，C いずれかの或るノーマルフォノンモードの固有振動数に一致し 共鳴が起こるため，その結晶 格子を構成する各受動粒子の強制振動による変位が大きくなり，粒子間ポテンシャルの非線形性を介したエネルギー散逸が起こる．一方，滑り速度を閾値より遅く設定した場合には，滑り速度に対応する強制振動の振動数が低いため結晶格子のノーマルフォノンモードとの共鳴が起

こらず，各受動粒子の強制振動による変位は小さい．この場合，荷重が小さい時は，粒子間ポテンシャルを調和ポテンシャルで良く近似することができ，粒子間ポテンシャルの非線形性を介したエネルギー散逸は殆ど起こらず，結果として動摩擦力は極めて小さい値をとる．強制振動による粒子の変位が小さい場合でも，荷重が大きくなる（すなわち，系の平均温度が高く受動粒子の熱運動による変位が大きくなる かまたは 平衡粒子間距離が短くなる）と，粒子間ポテンシャルの非線形性が顕著になり，エネルギー散逸レートが増加し，動摩擦力が強くなる．

3.2 同位体層の挿入による動摩擦特性の変化

弾性接触摩擦系の原子論的な単純化した二次元モデルで，結晶格子 A において 界面から n 層目と n+1 層目のすべての A 粒子を質量 m_I の同位体 I で置換することにより，動摩擦特性にどのような変化が見られるかを系統的に調べた．なお本稿では，$\sigma_{II}=\sigma_{AI}=\sigma_{AA}$，$\varepsilon_{II}=\varepsilon_{AI}=\varepsilon_{AA}$，$\sigma_{IC}=\sigma_{AC}$，$\varepsilon_{IC}=\varepsilon_{AC}$ なる粒子種 I を A の同位体と呼ぶ．図 2 および図 3 に示した例のシミュレーション条件において 結晶格子 A を構成する受動粒子 全 40 層の内の一部を質量 m_I=16 の同位体で置換した条件における 動摩擦力の滑り速度依存性の例を図 4 に示す．図 4(a)は n=37 の場合，図 4(b)は n=19 の場合で，界面から同位体層までの距離が異なる．各図中で 破線は，同位体置換を行なわなかった場合の結果である．図 4(a)，(b)において，閾値より遅い滑り速度領域では，動摩擦力に 同位体層の有無による有意の差は認められない．これに対し，閾値より速い滑り速度領域では，同位体層の存在によって 動摩擦力が概して減少するとともに 動摩擦力の滑り速度依存性に見られる閾値近傍のピークが高速度側にシフトする．しかも この傾向は，同位体層が界面に近いほど顕著である．弾性接触動摩擦の発現機構に関する前述の解釈に基づくと，このシミュレーション結果は次のように説明できる．ノーマルフォノンモードのうち，閾値近傍の滑り速度に対応する振動数での強制振動に共鳴するものは，固有振動数が低く しかも 界面近傍で大きい振幅を持つ モードである．図 4 に示した例のシミュレーション条件では，周期境界条件を課した結晶格子の左右両端間の距離は L_Z に比べ短いので，低い固有振動数を持つノーマルフォノンモードの波数ベクトルは界面に垂直な方向を持つと考えられる．さて，簡単のために 置換した同位体の質量が重い極限を考えると，同位体層を固定壁と見なすことができ，従って，界面から同位体層までの距離を変えることは，実質的に，界面から固定壁までの距離を変えることに相当する．よって，界面から同位体層までの距離が短いほど，界面に垂直な方向の波数ベクトルを持ちかつ界面近傍で大きい振幅を持つ 結晶格子 A のノーマルフォノンモードの固有振動数は高くなる．このため，同位体層が界面に近いほど，閾値近傍の滑り速度に対応する振動数での強制振動に共鳴するノーマルフォノンモードの状態密度が減少し，図 4 の結果がシミュレーション実験で得られたと考えられる．

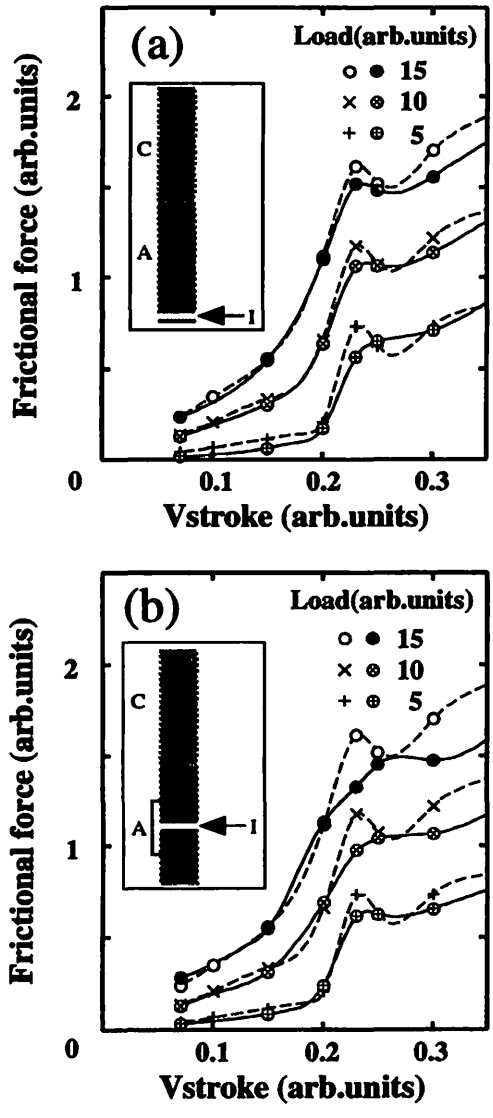


Fig. 4 Effect of isotope intercalation on sliding-velocity dependence of frictional force.

4. まとめ

弾性接触摩擦系の原子論的な単純化した二次元モデルに基づいて，格子欠陥の無い二つの結晶格子を定常的な相対速度で滑らせる分子動力学シミュレーションを 様々な実験条件のもとで系統的に行い，サブマイクロメートルスケールの系における動摩擦現象について成り立つ法則性を調べた．さらに，弾性接触動摩擦現象の微視的な発現機構について検討し，動摩擦力の大小が 粒子間ポテンシャルの非調和性を介したエネルギー散逸が起こ

り易いか否かによって決まるという仮説によってシミュレーション実験で見出された法則性を矛盾なく説明できることを示した.

この仮説に基づくと,弾性接触摩擦系について,構成する両方の結晶格子のサイズが サブマイクロメートルスケールの場合 と 巨視的な場合 とで動摩擦特性を比較したとき,二つの観点において 定性的な相違が見られること すなわち“サイズ効果”が現れることが予測される.

第一の“サイズ効果”は,弾性接触動摩擦に伴うエネルギー散逸は二段階のエネルギー移動過程を経て進むが そのうちの律速過程が摩擦系のサイズに依存して異なること に由来する.本稿ではエネルギー散逸が 二段階のエネルギー移動過程 すなわち

(i) 相対並進運動に伴って二つの結晶格子が互いに強制振動を受け,それぞれの結晶格子において基準振動がコヒーレントに励起される過程

および これに引き続く

(ii) 強制振動によってコヒーレントに励起された基準振動のエネルギーが,例えば粒子間ポテンシャルの非調和性を介して すべての基準振動モードにインコヒーレントに分配され,フォノンが熱分布する過程

を経て起こると捉え,シミュレーション実験から見出された動摩擦法則の由来を解釈した.留意すべき点として,例えば結晶格子 A, C が調和格子である場合を想定すると分かりやすいように,(ii) の過程によるエネルギー移動パスが存在しない限り,(i) の過程での強制振動に伴うエネルギー移動は,過渡的には起こるものの定常振幅に至るにつれレートがゼロになり,定常的なエネルギー散逸様現象は現れない.なお,強制振動が定常振幅に至る時定数は,サイズの大きい結晶格子ほど長くなる.このことを踏まえると,サブマイクロメートルサイズの機構部品のような少数多体系において,しかも系外へのエネルギー散逸が無視できる状況で使用する場合,動摩擦力の大小は (ii) の過程でのエネルギー移動レートの大小に強く依存することが分かる.これとは対照的に,粒子数の充分多い巨視的な結晶格子間の弾性接触動摩擦では,動摩擦力の大小は (i) の過程でのエネルギー移動レートの大小に依存して決まると考えられる.弾性接触動条件下での少数多体系の動摩擦様現象 に見られる法則性 の原子論的な由来について,Fermi - Pasta - Ulam 問題の工学的な側面としての観点 [7] からのさらなる研究が望まれる.

第二の“サイズ効果”は,微視的な結晶格子においてフォノンスペクトルが離散化することに由来する.サブマイクロメートルサイズの機構部品に用いる材料の設計に際しては,この第二の“サイズ効果”を利用することにより,例えば半導体微細加工技術を応用した“phonon-band engineering”[8] によって,動摩擦特性の向上を図れるものと期待される.

謝辞:本研究は文部省の私立大学ハイテクリサーチセンター整備事業への指定に基づく助成による.

参考文献

[1] M. Weiss and F. J. Elmer, “Physics of Sliding Friction”, Ed. by B. N. J. Persson and E. Tosatti, Kluwer Academic Publishers, Netherlands (1996) pp.163-178.

[2] K. Hayashi, N. Sakudo, and T. Kawai, *Surface and Coatings Technology*, 83, 313 (1996).

[3] K. Hayashi, JST Forum for Multidisciplinary Research “FRICTION – BASICS OF TRIBOLOGY”, p.79, Hayama, Japan (1997).

[4] K. Hayashi, N. Sakudo, H. Hiromae, S. Abe, M. Fujiyama, and T. Kawai, 5th International Conference on Nanometer-Scale Science and Technology, SS.PTh.125, Birmingham, United Kingdom (1998).

[5] M. Hirano and K. Shinjo, *Phys. Rev.*, B41, 11837 (1990).

[6] S. Nose, *J. Chem. Phys.*, 81, 511 (1984).

[7] 林, 第 5 回 総合研究大学院大学 国際シンポジウム“複雑性への戦略 ー構成と記述ー”, 26 (1999).

[8] 半導体微細加工技術の分野で超格子構造などにより電子物性を制御する際に用いられる“band engineering”になぞらえた造語.

Multiroll Structure Induced by Maxwell's Thermal Creep

Koichiro Shida*, Wm. G. Hoover†, and Shin'ichi Yamada
Musashi Institute of Technology
1-28-1 Tamazutumi, Setagaya, 158-8557
† UCDavis/Livermore, and LLNL

Abstract

"Thermal Creep" is a steady streaming motion, induced by a temperature gradient parallel to a fluid boundary, in the absence of gravity. Thermal creep has been studied by Maxwell, analyzed by Kennard, and simulated by Ibsen, Soto, and Cordero. Here we report several two-dimensional simulations. We find that the boundary driven flow can exhibit four or more roll structure which is induced by the basic two roll structure. Hence the velocity of flow decrease exponentially from the driving wall, it may difficult to see more than six rolls.

1 Introduction

Maxwell explained thermal creep using ideas from kinetic theory [1]. Consider a gas with a temperature gradient parallel to a confining wall. Because hotter particles impart more parallel momentum to the wall than do colder, a shear stress is exerted on the wall, with the gas flowing from colder to hotter as a reaction force. The resulting stationary velocity of the gas is the "thermal creep" velocity, and is parallel to the wall. The flow velocity is directly proportional to the temperature gradient. Many works have been done from the stand point of Boltzmann equation [2]. Ibsen, Soto, and Cordero reported computer simulations of thermal creep for two-dimensional hard-disk gases [3], [4]. Though there were no gravitational forces, the flow patterns which they reported resemble the familiar two convection rolls found in the Rayleigh-Bénard problem [5].

In the previous paper [6], we found that the creep velocity is sensitive to the imposed macroscopic boundary conditions and that the agreement with existing theoretical predictions is only semiquantitative. In this paper, we find that four and six convection rolls in broad rectangle area instead of two in the case of square area. Our simulations are described in section 2, and the numerical results are shown in section 3.

2 Simulations

Our simulations were carried out all at the same overall number density, N/V, and with the same hard-disk diameter, $\sigma = 0.2236$, as in ref. 2, so that the dimensionless density,

$$n\sigma^2 \equiv N\sigma^2/V = 1/20, \tag{1}$$

provided a collisional contribution to the pressure tensor of less than ten percent. The corresponding value of Maxwell's mean free path is $\lambda = 5\sqrt{2}\sigma = 1.58$. To save computer time, particles follow straight-line trajectories for a fixed timestep dt. If the distance between any pair of particles is less than σ, this pair needs to interact during the time step. We use

*shida@cs.musashi-tech.ac.jp

exact post-elastic-collisional velocities for hard disk, interpolated to the precise time of each collision.

There are two kind of boundaries. At an insulated boundary, hard-wall collisions occur whenever the separation between the wall and the center of a disk is less than $\sigma/2$. In treating thermal boundaries, particles hitting a thermostatted wall are ejected with a normal velocity drawn from the proper one-sided Gaussian distribution, $v_\perp f_{eq}$.

We chose a timestep:

$$dt \equiv 0.15\sigma(m/kT_H)^{1/2}. \tag{2}$$

The resulting heat flux between two parallel walls bounding a periodic channel was accurate to within five percent. For convenience we chose numerical values of Boltzmann's constant k, the hard disk mass m, and the highest boundary temperature T_H all equal to unity. With a typical value of the timestep the mean distance covered by a high-temperature particle was about one-sixth the particle diameter. The finite time step reduces the effective collision rate by about one percent. A typical simulation consists of four handred thousand hard-disk collisions. The weaker inducted flows require such long simulation term thought much shorter run is enough for thermal creep flow itself.

We employ "Chile" geometry, which is identical to that of ref. 2 in this paper. The temperature difference between the cold left wall, at T_L, and the hot right wall, at T_H, causes uniform heat conduction. The top and bottom walls have a temperature gradient $\propto x^{-1/3}$, with the temperature continuous at the side walls. These boundaries cause thermal creep flow from left to right, parallel to the walls, returning to the left in the central region. Another possiblility is "insulated" geometory, which has insulated side walls, so that the flow is only driven by the temperature gradient along the top and bottom walls. In this geometry, two rolls is the only possible structure because the particle never forget the transversal velosity at hard-wall collision.

Our computed temperature profiles along the top and bottom walls differ only very slightly from those of ref. 2. The temperature profile on the boundary $T \propto x^{2/3}$ corresponds to a constant heat flux with $\kappa \propto T^{1/2}$. This functional form does not allow for the existence of a temperature jump between the walls and gas near. Heat flows, roughly, parallel to the top and bottom walls. The mass flow forms two rolls, with the upper clockwise, and the lower counterclockwise. The mass flux is smoothed, by using a weighting function,

$$w(x \equiv r/a) = (5/\pi a^2)(1-x)^3(1+3x), \tag{3}$$

where $a = 2(V/N)^{1/2}$, computing the mean velocity at grid points with a spacing equal to $a/2$. The fields are averaged over the final half of each simulation.

3 Results

We tried various L_x, which is the distance between cold and hot walls, fixing $L_y = 100$, which is the distance between top and bottom walls. Flow velocity vectors are shown in Fig. 3. Though this is is the case of four rolls (29×100), the two induced rolls is weak and hard to see. Figure 2 is one dimensional plot of the same case. The plots are essentially symmetry about $x = 50$. The pair of large positive velocities at the left and right side is Maxwell's thermal creep flows. The next pair of moderate negative velocities are the return flows. The small positive velocities at central region is the evidence of another pair of rolls induced by the basic two rolls. This positive peak can be identified also in other runs and different number of particle with the same aspect ratio thoug the graph may seem to be a little bit noisy.

But in the 30×100 configurations, one grid fatter, only two basic rolls appear. There are not enough space between them (Fig. 3). We see four roll structure also in 27×100. In 25×100 case, spliting the central positive peak in two, six roll structure occurs (Fig. 4). Each roll tend to be compact to circle because the memory of velocity transversal to a wall is lost at collision to the wall. With the insulated geometry, the roll becomes wider because

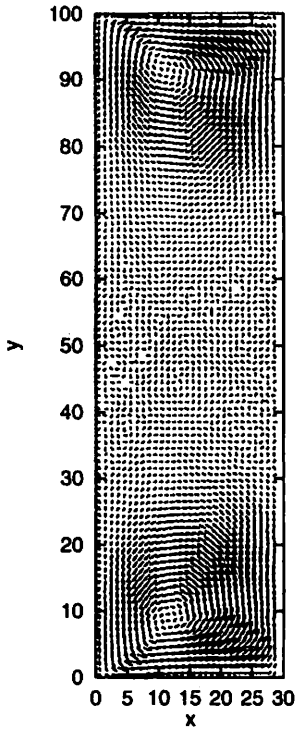


Figure 1: Flow velocity vectors for $N = 2900 = 100 \times 29$ particles with "Chile" boundary conditions. The temperature ratio for the two vertical walls is $T_H/T_L = 10$. The left wall are cold. Two apparent rolls are driven by thermal creep flows along with top and bottom walls.

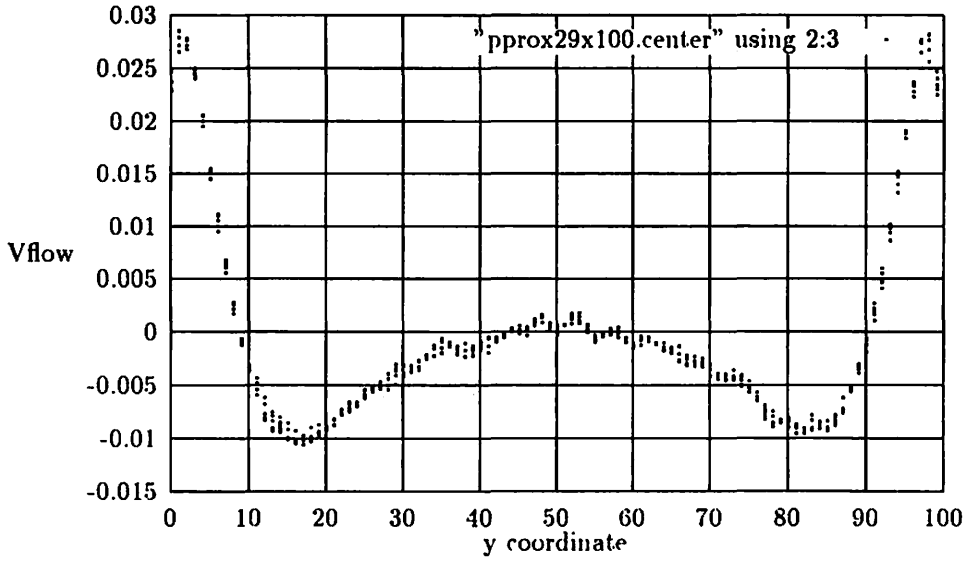


Figure 2: Flow velocity components parallel to the temperature gradient. The virtical axis shows the velocities and the horizontal axis is x coordinate from bottom to top. Central five grid points ($12 < y < 18$) where the flows are significant are shown for the same case of Fig. 1.

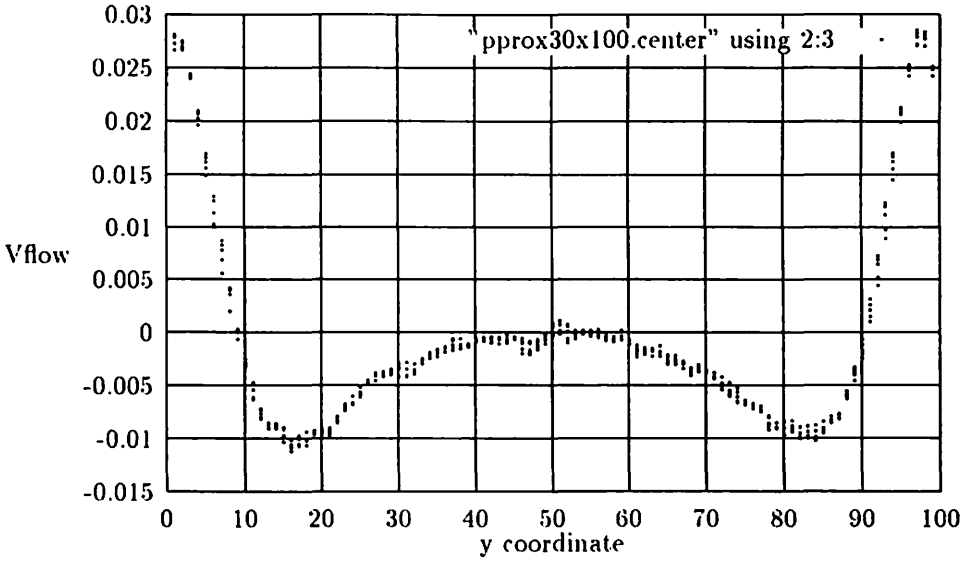


Figure 3: Flow velocity component parallel to the temperature gradient. We see only two rolls for $N = 3000 = 30 \times 100$ case.

the transversal velocity is conserved by elastic collision. We don't have the evidence of four roll structure upto 1:5 aspect ratio (25×125) with the geometry yet.

4 Conclusion

We found multiroll structure induced by thermal creep flow in wide rectangle boundary. The range of aspect ratio for four rolls is relatively narrow, and the wider aspect ratio leads six rolls. Though eight or more rolls may be possible, it will be hard to see because the flow velocity decays exponentially from the wall (source of motive force) to the center of the system.

Acknowledgements

Prof. Toshio Kawai at Chitose Institute of Science, Technology (Japan), Kazuo Haraoka, Kiyoshi Tsutusi at Keio University, and Keiji Hayashi at Kanazawa Institute of Technology for helpful discussions.

References

[1] E. H. Kennard: **Kinetic Theory,** (McGraw-Hill, New York, 1938).

[2] M. Wakabayashi, T. Ohwada, F. Golse, European J. Mech. **15**(1996)175.

[3] J. Ibsen, R. Soto, and P. Cordero: Phys. Rev. E **52**(1995)4533.

[4] P. Cordero, and D. Risso: *Fourth Granada Lectures in* **Computational Physics,** P. L. Garrido and J. Marro, eds, (Springer-Verlag, 1996).

[5] V. M. Castillo, Wm. G. Hoover, and C. G. Hoover: Phys. Rev. E **55**(1997)5546.

[6] K. Shida and Wm. G Hoover: J. Phys. Soc. Jpn. **67**(1998)2277.

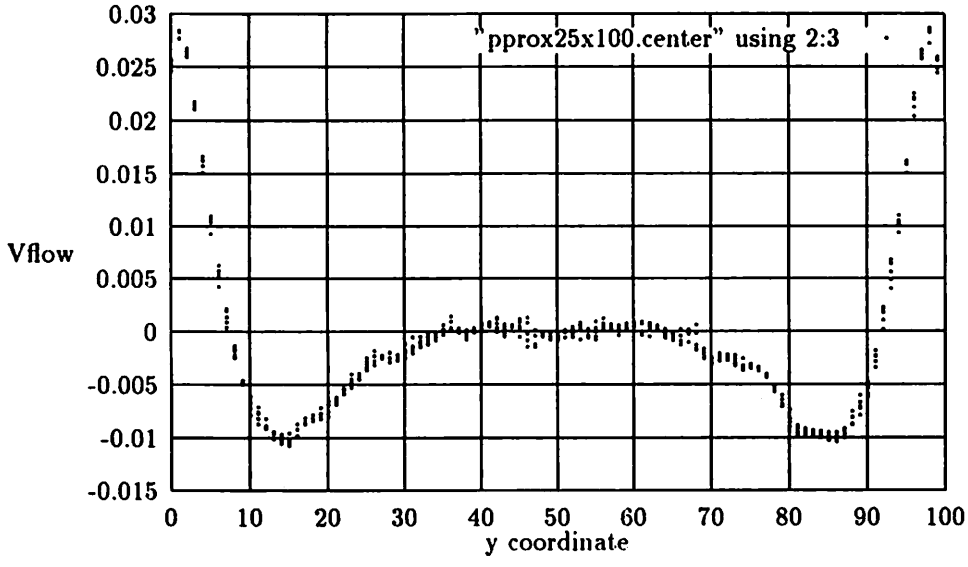


Figure 4: Flow velocity component parallel to the temperature gradient. This profile can be interpreted as six rolls. $N = 2500 = 25 \times 100$.

2次元ランダム環境中のランダムウォークにおけるレプリカ対称性の安定性

Stability of the replica symmetry of random walks in 2-D random environments

Kiyoshi Tsutsui†

Department of Electrical Engineering,
Keio University, Yokohama 223-8522, Japan

Abstract

We analyzed the asymptotic behavior of the random walks in a random environment. Interpreting the trajectory of random walker as domain wall defect of a Dirac fermion thory, the Callan and Harvey mechanism provide the relation between the bulk randomness and the zero mode localized on the domain wall. We calculated the dimension of this zero-mode nonperturbatively.

1 イントロダクション

ランダム環境中のランダムウォークは、統計力学では様々な状況で現われる．ここで言うランダム環境とは、ランダム変数が熱力学的変数ではない、いわゆるクエンチしたランダムネスを含む媒質である．例えば、ランダムウォークの軌跡を界面とみなせばランダムウォークの漸近的スケーリングは、ランダムボンド Ising 模型における domain wall のゆらぎや、不純物のある系での wetting 転移での界面ゆらぎや、やはり不純物が存在する吸着層における commensurate-incommensurate 転移での界面ゆらぎと密接に関連しているし [1]、またランダムウォークの軌跡を磁束線とみなせば、乱れたタイプ II 超伝導体における磁束線の depinning 転移での磁束線のゆらぎと関連する [2]．さらには、ランダムウォークの軌跡を高分子鎖とみなしたときには、ランダム環境中の高分子鎖の問題となる [3]．もし、non-random な環境中のランダムウォークであれば、intersection probability など完全には答えが求められていない部分も残ってはいるが、物理としてはかなり理解されている [4]．問題はランダムネスの効果であり、ランダムネスによってランダムネスがない状況でのスケーリング則がどのように変化するのかが興味の対象である．

ランダム系特有の問題として、サンプルによってランダムネスの realization が異なるゆえ、物理量を求めるときに熱力学的平均のみならず、ランダムネスの realization に関する平均もとる必要があることが挙げられる．後者の平均操作は、レプリカ法を用いて実行される．クエンチ・ランダム系を扱う際の難しさは、ランダム変数が熱力学的変数でないことから生じている．そして、ランダム変数について、平均をとることを考える．ランダム変数は熱力学変数ではないので、ランダム変数に関する平均操作を分配関数の中で行ってしまうことはできない．レプリカ法では、自由エネルギーまたは n 点の相関関数の平均を計算するために、オリジナルの系のコピー (レプリカ) を用意することによって、ランダム変数も熱力学的変数であるようなアニール系に書き換えてランダム平均操作を実行する．例えば、自由エネルギー F の平均を計算することを考えると、$F = \log Z$ であるので、

$$\overline{F} = \overline{\log Z} = \overline{\lim_{n\to 0} \frac{1}{n}(Z^n - 1)} = \lim_{n\to 0} \frac{1}{n}\left(\overline{Z^n} - 1\right) \tag{1}$$

†e-mail: tsutsui@nkmr.elec.keio.ac.jp

となる．(1) はどのようなことを意味しているのかといえば、自由エネルギーのランダム平均を求めるためには、オリジナルの系の n 個のレプリカからなる系の分配関数の平均を求めればよいことである．つまり、Hamiltonian が $\sum_\alpha \mathcal{H}_\alpha$ であるような系の分配関数についてランダム平均をとればよいのである．そして最後に $n \to 0$ の極限操作を行えばよい．(1) 中の $\overline{Z^n}$ を求める際には、場の理論でのテクニック、例えば摂動展開における Feynman ダイアグラムの方法やくりこみ群の方法 [5] を用いることが出来る．

レプリカ法で導入される各レプリカに添付されるレプリカ・インデックスの permutation に関する対称性であるレプリカ対称性は、破れることがある．フラストレーションの強いランダムスピン系の平均場模型で実現するスピングラス相ではレプリカ対称性が破れることが知られている [6]．またフラストレーションがそれほど強くなくても、低次元系であればレプリカ対称性が安定ではない可能性がある [7][8]．

2 モデル

以下では、2 次元ランダムウォークにおけるレプリカ対称性の安定性を、$(2+1)$-次元ゲージ場の理論の方法を用いて解析する．ランダムウォーカーの軌跡をランダム Ising 模型における domain wall とみなすと、問題は、空間に依存する質量を持つ Dirac フェルミオン模型の問題になる．domain wall は 2 次元のストリップ構造を考えたとき、ある方向で向かい合う境界に異なる境界条件、つまり片方が↑なら反対側は↓のような境界条件を課したときに、系が臨界点からそう離れていなければ、x 方向のどこかで隣り合うサイトの配置が、↑↓ となるところが現われるが、そのような↑領域と↓領域の壁となる点である．空間が 2 次元の場合の domain wall は、例えば、

$$\begin{array}{ccccc|ccccc} \uparrow & \uparrow & \cdots & \uparrow & \uparrow & \downarrow & \downarrow & \cdots & \downarrow & \downarrow \\ \uparrow & \uparrow & \cdots & \uparrow & \uparrow & \downarrow & \downarrow & \cdots & \downarrow & \downarrow \\ \uparrow & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \downarrow \\ \uparrow & \uparrow & \cdots & \uparrow & \uparrow & \downarrow & \downarrow & \cdots & \downarrow & \downarrow \\ \multicolumn{5}{c|}{fixed} & & & & & fixed \end{array} \tag{2}$$

の | に位置する．(2) に示した domain wall は真直ぐであるが、ランダムネスがあれば、domain wall の形状は直線のような単純な形状ではない．もし、バルク部分にランダムネスがなければ domain wall の形状はフラクタルであることが知られている [3]．今、調べたいことはバルク部分のランダムネスによって domain wall の形状はどのように影響されるか、ということである．作用は、

$$S_{2+1} = \int dx_3 \int dx_1 dx_2 \overline{\psi} \left(\gamma_\mu \partial_\mu + ie\gamma_\mu A_\mu + M(\boldsymbol{r}) \right) \psi \tag{3}$$

である．$\boldsymbol{r} = (x_1, x_2)$ は空間座標を表わすものとし、x_3 は時間変数であるとする．γ 行列は 2×2 の Hermite 行列であって、

$$\{\gamma_\mu, \gamma_\nu\} = 2\delta_{\mu\nu} \tag{4}$$

のように規格化されているとする．(3) に現われるゲージ場 A_μ は $U(1)$ ゲージ場で、フェルミオン数が保存されることに付随して現われる．ゲージを、

$$A_3 = 0 \tag{5}$$

のように選ぶと、ゲージ不変性の条件は、

$$\partial_3 A_\mu = 0, \qquad \mu = 1, 2 \tag{6}$$

のように書くことができる．

domain wall が存在するためには、Dirac 質量 $M(\boldsymbol{r})$ はパリティー不変性をあらわに破らなくてはならないことに注意する．そして、domain wall は、

$$M(\boldsymbol{r}) = 0 \tag{7}$$

を満たす点に位置する．

3 ランダム環境の効果

ランダム環境の効果は、作用 (3) にランダム項を付け加えることで考慮する．その前に、バルク部分のランダムネスが domain wall に果たしてどのように影響するのかを直観的に理解するには、Callan-Harvey メカニズム [9] が役に立つ．このメカニズムは、例えば domain wall のバックグラウンド場と結合する時空 $2n+1$ ($n \in \mathbf{Z}$) 次元 Dirac fermion を考えると、バルクのパリティーアノマリーを相殺するように、domain wall 上にカイラルゼロモードが誘起される、というものである．今、ランダムウォークの漸近的性質に興味があるので、

$$\psi(\boldsymbol{r}, x_3) = \int \frac{d\epsilon}{\sqrt{2\pi}} e^{-i\epsilon x_3} \psi_\epsilon(\boldsymbol{r}) \tag{8}$$

のように時間変数 x_3 について Fourier 変換してしまう．(8) を (3) に代入して、ゲージ固定条件 (5) を用いると、作用は、

$$S_{2-1} = \int d\epsilon \int d\boldsymbol{r} \overline{\psi}_\epsilon(\boldsymbol{r}) \left(\gamma_1 \partial_1 + ie\gamma_1 A_1 + \gamma_2 \partial_2 + ie\gamma_2 A_2 + M(\boldsymbol{r}) - i\gamma_3 \epsilon\right) \psi_\epsilon(\boldsymbol{r}) \tag{9}$$

のようになって、γ_3-行列を pauli 行列を使って $\gamma_3 = \sigma^3$ と選べば、

$$S_{2+1} = \int d\epsilon \int d\boldsymbol{r} \overline{\psi}_\epsilon(\boldsymbol{r}) \left(H_{eff} \mp i\epsilon\right) \psi_\epsilon(\boldsymbol{r}) \tag{10}$$

$$H_{eff} = \gamma_1 \partial_1 + ie\gamma_1 A_1 + \gamma_2 \partial_2 + ie\gamma_2 A_2 + M(\boldsymbol{r}) \tag{11}$$

のように、各 ϵ 成分は分離する．よって、Hamiltonian が (11) の H_{eff} で書かれるような 2 次元系を調べればよいことになる．さらに、Dirac fermion を、

$$\psi_\epsilon(\boldsymbol{r}) = \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} \tag{12}$$

のように 2 つの Majorana fermion で書き、$\boldsymbol{r} = (x_1, x_2)$ を

$$z = \frac{1}{\sqrt{2}} (x_1 + ix_2), \qquad \overline{z} = \frac{1}{\sqrt{2}} (x_1 - ix_2) \tag{13}$$

さらに、ゲージ場も、

$$A_z = \frac{1}{\sqrt{2}} (A_1 - iA_2), \qquad A_{\overline{z}} = \frac{1}{\sqrt{2}} (A_1 + iA_2) \tag{14}$$

のように書く．$\gamma_1 = \sigma^1$、$\gamma_2 = \sigma^2$ とすれば、$S_{eff} \equiv \int d\boldsymbol{r} \overline{\psi}_\epsilon(\boldsymbol{r}) H_{eff} \psi_\epsilon(\boldsymbol{r})$ は、

$$S_{eff} = \int \frac{d^2x}{2\pi} \left[\psi_1^\dagger \left(\partial_z - iA_z\right) \psi_2 + \psi_2^\dagger \left(\partial_{\overline{z}} - iA_{\overline{z}}\right) \psi_1 - iM(\boldsymbol{r}) \left(\psi_1^\dagger \psi_1 + \psi_2^\dagger \psi_2\right)\right] \tag{15}$$

となる．Callan-Harvey メカニズム [9] より (7) の解 $\boldsymbol{r}_0$ 上にゼロモードが現われる．fermion で書かれている作用 (15) をボゾン化する．スカラーボゾン場 ϕ を右向きボゾン場 ϕ_R と左向きボゾン場 ϕ_L の和 $\phi(z, \overline{z}) + \phi_L(z) + \phi_R(\overline{z})$ と書き、ボゾン化ルール、

$$\psi_1^\dagger = e^{-i\phi_R}, \quad \psi_1 = e^{-i\phi_L}, \quad \psi_2^\dagger = e^{+i\phi_L}, \quad \psi_2 = e^{+i\phi_R} \tag{16}$$

を用いる．すると、(15) は、

$$\phi(z, \overline{z}) = \phi_L(z) + \phi_R(\overline{z}) \tag{17}$$

を用いて、

$$S_{eff} = \int d^2x \left[\frac{1}{8\pi} \partial_\mu \phi \partial_\mu \phi - \frac{i}{2\pi} \varepsilon_{\mu\nu} \partial_\nu \phi A_\mu\right] - i \int \frac{d^2x}{2\pi} M(\boldsymbol{r}) \cos\phi \tag{18}$$

となる．$\varepsilon_{\mu\nu}$ は 2 次元完全反対称テンソルで、$\varepsilon_{z\overline{z}} = i$ とした．

ここで、ランダム環境の効果を導入する．最も簡単なランダムネスとして、ランダム・ゲージポテンシャルを導入する．ランダム・ゲージポテンシャルの分布 $P[A_\mu]$ は、

$$\begin{aligned} P[A_\mu] &= e^{-\frac{1}{g^2}\int \frac{d^2x}{2\pi} A_\mu(x) A_\mu(x)} \\ &\equiv e^{-S_{dis}} \end{aligned} \tag{19}$$

に従うものとする．g はランダムネスの強さを表わし、g が大きいほどランダムネスは強い．(19) では S_{dis} を定義したが、これを (18) の作用に加えてしまう．全作用は、

$$S = S_{eff} + S_{dis} \tag{20}$$

である．ボゾン化ルール (16) ではボゾン場のコンパクト化半径を 1 に固定してしまったが、その代わりボゾン場とゲージ場の結合定数 g を導入して、(18) の右辺第 2 項を、

$$\frac{i}{2\pi}\varepsilon_{\mu\nu}\partial_\nu\phi A_\mu = \frac{i}{2\pi g}\varepsilon_{\mu\nu}\partial_\nu\phi A_\mu \tag{21}$$

と書き換えておく．結局、作用は、

$$\int d^2x \left[\frac{1}{8\pi}\partial_\mu\phi\partial_\mu\phi - \frac{i}{2\pi g}\varepsilon_{\mu\nu}\partial_\nu\phi A_\mu + \frac{1}{2\pi g^2}A_\mu A_\mu\right] - i\int \frac{d^2x}{2\pi} M(r)\cos\phi \tag{22}$$

のようになる．A_μ が磁気モノポールの configuration に付随するような特異性を持っていなければ、Hodge 分解

$$A_\mu = \partial_\mu\Lambda \tag{23}$$

を用いることができる．すると、(22) は、

$$\int d^2x \left[\frac{1}{8\pi}\partial_\mu\phi\partial_\mu\phi - \frac{i}{2\pi g}\varepsilon_{\mu\nu}\partial_\nu\phi\partial_\mu\Lambda + \frac{1}{2\pi g^2}\partial_\mu\Lambda\partial_\mu\Lambda\right] - i\int \frac{d^2x}{2\pi} M(r)\cos\phi \tag{24}$$

となるが、これは 2 つのスカラー場 ϕ と Λ がトポロジカルに結合していることを表わしている．(24) で最後の cos-項がなければ、作用はセントラルチャージ c が $c=2$ の共形場理論の作用と一致する．しかし、今、我々が着目しているのは、$M(r)=0$ のライン上に局在しているゼロモードのダイナミクスである．このトポロジカル励起であるゼロモードのチャージ Q は、

$$\begin{aligned} Q &= -\oint dx_\mu \varepsilon_{\mu\nu} j_\nu = Q_L - Q_R \\ Q_L &= \frac{1}{2\pi g}\oint dz \partial_z \phi_L, \quad Q_R = \frac{1}{2\pi g}\oint d\bar{z}\partial_{\bar{z}}\phi_R \end{aligned} \tag{25}$$

であり、これらは、場 $e^{i\alpha_L\phi_L + i\alpha_R\phi_R}$ と、

$$\left[Q_{L/R}, e^{i\alpha_L\phi_L + i\alpha_R\phi_R}\right] = \alpha_{L/R} e^{i\alpha_L\phi_L + i\alpha_R\phi_R} \tag{26}$$

のような交換関係を満たす．この α_L と α_R は場 $e^{i\alpha_L\phi_L + i\alpha_R\phi_R}$ の次元と直接関係している．Λ のカレントと ϕ のカレントの間の関係、

$$\oint dx_\mu \partial_\mu\Lambda = \frac{\theta}{2\pi g}\oint dx_\mu\partial_\mu\phi \tag{27}$$

によってパラメーター θ を定義すると、

$$\alpha_{L/R} = \frac{1}{g}\left(m + \frac{\theta}{2\pi}n\right) \pm \frac{gn}{2} \tag{28}$$

となることを示すことが出来る．$m, n \in \mathbf{Z}$ であって、クーロンガス表示ではそれぞれ電荷、磁荷と呼ばれている量である [10]．しかし、(28) では θ 項が現われている．この項の存在によって、系の対称性はもともとの連続対称性から離散的な $SL(2,\mathbf{Z})$ まで落ちる．

求めたいのは、domain wall の漸近的振る舞いであり、それは (24) における $\cos\phi$ の次元によって特徴づけられる．domain wall の形状が直線になるように、共形変換 $z \to f(z)$ を施し、domain wall が (t,σ) 平面で $t=0$ の直線上に来るようにする．さらに folding によって、domain wall を境に系を折り畳む．

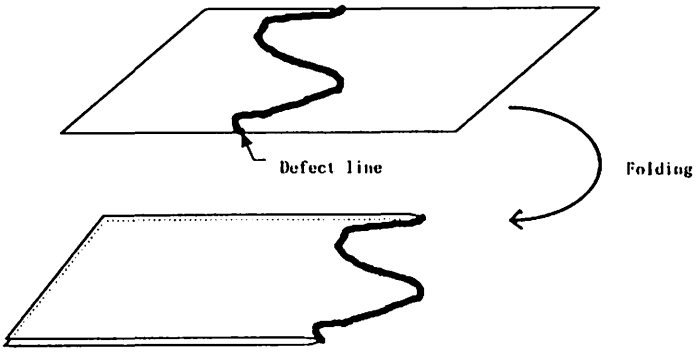


Fig. 1 Folding procedure

すると、バルク部分の自由度は元の系の 2 倍になっている．

$$\phi(t,\sigma) = \theta(t)\phi^{(+)}(t,\sigma) + \theta(-t)\phi^{(-)}(t,\sigma) \tag{29}$$

によって場 $\phi^{(\pm)}$ を定義する．$\theta(t)$ は Heviside 関数．さらに、$\phi^{\pm}$ それぞれを左向き成分 $\phi_L^{\pm}$ と右向き成分 $\phi_R^{\pm}$ に分けて、

$$\begin{aligned} \varphi_L(t,\sigma) &= \phi_L^{(+)}(t,\sigma), \qquad \varphi_R(t,\sigma) = \phi_L^{(-)}(-t,\sigma) \\ \tilde{\varphi}_L(t,\sigma) &= \phi_R^{(-)}(-t,\sigma), \qquad \tilde{\varphi}_R(t,\sigma) = \phi_R^{(+)}(t,\sigma) \end{aligned} \tag{30}$$

によって、$\varphi = \varphi_L + \varphi_R$ と $\tilde{\varphi} = \tilde{\varphi}_L + \tilde{\varphi}_R$ を定義する．同時に、

$$\frac{\varphi^{(e)}}{\sqrt{2}} = \varphi + \tilde{\varphi}, \qquad \frac{\varphi^{(o)}}{\sqrt{2}} = \varphi - \tilde{\varphi} \tag{31}$$

を定義すれば、

$$\begin{aligned} \cos\phi &= \cos\left(\frac{\phi^{(+)}+\phi^{(-)}}{\sqrt{2}}\right)\cos\left(\frac{\phi^{(+)}-\phi^{(-)}}{\sqrt{2}}\right) \\ &= \cos\left(\frac{\varphi_L - \varphi_R - \tilde{\varphi}_L + \tilde{\varphi}_R}{\sqrt{2}}\right) \end{aligned} \tag{32}$$

である．$\phi_L + \phi_R = 0$ を使った．

臨界理論を求めるには、Bernard[11] に従って、ゲージ変換によってランダムポテンシャル $\partial\Lambda$ を表面的に作用から消去してしまう．このゲージ変換は、今の場合 (28) に θ 項が存在することより、注意を要する．つまり、$\frac{\phi^{(+)}-\phi^{(-)}}{\sqrt{2}}$ は、ゲージ変換に伴って、

$$\Lambda \to \Lambda + \frac{\theta\phi}{2\pi g}, \quad \frac{\phi^{(+)}-\phi^{(-)}}{\sqrt{2}} \to \left(1+\frac{\theta\phi}{2\pi g}\right)\frac{\varphi_L}{\sqrt{2}} - \left(1-\frac{\theta\phi}{2\pi g}\right)\frac{\varphi_R}{\sqrt{2}} \tag{33}$$

のように変換される．ゲージ変換によって作用は、

$$S = \frac{1}{8\pi}\int_0^\infty dt \int d\sigma \left(\partial_\mu \varphi^{(o)}\right)^2 - iM(t,\sigma=0)\int \frac{d\sigma}{2\pi}\cos\left(\left(1+\frac{\theta}{2\pi g}\right)\frac{\varphi_L(t,0)}{\sqrt{2}} - \left(1-\frac{\theta}{2\pi g}\right)\frac{\varphi_R(t,0)}{\sqrt{2}}\right) \tag{34}$$

のようになる．この作用から cos 項の次元を求める．そのために、もう一度系を unfolding する．先に folding した後、$\phi_L + \phi_R = 0$ によって自由度を半分にしているので、unfolding 後の作用は、左向き成分だけで書ける．その作用から、ϕ 項の次元を求めると、結果は、

$$\frac{\left(1-\left(\frac{\theta}{2\pi g^2}\right)^2\right)^2}{1+\left(\frac{\theta}{2\pi g^2}\right)^2} \tag{35}$$

となる．これは、$g \to \infty$ の極限で次元 1 となるが、θ/g が有限ならば、次元は 1 より小さい．このように、バルク部分のランダムネスは、domain wall 上に局在するゼロモードの次元に大きく影響する．

4 ディスカッション

ランダム・ゲージ・ポテンシャル環境中の domain wall である Dirac fermion のゼロモードの臨界指数を非摂動論的に求めた．それは、ランダムネスの大きさによって連続的に変化する．ランダムネスがなければ、次元は 1 であり Gaussian 理論におけるボゾン場の次元と一致する．ランダムネスを強くして行くと、次元は単調に減少し、強さ $\to \infty$ の極限で 0 になる．議論の過程で現われたのは、$SL(2,\mathbf{Z})$ 対称な理論である．

レプリカ対称解の安定性の議論については、Fisher[12] の議論が参考になる．ランダムゲージポテンシャルを含む作用 (18) をレプリカ法によって解析すれば、レプリカ対称解は不安定であることを示すことが出来る．$SL(2,\mathbf{Z})$ 対称性とレプリカ対称解の安定性の関連を調べることは今後の課題である．

References

[1] M. Kardar, Nucl. Phys. B290 [FS] (1987) 582

[2] L. Balents and M. Kardar, Nucl. Phys. B393 (1993) 480

[3] M. Lässig, J. Phys.: Condens. Matter 10 (1998) 9905

[4] N.G. Van Kampen, *Stichastic Processes in Physics and Chemistry*, Revised and enlarged edition, (North-Holland, Amsterdam, 1992)

[5] P. Le Dousaal and T. Giamarchi, Phys. Rev. Lett. 74 (1995) 606

[6] M. Mézard, G. Parisi and M. A. Virasoro, *Spin Glass Theory and Beyond* (World Scientific, Singapore, 1987)

[7] V. S. Dotsenko, Vl. S. Dotsenko, M. Picco, Nucl. Phys. B 520 [FS] (1998) 633

[8] Marc-Andre Lewis, cond-mat/9710312

[9] C. G. Callan, Jr. and J. A. Harvey, Nucl. Phys. B250 (1985) 427

[10] B. Nienhus, in *Phase Transitions and Critical Phenomana*, Vol. 11, C. Domb and J. L. Lebowitz, eds. (Academic Press, London, 1987)

[11] D. Barnard, Nucl. Phys. B 441 [FS] (1995) 471

[12] M. P. A. Fisher, Phys. Rev. Lett. 62 (1989) 1415

Improvement of the triple-encounter criterion

Hiroaki UMEHARA[1] and Kiyotaka TANIKAWA[2]

[1] Kashima Space Research Center, Communications Research Laboratory

Hirai, Kashima, Ibaraki 314-0012, Japan (ume@crl.go.jp, http://www.crl.go.jp/ka/)

[2] National Astronomical Observatory, Mitaka, Tokyo, 181-8588, Japan

Abstract. The criterion for classifying a triple encounter is improved in order to cover the so-called wide triple-encounter. Its validity is confirmed by systematic numerical observations of orbits in the three-body problem with zero initial-velocities and equal masses. The physical meaning of this criterion is discussed.

1 Introduction

Judging triple encounters is important for analyzing the scattering process and chaotic dynamics in the three-body problem. In the negative-energy system, Szebehely (1971) formulated the classification of the possible states. Agekian and Martynova (1973) divided *Interplay* (one of the states) into two categories: *Close triple-encounter* and *Simple interplay.* Agekian and Anosova (1990) defined an index of triple encounter and evaluated a probability of close triple-encounters leading to escape. Therefore, the quantification of close encounters has been paid attention to in the unstable three-body problem.

On the other hand, there are chaos analyses which are indispensable for counting the frequency of the approach to three masses regardless of escape or ejection. Tanikawa and Umehara (1998) deduced the existence and the distribution of the oscillatory solutions in the free-fall problem with equal masses ("*free fall*" means zero initial-velocities). They found distribution rules of escape orbits after the n-th triple-encounter for each positive integer n. This implies the importance of an accurate triple-encounter criterion that covers a state where three masses approaches each other at a distance.

The following is our improvement of the triple-encounter criterion.

Our criterion (hereafter refered as [UT99]). *Let $I(t)$ be the moment of inertia of three particles with masses m_i, $i = 1, 2, 3$. The system is said to be in a triple encounter if it satisfies either of the following two conditions:*

(i) $I(t) \leq I_0$, *where*

$$I_0 = \frac{M_*^6}{4Mh^2}, \quad M = m_1 + m_2 + m_3, \quad M_* = \sqrt{m_1m_2 + m_2m_3 + m_3m_1}, \tag{1}$$

and h *is the total energy of the system.*

(ii) $\ddot{I}(t) > 0$ *and* $\rho(t) \leq \rho_0$, *where* ρ *is the distance between the third particle and the gravity center of the nearest two particles, and*

$$\rho_0 = \frac{M_*^2}{|h|}. \tag{2}$$

Tanikawa and Umehara (1998) use [UT99]. Some criteria of a triple encounter were formulated by Agekian and Martynova (1973), Anosova and Zavalov (1989), and Junzo Yoshida (1997) (see Section 2). However, it is found that these existing definitions have erroneous results concerning the number of triple encounters (Section 3). The validity and the meaning of our criterion are discussed in Section 4.

2 Existing definitions

The criterion suggested by Agekian and Martynova (1973, hereafter [AM73]) detects a concave state of $I(t)$, i.e., $\ddot{I} \geq 0$. The following is given by the Lagrange-Jacobi equation $\ddot{I} = 2(U + 2h)$ with potential energy U.

Criterion [AM73]. *The system with negative energy is said to be in a triple encounter while* $U(t) > 2|h|$ *and if* $\dot{I} = 0$ *or triple collision occurs in the time interval.*

In the next section, it will be shown that [AM73] has a problem since it is too weak. In some cases, one particle is ejected from the others although $\ddot{I} \geq 0$.

Anosova and Zavalov (1989) formulated another definition (hereafter [AZ89]) in terms of the perimeter of the configuration triangle.

Criterion [AZ89]. *The system with negative energy is said to be in a triple encounter if* $\sigma(t) \leq \sigma_0$, *where* σ *is the perimeter and*

$$\sigma_0 = \frac{3}{2}\rho_0. \tag{3}$$

We find that [AZ89] has a relationship with $I(t)$ such that $\ddot{I}(t) > 0$ if $\sigma(t) < \sigma_0$ in the equal-mass case (Umehara 1997).

The following definition is developed by Yoshida (1997, hereafter [Y97]).

Criterion [Y97] *The negative-energy system is said to be in a triple encounter if* $I(t) \leq I_0$.

Yoshida (1997) proved that $\ddot{I}(t) > 0$ if $I(t) < I_0$. The assumption yields $U(t) > 2|h|$ since

$$U \geq \sqrt{\frac{M_*^6}{MI}} \qquad \text{(Sibahara and Yoshida, 1963).} \tag{4}$$

It is noticed that [Y97] includes a corollary by Zare (1981, eq.(15)) which implies that $\ddot{I}(t) > 0$ if $I(t) < I_1$, where

$$I_1 = \frac{m_*^4 M_*^2}{4Mh^2} < I_0, \quad m_* = \min\{\sqrt{m_1 m_2}, \sqrt{m_2 m_3}, \sqrt{m_3 m_1}\}. \tag{5}$$

Nevertheless, [Y97] is too strong, i.e., [Y97] does not contain a case of wide triple-encounter. This is discussed in the next section.

3 Numerical experiments

Here we survey distributions of escape orbits until the third triple-encounter under the respective criteria in the free-fall problem with equal masses. All possible initial configurations with equal masses are obtained if m_1, m_2, and m_3 are put at $\{(x,y)|x \geq 0, y \geq 0, (x+0.5)^2+y^2 \leq 1\}$, $(-0.5,0)$, and $(0.5,0)$, respectively. The bound arc $(x+0.5)^2+y^2 = 1$ is denoted by $I^{(2)}$. See Tanikawa *et al.* (1995), and Umehara and Tanikawa (1997, 1999) for a precise explanation. The *n-th-escape point* is defined as the initial point of orbits satisfying the escape criterion (Yoshida, 1972, 1974) within an ejection distance $10\rho_0$ after the n-th triple-encounter. A region formed with n-th-escape points will be called an *n-th-escape region.*

In Figs. 1(a), (b), (c), and (d), full-tone and half-tone squares represent grid elements of the first and the second-escape regions by using [AM73], [AZ89], [Y97], and [UT99], respectively. We will compare the distributions with each other by answering the following propositions.

Proposition

(I) There is a infinite sequence of band-like first-escape regions crossing $I^{(2)}$ *such that it accumulates to the corner* $(0.5,0)$.

(II) Band-like escape regions which do not cross $I^{(2)}$ *do not have any first-escape points.*

If [AM73] is adopted, (I) is false and (II) is true. Both [AZ89] and [Y97] show that (I) is true and (II) is false. Only our definition [UT99] satisfies both (I) and (II).

Proposition (I) is true since triple-collision initial-points T_i, $i \in \mathbf{N}$, have been found on $I^{(2)}$ by Tanikawa *et al.* (1995). Escape regions are distributed around each T_i (Umehara and

Tanikawa, 1999). Only Figure 1(a) shows the contrary result. A typical orbit judged as the second-escape point is observed in the lower-right part of the initial-value space. See Fig. 2(a). The upper and the lower frames show the time-dependent r_{ij} and $I(t)$. Here, r_{ij} is the distance between m_i and m_j, $\{i \neq j\} \in \{1,2,3\}$. In the lower frame, a bold arc represents a time-interval with $\ddot{I}(t) \geq 0$. In these intervals, the time indicated by an arrow satisfies $\dot{I} = 0$. A wrong interval is detected during $t \in (0.03, 0.05)$ while one particle is far from the others. From this result, we consider that (I) is true. Figure 2(b) shows the other example of erroneous detection during the ejection of one particle.

Contrary to the results in Figs. 1(b) and (c), Proposition (II) is confirmed by the numerical observations of orbits. In both figures, however, band-like first-escape region containing a segment ℓ $\{(x,y)|x = 0.11, y < 0.1\}$ bends around $y \cong 0.32$. Figure 3(a) shows that the system fails to satisfy both [AZ89] and [Y97] during the true first triple-encounter. The minimal value is greater than I_0 represented by a horizontal line. On the other hand, Figure 3(b) shows that the detection succeeds during the first triple-encounter. The initial point lies in the upper part of ℓ. So, ℓ should be embedded in the second-escape region, and (II) is considered to be true.

4 Additional condition

It is necessary to adapt [AZ89] or [Y97] for wide triple-encounters. During $\ddot{I} > 0$, the size of a system is considered to be smaller than the length scale $\rho_0/2$ in virial equilibrium. So, another larger scale of triple encounter is required.

We consider ρ_0 as the second smallest scale. In the negative-energy system, the distance between the closest two particles is less than ρ_0 even if such a pair may change successively. So we consider that the system is also in a triple encounter when the distant particle approaches the nearest particle within ρ_0.

Acknowledgement

The authors would like to thank Dr. Junzo Yoshida and Dr. Takeshi Nakamura for their helpful discussions.

References

Agekian and Martynova, 1973, Classification of state in the three-body problem (in Russian), Vestn. Leningrad Univ., 122-126.

Agekian, T. A. and Anosova, J. P., 1990, Distribution of minima of the current size of the triple systems and their disruption, *Celestial Mechanics and Dynamical Astronomy*, **49**, 145-149.

Anosova, J. P. and Zavalov, N. N., 1989, States of strong gravitational interaction in the general three-body problem, *Astron. Zh.*, **66**, 152-160; translated into English in *Soviet Astronomy*, **33**, 79-83 (1989).

Sibahara, R. and Yoshida, J., 1963, Erweiterungen der Birkhoff-Mermanschen Methoden im Dreikörperperproblem, *Publ. Astron. Soc. Japan*, **15**, 86-99.

Szebehely, V., 1971, Classification of the motions of three bodies in a plane, *Celestial Mechanics*, **4**, 116-118.

Tanikawa, K., Umehara, H., and Abe, H., 1995, A search for collision orbits in the free-fall three-body problem. I. numerical procedure, *Celestial Mechanics and Dynamical Astronomy*, **62**, 335-362.

Tanikawa, K. and Umehara, H., 1998, Oscillatory orbits in the planar three-body problem with equal masses, *Celestial Mechanics and Dynamical Astronomy*, **70**, 167-180.

Umehara, H., 1997, The free-fall three-body problem: escape and collision, PhD Thesis, The Graduate University for Advanced Studies, Mitaka, Tokyo, 181-8588 Japan.

Umehara, H. and Tanikawa, K., 1997, Orbital distribution arbitrarily close to the homothetic equilateral triple collision in the free-fall three-body problem with equal masses, submitted to *Celestial Mechanics and Dynamical Astronomy.*

Umehara, H. and Tanikawa, K., 1999, Binary and triple collisions causing instability in the free-fall three-body problem, submitted to *Celestial Mechanics and Dynamical Astronomy.*

Yoshida, J., 1972, Improved criteria for hyperbolic-elliptic motion in the general three-body problem, *Publ. Astron. Soc. Japan*, **24**, 391-408.

Yoshida, J., 1974, Improved criteria for hyperbolic-elliptic motion in the general three-body problem. II, *Publ. Astron. Soc. Japan*, **26**, 367-377.

Junzo Yoshida, 1997, private communications.

Zare, K., 1981, Properties of the moment of inertia in the problem of three bodies, *Celestial Mechanics*, **24**, 345-354.

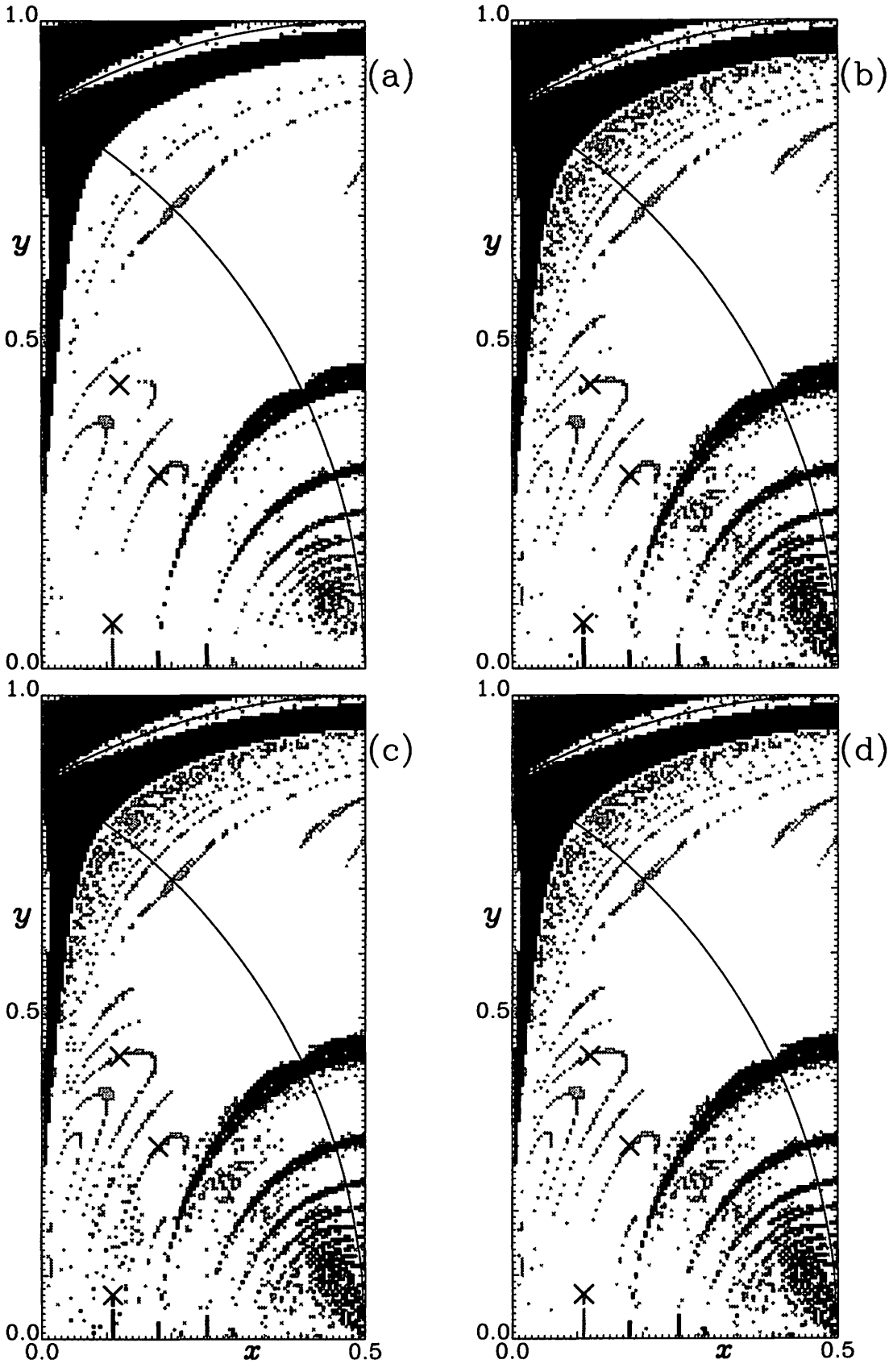


図 1: Initial-value distributions of escape orbits until the second triple-encounter under various definitions by (a) Agekian and Martynova (1973), by (b) Anosova and Zavalov (1989), by (c) J.Yoshida (1997), and by (d) our criterion.

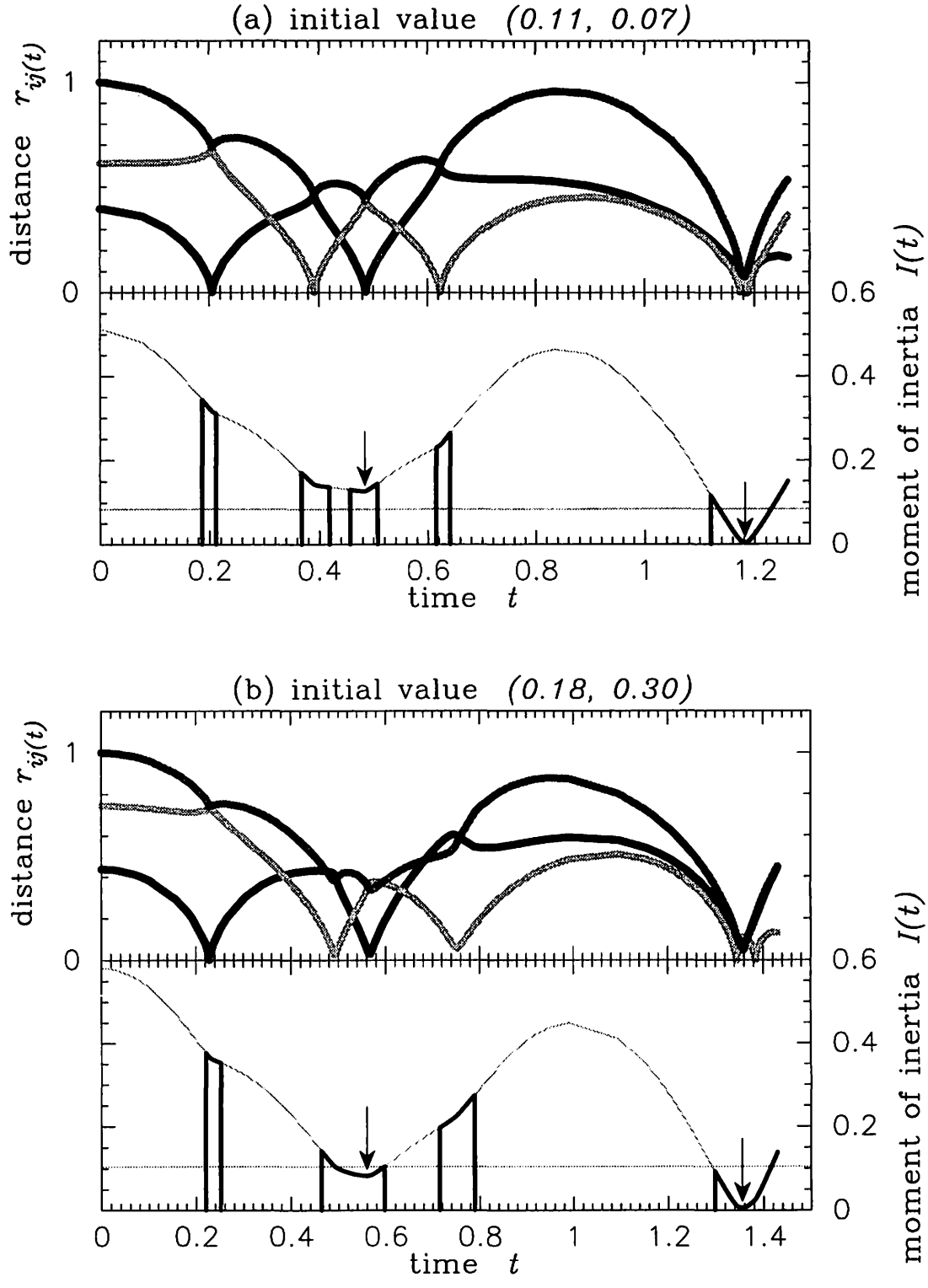


図 2: Time evolution of the distances between the respective particles and the moment of inertia. A full-tone curve stands for r_{31} in the upper frame. A light-gray curve and a dark-gray curve represent r_{12} and r_{23}, respectively. In the lower frame, a bold arc shows a time interval satisfying $\ddot{I}(t) > 0$. An arrow indicates the time of minimal $I(t)$.

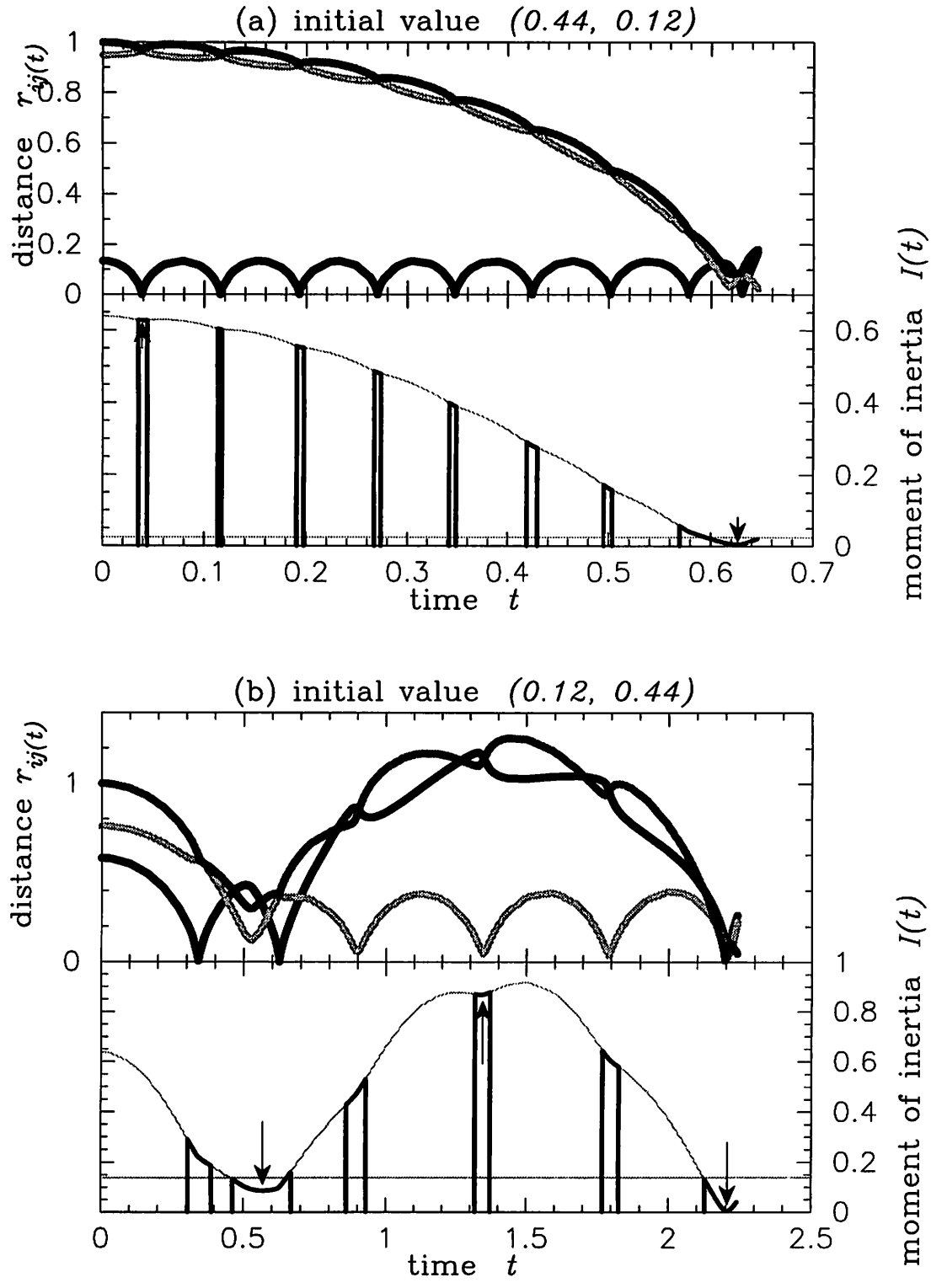


図 3: Time evolution of the distances between the respective particles and the moment of inertia. The initial values are located in the additional band containing the segment ℓ. Notations are the same as those of the preceding figure. In a lower frame, a horizontal line represents a critical value I_0.

Escape Regions near an Equilateral Triangular Equilibrium Point in the Free-Fall Three-Body Problem

Takeshi Nakamura
Okayama Shoka University, 2-10-1 Kyomachi Tusima Okayama, 700-8601, Japan

Abstract. We give the escape regions near an equilateral triangular equilibrium point in the free-fall three-body problem with masses $m_1=m_2=1$ and $m_3=0.5$, 1, 2 and 3, where all orbits in the escape regions have only one triple close approach. When $m_3=3$, between two large escape regions there exist ten kinds of very small escape regions.

1. Introduction

In [1], we found that there existed three kinds of very small escape regions near the equilateral triangular equilibrium point L_4 in the free-fall three-body problem with masses $m_1=3$, $m_2=4$ and $m_3=5$. Here we give the results for the masses $m_1=m_2=1$ and $m_3=0.5$, 1, 2 and 3. Although the case $m_3=1$ was treated by Umehara and Tanikawa [2], we give it for comparison. In spite of the two same masses, we get more escape regions than masses $m_1=3$, $m_2=4$ and $m_3=5$.

2. Initial positions of the three bodies

We use the same initial positions and the same escape condition as given in [1]. Namely, m_1, m_2 are initially placed at the base of an equilateral triangle with side 1 and they stay there. The top vertex of this triangle is Lagrange point L_4. The body m_3 is initially placed on the circle with center L_4 and radius 0.01 (see figure 1). We examine the escape regions on this circle. Let the position of m_3 be (x_1, x_2) and be measured by the angular variable θ. Then

$$x_1=0.01\cos\theta, \qquad x_2=0.01\sin\theta+\sqrt{3}/2. \tag{1}$$

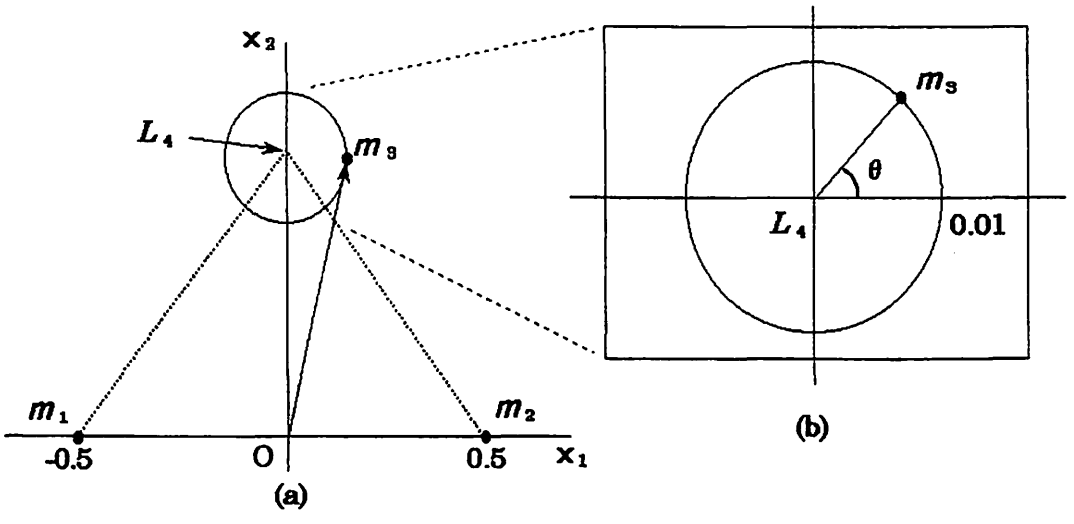


Fig. 1. The initial positions of m_1, m_2 and m_3. In (a), the circle around L_4 is intentionally drawn extendedly. The (b) is more extension.

Since $m_1 = m_2$, it follows that the escape regions are symmetric with respect to the x_2 axis. Therefore, we consider only the right half of this circle and the range of θ is $-90 \leqq \theta \leqq 90$, where θ is measured by degree.

We use also the same condition of the triple close approach as given in [1]. Namely, it is

$$I < I_c \ , \qquad (2)$$

where I is the half of the moment of inertia and $I_c = 6.1274$, 6.1270, 6.1249 and 6.1251 for $m_3 = 0.5, 1, 2$ and 3, respectively. All orbits in our escape regions enter into the region defined by the condition (2) once before they escape.

The initial value of I is $I_0 = 24.5$.

3. Results of the numerical integrations

We show here the escape regions according to $m_3 = 0.5$, 1, 2 and 3.

$m_3 = 0.5$

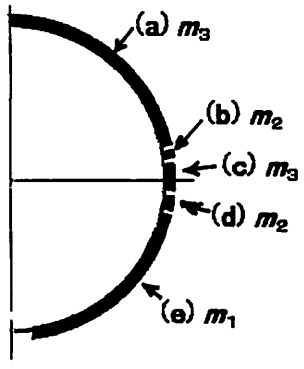


Fig. 2.

Table 1.

Escape regions / Escape bodies	θ (degree)	Escape region/ Half circle
(a) m_3	90.00 ~ 11.35	43.7%
(b) m_2	11.27 ~ 7.06	2.3%
(c) m_3	7.00 ~ -4.00	6.1%
(d) m_2	-6.00 ~ -10.00	2.2%
(e) m_1	-12.00 ~ -86.50	41.4%

$m_3 = 1$

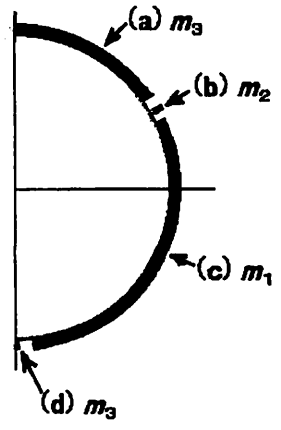


Fig. 3.

Table 2.

Escape regions / Escape bodies	θ (degree)	Escape region/ Half circle
(a) m_3	90.00 ~ 34.00	31.1%
(b) m_2	30.30 ~ 29.30	0.6%
(c) m_1	25.50 ~ -85.50	61.7%
(d) m_3	-89.50 ~ -90.00	0.3%

$m_3 = 2$

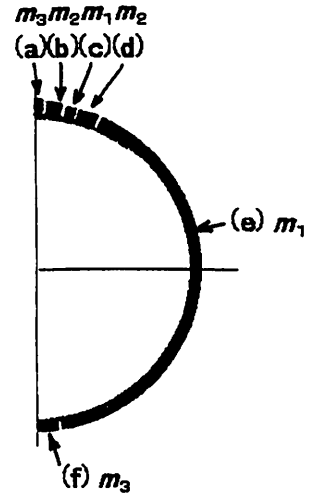


Fig. 4.

Table 3.

Escape regions Escape bodies		θ (degree)	Escape region/ Half circle
(a)	m_3	90.00 ~ 88.30	0.9%
(b)	m_2	88.05 ~ 81.30	3.8%
(c)	m_1	80.90 ~ 75.34	3.1%
(d)	m_2	75.30 ~ 67.74	4.2%
(e)	m_1	67.77 ~ -83.00	38.8%
(f)	m_3	-84.50 ~ -90.00	3.1%

$m_3 = 3$

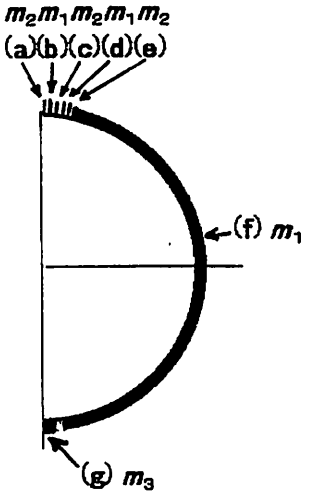


Fig. 5.

Table 4.

Escape regions Escape bodies		θ (degree)	Escape region/ Half circle
(a)	m_2	89.99 ~ 88.04	1.1%
(b)	m_1	87.98 ~ 85.817	1.2%
(c)	m_2	85.814~ 85.465	0.2%
(d)	m_1	85.460~ 33.660	1.0%
(e)	m_2	83.643~ 82.388	0.7%
(f)	m_1	82.385~ -84.70	92.8%
(g)	m_3	-86.80 ~ -90.00	1.8%

Although when $m_3 = 1$, then there exists only one small escape region between two large escape regions, when $m_3 = 0.5$, 2 or 3, then the number of the small escape regions increase. In Figure 5, there exist ten kinds of small escape regions near $\theta = 90$ by symmetry.

In Appendix, we give the orbits in the escape regions.

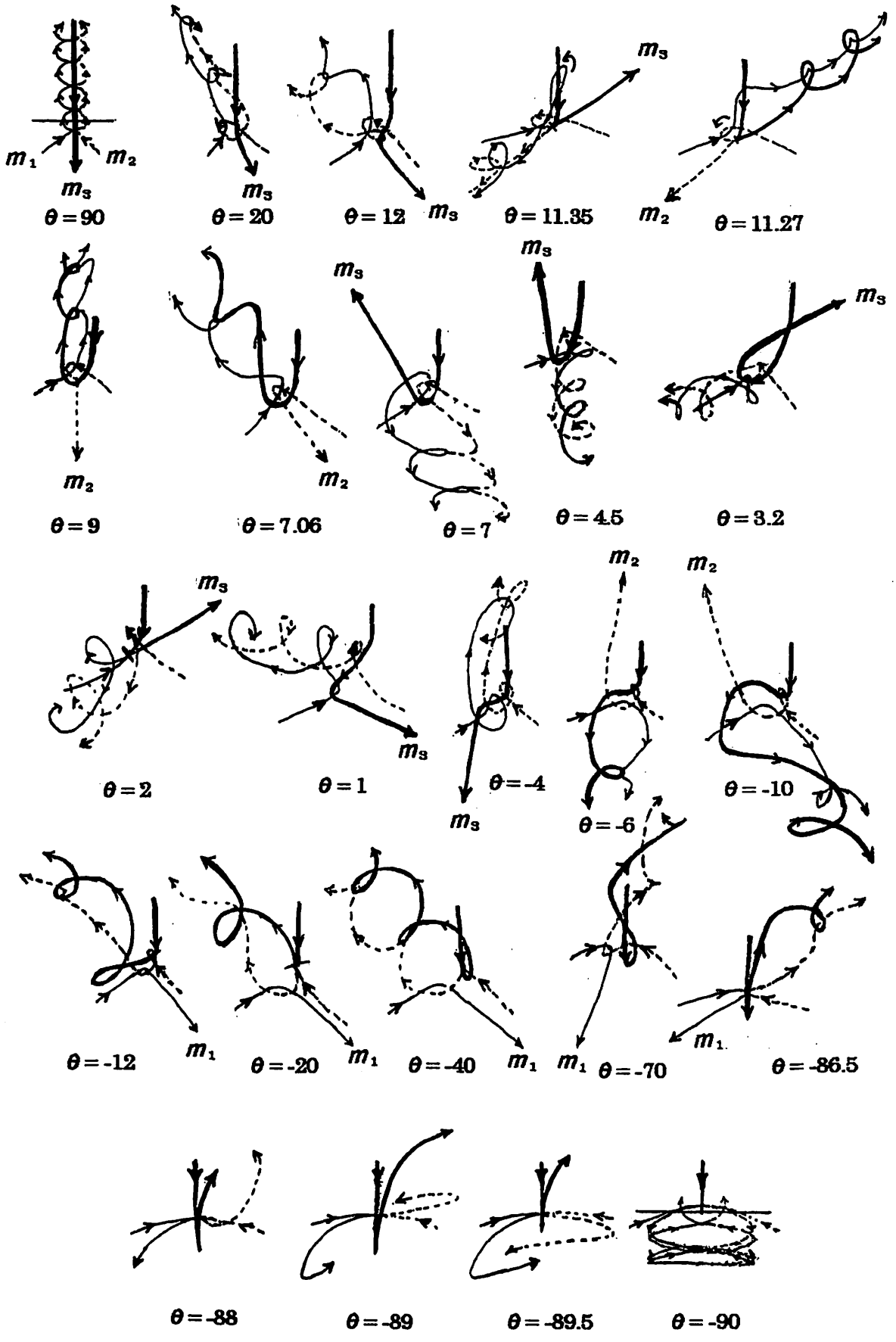
References

[1] Nakamura, T.: 1998, New regions of escape in the free-fall three-body problem, in *Proceedings of the 30th symposium on celestial mechanics*, 147

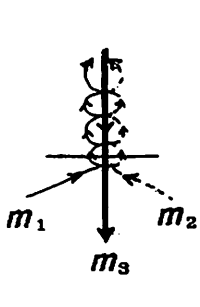
[2] Umehara, H and K. Tanikawa : 1995, Escape probability at the first collapse in the free-fall three-body problem, in *Proceedings of the 27th symposium on celestial mechanics*, 219

Appendix

m_3=0.5



$m_3 = 1$



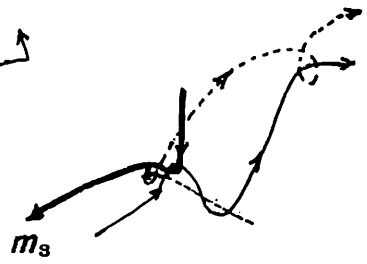
$\theta = 90$

$\theta = 80$

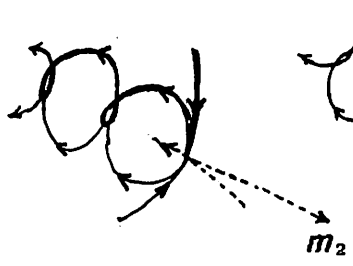
$\theta = 60$

$\theta = 50$

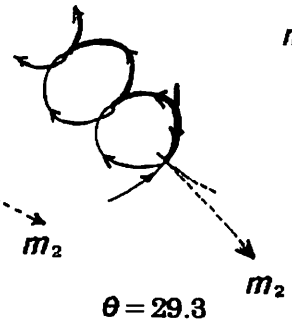
$\theta = 40$



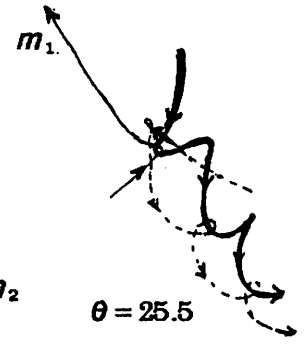
$\theta = 34$



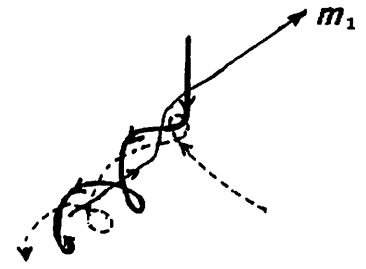
$\theta = 30.3$



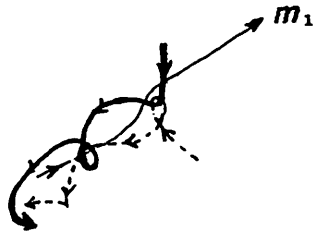
$\theta = 29.3$



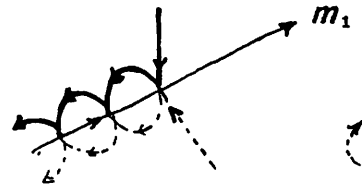
$\theta = 25.5$



$\theta = 10$



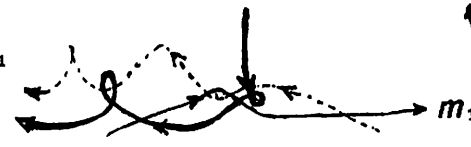
$\theta = 0$



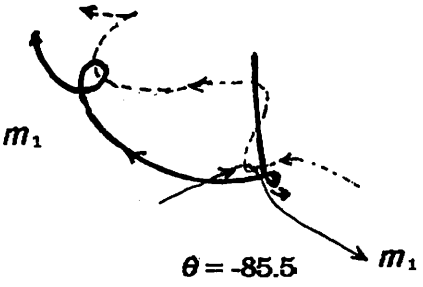
$\theta = -30$

$\theta = -60$

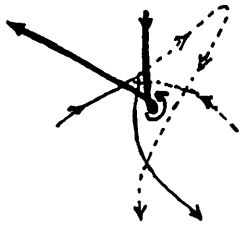
$\theta = -70$



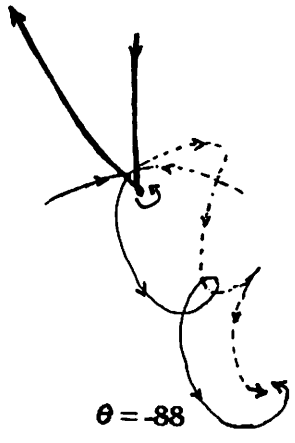
$\theta = -80$



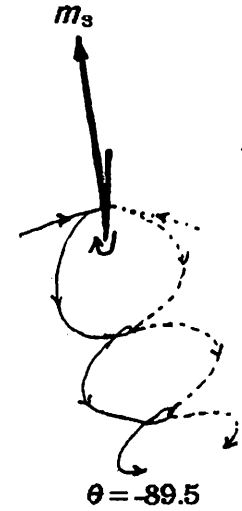
$\theta = -85.5$



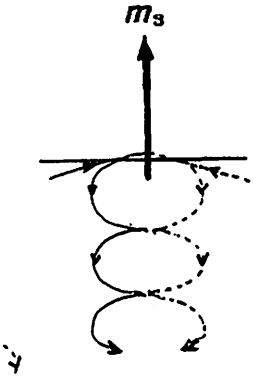
$\theta = -87$



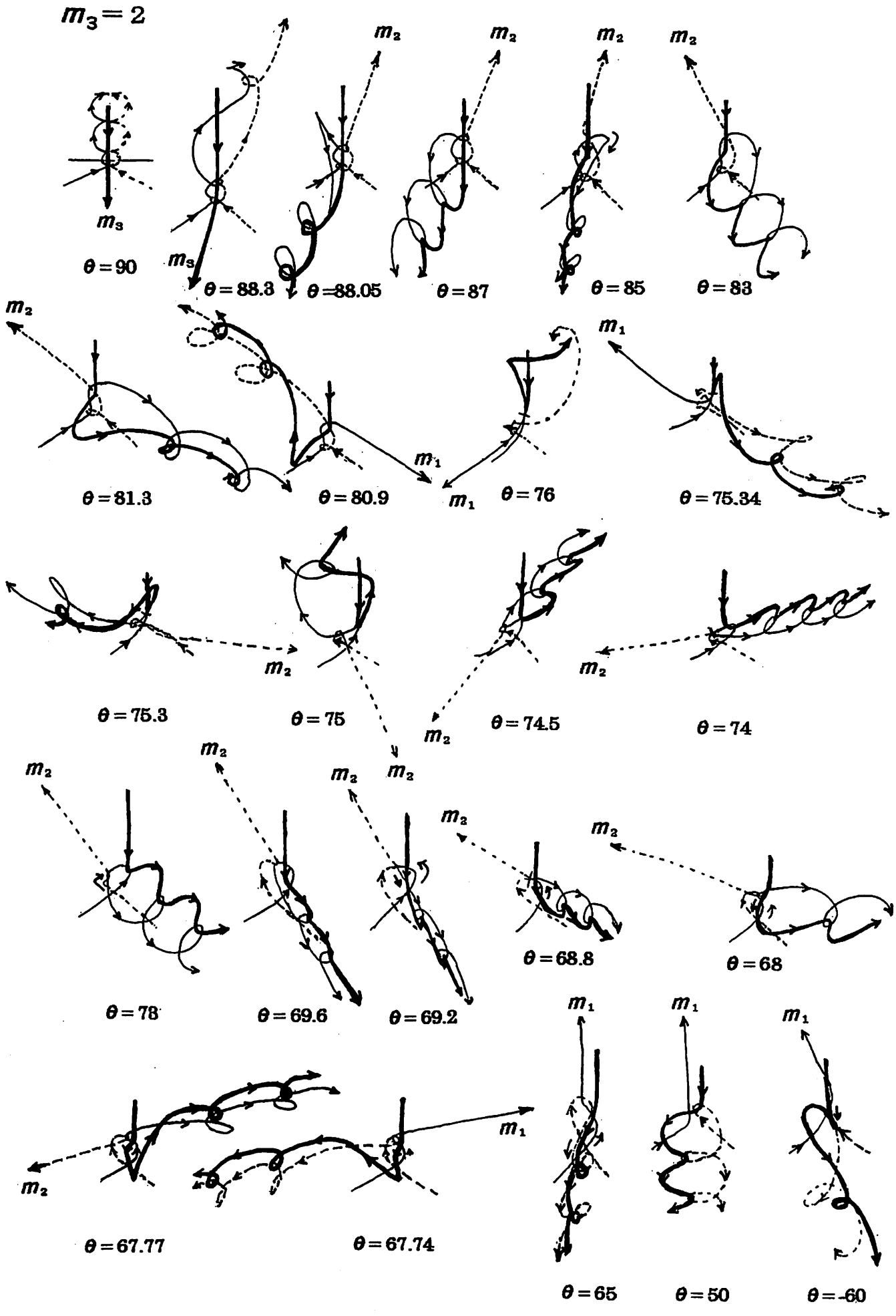
$\theta = -88$

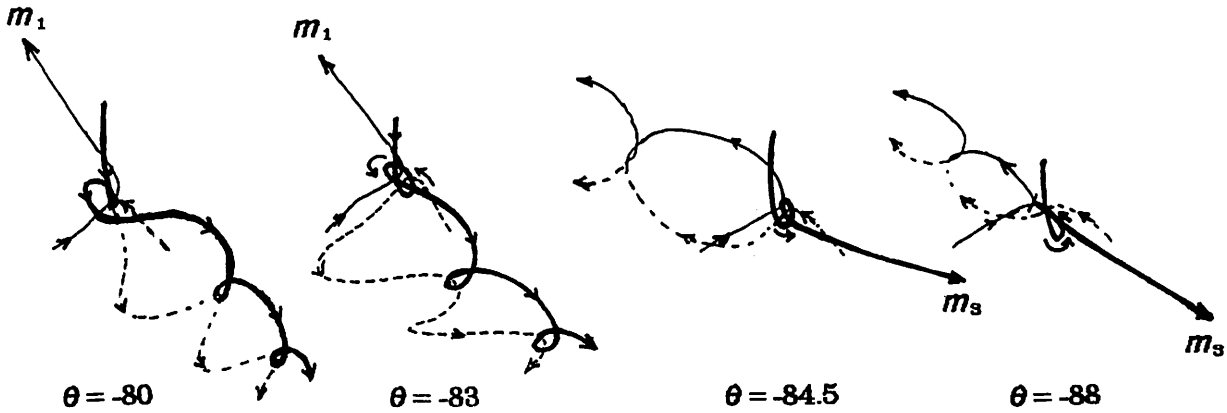


$\theta = -89.5$

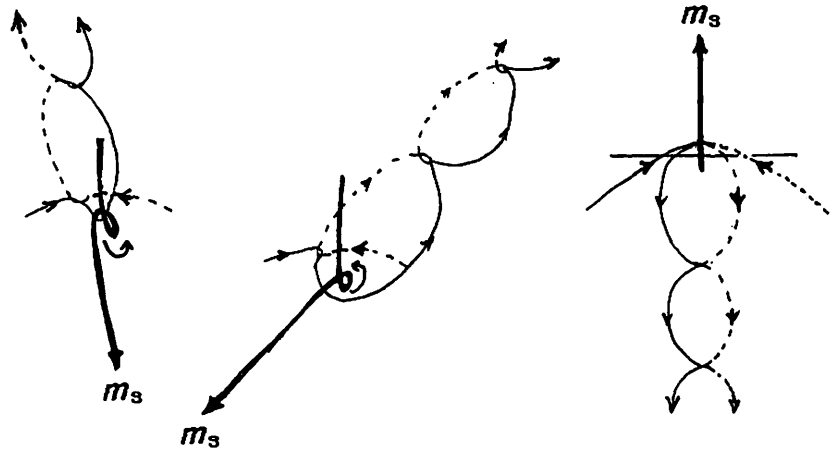


$\theta = -90$


$m_3 = 2$
m_3
$\theta = 90$
m_2
m_3
$\theta = 88.3$
m_2
$\theta = 88.05$
m_2
$\theta = 87$
m_2
$\theta = 85$
m_2
$\theta = 83$
m_2
$\theta = 81.3$
m_1
$\theta = 80.9$
m_1
$\theta = 76$
m_1
$\theta = 75.34$
m_2
$\theta = 75.3$
m_2
$\theta = 75$
m_2
$\theta = 74.5$
m_2
$\theta = 74$
m_2
$\theta = 78$
m_2
$\theta = 69.6$
m_2
m_2
$\theta = 69.2$
m_2
$\theta = 68.8$
m_2
$\theta = 68$
m_2
m_1
$\theta = 67.77$
$\theta = 67.74$
m_1
$\theta = 65$
m_1
$\theta = 50$
m_1
$\theta = -60$

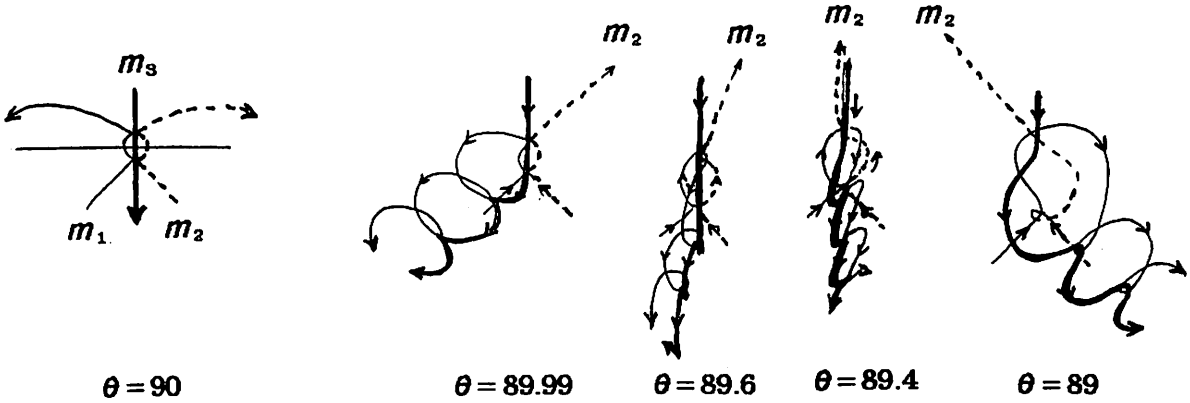


$\theta = -80$ $\theta = -83$ $\theta = -84.5$ $\theta = -88$

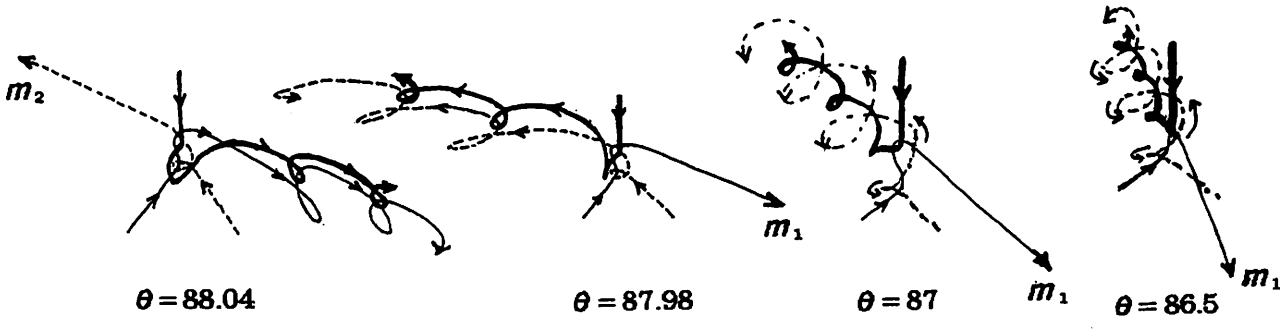


$\theta = -89$ $\theta = -89.5$ $\theta = -90$

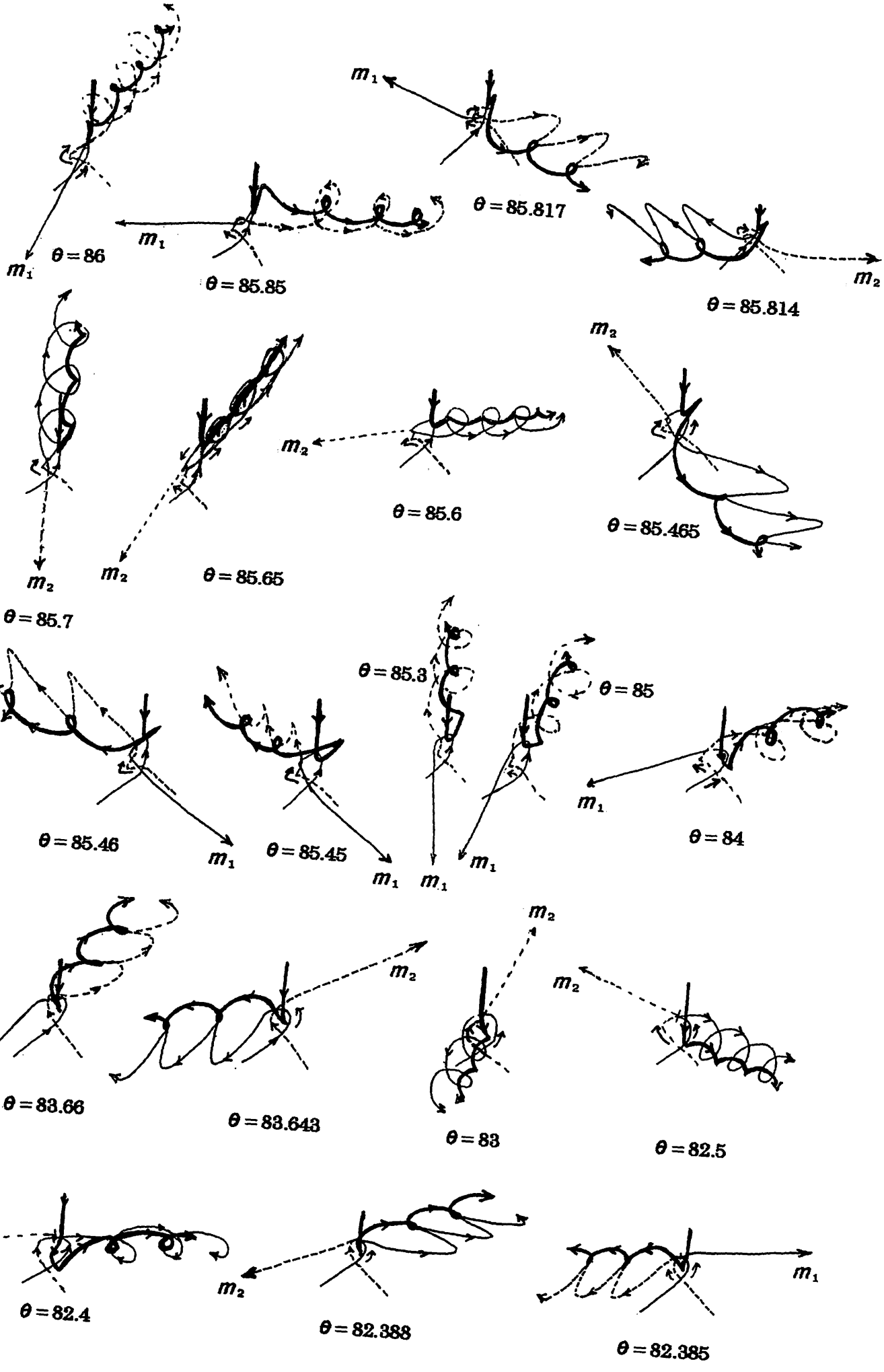
$m_3 = 3$

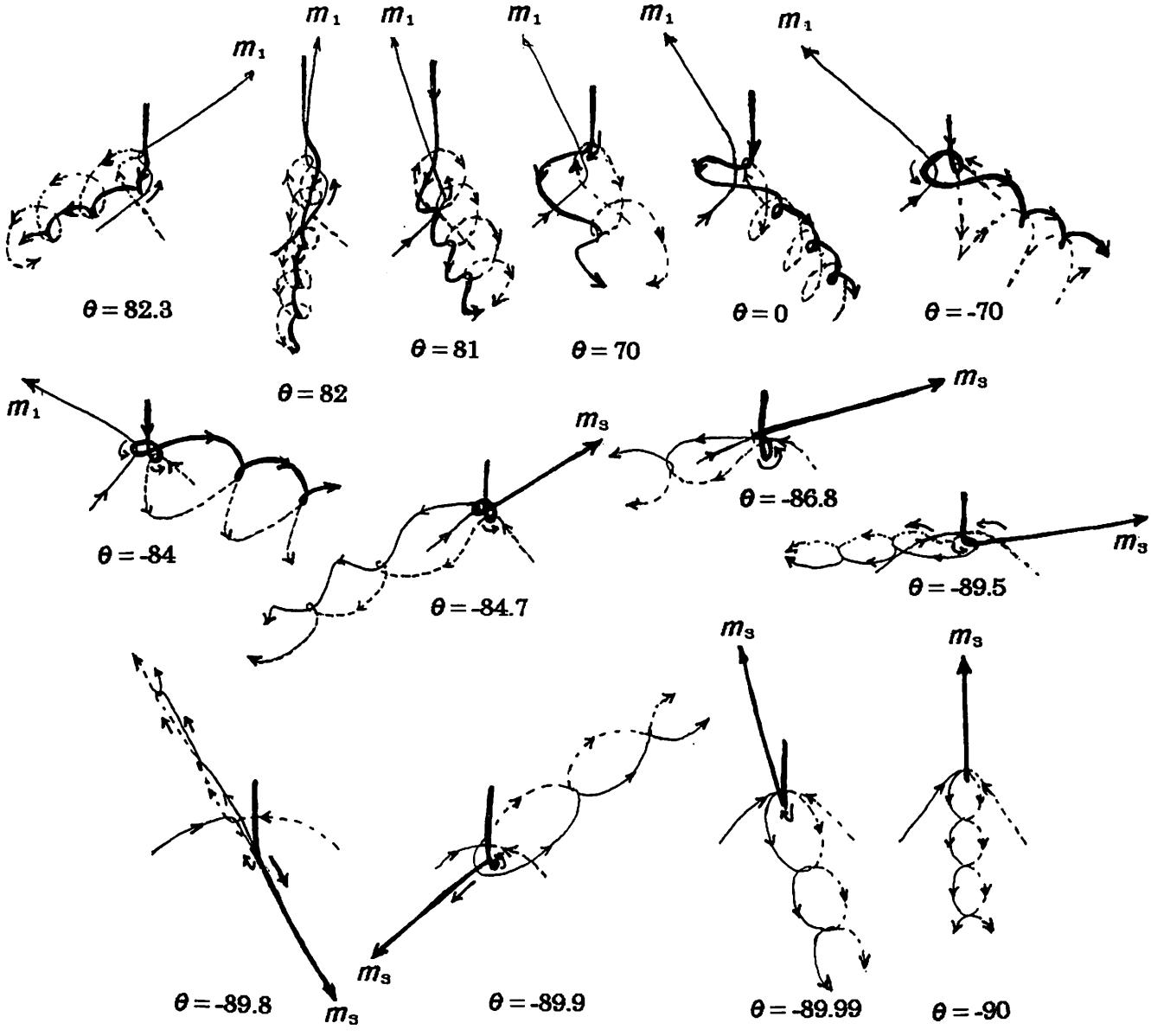


$\theta = 90$ $\theta = 89.99$ $\theta = 89.6$ $\theta = 89.4$ $\theta = 89$

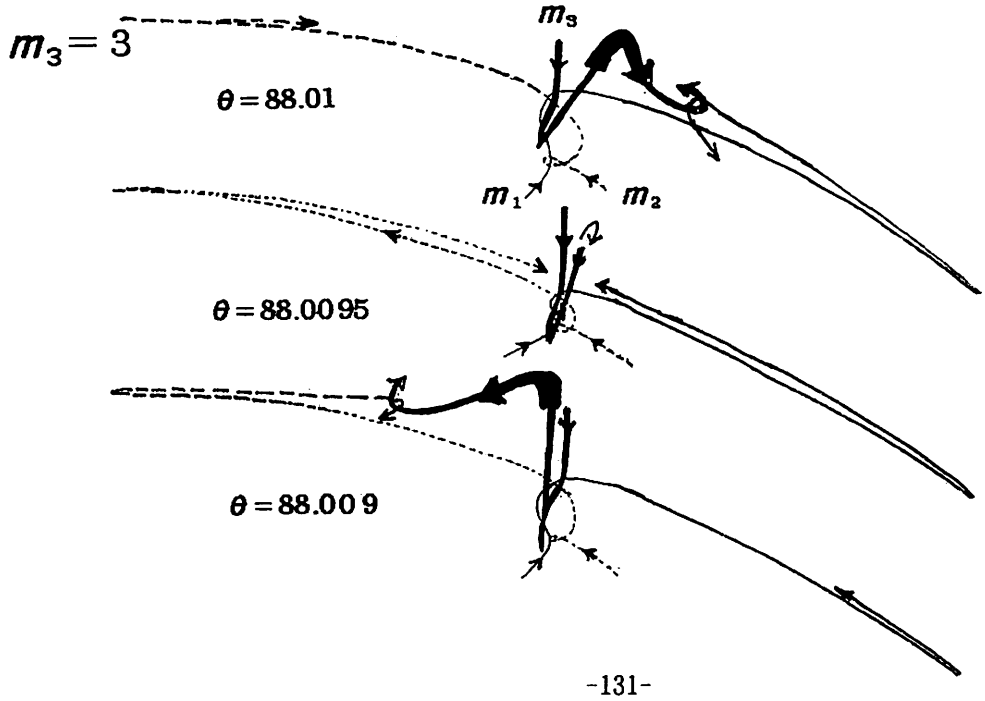


$\theta = 88.04$ $\theta = 87.98$ $\theta = 87$ $\theta = 86.5$


θ=86
θ=85.85
θ=85.817
θ=85.814
θ=85.7
θ=85.65
θ=85.6
θ=85.465
θ=85.46
θ=85.45
θ=85.3
θ=85
θ=84
θ=83.66
θ=83.643
θ=83
θ=82.5
θ=82.4
θ=82.388
θ=82.385



When the escaping bodies change, there exists a second triple close approach. The next figures show this change of the orbits.



New variables for the Free-Fall Three-Body Problem

Masayoshi SEKIGUCHI
Kisarazu National College of Technology,
Kiyomidai-Higashi 2-11-1, 292-0041, JAPAN
masa@yana.kisarazu.ac.jp

Kiyotaka Tanikawa
National Astronomical Observatory
Mitaka, Tokyo, 181-8588, JAPAN
tanikawa@exodus.mtk.nao.ac.jp

Abstract

We introduced a set of new variables describing the planar three-body problem in order to investigate the structure of set of binary collision orbits in the neighbourhood of triple collision orbits in the case of "Free-Fall" initial condition. Our variables blew up the triple-collision singularity and continued analytically the flow beyond singularities like McGehee variables.

1 Introduction

There are observed many interesting phenomena in the free-fall three-body problem by numerical integration ([4][5][6]). Binary collision orbits play an important role to understand the mechnism to create almost interesting features. Our main purpose to introduce a set of new variables is to analyze the differential equations of motion theoretically, especially relation between binary collision and triple collision in the case of non-isosceles and non-collinear problem.

Many methods for theoretical study of singularity in the three-body problem are well-known([1],[2],[3],[7]). However, tools for non-isosceles and non-collinear problem([3],[7]) are confusing to handle. Therefore we propose a new method for study of the planar three-body problem. Remarkable feature of our method is that it regularize triple collision and one binary collision of three ones.

2 Canonical transformations

Hamiltonian function in the planar three-body problem is as below.

$$\mathrm{H}(q,p)=\frac{1}{2}pA^{t}p-\mathrm{U}(q)$$

where $A = \mathrm{diag}\left(\frac{1}{m_1}, \frac{1}{m_1}, \frac{1}{m_2}, \frac{1}{m_2}, \frac{1}{m_3}, \frac{1}{m_3}\right)$

and $\mathrm{U}(q) = \sum_{j=i+1}^{3}\left(\sum_{i=1}^{2}\frac{m_i m_j}{\sqrt{(q_{i,x}-q_{j,x})^2+(q_{i,y}-q_{j,y})^2}}\right)$.

Hamiltonian equations are

$$\frac{dq}{dt} = -\left(\frac{\partial}{\partial p}H\right),\ \frac{dp}{dt} = \frac{\partial}{\partial q}H$$

where $q = (q_{1,x}, q_{1,y}, q_{2,x}, q_{2,y}, q_{3,x}, q_{3,y})$ and $p = (p_{1,x}, p_{1,y}, p_{2,x}, p_{2,y}, p_{3,x}, p_{3,y})$. In this system degree of freedom is 6. This system of differential equations has five integrals independent to Hamiltonian which are coordinate and momentum of their gravity center and total angular momentum. The behavior of ϕ : angle variable conjugate to angular momentum is determined by behavior of other variables, but does not contribute to the other's behavior. Therefore differential equation for ϕ can be seperated from the remaining equations. We assume that the angular momentum is zero in this paper. Thus we obtain a Hamiltonian system with 3 degrees of freedom from original system through canonical transformations.

Hamiltonian function is

$$\mathrm{H}(Q,P) = \frac{1}{2}P\,\mathrm{A}(Q)^t P - \mathrm{U}(Q)$$

where

$$\mathrm{A}(Q) = \begin{bmatrix} \nu_1+\nu_3 & -\frac{\nu_3\sin(\theta)}{r_2} & \nu_3\cos(\theta) \\ -\frac{\nu_3\sin(\theta)}{r_2} & \frac{\nu_1+\nu_3}{r_1^2}+\frac{\nu_2+\nu_3}{r_2^2}-\frac{2\nu_3\cos(\theta)}{r_1 r_2} & -\frac{\nu_3\sin(\theta)}{r_1} \\ \nu_3\cos(\theta) & -\frac{\nu_3\sin(\theta)}{r_1} & \nu_2+\nu_3 \end{bmatrix},$$

$$\mathrm{U}(Q) = m\left(\frac{\nu_2}{r_1}+\frac{\nu_1}{r_2}+\frac{\nu_3}{\sqrt{r_1^2+r_2^2-2r_1 r_2\cos(\theta)}}\right),$$

with canonical variables $Q = (r_1, \theta, r_2)$ and canonical momentums $P = (p_{r_1}, p_\theta, p_{r_2})$. In above expressions, ν_i means mass ratio $(m_1+m_2+m_3)/m_i$, m means $\sqrt{m_1+m_2+m_3}$ and r_i means distance between m_i and m_3.

Next we introduce a set of variables r and ψ defined by the following equations, i.e.

$$r = \sqrt{r_1^{\,2} + r_2^{\,2}},$$
$$r_1 = r\cos(\psi),$$
$$r_2 = r\sin(\psi).$$

View of these variables is given as below.

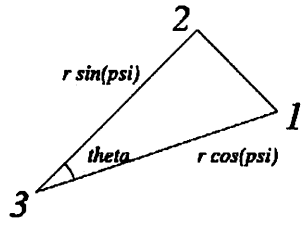


Fig.1 View of Variables

If r becomes zero, then there occurs simultaneous colision with three particles, which we call "triple collision". If particle 3 goes far away from other particles, then r tends to infinity. In this meaning, r reflects size of the triangle formed by three bodies. If ψ becomes zero(right angle), then there occurs collision with particles 1(2) and 3, which we call "binary collisoin" of 1(2) and 3. If $\theta = \pi/3$ and $\psi = \pi/4$, then the triangle is equilateral. θ and ψ determines form of the triangle in this manner. We present a diagram in which θ and ψ indicates form of the triangle as below. This diagram is a projection of configuration space onto $\theta - \psi$ plane.

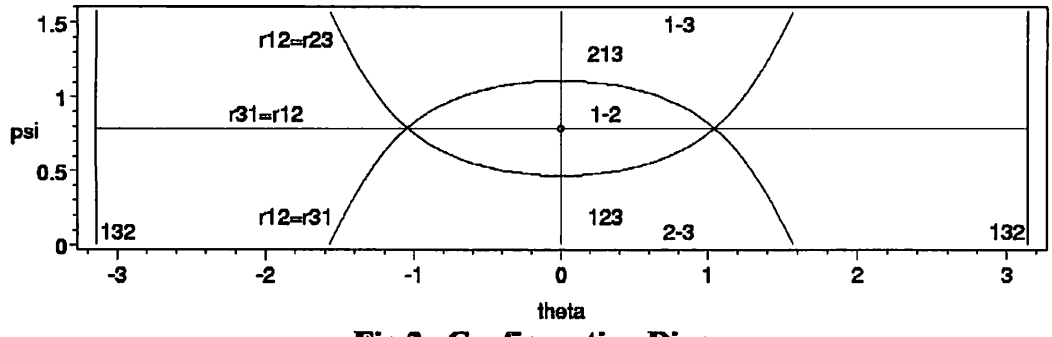


Fig.2 Configuration Diagram

Curve ($\cos(\psi) = 2\cos(\theta)\sin(\psi)$) near string "*r12 = r23*" indicates isosceles subsystem when $v_1 = v_3$. The other strigs "*rij = rjk*" are used in the same way. Segment ($\theta = \pi$ and $\theta = -\pi$) near string "132" indicates collinear subsystem in the order of m_1, m_3, m_2. The other strings "*ijk*" are used in the same way. Segment near string "1-3"($\psi = \pi/2$), segment near string "2-3" ($\psi = 0$) and point ($\theta = 0, \psi = \pi/4$) indicate binary collision respectively.

Canonical momentums p_r and p_ψ conjugate to r and ψ can be found easily. The definition is as follows.

$$\begin{bmatrix} p_r \\ p_\psi \end{bmatrix} = \begin{bmatrix} \cos(\psi) & \sin(\psi) \\ -r\sin(\psi) & r\cos(\psi) \end{bmatrix} \begin{bmatrix} p_{r_1} \\ p_{r_2} \end{bmatrix}$$

Hence we obtain the following Hamiltonian , i.e.

$$H = \frac{1}{2}[p_r \quad p_\theta \quad p_\psi]\, A(r,\theta,\psi) \begin{bmatrix} p_r \\ p_\theta \\ p_\psi \end{bmatrix} - \frac{V(\theta,\psi)}{r\sin(\psi)\cos(\psi)},$$

where

$$A(r,\theta,\psi) = \begin{bmatrix} \rho_1 & -\dfrac{2\nu_3\sin(\theta)}{r\sin(2\psi)} & \dfrac{\rho_2}{r} \\ -\dfrac{2\nu_3\sin(\theta)}{r\sin(2\psi)} & \dfrac{4\rho_3}{r^2\sin(2\psi)^2} & 0 \\ \dfrac{\rho_2}{r} & 0 & \dfrac{\rho_3}{r^2} \end{bmatrix},$$

$$V(\theta,\psi) = m\left(\nu_1\cos(\psi) + \nu_2\sin(\psi) + \frac{\nu_3\sin(\psi)\cos(\psi)}{\sqrt{1-\cos(\theta)\sin(2\psi)}}\right),$$

$$\rho_1 = \nu_1\cos(\psi)^2 + \nu_2\sin(\psi)^2 + \nu_3(1+\cos(\theta)\sin(2\psi)),$$

$$\rho_2 = \frac{(\nu_2-\nu_1)\sin(2\psi)}{2} + \nu_3\cos(\theta)\cos(2\psi),$$

$$\rho_3 = \nu_1\sin(\psi)^2 + \nu_2\cos(\psi)^2 + \nu_3(1-\cos(\theta)\sin(2\psi)).$$

3 Regularization

We show the last stage of transformation: McGhee-like regularization. Triple collision singularity and two of binary collision singularity are simultaneously blew up by our transformation. New variable $z = (z_1, z_2, z_3)$ and regularized time are defined as below.

$$z_1 = p_r\sqrt{r\sin(\psi)\cos(\psi)},$$

$$z_2 = \frac{p_\theta}{\sqrt{r \sin(\psi) \cos(\psi)}},$$

$$z_3 = p_\psi \sqrt{\frac{\sin(\psi) \cos(\psi)}{r}},$$

$$ds = dt \sqrt{r \sin(\psi) \cos(\psi)}^{(-3)}.$$

These variables are no longer canonical. We can obtain the follwing system of differential equations after enormous calculations.

(1.1) $$\frac{dr}{ds} = r \sin(\psi) \cos(\psi) (\rho_1 z_1 - \nu_3 \sin(\theta) z_2 + \rho_2 z_3),$$

(1.2) $$\frac{d\theta}{ds} = -\nu_3 \sin(\theta) z_1 + \rho_3 z_2,$$

(1.3) $$\frac{d\psi}{ds} = \sin(\psi) \cos(\psi) (\rho_2 z_1 + \rho_3 z_3),$$

(1.4) $$\frac{dz_1}{ds} = \frac{1}{2} z Z_1^{\,t} z - V(\theta, \psi) \sin(\psi) \cos(\psi),$$

(1.5) $$\frac{dz_2}{ds} = \frac{1}{2} z Z_2^{\,t} z + \left(\frac{\partial}{\partial \theta} V(\theta, \psi) \right),$$

(1.6) $$\frac{dz_3}{ds} = \frac{1}{2} z Z_3^{\,t} z + \left(\frac{\partial}{\partial \psi} V(\theta, \psi) \right) \sin(\psi) \cos(\psi) - V(\theta, \psi) (\cos(\psi)^2 - \sin(\psi)^2),$$

$$V(\theta, \psi) = m \left(\nu_1 \cos(\psi) + \nu_2 \sin(\psi) + \frac{\nu_3 \sin(\psi) \cos(\psi)}{\sqrt{1 - \cos(\theta) \sin(2\psi)}} \right),$$

$$Z_1 = \begin{bmatrix} \frac{\rho_3 \sin(2\psi) + 2\nu_3 \cos(\theta)}{2} & -\frac{3 \sin(2\psi) \nu_3 \sin(\theta)}{4} & \frac{\rho_2 \sin(2\psi) - \rho_4}{4} \\ -\frac{3 \sin(2\psi) \nu_3 \sin(\theta)}{4} & \rho_3 \sin(2\psi) & 0 \\ \frac{\rho_2 \sin(2\psi) - \rho_4}{4} & 0 & \rho_3 \sin(2\psi) \end{bmatrix},$$

$$Z_2 = \begin{bmatrix} \nu_3 \sin(2\psi)\sin(\theta), & \dfrac{(\rho_3 - 2\rho_1)\sin(2\psi) + 6\nu_3\cos(\theta)}{4}, & \nu_3 \sin(\theta)\cos(2\psi) \\ \dfrac{(\rho_3 - 2\rho_1)\sin(2\psi) + 6\nu_3\cos(\theta)}{4}, & -\dfrac{\nu_3 \sin(2\psi)\sin(\theta)}{2}, & \dfrac{\rho_4 - 3\rho_2 \sin(2\psi)}{4} \\ \nu_3 \sin(\theta)\cos(2\psi), & \dfrac{\rho_4 - 3\rho_2 \sin(2\psi)}{4}, & -\nu_3 \sin(2\psi)\sin(\theta) \end{bmatrix},$$

$$Z_3 = \begin{bmatrix} -\rho_2 \sin(2\psi), & -\nu_3 \sin(\theta)\cos(2\psi), & \dfrac{(2\rho_1 - 3\rho_3)\sin(2\psi) - 2\nu_3\cos(\theta)}{4} \\ -\nu_3 \sin(\theta)\cos(2\psi), & \rho_3\cos(2\psi) - \dfrac{\rho_4}{2}, & \dfrac{\nu_3 \sin(\theta)\sin(2\psi)}{4} \\ \dfrac{(2\rho_1 - 3\rho_3)\sin(2\psi) - 2\nu_3\cos(\theta)}{4}, & \dfrac{\nu_3 \sin(\theta)\sin(2\psi)}{4}, & \dfrac{3\rho_2 \sin(2\psi) + \rho_4}{2} \end{bmatrix},$$

where $\rho_1 = \nu_1 \cos(\psi)^2 + \nu_2 \sin(\psi)^2 + \nu_3 (1 + \cos(\theta)\sin(2\psi))$,

$$\rho_2 = \frac{(\nu_2 - \nu_1)\sin(2\psi)}{2} + \nu_3 \cos(\theta)\cos(2\psi),$$

$\rho_3 = \nu_1 \sin(\psi)^2 + \nu_2 \cos(\psi)^2 + \nu_3 (1 - \cos(\theta)\sin(2\psi))$,

$\rho_4 = 2((\nu_1 + \nu_3)\sin(\psi)^2 - (\nu_2 + \nu_3)\cos(\psi)^2)$

If value of the Hamiltonian is h, then energy relation becomes

$$(2)\quad 2h\, r \sin(\psi)\cos(\psi) = [z_1 \quad z_2 \quad z_3] \begin{bmatrix} \rho_1 & -\nu_3 \sin(\theta) & \rho_2 \\ -\nu_3 \sin(\theta) & \rho_3 & 0 \\ \rho_2 & 0 & \rho_3 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} - 2\,V(\theta, \psi).$$

This equation define 5-dimensional manifold in the phase space for some values of h.

4 Subsystems

There exist some subsystems in the regularized system (1.1-6) and (2).

4.1 Isosceles system

The system (1.1-6) is called isosceles problem if $\nu_1 = \nu_3$, $\cos(\psi) = 2\cos(\theta)\sin(\psi)$ and $z_1 = 2(z_2 \sin(\theta) - z_3 \cos(\theta))$, or if $\nu_2 = \nu_3$, $\sin(\psi) = 2\cos(\theta)\cos(\psi)$ and $z_1 = 2(z_2 \sin(\theta) - z_3 \cos(\theta))$, In both case the system (1.1-6) becomes 4 order differential equations. If value of h is fixed, energy relation (2) and these isosceles condition restrict $(r, \theta, \psi, z_1, z_2, z_3)$ to 3-manifold on which (z_1, z_2, z_3) is a point of 1-dimensional circle S

for each value of (r, θ) .Variable r is on a semi-infinite interval $(0,\infty]$, variable θ is on an open interval $(-\pi/2, \pi/2)$. Therefore equi-energy surface is homeomorphic to the manifold which is a direct product of sphere S^2 and $(0,\infty]$.

We omit the case $\nu_1 = \nu_2$ because there neccesarily occur binary collision between m_1 and m_2 which is not regularized in this work.

4.2 Collinear system

The system (1.1-6) is called collinear problem if $(\theta, z_2) = (0, 0)$ or $(\theta, z_2) = (\pi, 0)$, but we omit the case $\theta = 0$ because there neccesarily occur binary collision between m_1 and m_2. When $(\theta, z_2) = (\pi, 0)$, ψ is on an open interval $(0,\pi/2)$. Variable r is on a semi-infinite interval $(0,\infty]$. (z_1, z_3) is on 1-dimensional circle S for each value of (r, ψ) which is defined by energy relation (2) restricted to a hyper-surface $(\theta, z_2) = (\pi, 0)$. Therefore equi-energy surface is homeomorphic to the manifold which is a direct product of circle S, $(0,\infty]$ and $(0,\pi/2)$.

4.3 Binary collsions

The system (1.1-6) is no longer singular on binary collisions except for one between m_1 and m_2. If $\psi = 0$ or π, then $\frac{d\psi}{ds} = 0$ and $\frac{dr}{ds} = 0$. In these cases the system (1.1-6) is reduced to a fictitious system with order 4. $V(\theta, \psi)$ is constant for regularized time s. (z_1, z_2, z_3) is on a 2-sphere for each value of θ. Therefore the energy relation (2) is reduced to equation of 3-manifold which is a direct product of 2-sphere and $[-\pi, \pi)$. This manifold is called binary collision manifold.

4.4 Triple collision

The system (1.1-6) is no longer singular on triple collision $r = 0$ (except for $\theta = 0$) and is reduced to a fictitious system with order 5. The energy relation (2) is reduced to an equation of a 4-manifold which is called tripple collision manifold, i.e. a direct product of $(0,\pi/2)$,$(0,2\pi)$ and a 2-sphere. Another manifold obtained as $h = 0$ is completely same as the triple collision manifold.

5 Zero Velocity Manifold

The initial condition for free-fall three-body problem is $(z_1, z_2, z_3) = (0, 0, 0)$ which defines a 2-manifold in the phase-space. We present a view of the manifold(hereafter referred to as *ZVM*) in Fig.3.

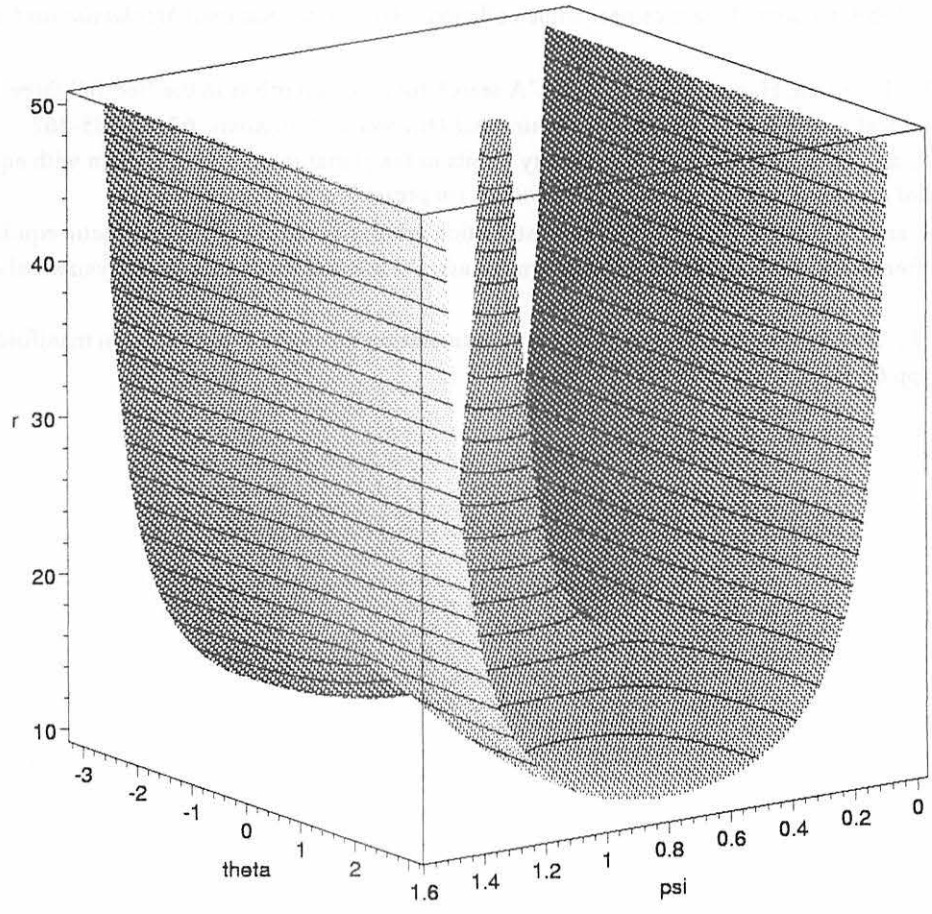


Fig.3. Zero velocity manifold

In the equi-mass case. Solid curves indicate contours.
r goes to infinity where (θ, ψ) implies binary collision.

Let us call a set of orbits which experience binary collision between m_i and m_j in the past or the future B_{ij}. Let us call a set of orbits which experience triple collision in the past or the future T. Our main concern is a geometrical relation between B_{ij} and T, especially in non-isosceles and non collinear region. Intuitionally it is trivial that T is an intersection of B_{ij} s. However, it is not easy to prove. Moreover, there are observed many odd phenomena by numerical investigations([4],[5],[6]) , for instance fractal structure of B_{ij} accumulating to Ton ZVM and escape orbits observed in the fractal structure. Now we continue to study analytically the system of differential equations (1.1-6) in order to make the structure clear.

References

[1] Devaney, R.L., 1980,'Triple collision in the planar isosceles three body problem',*Inventiones Mathematicae*, **60**, pp.249-267.

[2] McGehee, R., 1974,'Triple Collision in the collinear three-body problem',*Inventiones Mathematicae*, **27**, pp.191-227.
[3] Moeckel, R., 1989,'Chaotic dynamics near triple collision', *Archive for Rational Mechanics and Analysis*, **107**, pp.37-69.
[4] Tanikawa, K., Umehara, H. and Abe, H., 1995, 'A search for collision orbits in the free-fall three-body problem. I. numerical procedure', *Celestial Mechanics and Dynamical Astronomy*, **62**, pp.335-362.
[5] Tanikawa, K. and Umehara, H.,1998,'Oscillatory Orbits in the planar three-body problem with equal masses', *Celestial Mechanics and Dynamical Astronomy*, (in press).
[6] Umehara, H. and Tanikawa, K. 1997,'Orbital distribution arbitrarily close to the homothetic equilateral triple collision in the free-fall three-body problem with equal masses',*Celestial Mechanics and Dynamical Astronomy*, (in press).
[7] Waldvogel, J., 1982,'Symmetric and regularized coordinates on the plane triple collision manifold',*Celestial Mechanics*, **28**, pp.69-82.

Symbolic dynamics in one-dimensional three-body problem

Kiyotaka TANIKAWA [1] **and Seppo MIKKOLA** [2]

National Astronomical Obsevatory, Mitaka, Tokyo, 181 Japan

Abstract

Symbolic dynamics is applied to the one-dimensional three-body problem with equal masses. The sequence of collisions along an orbit is expressed as a symbol sequence of three symbols. Triple-collision orbits are found as the boundary of sets of orbits with different symbol sequences in the surface of section. Orbits which start and end in triple collision are obtained. Triple-collision orbits turn out to form the skeleton of the structure of the phase space.

1. Introduction and Idea

One-dimensional three-body problem is one of general three-body problems where three mass points are confined in a line fixed in space. The problem is of two degrees freedom and is known to be non-integrable. Triple systems exhibit chaotic behavior due to the existence of triple collision. The behavior of orbits near triple collision has been analysed via McGehee's regularizing variables (McGehee, 1974).

One-dimensional three-body systems starting from general initial conditions have been extensively studied by Mikkola and Hietarinta(1989, hereafter referred to as MH89; 1990; 1991). They organized numerical results as functions of continuous time until escape. Here, escape means the disruption of a triple system. They obtained rather random than chaotic orbit structure in their surface of section. The sea of chaos remained unresolved(see Fig.1 of MH89).

Particles in one-dimensional three-body problem necessarily repeat binary collisions unless they end in triple collision. This is in contrast to higher dimensional cases(see, for the treatment of collision in two-dimensional cases, Tanikawa *et al.*, 1995; Tanikawa & Umehara, 1998). We may classify the history of motion of three particles referring to the itinerary of binary collisions along orbits. If we denote collisions of pairs of particles by symbols, we get a symbol sequence for each orbit. Then we may apply the standard technique of symbolic dynamics to our problem and may get a meaningful result. This is the main idea. We are going to show that this is actually the case to a certain extent.

One symbol sequence specifies one type of orbits which have the same sequence of collisions. Let us say that orbits with a symbol sequence belongs to the same *topological type*. It is obvious that an infinite number of orbits belong to a topological type. With

[1]email: tanikawa@cc.nao.ac.jp

[2]Permanent address: Tuorla Observatory, University of Turku, 21500 Piikkiö, Finland

regard to the fineness of classification of orbits, topological type stands bewteen the finest classification of orbits as a function of continuous t and the classification of orbits on surface of section. This can be seen by looking at the treatment of independent variable t in these classifications. Both in symbol sequence and in surface of section, time is discrete. As we will see later, not every symbol in a sequence corresponds to a point in the surface of section.

The notion of final motion itself is of topological character. So our approach is expected to work well because we are interested mainly in the final motion of triple systems.

In Section 2, we briefly introduce the Hamiltonian and the variables to be used in later sections. In section 3, symbol sequences are introduced. Then the relationship between the surface of section and symbol sequences is discussed. In section 4, numerical resutls are exhibited. Triple-collision orbits are found automatically as boundaries of orbits with different symbol sequences. Section 5 gives the conclusion.

2. Formulation of the Problem

We put three mass points m_1, m_0, and $m_2 (m_0 = m_1 = m_2)$ in this order on a line. Fix the masses and the gravitational constant to one. Then the Hamiltonian of the problem is given (MH89) by

$$H = \frac{1}{2}\sum_{i=0}^{2} w_i^2 - \sum_{i<j} \frac{1}{|x_i - x_j|}. \tag{1}$$

where w_i are momenta conjugate to the coordinates x_i on the line. We have $x_1 \leq x_0 \leq x_2$. Introducing new coordinates by

$$\begin{aligned} q_1 &= x_0 - x_1, \\ q_2 &= x_2 - x_0, \end{aligned} \tag{2}$$

we get the new Hamiltonian as

$$H = p_1^2 + p_2^2 - p_1 p_2 - \frac{1}{q_1} - \frac{1}{q_2} - \frac{1}{q_1 + q_2}. \tag{3}$$

We fix the total energy to -1 and start the integration at $q_1(0) = q_2(0) = R$. This means that two outer particles are placed in an equal distance from the central. Then the value of potential is fixed to $2.5/R$. The kinetic energy T is determined by

$$T = 2.5/R - 1. \tag{4}$$

If we introduce a parametrization

$$\begin{aligned} \sqrt{3}(p_1 - p_2) &= 2\sqrt{T}\sin\theta, \\ (p_1 + p_2) &= 2\sqrt{T}\cos\theta, \end{aligned} \tag{5}$$

then (θ, R) specify the initial value. Velocities are given by R and θ as

$$\begin{aligned} \dot{q}_1 &= 2p_1 - p_2 = 2\sqrt{T}\cos(\theta - \pi/3) \\ \dot{q}_2 &= 2p_2 - p_1 = 2\sqrt{T}\cos(\theta + \pi/3). \end{aligned} \tag{6}$$

3. Symbolic Dynamics

3.1. Symbol sequence and shift operator

There are three types of collision: binary collisions between m_1 and m_0 and between m_0 and m_2 and triple collision. Let us denote these collisions by symbol $1, 2$, and 0, respectively. Let us express an orbit as a sequence of $0, 1$ and 2 as follows:

$$(\cdots n_{-3}n_{-2}n_{-1}.n_1n_2n_3\cdots)$$

where $n_i, i \in \mathbf{Z}, i \neq 0$ are either $0, 1$, or 2.

We follow the orbit starting from the initial condition defined in §2 to the future and to the past. n_1 represents the first binary collision, n_2 the second, n_3 the third, and so on. Similarly, $n_{-1}, n_{-2}, n_{-3}, \ldots$ represent the type number of the past binary collisions. For triple-collision orbits, symbol sequences terminate within finite digits. Then, symbol sequences corresponding to these orbits read either $(0\cdots n_{-3}n_{-2}n_{-1}.n_1n_2n_3\cdots)$, $(\cdots n_{-3}n_{-2}n_{-1}.n_1n_2n_3\cdots 0)$, or $(0\cdots n_{-3}n_{-2}n_{-1}.n_1n_2n_3\cdots 0)$. For convenience, we regard that 0 continues to the right and/or to the left indefinitely. Then, any orbit is expressed as a bi-infinite symbol sequence.

Let Σ denote the set of all bi-infinite sequences $s = (\ldots n_{-3}n_{-2}n_{-1}.n_1n_2n_3\ldots)$. We as usual define a metric on Σ setting $\mathrm{d}(s, s) = 0$ and $\mathrm{d}(s^1, s^2) = 3^{-|m|}$ if $s^1 \neq s^2$ and $|m|$ is the least integer such that $n^1_m \neq n^2_m$. Then Σ is a compact metric space.

Let Σ^+ denote the set of all infinite sequences of the form $s = (.n_1n_2n_3\ldots)$. This set is also a compact metric space with a metric defined on Σ^+ by $\mathrm{d}(s, s) = 0$ and $\mathrm{d}(s^1, s^2) = 3^{-m}$ if $s^1 \neq s^2$ and m is the least positive integer such that $n^1_m \neq n^2_m$.

The shift operator σ is defined by

$$\sigma(\ldots n_{-3}n_{-2}n_{-1}.n_1n_2n_3\ldots) = (\ldots n_{-3}n_{-2}n_{-1}n_1.n_2n_3\ldots)$$

on Σ and by

$$\sigma(.n_1n_2n_3\ldots) = (.n_2n_3\ldots)$$

on Σ^+.

3.2. Surface of section and symbol sequences

We take surface of section H defined by $q_1 = q_2$. This surface H corresponds to non-collision states except at $q_1 = q_2 = 0$ which is a triple collision. Our surface of section H has θ as its abscissa and R as its ordinate.

$$H = \{(\theta, R) : 0 \leq \theta \leq 2\pi, 0 \leq R \leq 2.5\}.$$

For $0 < \theta < \pi$, trajectories in (q_1, q_2)-plane cross the *homographic line*$(q_1 = q_2)$ transversely to the $(q_1 > q_2)$-side. For $\pi < \theta < 2\pi$, trajectories in (q_1, q_2)-plane cross the homographic line transversely to the $(q_1 < q_2)$-side. $\theta = 0, \pi$, and 2π correspond to special orbits. $\theta = 0$ and 2π should be identified. These are all homographic solutions: $\theta = 0$ corresponds to the expanding initial condition and $\theta = \pi$ corresponds to the contracting condition.

Since three particles are of the same mass, there are kinds of degeneracy. Let $H_2 = \{(\theta, R) : 0 \leq \theta < \pi, 0 \leq R \leq 2.5\}$, $H_1 = \{(\theta, R) : \pi \leq \theta < 2\pi, 0 \leq R \leq 2.5\}$, and $H_2^t = \{(\theta, R) : (\pi - \theta, R) \in H_2;$ names of m_1 and m_2 exchanged.$\}$. First, the future of the orbit starting in H_1(resp. H_2) is the past of the orbit starting in H_2(resp. H_1). Second, the future of the orbits starting at H_2^t is the past of the orbit starting at H_2. In the words of symbol sequence, if $(.n_1n_2n_3\cdots) \in \Sigma^+$ be the symbol sequence corresponding to the orbit starting at $(\theta, R) \in H_2$ and if $(.m_1m_2m_3\cdots) \in \Sigma^+$ be the symbol sequence corresponding to the orbit starting at $(\theta, R) \in H_2^t$, then $(\cdots m_3m_2m_1.n_1n_2n_3\cdots) \in \Sigma$ is the bi-infinite sequence corresponding to the full orbit starting in H_2. Or equivalently, if $(.n_1n_2n_3\cdots) \in \Sigma^+$ be the symbol sequence corresponding to the orbit starting at $(\theta, R) \in H_2$ and if $(.m_1m_2m_3\cdots) \in \Sigma^+$ be the symbol sequence corresponding to the orbit starting at $(\pi - \theta, R) \in H_2$, then $(\cdots m_3^t m_2^t m_1^t.n_1n_2n_3\cdots) \in \Sigma$ is the bi-infinite sequence corresponding to the full orbit starting in H_2, where $m_i^t = 2$ if $m_i = 1$ and $m_i^t = 1$ if $m_i = 2$. Summarizingly, we only need to study the structure of H_2.

Now, let us look at particular parts of H.

(1) The line $\theta = 0$ and π.
We have $\dot{q}_1 = \dot{q}_2$ for $\theta = 0$. The orbit is homographic and is always on H. Different points on this line correspond to different phases of the same orbit. Thus all points on this line correspond to a single orbit. The symbol sequence is (0.0). Similarly, we have $\dot{q}_1 = \dot{q}_2$ for $\theta = 0$. This gives the homographic solution. All points on this line correspond to a single orbit. The orbit is continuously on H. The symbol sequence is (0.0).

(2) The line $R = 2.5, 0 \leq \theta \leq \pi$.
On this line, we have $T = 0$, that is, initial zero-velocities. Strictly speaking, θ is not defined. This essentially is one orbit. The symbol sequence is (0.0).

(3) The line $R = 0$.
T is infinite. A triple collision. This is a triple collision manifold. The symbol sequence is $(\cdots.0)$.

(4) The Schubart orbit.
It has been known[MH89] that a periodic orbit called the Schubart orbit exists. This has symbol sequence $(\cdots 121.212\cdots)$ and has starting point on the line $\theta = \pi/2$. The orbit is stable and has a stable region around it. All the orbit starting inside the stable region has the same symbol sequence as that of the Schubart orbit.

Next, let us consider a relation between surface of section H_2 and symbol sequences.

Proposition 3.1. A trajectory in the (q_1, q_2)-plane crosses transversely the homographic line except at $(0,0)$, if it does at all.

Proof. If $q_1 = q_2$ and $\dot{q}_1 = \dot{q}_2$, the solution is homographic and always on the homographic line. So the trajectory initially outside the homographic line never become tangent to the line except the case of triple collision. □

Proposition 3.2. If a trajectory crosses the homographic line on the (q_1, q_2)-plane, a binary collision occurs before the trajectory again crosses it.

Proof. Suppose that a trajectory crosses the homographic line from $(q_1 < q_2)$-side to $(q_1 > q_2)$-side at $t = t_0$, i.e., $q_1 = q_2$ and $\dot{q}_1 > \dot{q}_2$ at $t = t_0$. Then there exists a small $\varepsilon > 0$ such that $q_1 > q_2$ and $\dot{q}_1 > \dot{q}_2$ at $t = t_0 + \varepsilon$. This implies that there needs a finite time in order to again cross the homographic line. Suppose that the first crossing occurs at $t = t^* > t_0$ without binary collision during (t_0, t^*). We can use the continuity argument in this case. At $t = t^*$, we should have $q_1 = q_2$ and $\dot{q}_1 < \dot{q}_2$ by Proposition 3.1. In order to have this, we should have $q_1 > q_2$ and $\dot{q}_1 = \dot{q}_2$ at some $t^{**}, t_0 < t^{**} < t^*$. We should have in turn $\ddot{q}_2 > \ddot{q}_1$ at least at some time between t_0 and t^{**}. This implies $q_2 < q_1$ at the same instant, which is a contradiciton. □

From Proposition 3.2, we see that if the orbit appears again on the surface H, then the next binary collision occurs with a different pair. Conversely, an orbit does not appear on H as long as binary collisions between the same particles are repeated.

Examples. (1) A symbol sequence $(\cdots 21 \cdots)$. The transition from symbols 2 to 1 means the change of pair in binary collision. This in turn means that the trajectory on (q_1, q_2)-plane passes through the homographic line, i.e., the orbit crosses H_2.

(2) A symbol sequence $(\cdots 12 \cdots)$. By a similar reasoning, the trajectory on (q_1, q_2)-plane passes through the homographic line, i.e., the orbit crosses H_1.

(3) Periodic orbits. If periodic orbits exist, their symbolic sequence should repeat. Of course, the converse is not true. Consider a sequence $s = (.221221\ldots) \in \Sigma^+$. Then, $(s)_3 = (.221)$, $(\sigma s)_3 = (.212)$, $(\sigma^2 s)_3 = (.122)$, $(\sigma^3 s)_3 = (.221)$ where the suffix 3 means the truncation of sequence within the first 3 digits. If we look at the first three digits in the surface H_2, the corresponding sequences are $(.221), (.212)$, [(.122)](this does not appear in H_2), and $(.221)$.

4. Numerical Results: Triple-Collision Orbts

In this section, we are going to resolve the sea of chaos. In order to do this, we classify the points in surface of section H_2 according to different symbol sequences. We will get triple-collision orbits.

To obtain the global structure, we cover surface of section H_2 with mesh of grid size $(\Delta R, \Delta\theta) = (0.0025, 0.^\circ 1)$ and integrate orbits starting at each vertices of grids forward until the thirty-sixth binary collision is obtained. The total number of orbits amounts to 1.8×10^6. It turns out that the structure becomes quickly fine in some part

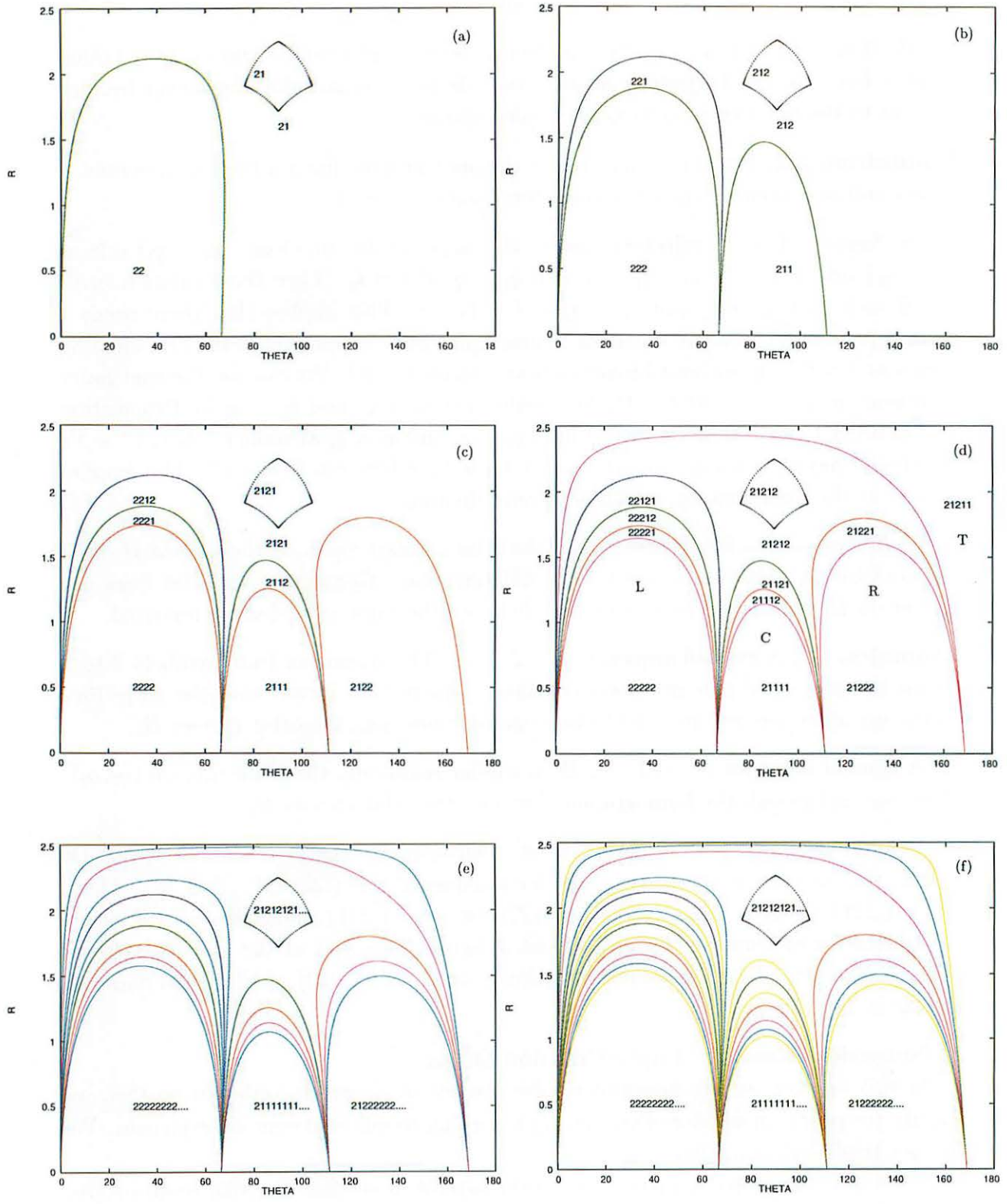


Fig.1. Division of surface of section H_2 by symbol sequences. Symbol '2' denotes binary collision between m_0 and m_2, whereas symbol '1' denotes the one between m_0 and m_1. Boundary curves are triple-collision curves. The Schubart region is added for reference.

of H_2, so we stop the analysis of the data at about the 15th digit.

We already noted the existence of the Schubart orbit in Section 3. It has $(.\overline{21})$ as a symbol sequence where the overline indicates the infinite repetition of numbers. For later convenience, let us call the stable region of Schubart orbit the *Schubart region.* Let us call the area of H_2 for which the symbol sequence is the same as that of the Schubart orbit the *extended Schubart region.*

Figures 1(a) through (f) show the process of division with increasing digits of symbol sequences. In Fig.1(a), H_2 is divided into two areas by symbol sequences (.21) and (.22). Then, their boundary should correspond to (.20), i.e., the set of initial conditions for orbits which end in triple-collision after the binary collision between m_0 and m_2. The boundary forms a curve extending to the boundary of H_2. We call this a *triple-collision curve* or simply a *collision curve.* The triple-collision curve starts at $T_0 = (\theta, R) = (0, 0)$ and ends at $T_1 = (66.°5\ldots, 0)$. One sees a diamond-shaped area in the figure. This is the Schubart region. We added it for reference.

In Fig.1(b), two new collision curves appear in three digits. One of them starts at T_0 and ends at T_1, and the other starts at T_1 and ends at $T_2 = (110.°9\ldots, 0)$. In Fig.1(c), three new curves appear in four digits. One of them starts at T_2 and ends at $T_3 = (168.°7\ldots, 0)$. Finally, in Fig.1(d), one of four new collsion curves starts at T_3 and ends at T_0. For longer digits, all collision curves start at T_i and end at T_{i+1} for some $i, 0 \leq i \leq 3$ where $T_4 = T_0$.

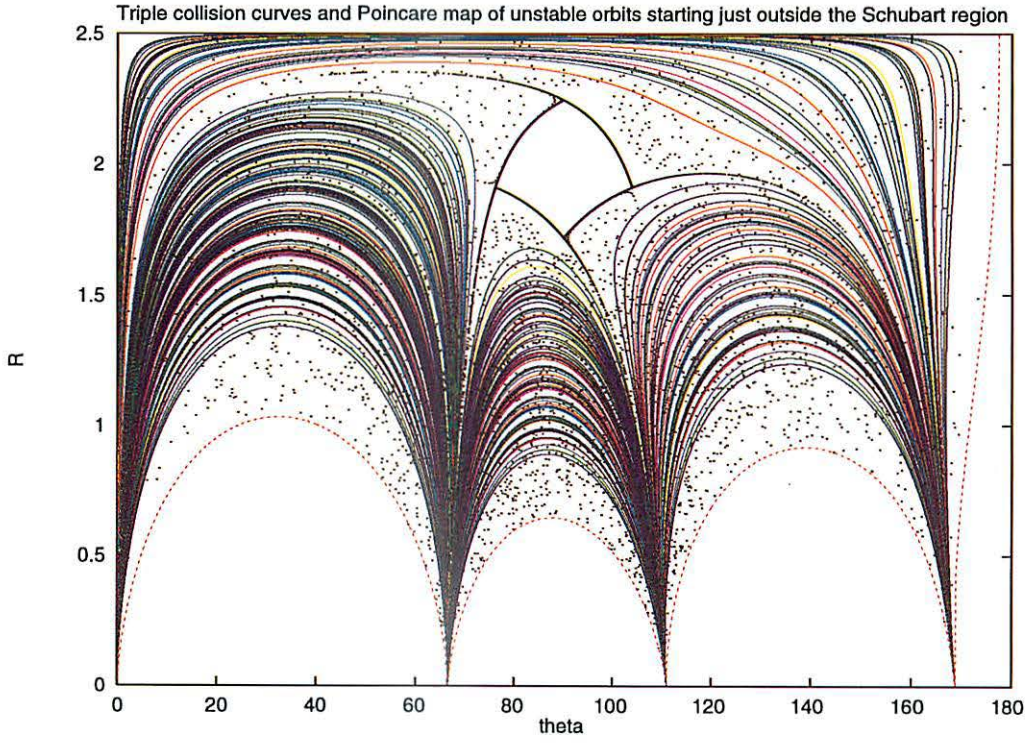


Fig.2. Triple-collision curves, the Schubart region, and general orbits.

For convenience, let us denote the regions of H_2 related to collision curves starting at T_i and ending at T_{i+1}, i=0,1,2,3 by 'L'(left), 'C'(center), 'R'(right), and 'T'(top), repsectively. Table I shows how the number of triple-collision curves increases with increasing digits. It is interesting to note that new number appears first in 'L' and is transmitted in turn to 'C', 'R', and 'T'. A rule seems to exist, but it is not apparent.

In Fig.2, a larger number of triple-collision curves are shown together with the Schubart region. Symbol sequences up to and including 12 digits are plotted, that is, triple-collision orbits which end after 11 binary collisions are shown. Here four blank regions bounded by red broken curves represent immediate escape regions. Corresponding symbol sequences are $(.\bar{2}), (.2\bar{1}), (.21\bar{2})$, and $(.212\bar{1})$. Scattered points are orbits starting just outside the Schubart region. These points move for a while along the boundary of Schubart region, leave the region along the unstable manifolds of the unstable period-4 points at the corner of the Schubart region, and finally move almost randomly in H_2. Thick curves starting at each corner of the Schubart region are part of these scattered points. In superposing scattered points on H_2, we used a trick, i.e. we changed θ into $\pi - \theta$. This is equivalent to change t to $-t$. So actually in the figure, the points start in scattered state, gradually gather into thick balck curves, and then come close to the boundary of the Schubart region. Black thick curves are the stable manifolds of period-4 unstable points. These stable manifolds seem to extend to points $T_i, i = 0, 1, 2, 3$ and divide H_2 into four.

Table I. The number of triple-collision curves newly added. L, C, R, T denote the regions shown in Fig.1.

Number of digits	L	C	R	T	Total
2	1	0	0	0	1
3	1	1	0	0	2
4	1	1	1	0	3
5	1	1	1	1	4
6	3	1	1	1	6
7	5	3	1	1	10
8	7	5	3	1	16
9	9	7	5	3	24
10	15	9	7	5	36
11	25	15	9	7	56
12	39	25	15	9	88
13	57	39	25	15	136
14	87	57	39	25	208

Now let us summarize the above observations. First, the stable manifolds of period-4 points outside the Schubart region divide the surface except the Schubart region into four: regions L, C, R, and T. Each of these regions contains immediate escape regions. Second, the set of triple-collision curves approach both to the boundary of the Schubart region and to the boundary of immediate escape regions with increasing digits of symbol sequences. We can say that the Schubart orbit, stable manifolds of

period-4 points accompanying the Schubart orbit, and triple-collision curves form the skeleton of the structure of phase space of the 1-D three-body problem.

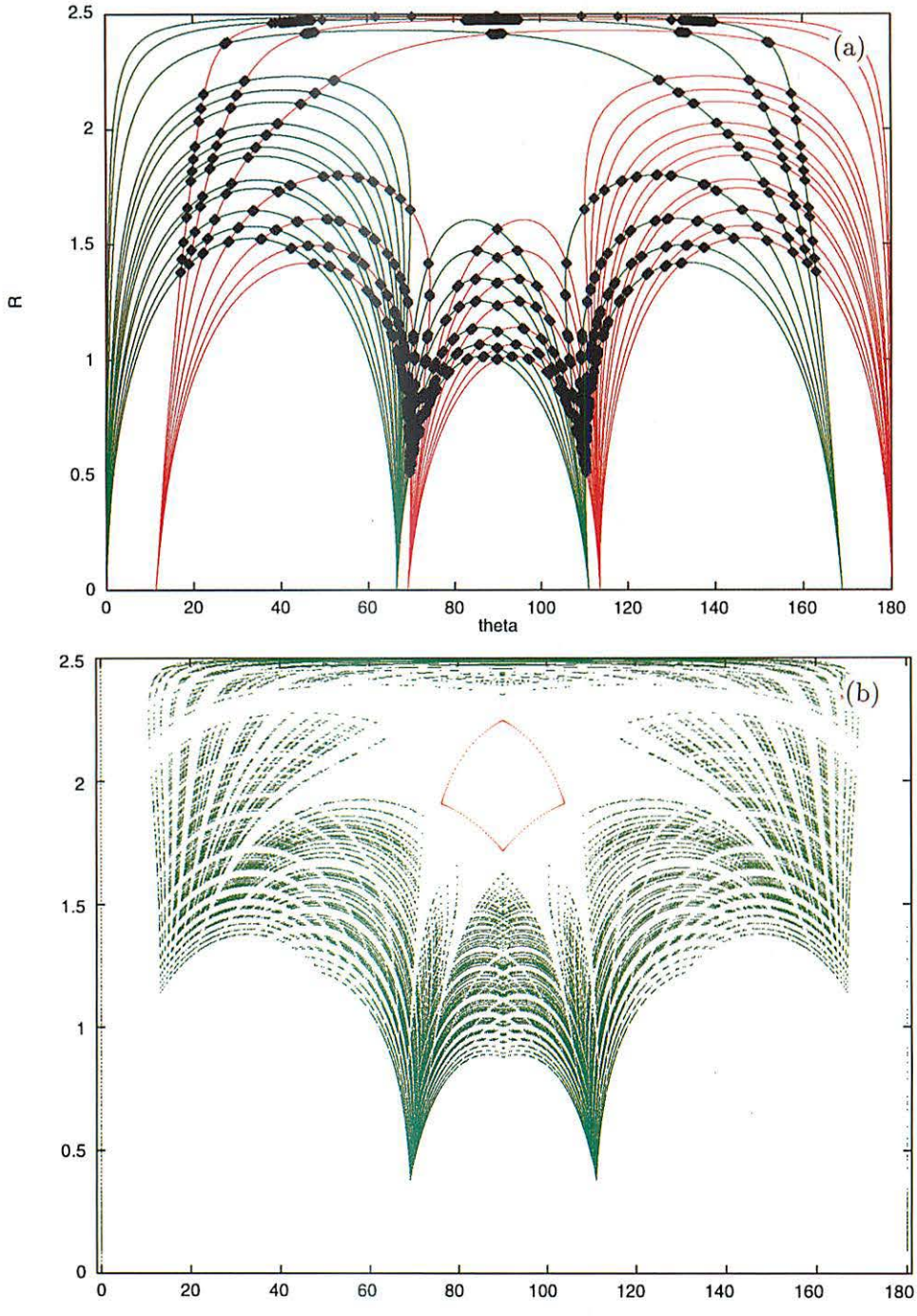


Fig.3. Orbits that start and end in triple collision. (a) Red curves: triple-collsion curves in the future. Green curves: triple collision curves in the past. Black marks correspond to orbits that end in triple collision after seven binary collisions both in the past and future. (b) Green points correspond to orbits that end in triple collision after eleven binary collisions both in the past and future.

Figure 3(a) shows the superposition of H_2 and H_2^t. Curves are all triple-collision curves up to and including the first seven digits of symbol sequences. Green curves indicate triple collisions in the future, whereas red curves indicate triple collisions in the past. Therefore, cross points are the initial conditions of orbits which start and end in triple collision. Figure 3(b) shows a deeper survey of these orbits until 12 digits. We see quite a many of these orbits. Apparently, the distribution is fractal.

Remark. We see how the points move on surface H_2. As an example, let us take a symbol sequence 21221 in Fig.1(d). We have $\sigma^2(.21221) = (.221)$. This means that the orbits passing through '21221' will next intersect H_2 in '221'(Fig.1(b)). Thus, in principle, using symbol sequences and shift operator, we can follow any orbit on the surface of section.

5. Conclusions

We have demonstrated that symbolic dynamics is effective in the one-dimensional three-body problem. Main results are

(i) Triple-collision orbits are obtained systematically and automatically with the aid of symbolic sequences.

(ii) The Schubart orbit, stable manifolds of the period-4 orbit, and triple-collision curves form the skeleton of the phase space of the 1-D three-body problem.

There remained an important question: Does the set of symbol sequences form a Cantor set? To answer this question, we need a careful analysis of symbol sequences which are not realized.

References

[1] McGehee, R., 1974, *Inventiones Mathematicae* **27**, 191-227.

[2] Mikkola, S. and Hietarinta, J., 1989, *Celestial Mechanics and Dynamical Astronomy* **46**, 1-18.

[3] Mikkola, S. and Hietarinta, J., 1990, *Celestial Mechanics and Dynamical Astronomy* **47**, 321-331.

[4] Mikkola, S. and Hietarinta, J., 1991, *Celestial Mechanics and Dynamical Astronomy* **51**, 379-394.

[5] Tanikawa, K., Umehara, H. and Abe, H., 1995, *Celestial Mechanics and Dynamical Astronomy* **62**, 335-362.

[6] Tanikawa, K. and Umehara, H., 1998, *Celestial Mechanics and Dynamical Astronomy* **70**, 167-180.

[7] Block, L.S. and Coppel, W.A., 1992, *Dynamics in One Dimension*, Lecture Notes in Mathematics Vol. 1513, Springer-Verlag.

Synchronization of orbital elements and stability of protoplanet systems

Takashi Ito and Kiyotaka Tanikawa

National Astronomical Observatory, Mitaka, Tokyo 181–8588, Japan

(E-mail) tito@cc.nao.ac.jp

Recent numerical studies indicate that so called secular orbit-orbit resonances which have been found in planet-satellite or satellite-ring systems can occur in systems like planetary and protoplanetary systems. In this paper, we investigate the dynamical character of such resonances: synchronization of eccentricities and longitudes of perihelion in terrestrial protoplanet systems perturbed by massive jovian planets. Orbital synchronization in protoplanet systems with plural perturbed bodies is qualitatively explained by using a secular disturbing function and superposition of "mutual shepherding" among neighboring particles. Orbital synchronization serves as a stabilizer of the system by preventing particles from making close encounters. Systems appear most stable where tidal torque by perturbers and mutual gravitational interaction among particles balance; in such cases, the orbital synchronization is realized.

1. Introduction

So called orbit-orbit resonances in planet-satellite or satellite-ring systems have been investigated for decades (Peale, 1976). The orbital stability of several satellites around Jupiter, Saturn, and Uranus seems to be sustained through various types of mean motion commensurabilities (Allan, 1969; Sinclair, 1972; Sinclair, 1975; Greenberg, 1973; Greenberg, 1975; Greenberg, 1976). Among the resonances, secular orbital resonances in particular are considered to play important roles in maintaining stability of systems. Some Uranian rings are thought to remain elliptic through mutual gravitational interaction and/or through the influence of nearby shepherding satellites (Goldreich and Tremaine, 1979b; Porco and Goldreich, 1987; Goldreich and Porco, 1987; Kozai, 1992; Kozai, 1993). Saturnian satellites Rhea and Titan are also in such a secular synchronization state (Greenberg, 1977; Pauwels, 1983).

Recent numerical calculations indicate that similar synchronization phenomena appear not only in planet-satellite or satellite-ring systems, but also in planetary and protoplanetary systems. Innanen et al. (1997) discovered in their numerical integrations that the nodes and mutual inclinations of several planets around a hypothetical primary star with a highly inclined secondary star show strong synchronization as if they compose a rigid disk. Ito and Tanikawa (1998, hereafter called Paper I) showed that perturbation by giant jovian planets could cause the coherent oscillation of eccentricities and longitudes of perihelion of terrestrial protoplanets. Although spatial scales and timescales of such phenomena are all

different, the essential mechanism of orbital synchronization is believed to be common to all of such planetary and protoplanetary systems.

In this paper we emphasize that the orbital synchronization in terrestrial protoplanet systems discussed in Paper I is not only interesting in itself from the viewpoint of celestial mechanics, but also plays an important role in stability of the system. When distances between particles are small, strong mutual interaction renders the system unstable, causing rapid orbital crossings and close encounters. On the other hand, the large orbital distances and wide spatial range of the system also work as destabilizers of the system, provoking large tidal torque by perturbers. Synchronization of orbital elements occurs when mutual gravitational interaction and tidal torque by perturbers balance. The system's stability is maintained by this orbital synchronization.

In Section 2, we briefly review Greenberg's (1977) method of explaining the dynamical mechanism of orbital synchronization using a secular disturbing function. Greenberg developed his method to explain the orbital synchronization of particle systems containing only one perturbed body.

Section 3 describes the dynamical characters of the terrestrial protoplanet system which we will base our argument on. Typical examples of the orbital synchronization in protoplanet systems are shown. We find that the orbital synchronization in point-mass systems including several perturbed particles can be qualitatively explained by superposition of gravitational interaction among neighboring particles, "mutual shepherding".

Section 4 is devoted to an investigation of the relationship between orbital synchronization and stability. We introduce a measure of the stability of protoplanet systems, a so called "CWB diagram" (Chambers et al. 1996, hereafter called CWB). CWB diagrams show that the orbital synchronization of protoplanets occurs in the most stable region, where mutual gravitational interaction and tidal torque by perturbers balance. We show that protoplanet systems having synchronized initial orbital elements remain stable longer than those having randomly chosen initial conditions. We remark that orbital synchronization in the terrestrial protoplanet system is a transient process, gradually breaking down toward instability over a $O(10^6)$-year timescale.

Section 5 goes on to discuss the implications of these results. Stability of Uranian ϵ-rings and a hypothetical planetary system in a binary system advocated by Innanen et al. (1997) are reconsidered in terms of orbital synchronization.

2. Mechanisms of orbital synchronization

In this section, we briefly explain the dynamical mechanism of orbital synchronization by utilizing a secular disturbing function following Greenberg (1977). Greenberg's method is intended to explain the orbital synchronization of a particle system containing only one perturbed body. In later sections we will discuss how orbital synchronization in a system

with two or more perturbed particles can be qualitatively understood as a superposition of the mutual gravitational interaction between neighboring particles.

In low eccentricity and small inclination systems, the orbital synchronization between a perturber and a perturbed body can be qualitatively explained by the motion of eccentricity vectors (Laplace-Runge-Lentz vectors) using a secular disturbing function. Here we concentrate on the synchronized oscillation of eccentricities e and longitudes of perihelion ϖ in planar orbits. We can also show that inclinations I and longitudes of ascending node Ω also synchronize in low eccentricity and small inclination systems, since the secular equations are decoupled in similar forms into components of (e, ϖ) pairs and (I, Ω) pairs in such systems (cf. Brouwer and Clemence 1969).

Let us consider a system consisting of two particles on planar orbits around a central mass. The inner one (indicated by subscript 1) is a perturbed body, and the outer one (subscript 2) is a perturber. a and m denote semimajor axis and mass, respectively. Secular part $\mathcal{R}_1$ of the disturbing function at the inner perturbed body without using higher order terms in eccentricity becomes

$$\mathcal{R}_1 = \frac{3GMJ_2R^2{e_1}^2}{8{a_1}^3} + Gm_2\left(\frac{1}{2a_2}b_{1/2}^{(0)} + N_{12}\left(e_1^2+e_2^2\right) - 2P_{12}e_1e_2\cos(\varpi_1-\varpi_2)\right), \quad (1)$$

where G is the gravitational constant, J_2 denotes the gravitational oblateness of the central mass, M and R are the mass and equatorial radius of the central mass, $b_{3/2}^{(0)}$, $b_{3/2}^{(1)}$, $b_{3/2}^{(2)}$ are the Laplace coefficients, and

$$N_{12} = \frac{\alpha_{12}b_{3/2}^{(1)}}{8a_2}, \quad P_{12} = \frac{\alpha_{12}b_{3/2}^{(2)}}{8a_2}, \quad \alpha_{12} = \frac{a_1}{a_2} < 1. \quad (2)$$

The first term of the right-hand side in (1) comes from the gravitational oblateness of the central mass. In the terrestrial protoplanet system we do not take this J_2 term into account since the Sun's J_2 does not significantly influence the terrestrial region.

Lagrange's planetary equations are transformed into differential equations for eccentricity vector (h, k) as

$$\frac{dh_1}{dt} = -\frac{1}{n_1a_1^2}\frac{\partial\mathcal{R}_1}{\partial k_1}, \quad \frac{dk_1}{dt} = \frac{1}{n_1a_1^2}\frac{\partial\mathcal{R}_1}{\partial h_1}. \quad (3)$$

The eccentricity vector is usually defined as $(e\cos\omega, e\sin\omega)$, using the argument of perihelion ω. But in this paper we define the vector using the longitude of perihelion ϖ as $(h, k) = (e\cos\varpi, e\sin\varpi)$.

Inserting the secular disturbing function $\mathcal{R}_1$ (1) into Lagrange's planetary equations (3), we get

$$\frac{dh_1}{dt} = -Ak_1 + Bk_2, \quad \frac{dk_1}{dt} = Ah_1 - Bh_2, \quad (4)$$

where

$$A = \frac{2Gm_2N_{12}}{n_1{a_1}^2} \simeq \frac{n_1m_2{\alpha_{12}}^2b_{3/2}^{(1)}}{4M}, \quad (5)$$

$$B = \frac{2Gm_2P_{12}}{n_1{a_1}^2} \simeq \frac{n_1 m_2 {\alpha_{12}}^2 b_{3/2}^{(2)}}{4M}, \tag{6}$$

neglecting the term derived from the gravitational oblateness of the central mass in $\mathcal{R}_1$. n_1 denotes the mean motion of the particle 1. Since the mass of the perturbed body m_1 is much smaller than that of the central mass M, we consider $m_2/(M+m_1)$ in (5) and (6) as m_2/M.

As the motion of the perturber, we assume a slow variation of eccentricity vector (h_2, k_2) as

$$h_2 = e_2\cos(gt+\beta), \quad k_2 = e_2\sin(gt+\beta), \tag{7}$$

where g is the perturber's constant precession rate and β is a phase constant of the eccentricity vector. e_2 is treated as a constant here. Substituting (7) into (4), we obtain a solution to equation (4) as

$$h_1 = c\cos(At+\delta) + Kh_2, \quad k_1 = c\sin(At+\delta) + Kk_2, \tag{8}$$

where

$$K \equiv \frac{B}{A-g} \tag{9}$$

and c and δ are arbitrary constants of integration. Solution (8) is illustrated in the eccentricity vector space (h,k) as in Figure 1. This solution indicates that the eccentricity vector of the perturbed particle (h_1, k_1) moves in a circle around the vector (Kh_2, Kk_2) with an angular velocity A. If $|c| < |K|e_2$, $\varpi_2 - \varpi_1$ librates, the synchronization of perihelia is achieved. Otherwise, $\varpi_2 - \varpi_1$ circulates, and no synchronization is seen. Ke_2 is the amplitude of forced oscillation of (h_1, k_1), and c is the amplitude of free oscillation determined by the initial orbital elements of the perturbed particle.

The achievement of orbital synchronization described in solution (8) does not depend on whether or not the perturber's orbit is fixed in inertial space. When the perihelion of the perturber does not move at all, the perihelion of the perturbed particle librates around that of the perturber if conditions for orbital synchronization ($|c| < |K|e_2$) are satisfied.

One of the typical examples of orbital synchronization is seen in Saturnian satellites, in the Rhea-Titan system (Pauwels, 1983). We have to take J_2 term $3J_2R^2/4{a_1}^2$ into account in the expression of A (5), since the motion of the Rhea-Titan system is strongly affected by the gravitational oblateness of Saturn. Substituting the actual value of the physical parameters of the system, we get $|K|e_2/|c| \sim 3.3$ which leads to the synchronization of Rhea's perihelion with that of Titan (Greenberg, 1977).

3. Terrestrial protoplanet systems

In this section we briefly describe the dynamical character of the terrestrial protoplanet system which we will discuss further in later sections. The numerical model of terrestrial protoplanet systems is the same as in Paper I.

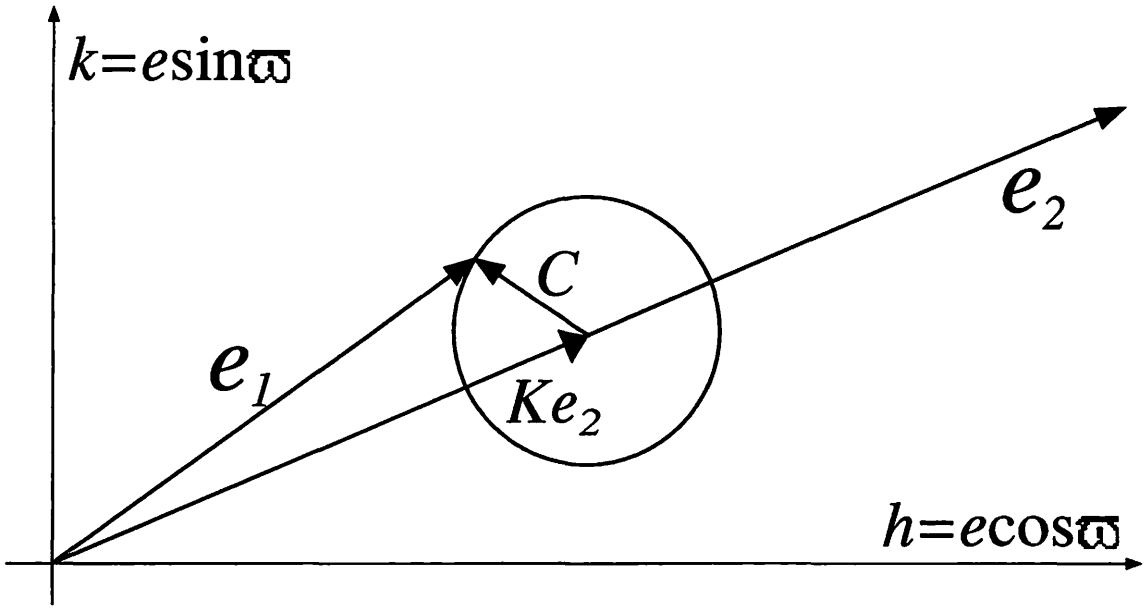


Figure 1. Schematic illustration of orbital synchronization in eccentricity vector space. e_1 and e_2 represent the eccentricity vector of the perturbed body and the perturber, respectively. Adapted from Greenberg (1977).

3.1 Model of terrestrial protoplanets and numerical method

Protoplanets, also called "planetary embryos," are regarded as precursors of the present planets. Venus, Earth, and other terrestrial planets are thought to have formed via the collisional evolution of protoplanets. Protoplanets themselves would have formed through the runaway growth of smaller planetesimals. According to the latest numerical results, a number of protoplanets would form separated by nearly equal distances with slightly eccentric orbits, through runaway growth in the swarm of planetesimals (Kokubo and Ida, 1996; Kokubo and Ida, 1998). Once all the planetesimals have accumulated into protoplanets, the typical mass of the protoplanets becomes $\sim 0.2M_\oplus$ at $a = 1$AU, and $\sim 7M_\oplus$ at $a = 7$AU where $M_\oplus$ is the mass of Earth. Typical separations among the protoplanets are around ten Hill radii, which is equivalent to 0.02AU $\sim$ 0.04AU at $a = 1$AU. Orbits of protoplanets formed through runaway growth are nearly circular due to the equipartition of energy (dynamical friction), and they remain globally stable till the end of the formation of the cores of Jupiter and Saturn within $10^6 \sim 10^7$ years. The accretion of the cores of jovian planets, i.e. the runaway growth of the protoplanets in Jupiter-Saturn region, is thought to proceed in solar nebula gas. The core accretion of Jupiter and Saturn, especially, can almost simultaneously accompany drastic collapse of nebula gas onto the surfaces of the cores due to their strong gravitational attraction (Mizuno *et al.*, 1978). Then these two planets acquire their present masses within 10^6 years, or at most 10^7 years (Pollack *et al.*, 1996; Tajima, 1998). At this point, growth of the jovian planets may catch

up with, or even surpass, growth of the terrestrial planets.

For the numerical model of the terrestrial protoplanet system, we distribute about ten to twenty particles whose total mass is about $2M_\oplus$ (which is close to the total mass of the present terrestrial planets) in the region of the present Venus to Earth. Jupiter and Saturn are assumed to possess their present masses and orbital elements outside of the terrestrial protoplanets. The terrestrial protoplanets we treat in this paper range upward from $a = a_1 = 0.5$AU with equal separations of $\Delta \times R_H$ using the mutual Hill radius R_H, where a_1 is the semimajor axis of the innermost protoplanet. A parameter Δ controls the planetary separation. The semimajor axis of the $i+1$-th protoplanet is determined by an implicit relationship $a_{i+1} = a_i + \Delta \times R_H^{i,i+1}$, where $R_H^{i,i+1}$ is the mutual Hill radius between the i-th and the $i+1$-th protoplanets (Gladman 1993, CWB). Spatial distributions of protoplanets differ from model to model. The number of protoplanets is denoted by N.

We utilize two kinds of numerical integration method: one is second-order symplectic mapping (Wisdom and Holman, 1991) with a special "warm start" procedure (Saha and Tremaine, 1992; Saha and Tremaine, 1994), and the other is a fourth-order time-symmetric Hermite integrator (Makino and Aarseth, 1992; Kokubo *et al.*, 1998). We employ the Hermite integrator when either of two bodies approaches the other within five Hill radii. Otherwise we use symplectic mapping.

We assume that the eccentricities and inclinations of protoplanets obey Rayleigh distribution with dispersions $\langle e^2 \rangle^{\frac{1}{2}}$ and $\langle I^2 \rangle^{\frac{1}{2}}$, and satisfy $\langle e^2 \rangle^{\frac{1}{2}} = 2 \langle I^2 \rangle^{\frac{1}{2}}$ (Ida and Makino, 1992). We mainly assume $\langle e^2 \rangle^{\frac{1}{2}} \leq 0.005$ according to the results of numerical N-body simulations of runaway growth of planetesimals. The initial values of the mean anomalies are selected randomly from a uniform distribution subject to the constraint that the longitudes of bodies on adjacent orbits differ by at least 20°. The initial values of the longitudes of ascending node Ω, and perihelion ϖ are randomly selected from 0° to 360°. We do not consider any variation of masses nor separations in protoplanet systems as we did in Paper I.

3.2 Orbital synchronization in protoplanet systems

Typical terrestrial protoplanet systems satisfy the condition of orbital synchronization described in the previous section, $|c| < |K|e_2$. As an example, we consider Jupiter as a perturber with $m_2 = 1.0 \times 10^{-3} M_\odot$, $a_2 = 5.2$AU, and $e_2 = 0.049$. $M_\odot (= M)$ is a solar mass. In the following sections we use a subscript J to denote the quantity for Jupiter (e.g. $e_J, \varpi_J, a_J, \ldots$). We adopt $g = g_J = 2\pi/(3 \times 10^5)$ (rad/year) as the constant precession rate of Jupiter's perihelion in (7), since the circulation period of the present Jupiter's longitude of perihelion is about 3×10^5 years. Let us take a protoplanet as the perturbed particle with semimajor axis $a = 0.8$AU and mass $m = 0.1 M_\oplus = 3.0 \times 10^{-7} M_\odot$. The amplitude of free oscillation c in the protoplanet system is quite small, since random velocity of the system would have been reduced through dynamical friction among the

swarm of planetesimals. When we assume $c = 0.005$ and calculate the Laplace coefficients, $A = 2.63 \times 10^{-5}$ (rad/year), $B = 5.36 \times 10^{-6}$ (rad/year), so $K = B/(A - g) \sim 1.0$. Thus $|c| = 0.005 < |K|e_2 \sim 0.05$, which means that the orbital synchronization occurs between Jupiter and a perturbed protoplanet in this case.

In addition, the orbital synchronization in protoplanet systems is characterized by the fact that the orbits of several perturbed protoplanets are synchronized. One of the results of the orbital evolution of 14 protoplanets ($\Delta = 10.4$) perturbed by Jupiter and Saturn is shown as Figure 2. Perihelia of protoplanets synchronize and apparently follow that of Jupiter. Note that apparent circulation of protoplanets' perihelia when their orbits are nearly circular ($e \sim 0$) is caused by the fact that perihelia themselves are only weakly defined in nearly circular orbits. Synchronization of the protoplanets' eccentricities is also seen in Figure 2(b).

The orbital synchronization in terrestrial protoplanets is qualitatively reproduced by a first-order secular perturbation theory (Brouwer and Clemence, 1961) as in Figure 3. Solutions by perturbation theory well reproduce those given by numerical integration. Henceforward, we will use secular perturbation theory to explain the mechanism of orbital synchronization among several perturbed particles. Note that in Figure 3(b), the circulation period of Jupiter's longitude of perihelion differs somewhat from numerical integration results in Figure 2(b). This is because we have neglected the terms derived from the 5:2 mean motion commensurability between Jupiter and Saturn — called "the great inequality" — in the original disturbing function in our calculation of basic planetary frequencies (cf. Brouwer and van Woerkom 1950).

The orbital synchronization of protoplanet systems is also seen when we specify only Jupiter as a perturber (Figure 4). This is a plausible condition in the terrestrial protoplanet system since there must be some time-lag between the formation of Jupiter and that of Saturn, due to the difference in their distances from the Sun. In the meantime, only Jupiter would exist as a perturber for the terrestrial protoplanets. In Figure 4(a), Jupiter's perihelion is nearly fixed in inertial space, and perihelia of the protoplanets librate around that of Jupiter. Synchronization of eccentricities is also apparent (Figure 4(b)). It should also be noted here that apparent circulation of protoplanets' perihelia when their orbits are nearly circular ($e \sim 0$) is caused by the fact that perihelia themselves are only weakly defined in nearly circular orbits.

3.3 Interaction among neighboring particles

Even when there are several perturbed particles, we can also get linear solutions similar to solution (8) which express orbital synchronization among perturbed particles. But here our aim is to provide a simpler and more intuitive explanation of orbital synchronization.

When there exist more than one perturbed particle, libration frequencies of their perihelia around that of Jupiter are all different if there is no mutual gravitational interaction. This

Synchronization of orbital elements and stability of protoplanet systems

Takashi Ito and Kiyotaka Tanikawa
National Astronomical Observatory, Mitaka, Tokyo 181-8588, Japan
(E-mail) tito@cc.nao.ac.jp

Recent numerical studies indicate that so called secular orbit-orbit resonances which have been found in planet-satellite or satellite-ring systems can occur in systems like planetary and protoplanetary systems. In this paper, we investigate the dynamical character of such resonances: synchronization of eccentricities and longitudes of perihelion in terrestrial protoplanet systems perturbed by massive jovian planets. Orbital synchronization in protoplanet systems with plural perturbed bodies is qualitatively explained by using a secular disturbing function and superposition of "mutual shepherding" among neighboring particles. Orbital synchronization serves as a stabilizer of the system by preventing particles from making close encounters. Systems appear most stable where tidal torque by perturbers and mutual gravitational interaction among particles balance; in such cases, the orbital synchronization is realized.

1. Introduction

So called orbit-orbit resonances in planet-satellite or satellite-ring systems have been investigated for decades (Peale, 1976). The orbital stability of several satellites around Jupiter, Saturn, and Uranus seems to be sustained through various types of mean motion commensurabilities (Allan, 1969; Sinclair, 1972; Sinclair, 1975; Greenberg, 1973; Greenberg, 1975; Greenberg, 1976). Among the resonances, secular orbital resonances in particular are considered to play important roles in maintaining stability of systems. Some Uranian rings are thought to remain elliptic through mutual gravitational interaction and/or through the influence of nearby shepherding satellites (Goldreich and Tremaine, 1979b; Porco and Goldreich, 1987; Goldreich and Porco, 1987; Kozai, 1992; Kozai, 1993). Saturnian satellites Rhea and Titan are also in such a secular synchronization state (Greenberg, 1977; Pauwels, 1983).

Recent numerical calculations indicate that similar synchronization phenomena appear not only in planet-satellite or satellite-ring systems, but also in planetary and protoplanetary systems. Innanen et al. (1997) discovered in their numerical integrations that the nodes and mutual inclinations of several planets around a hypothetical primary star with a highly inclined secondary star show strong synchronization as if they compose a rigid disk. Ito and Tanikawa (1998, hereafter called Paper I) showed that perturbation by giant jovian planets could cause the coherent oscillation of eccentricities and longitudes of perihelion of terrestrial protoplanets. Although spatial scales and timescales of such phenomena are all

different, the essential mechanism of orbital synchronization is believed to be common to all of such planetary and protoplanetary systems.

In this paper we emphasize that the orbital synchronization in terrestrial protoplanet systems discussed in Paper I is not only interesting in itself from the viewpoint of celestial mechanics, but also plays an important role in stability of the system. When distances between particles are small, strong mutual interaction renders the system unstable, causing rapid orbital crossings and close encounters. On the other hand, the large orbital distances and wide spatial range of the system also work as destabilizers of the system, provoking large tidal torque by perturbers. Synchronization of orbital elements occurs when mutual gravitational interaction and tidal torque by perturbers balance. The system's stability is maintained by this orbital synchronization.

In Section 2, we briefly review Greenberg's (1977) method of explaining the dynamical mechanism of orbital synchronization using a secular disturbing function. Greenberg developed his method to explain the orbital synchronization of particle systems containing only one perturbed body.

Section 3 describes the dynamical characters of the terrestrial protoplanet system which we will base our argument on. Typical examples of the orbital synchronization in protoplanet systems are shown. We find that the orbital synchronization in point-mass systems including several perturbed particles can be qualitatively explained by superposition of gravitational interaction among neighboring particles, "mutual shepherding".

Section 4 is devoted to an investigation of the relationship between orbital synchronization and stability. We introduce a measure of the stability of protoplanet systems, a so called "CWB diagram" (Chambers et al. 1996, hereafter called CWB). CWB diagrams show that the orbital synchronization of protoplanets occurs in the most stable region, where mutual gravitational interaction and tidal torque by perturbers balance. We show that protoplanet systems having synchronized initial orbital elements remain stable longer than those having randomly chosen initial conditions. We remark that orbital synchronization in the terrestrial protoplanet system is a transient process, gradually breaking down toward instability over a $O(10^6)$-year timescale.

Section 5 goes on to discuss the implications of these results. Stability of Uranian ϵ-rings and a hypothetical planetary system in a binary system advocated by Innanen et al. (1997) are reconsidered in terms of orbital synchronization.

2. Mechanisms of orbital synchronization

In this section, we briefly explain the dynamical mechanism of orbital synchronization by utilizing a secular disturbing function following Greenberg (1977). Greenberg's method is intended to explain the orbital synchronization of a particle system containing only one perturbed body. In later sections we will discuss how orbital synchronization in a system

with two or more perturbed particles can be qualitatively understood as a superposition of the mutual gravitational interaction between neighboring particles.

In low eccentricity and small inclination systems, the orbital synchronization between a perturber and a perturbed body can be qualitatively explained by the motion of eccentricity vectors (Laplace-Runge-Lentz vectors) using a secular disturbing function. Here we concentrate on the synchronized oscillation of eccentricities e and longitudes of perihelion ϖ in planar orbits. We can also show that inclinations I and longitudes of ascending node Ω also synchronize in low eccentricity and small inclination systems, since the secular equations are decoupled in similar forms into components of (e, ϖ) pairs and (I, Ω) pairs in such systems (cf. Brouwer and Clemence 1969).

Let us consider a system consisting of two particles on planar orbits around a central mass. The inner one (indicated by subscript 1) is a perturbed body, and the outer one (subscript 2) is a perturber. a and m denote semimajor axis and mass, respectively. Secular part $\mathcal{R}_1$ of the disturbing function at the inner perturbed body without using higher order terms in eccentricity becomes

$$\mathcal{R}_1 = \frac{3GMJ_2R^2{e_1}^2}{8{a_1}^3} + Gm_2\left(\frac{1}{2a_2}b_{1/2}^{(0)} + N_{12}\left(e_1^2+e_2^2\right) - 2P_{12}e_1e_2\cos(\varpi_1-\varpi_2)\right), \quad (1)$$

where G is the gravitational constant, J_2 denotes the gravitational oblateness of the central mass, M and R are the mass and equatorial radius of the central mass, $b_{3/2}^{(0)}$, $b_{3/2}^{(1)}$, $b_{3/2}^{(2)}$ are the Laplace coefficients, and

$$N_{12} = \frac{\alpha_{12}b_{3/2}^{(1)}}{8a_2}, \quad P_{12} = \frac{\alpha_{12}b_{3/2}^{(2)}}{8a_2}, \quad \alpha_{12} = \frac{a_1}{a_2} < 1. \quad (2)$$

The first term of the right-hand side in (1) comes from the gravitational oblateness of the central mass. In the terrestrial protoplanet system we do not take this J_2 term into account since the Sun's J_2 does not significantly influence the terrestrial region.

Lagrange's planetary equations are transformed into differential equations for eccentricity vector (h, k) as

$$\frac{dh_1}{dt} = -\frac{1}{n_1a_1^2}\frac{\partial\mathcal{R}_1}{\partial k_1}, \quad \frac{dk_1}{dt} = \frac{1}{n_1a_1^2}\frac{\partial\mathcal{R}_1}{\partial h_1}. \quad (3)$$

The eccentricity vector is usually defined as $(e\cos\omega, e\sin\omega)$, using the argument of perihelion ω. But in this paper we define the vector using the longitude of perihelion ϖ as $(h, k) = (e\cos\varpi, e\sin\varpi)$.

Inserting the secular disturbing function $\mathcal{R}_1$ (1) into Lagrange's planetary equations (3), we get

$$\frac{dh_1}{dt} = -Ak_1 + Bk_2, \quad \frac{dk_1}{dt} = Ah_1 - Bh_2, \quad (4)$$

where

$$A = \frac{2Gm_2N_{12}}{n_1{a_1}^2} \simeq \frac{n_1m_2{\alpha_{12}}^2b_{3/2}^{(1)}}{4M}, \quad (5)$$

$$B = \frac{2Gm_2P_{12}}{n_1{a_1}^2} \simeq \frac{n_1 m_2 {a_{12}}^2 b_{3/2}^{(2)}}{4M}, \tag{6}$$

neglecting the term derived from the gravitational oblateness of the central mass in $\mathcal{R}_1$. n_1 denotes the mean motion of the particle 1. Since the mass of the perturbed body m_1 is much smaller than that of the central mass M, we consider $m_2/(M+m_1)$ in (5) and (6) as m_2/M.

As the motion of the perturber, we assume a slow variation of eccentricity vector (h_2, k_2) as

$$h_2 = e_2\cos(gt+\beta), \quad k_2 = e_2\sin(gt+\beta), \tag{7}$$

where g is the perturber's constant precession rate and β is a phase constant of the eccentricity vector. e_2 is treated as a constant here. Substituting (7) into (4), we obtain a solution to equation (4) as

$$h_1 = c\cos(At+\delta) + Kh_2, \quad k_1 = c\sin(At+\delta) + Kk_2, \tag{8}$$

where

$$K \equiv \frac{B}{A-g} \tag{9}$$

and c and δ are arbitrary constants of integration. Solution (8) is illustrated in the eccentricity vector space (h,k) as in Figure 1. This solution indicates that the eccentricity vector of the perturbed particle (h_1, k_1) moves in a circle around the vector (Kh_2, Kk_2) with an angular velocity A. If $|c| < |K|e_2$, $\varpi_2 - \varpi_1$ librates, the synchronization of perihelia is achieved. Otherwise, $\varpi_2 - \varpi_1$ circulates, and no synchronization is seen. Ke_2 is the amplitude of forced oscillation of (h_1, k_1), and c is the amplitude of free oscillation determined by the initial orbital elements of the perturbed particle.

The achievement of orbital synchronization described in solution (8) does not depend on whether or not the perturber's orbit is fixed in inertial space. When the perihelion of the perturber does not move at all, the perihelion of the perturbed particle librates around that of the perturber if conditions for orbital synchronization ($|c| < |K|e_2$) are satisfied.

One of the typical examples of orbital synchronization is seen in Saturnian satellites, in the Rhea-Titan system (Pauwels, 1983). We have to take J_2 term $3J_2R^2/4{a_1}^2$ into account in the expression of A (5), since the motion of the Rhea-Titan system is strongly affected by the gravitational oblateness of Saturn. Substituting the actual value of the physical parameters of the system, we get $|K|e_2/|c| \sim 3.3$ which leads to the synchronization of Rhea's perihelion with that of Titan (Greenberg, 1977).

3. Terrestrial protoplanet systems

In this section we briefly describe the dynamical character of the terrestrial protoplanet system which we will discuss further in later sections. The numerical model of terrestrial protoplanet systems is the same as in Paper I.

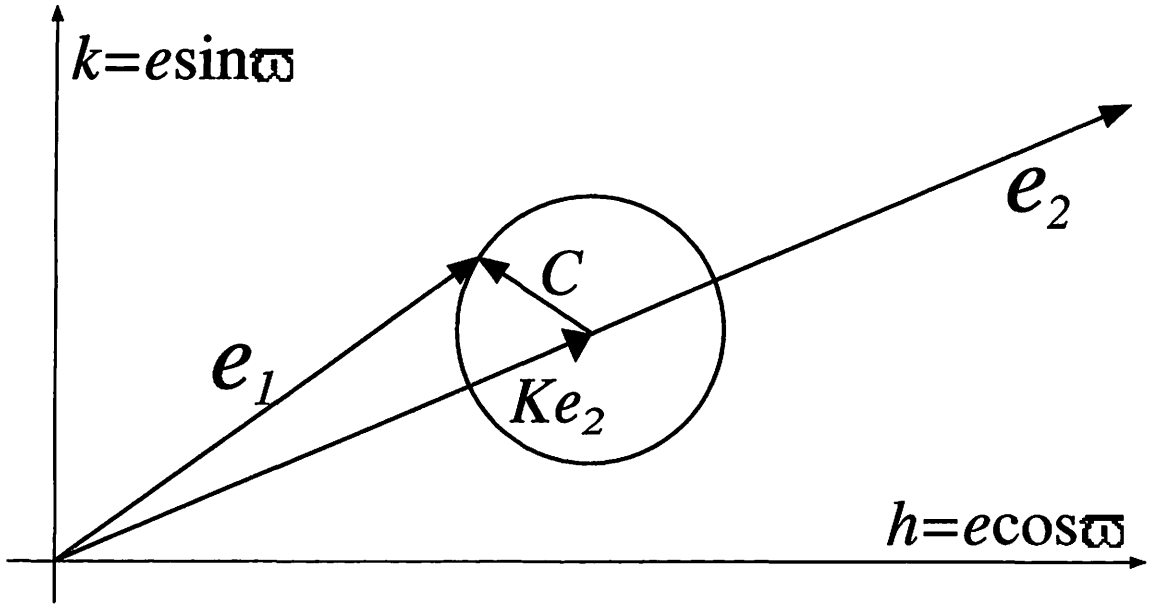


Figure 1. Schematic illustration of orbital synchronization in eccentricity vector space. e_1 and e_2 represent the eccentricity vector of the perturbed body and the perturber, respectively. Adapted from Greenberg (1977).

3.1 Model of terrestrial protoplanets and numerical method

Protoplanets, also called "planetary embryos," are regarded as precursors of the present planets. Venus, Earth, and other terrestrial planets are thought to have formed via the collisional evolution of protoplanets. Protoplanets themselves would have formed through the runaway growth of smaller planetesimals. According to the latest numerical results, a number of protoplanets would form separated by nearly equal distances with slightly eccentric orbits, through runaway growth in the swarm of planetesimals (Kokubo and Ida, 1996; Kokubo and Ida, 1998). Once all the planetesimals have accumulated into protoplanets, the typical mass of the protoplanets becomes $\sim 0.2M_\oplus$ at $a = 1$AU, and $\sim 7M_\oplus$ at $a = 7$AU where $M_\oplus$ is the mass of Earth. Typical separations among the protoplanets are around ten Hill radii, which is equivalent to 0.02AU $\sim$ 0.04AU at $a = 1$AU. Orbits of protoplanets formed through runaway growth are nearly circular due to the equipartition of energy (dynamical friction), and they remain globally stable till the end of the formation of the cores of Jupiter and Saturn within $10^6 \sim 10^7$ years. The accretion of the cores of jovian planets, i.e. the runaway growth of the protoplanets in Jupiter-Saturn region, is thought to proceed in solar nebula gas. The core accretion of Jupiter and Saturn, especially, can almost simultaneously accompany drastic collapse of nebula gas onto the surfaces of the cores due to their strong gravitational attraction (Mizuno *et al.*, 1978). Then these two planets acquire their present masses within 10^6 years, or at most 10^7 years (Pollack *et al.*, 1996; Tajima, 1998). At this point, growth of the jovian planets may catch

up with, or even surpass, growth of the terrestrial planets.

For the numerical model of the terrestrial protoplanet system, we distribute about ten to twenty particles whose total mass is about $2M_\oplus$ (which is close to the total mass of the present terrestrial planets) in the region of the present Venus to Earth. Jupiter and Saturn are assumed to possess their present masses and orbital elements outside of the terrestrial protoplanets. The terrestrial protoplanets we treat in this paper range upward from $a = a_1 = 0.5$AU with equal separations of $\Delta \times R_H$ using the mutual Hill radius R_H, where a_1 is the semimajor axis of the innermost protoplanet. A parameter Δ controls the planetary separation. The semimajor axis of the $i+1$-th protoplanet is determined by an implicit relationship $a_{i+1} = a_i + \Delta \times R_H^{i,i+1}$, where $R_H^{i,i+1}$ is the mutual Hill radius between the i-th and the $i+1$-th protoplanets (Gladman 1993, CWB). Spatial distributions of protoplanets differ from model to model. The number of protoplanets is denoted by N.

We utilize two kinds of numerical integration method: one is second-order symplectic mapping (Wisdom and Holman, 1991) with a special "warm start" procedure (Saha and Tremaine, 1992; Saha and Tremaine, 1994), and the other is a fourth-order time-symmetric Hermite integrator (Makino and Aarseth, 1992; Kokubo *et al.*, 1998). We employ the Hermite integrator when either of two bodies approaches the other within five Hill radii. Otherwise we use symplectic mapping.

We assume that the eccentricities and inclinations of protoplanets obey Rayleigh distribution with dispersions $\langle e^2 \rangle^{\frac{1}{2}}$ and $\langle I^2 \rangle^{\frac{1}{2}}$, and satisfy $\langle e^2 \rangle^{\frac{1}{2}} = 2 \langle I^2 \rangle^{\frac{1}{2}}$ (Ida and Makino, 1992). We mainly assume $\langle e^2 \rangle^{\frac{1}{2}} \leq 0.005$ according to the results of numerical N-body simulations of runaway growth of planetesimals. The initial values of the mean anomalies are selected randomly from a uniform distribution subject to the constraint that the longitudes of bodies on adjacent orbits differ by at least 20°. The initial values of the longitudes of ascending node Ω, and perihelion ϖ are randomly selected from 0° to 360°. We do not consider any variation of masses nor separations in protoplanet systems as we did in Paper I.

3.2 Orbital synchronization in protoplanet systems

Typical terrestrial protoplanet systems satisfy the condition of orbital synchronization described in the previous section, $|c| < |K|e_2$. As an example, we consider Jupiter as a perturber with $m_2 = 1.0 \times 10^{-3} M_\odot$, $a_2 = 5.2$AU, and $e_2 = 0.049$. $M_\odot (= M)$ is a solar mass. In the following sections we use a subscript J to denote the quantity for Jupiter (e.g. $e_J, \varpi_J, a_J, \ldots$). We adopt $g = g_J = 2\pi/(3 \times 10^5)$ (rad/year) as the constant precession rate of Jupiter's perihelion in (7), since the circulation period of the present Jupiter's longitude of perihelion is about 3×10^5 years. Let us take a protoplanet as the perturbed particle with semimajor axis $a = 0.8$AU and mass $m = 0.1 M_\oplus = 3.0 \times 10^{-7} M_\odot$. The amplitude of free oscillation c in the protoplanet system is quite small, since random velocity of the system would have been reduced through dynamical friction among the

swarm of planetesimals. When we assume $c = 0.005$ and calculate the Laplace coefficients, $A = 2.63 \times 10^{-5}$ (rad/year), $B = 5.36 \times 10^{-6}$ (rad/year), so $K = B/(A - g) \sim 1.0$. Thus $|c| = 0.005 < |K|e_2 \sim 0.05$, which means that the orbital synchronization occurs between Jupiter and a perturbed protoplanet in this case.

In addition, the orbital synchronization in protoplanet systems is characterized by the fact that the orbits of several perturbed protoplanets are synchronized. One of the results of the orbital evolution of 14 protoplanets ($\Delta = 10.4$) perturbed by Jupiter and Saturn is shown as Figure 2. Perihelia of protoplanets synchronize and apparently follow that of Jupiter. Note that apparent circulation of protoplanets' perihelia when their orbits are nearly circular ($e \sim 0$) is caused by the fact that perihelia themselves are only weakly defined in nearly circular orbits. Synchronization of the protoplanets' eccentricities is also seen in Figure 2(b).

The orbital synchronization in terrestrial protoplanets is qualitatively reproduced by a first-order secular perturbation theory (Brouwer and Clemence, 1961) as in Figure 3. Solutions by perturbation theory well reproduce those given by numerical integration. Henceforward, we will use secular perturbation theory to explain the mechanism of orbital synchronization among several perturbed particles. Note that in Figure 3(b), the circulation period of Jupiter's longitude of perihelion differs somewhat from numerical integration results in Figure 2(b). This is because we have neglected the terms derived from the 5:2 mean motion commensurability between Jupiter and Saturn — called "the great inequality" — in the original disturbing function in our calculation of basic planetary frequencies (cf. Brouwer and van Woerkom 1950).

The orbital synchronization of protoplanet systems is also seen when we specify only Jupiter as a perturber (Figure 4). This is a plausible condition in the terrestrial protoplanet system since there must be some time-lag between the formation of Jupiter and that of Saturn, due to the difference in their distances from the Sun. In the meantime, only Jupiter would exist as a perturber for the terrestrial protoplanets. In Figure 4(a), Jupiter's perihelion is nearly fixed in inertial space, and perihelia of the protoplanets librate around that of Jupiter. Synchronization of eccentricities is also apparent (Figure 4(b)). It should also be noted here that apparent circulation of protoplanets' perihelia when their orbits are nearly circular ($e \sim 0$) is caused by the fact that perihelia themselves are only weakly defined in nearly circular orbits.

3.3 Interaction among neighboring particles

Even when there are several perturbed particles, we can also get linear solutions similar to solution (8) which express orbital synchronization among perturbed particles. But here our aim is to provide a simpler and more intuitive explanation of orbital synchronization.

When there exist more than one perturbed particle, libration frequencies of their perihelia around that of Jupiter are all different if there is no mutual gravitational interaction. This

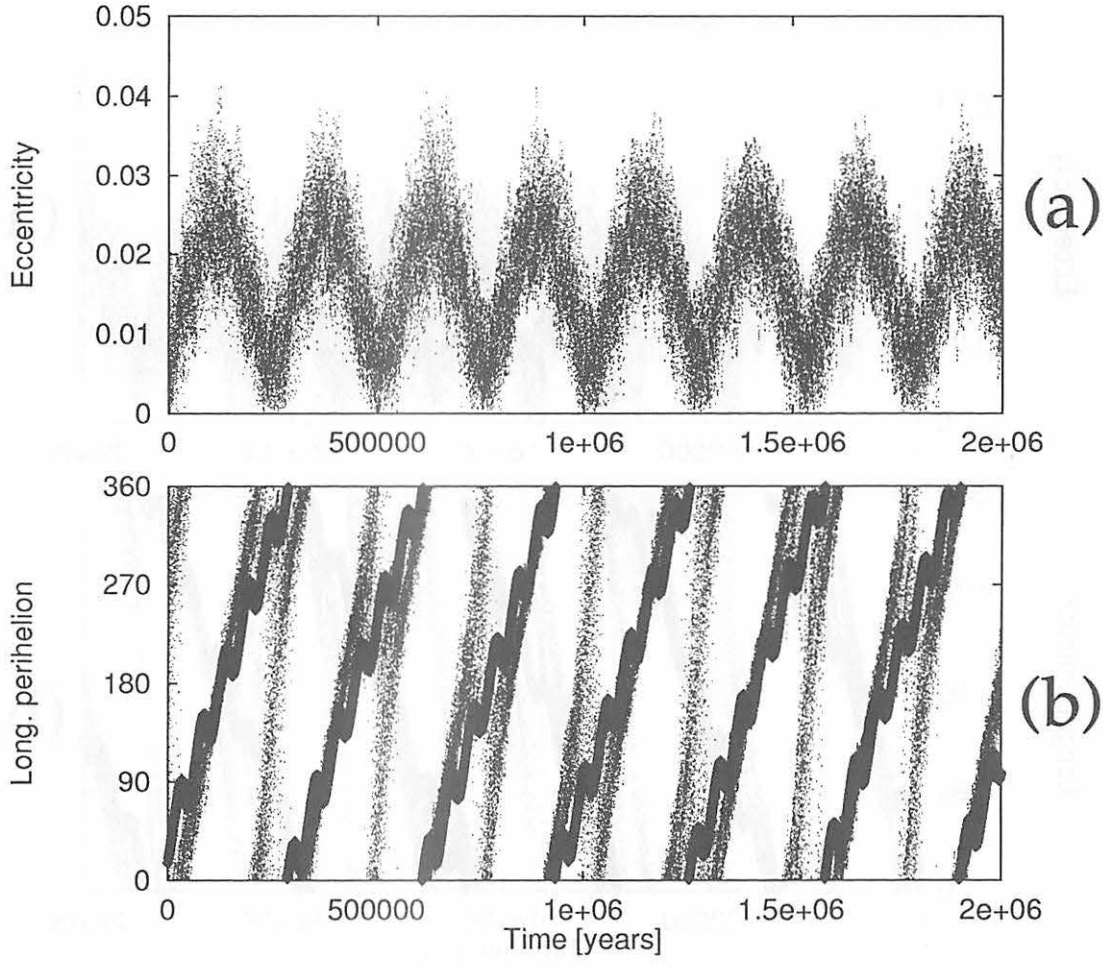


Figure 2. Typical example of orbital synchronization in a terrestrial protoplanet system perturbed by Jupiter and Saturn for two million years. $N = 14$, $m = 5 \times 10^{-7} M_{\odot}$, $\Delta = 10.4$, and $\langle e^2 \rangle^{\frac{1}{2}} = 2 \langle I^2 \rangle^{\frac{1}{2}} = 0.005$. (a) eccentricities and (b) longitudes of perihelion. The thick line in (b) denotes the variation of Jupiter's longitude of perihelion. All data of 14 protoplanets are plotted together.

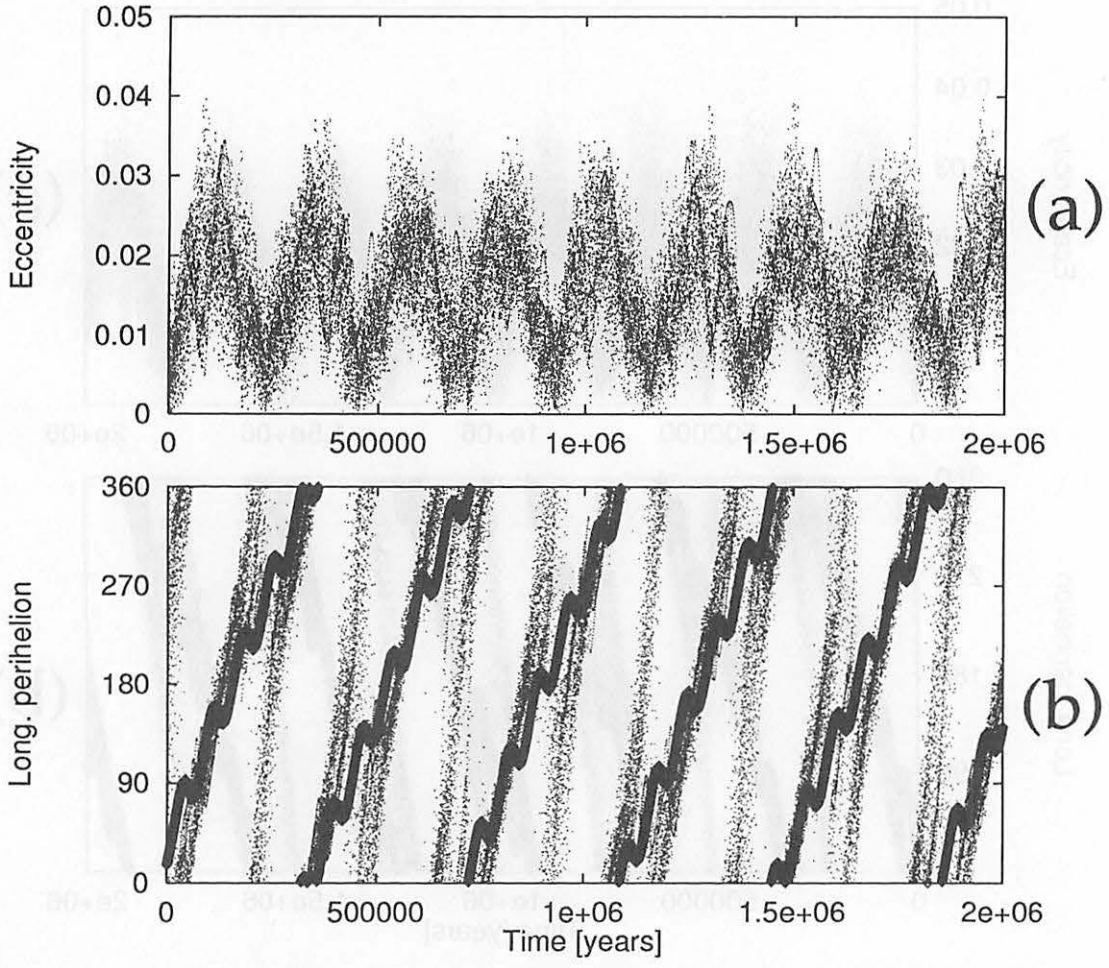


Figure 3. Typical example of orbital synchronization using a first-order secular perturbation theory for two million years. Initial orbital elements of protoplanets are the same as in Figure 2.

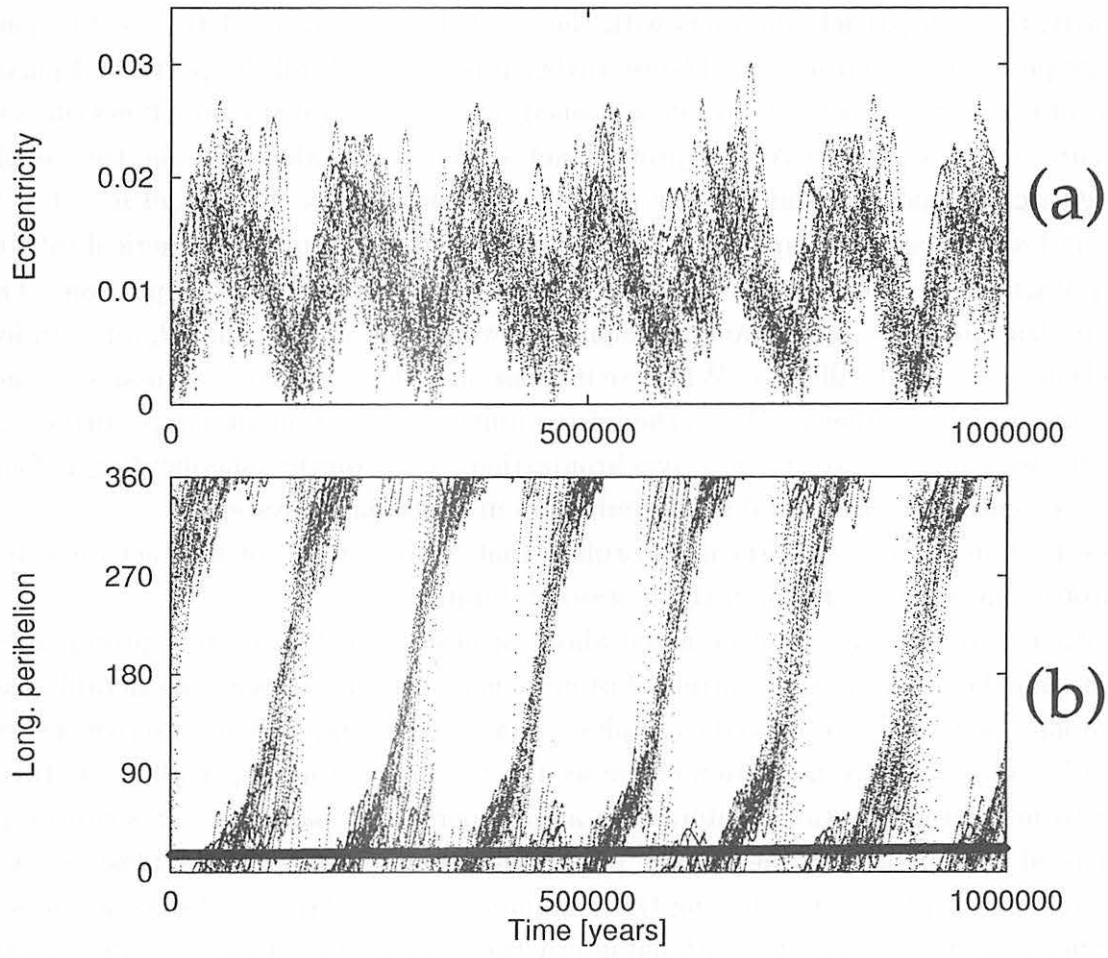


Figure 4. Typical example of orbital synchronization in a terrestrial protoplanet system perturbed only by Jupiter for one million years. $N = 14$, $m = 5 \times 10^{-7} M_{\odot}$, $\Delta = 11.7$, and $\langle e^2 \rangle^{\frac{1}{2}} = 2 \langle I^2 \rangle^{\frac{1}{2}} = 0.005$. (a) eccentricities and (b) longitudes of perihelion.

is because libration frequency is basically a function of two things, mass and the semimajor axis of the perturbed particle in restricted cases. However in an actual system with mutual gravitational interaction, neighboring particles can perturb one another, acting both as perturbers and as perturbed bodies. Hence, given some appropriate initial parameters, the orbital elements of several neighboring particles can synchronize due to this "mutual shepherding". To confirm this, we have devised a secular disturbing function which includes only a limited number (n_i) of mutual interactions. For example, when $n_i = 1$, the i-th particle interacts with only the $i-1$-th and $i+1$-th particles, in addition to the perturber. Similarly, the i-th particle interacts with the $i-2$-th, $i-1$-th, $i+1$-th, $i+2$-th particles and the perturber when $n_i = 2$. The perturber interacts with all the perturbed particles.

One of the resulting orbital motions obtained by this secular disturbing function is shown in Figure 5. We chose an $N = 9$ protoplanet system, and calculated the time evolution of eccentricities and longitudes of the perihelia of protoplanets. Results of $n_i = 0, 1, 2, 3, 6$ cases and a normal (fully interactive) case are plotted. A result of numerical integration which starts from the same initial condition is also provided for comparison. Orbital synchronization of the protoplanets is achieved even when $n_i = 1$, though amplitudes and periods are somewhat different. When we increase n_i, synchronization approaches a normal (fully interactive) condition. When there is no mutual interaction among perturbed bodies ($n_i = 0$), we cannot recognize any synchronization. Thus mutual shepherding is found to play an essential role in orbital synchronization in protoplanet systems.

Note that in Figure 5(a)$\sim$(f) we postulate that Saturn does not interact directly with the protoplanets except through the motion of Jupiter.

Another thing which should be noted about orbital synchronization in protoplanet systems is that the oscillatory amplitude of protoplanets' eccentricities is considerably reduced in comparison with the restricted examples (cf. $n_i = 0$ (a) and fully interactive (f) cases in Figure 5). This is due to migration of the secular resonance zone, especially ν_5 in this case, owing to mutual gravitational interaction among perturbed particles. This can be partly understood by calculating the secular precessional frequency $g(a)$ of a massless particle in the terrestrial planet zone having the semimajor axis a (Paper I). Migration of secular resonance through mutual gravitational interaction is considered to play an important role in the stability of present terrestrial planets (Innanen *et al.*, 1998).

4. Orbital synchronization and stability

4.1 CWB diagrams

Orbital synchronization acts as a stabilizer of protoplanet systems, preventing close encounters between particles, as we can see geometrically (Figure 6). Weakening of synchronization increases the possibility of close encounters between neighboring protoplanets.

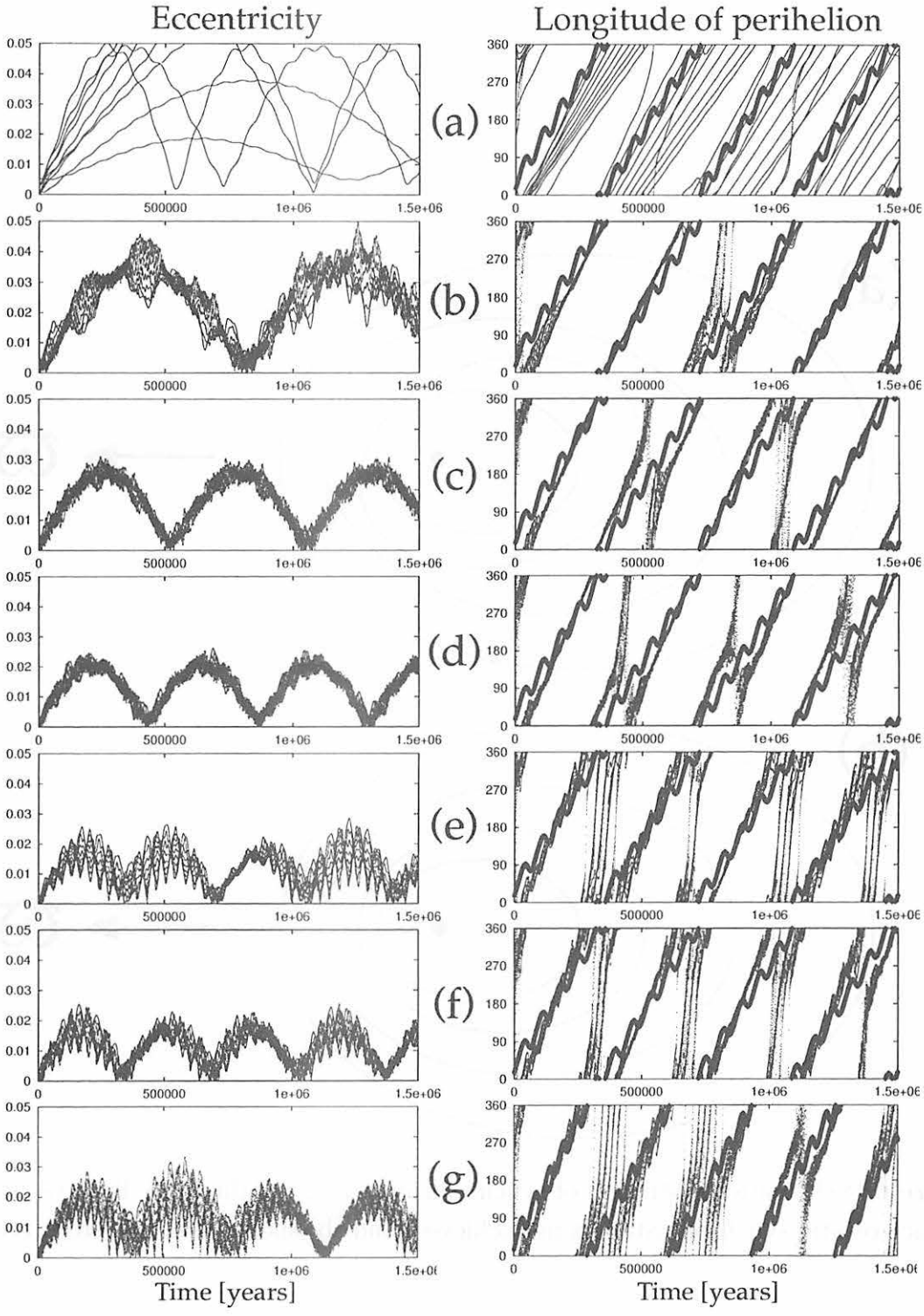


Figure 5. The orbital evolution of a protoplanet system of $N = 9$, $m = 5 \times 10^{-7} M_{\odot}$, $\Delta = 10.4$, $a_1 = 0.5$AU, and $\langle e^2 \rangle^{\frac{1}{2}} = 2 \langle I^2 \rangle^{\frac{1}{2}} = 0.002$ calculated with a secular disturbing function containing a limited number of interactions (n_i) among protoplanets. (a) $n_i = 0$, (b) $n_i = 1$, (c) $n_i = 2$, (d) $n_i = 3$, (e) $n_i = 6$, (f) normal (fully interactive), (g) a result by numerical integration. Thick lines in right panels denote the variation of Jupiter's longitude of perihelion.

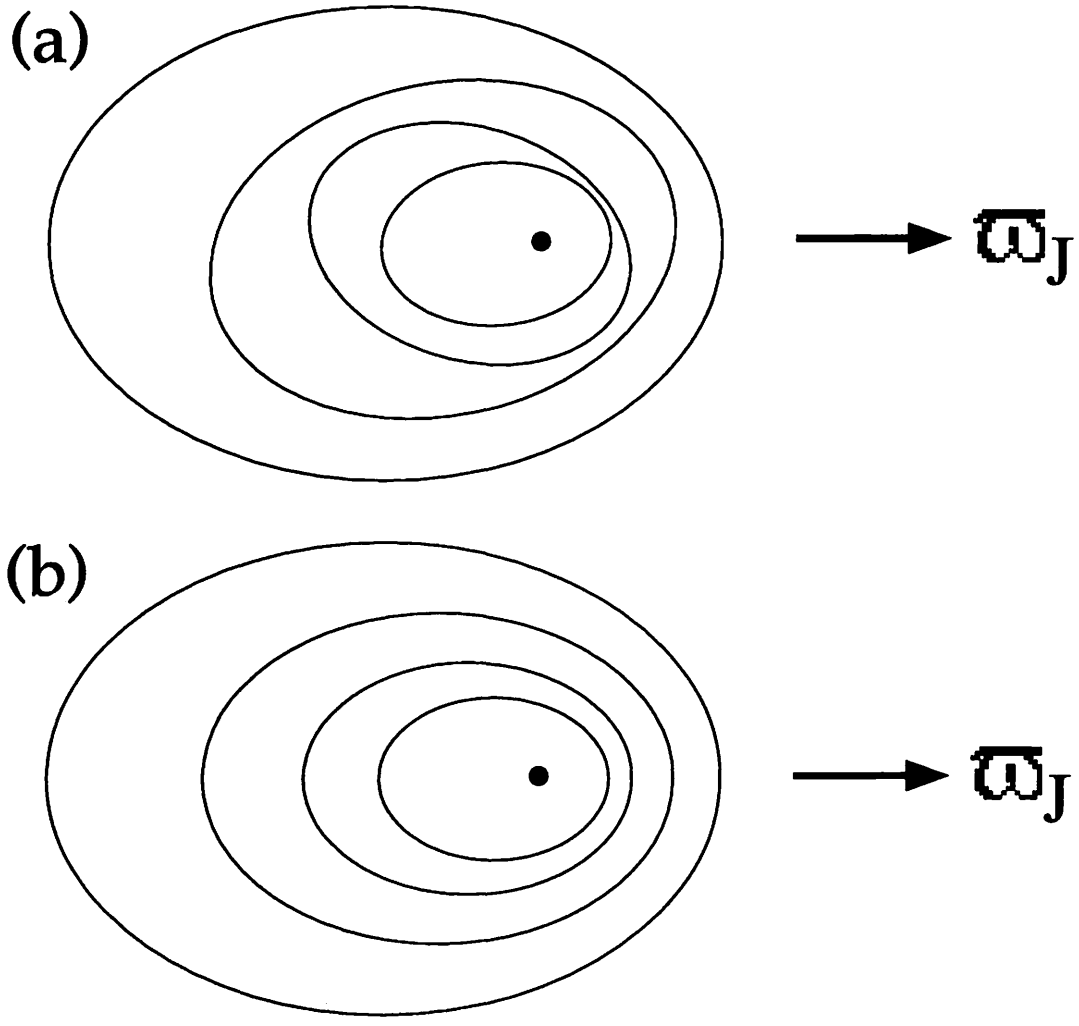


Figure 6. Schematic illustration of synchronization and stability. (a) displays an example in which orbital synchronization is not achieved, and (b) shows a case where it is achieved.

We estimate the instability timescale T_I of the protoplanet system by measuring the time elapsing till a first close encounter takes place, starting from a stable initial condition. Here, a system is defined as becoming unstable when a first close encounter occurs somewhere in the system. We consider that a system experiences a close encounter when either of two bodies approaches the other within the larger Hill radius.

The relation between the instability timescale T_I and normalized separation Δ of protoplanet systems is expressed as $(\Delta, \log T_I)$-plot called the CWB diagram advocated first by CWB. Making a CWB diagram is merely equivalent to taking snapshots of the dynamical evolution of the system. Detailed numerical N-body simulations are required to describe the actual evolution of protoplanet systems (Levison *et al.*, 1998; Chambers and Wetherill, 1999). But it is convenient to utilize CWB diagrams to investigate some of the basic dynamical characteristics of protoplanet systems with fixed boundary conditions,

When no perturbers such as jovian planets are taken into account, $\log T_I$ of the terrestrial protoplanet system is found to be proportional to Δ (Figure 7(a)). Instability timescale T_I seems to be limitlessly longer when Δ increases. But once we introduce massive perturbers outside of terrestrial protoplanets, this is not the case. Figure 7(b) is a typical CWB diagram of 14-protoplanet systems with equal mass ($5 \times 10^{-7} M_\odot$) and small random velocity $\langle e^2 \rangle^{\frac{1}{2}} = 2 \langle I^2 \rangle^{\frac{1}{2}} = 0.005$. The instability timescale T_I is longest in the modest-Δ region, and shorter in the small- and large-Δ regions. T_I is limited to a maximum of 10^6 to 10^7 years.

Note that in small-Δ regions such as $\Delta < 8$, the logarithmic average of T_I is somewhat longer in perturbed protoplanet systems (Figure 7(b)) than in systems that are not perturbed (Figure 7(a)). This is probably because orbital synchronization occurs in the perturbed protoplanet system, and prevents particles from making close encounters for a slightly longer period than in the unperturbed protoplanet system, as we see in the next subsection and Figure 8(a)(b).

The stabilizing effect of orbital synchronization is more apparent in protoplanet systems with larger random velocity. Larger random velocities such as $\langle e^2 \rangle^{\frac{1}{2}} = 2 \langle I^2 \rangle^{\frac{1}{2}} = 0.015$ are plausible, since the collisional evolution of protoplanet systems tends to enhance the eccentricities and inclinations of particles. In large random velocity systems such as $\langle e^2 \rangle^{\frac{1}{2}} = 2 \langle I^2 \rangle^{\frac{1}{2}} = 0.015$, instability timescale T_I varies considerably depending on initial conditions (Figure 7(c)). We performed some numerical experiments in which all the orbits initially synchronize ($e_j = 0.015, \varpi_j = \varpi_J$ where $j = 1, \ldots, N$). One of the resulting CWB diagrams is shown as Figure 7(d). Instability timescale T_I becomes much longer than in Figure 7(c) due to the initial orbital synchronization.

4.2 Mutual interaction and tidal torque

The relation between orbital synchronization and stability depends on the orbital separation and spatial range of the system which is described by the parameter, Δ. Δ controls

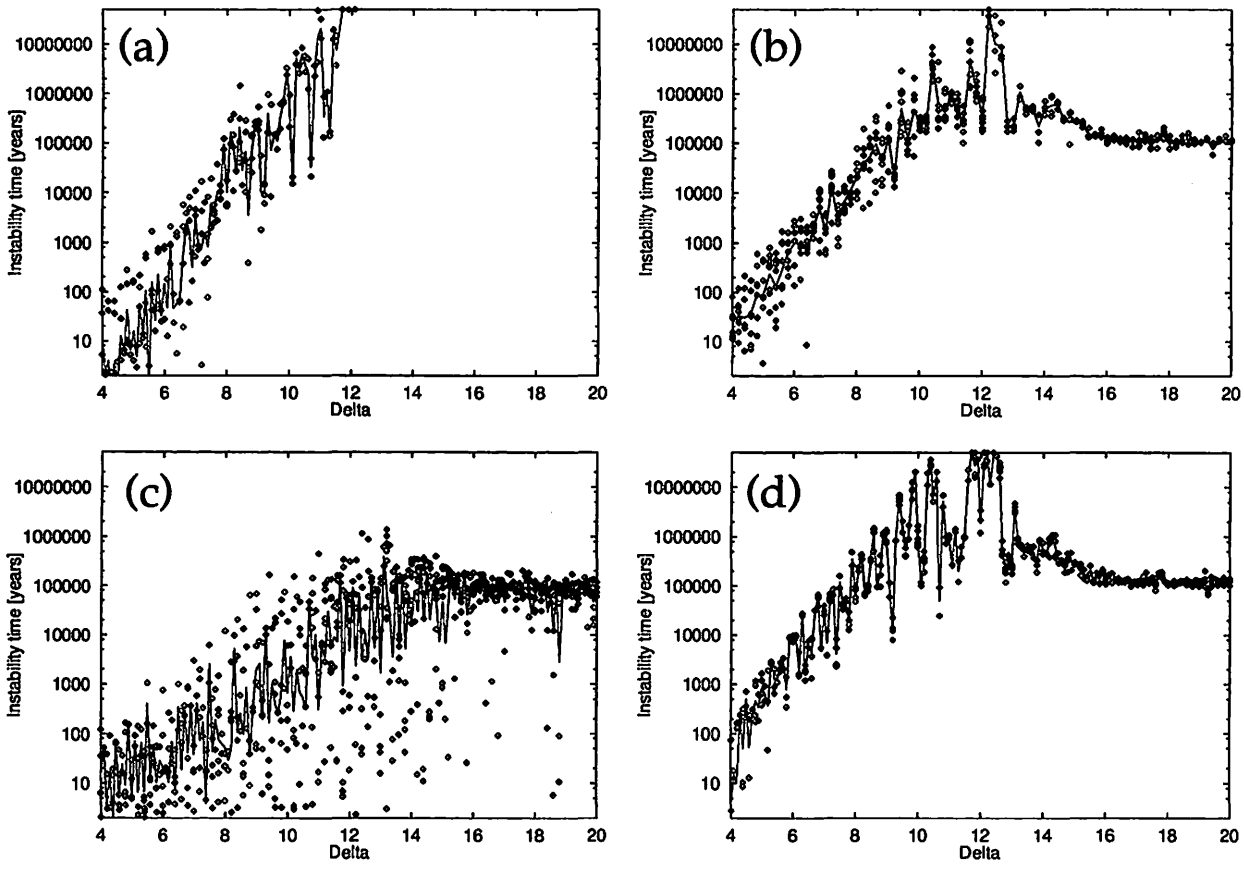


Figure 7. Typical examples of CWB diagrams. "Delta" in the abscissa denotes Δ. Several data for a Δ bin from different initial conditions are plotted by squares, together with their logarithmic averages by solid lines. In all panels, $N = 14$ and $m = 5 \times 10^{-7} M_{\odot}$. (a) $\langle e^2 \rangle^{\frac{1}{2}} = 2 \langle I^2 \rangle^{\frac{1}{2}} = 0.005$ protoplanet systems which are not disturbed by any perturber outside. (b) $\langle e^2 \rangle^{\frac{1}{2}} = 2 \langle I^2 \rangle^{\frac{1}{2}} = 0.005$ protoplanet systems perturbed by Jupiter and Saturn. (c) $\langle e^2 \rangle^{\frac{1}{2}} = 2 \langle I^2 \rangle^{\frac{1}{2}} = 0.015$ protoplanet systems perturbed by Jupiter and Saturn. (d) $\langle e^2 \rangle^{\frac{1}{2}} = 2 \langle I^2 \rangle^{\frac{1}{2}} = 0.015$ protoplanet systems perturbed by Jupiter and Saturn, in which the initial orbital elements of all the protoplanets are synchronized.

not only the strength of the mutual gravitational interaction among protoplanets, but also the strength of disturbance by perturbers. The disturbing force on a protoplanet by outer perturbers is a function of distance between each protoplanet and the perturber; it is proportional to the inverse cube of the distance. This means that the disturbance by the outer perturber serves as a tidal force (cf. Danbby 1992). Differences in the distance from the perturber result in differences in the strength of the disturbing force exerted on perturbed bodies. These differential forces work as a kind of torque on perturbed particles. We hereafter refer to this differential force as "tidal torque".

Tidal torque creates differences in the natural precession rate of each perturbed particle. In small-Δ systems, tidal torque is small, hence the difference in the natural precession rate of each orbit is small. Although it can lead to the synchronization of orbital elements, strong mutual gravitational interaction in small-Δ systems threatens their stability. In a small-Δ system ($\Delta = 8.2$), orbital synchronization seems to be well established (Figure 8(a)). But within 3×10^5 years, instability occurs due to too great a mutual interaction. On the other hand, mutual gravitational interaction in large-Δ systems is weaker than in small-Δ systems. But tidal torque produced by a perturber is stronger, and the natural precession rate of each protoplanet's orbit differs a great deal. This can lead to orbital crossings, and finally close encounters occur. In a large-Δ system ($\Delta = 20.0$), no orbital synchronization is seen (Figure 8(b)) since the natural precession rates are largely different in each particle, and the mutual interaction among particles is weak. Orbital crossing provokes rapid close encounters in this case. In the largest-Δ area in Figure 7(b), T_I is bound to several 10^5 years. This is a limit imposed by the precession period of Jupiter's perihelion, about 3×10^5 years. This means that the stability of protoplanet systems with large-Δ cannot be maintained even during a single circulation of Jupiter's perihelion.

Thus the system is relatively stable in the modest-Δ area where the mutual gravitational interaction and tidal torque balance.

4.3 Gradual break down of orbital synchronization

The orbital synchronization in protoplanet systems keeps them stable. However, a protoplanetary system would have to become unstable to evolve toward the present planetary system; protoplanets must collide with one another and reduce their numbers.

The orbital synchronization, which seems to be well maintained at the modest-Δ region such as $10 \leq \Delta \leq 12$ in Figure 7(b), is a transient process having at most a $O(10^7)$-year timescale. The initial configuration of the terrestrial protoplanet system (i.e. nearly equal masses and separations such as $8 \leq \Delta \leq 10$, and small random velocity) is likely to induce the orbits to synchronize. But the synchronization gradually breaks down due to short-periodic or higher-order perturbation among protoplanets and/or by perturbers in $O(10^6)$ to $O(10^7)$-year timescales. The first-order secular perturbation theory we use in this paper does not contain these effects.

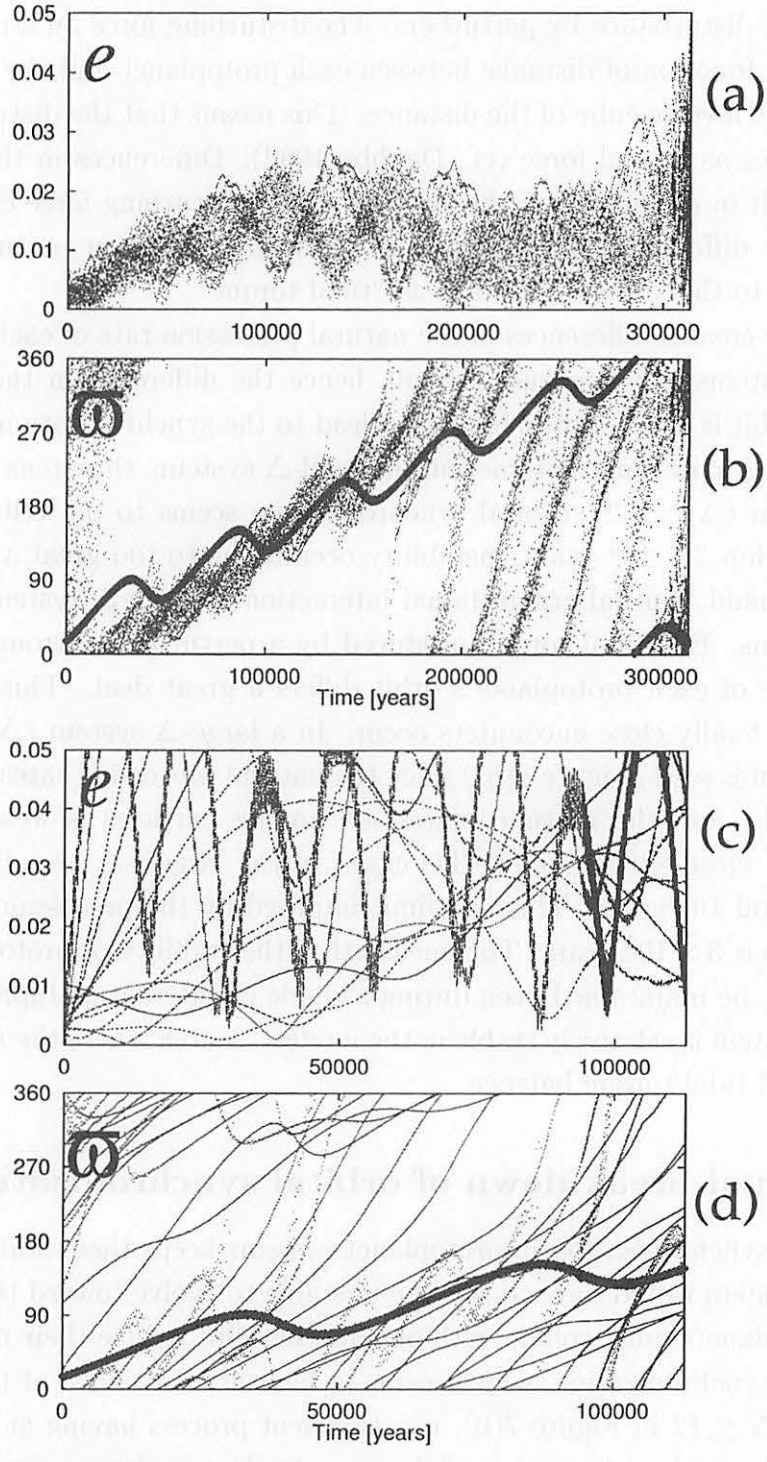


Figure 8. The orbital evolution of protoplanets in a smaller- and a larger-Δ systems. (a) the eccentricities and (b) longitudes of perihelion of a $\Delta = 8.2$, $N = 14$, $m = 5 \times 10^{-7} M_{\odot}$, and $\langle e^2 \rangle^{\frac{1}{2}} = 2 \langle I^2 \rangle^{\frac{1}{2}} = 0.005$ system, and (c) the eccentricities and (d) longitudes of perihelion of a $\Delta = 20.0$, $N = 14$, $m = 5 \times 10^{-7} M_{\odot}$, and $\langle e^2 \rangle^{\frac{1}{2}} = 2 \langle I^2 \rangle^{\frac{1}{2}} = 0.005$ system. In the right side of (c) and (d), the density of dots becomes higher due to the automatic change of integration method from symplectic mapping to Hermite integrator, following the change of minimum distance among protoplanets.

We define some parameters which indicate the degree of orbital synchronization using eccentricity vector $\boldsymbol{e}_j = (e_j \cos \varpi_j, e_j \sin \varpi_j)$. The root mean square of eccentricity vectors of perturbed protoplanets is defined here as

$$\left|\boldsymbol{e}^2\right|^{\frac{1}{2}} \equiv \sqrt{\frac{\sum_{j=1}^{N} \left|\bar{\boldsymbol{e}} - \boldsymbol{e}_j\right|^2}{N}}, \tag{10}$$

and maximum deviation from the averaged value as

$$\delta_{max}\boldsymbol{e} \equiv \max \left|\bar{\boldsymbol{e}} - \boldsymbol{e}_j\right|, \tag{11}$$

where N is the number of protoplanets and $\bar{\boldsymbol{e}}$ is the average of their eccentricity vectors,

$$\bar{\boldsymbol{e}} = \frac{\sum_{j=1}^{N} \boldsymbol{e}_j}{N}. \tag{12}$$

We selected a numerical example from the calculations in Figure 7(b), and showed their $\left|\boldsymbol{e}^2\right|^{\frac{1}{2}}$, $\delta_{max}\boldsymbol{e}$, and minimum distances among all particles in Figure 9. We can see that orbital synchronization gradually breaks down, finally resulting in a close encounter.

5. Implications for other celestial systems

Finally we reconsider two other celestial systems which consist of a central mass, several perturbed bodies, and distant perturbers, from the viewpoint of orbital synchronization and stability: the first is a hypothetical planetary system in a binary star system, and the second is the Uranian ϵ-rings.

Innanen et al. (1997) showed that in a highly inclined planetary system around the primary star of a binary, nodes and inclinations of the planets show strong synchronization as if they compose a rigid disk. This phenomenon seems to be a variant of the orbital synchronization described in this paper, occurring at a region where the mutual gravitational interaction and the tidal torque balance. A difference is that the "dynamical rigidity" in Innanen et al. (1997) is a process specific to three dimensional space. In such a case, orbital synchronization occurs only in (I, Ω) pairs, not in (e, ϖ) pairs. We performed some extensional numerical calculations of Innanen and collaborators and confirmed that orbital synchronization in (e, ϖ) pairs does not occur in the system proposed by Innanen et al. With the exception of this feature, this "dynamical rigidity" shows characteristics common with the orbital synchronization seen in protoplanet systems. When the mutual gravitational interaction among planets is weak, planets suffer from Kozai resonance. Their eccentricities and inclinations oscillate at quite a large amplitude, which can lead to instability. When the mutual gravitational interaction among planets is strong, close encounters among planets are likely to occur. Although getting an analytical explanation is more difficult in a highly eccentric system with a large inclination than in a system with little

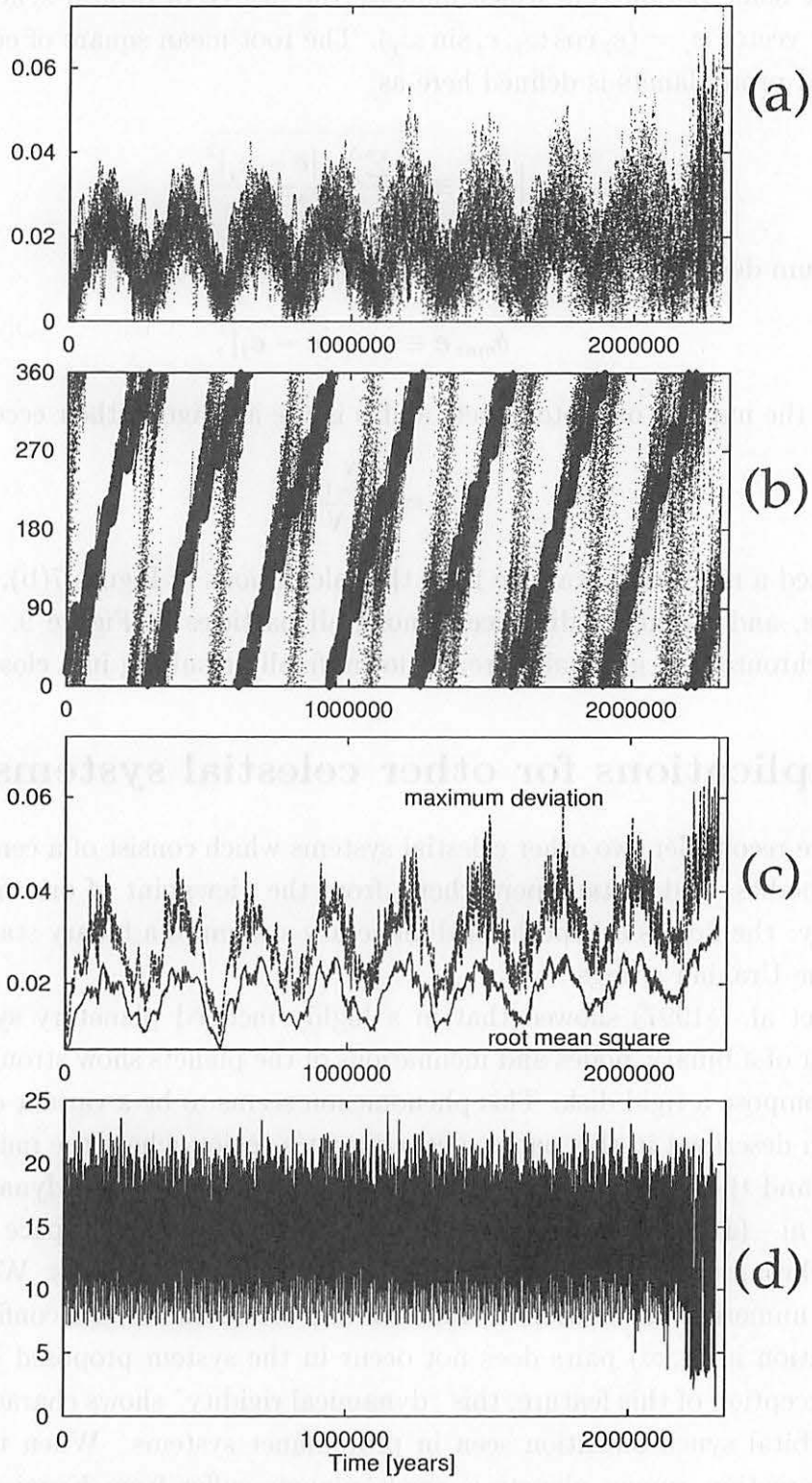


Figure 9. The orbital evolution of protoplanets in a $\Delta = 11.8$, $N = 14$, $m = 5 \times 10^{-7} M_{\odot}$, and $\langle e^2 \rangle^{\frac{1}{2}} = 2 \langle I^2 \rangle^{\frac{1}{2}} = 0.005$ system. (a) eccentricity, (b) longitude of perihelion, (c) maximum deviation of the eccentricity vector from the averaged value $\delta_{max} \boldsymbol{e}$ and root mean square of eccentricity vectors $|\boldsymbol{e}^2|^{\frac{1}{2}}$, (d) minimum distance among all the protoplanets normalized by the mutual Hill radius (R_H).

eccentricity and inclination such as the terrestrial protoplanet system, the mechanisms of orbital synchronization and "dynamical rigidity" seem to be essentially the same. According to Innanen's numerical integration, "dynamical rigidity" also collapses into instability after a long interval such as $\sim 10^8$ years ($\sim 10^7$ revolutions of Jupiter).

If planetary rings such as Uranian ϵ-rings consist of small ringlets having nearly equal mass and equal separation like the terrestrial protoplanets, they can be a typical example of orbital synchronization. Perturbation by shepherding satellites and the gravitational oblateness of Uranus causes ϵ-ringlets to precess, whereas mutual gravitational interaction among ringlets helps them to synchronize. Then, the rings are assumed to remain stably elliptic through the orbital synchronization of ringlets. Since the width of the ring is quite narrow compared with the orbital radius of the ring, natural precession rates due to perturbers are not so different at the innermost and outermost edges of the ring; i.e. the tidal torque by perturbers is relatively weak. This is the biggest difference between Uranian ϵ-rings and a terrestrial protoplanet system whose spatial width (such as $\sim$ 1AU) is comparable to its orbital radius.

However, actual Uranian ringlets are thought to have a wide range of mass distribution and be exposed to frequent physical collision (Goldreich and Tremaine, 1979a; Elliot *et al.*, 1984; French *et al.*, 1991; Esposito *et al.*, 1991; Hui *et al.*, 1991). In that case, their stability may not be maintained by the simple orbital synchronization described in this paper.

Acknowledgments

The authors have greatly benefited from stimulating discussions with and encouragement from Eiichiro Kokubo and Shigeru Ida. Detailed and constructive review by Yolande McLean has considerably improved the English presentation of this paper. This study was supported by the Decoding Earth Evolution Program (DEEP), an intensified study area program of the Ministry of Culture and Education (No. 259, 1995–1997), and the ADAC scientific simulation project, National Astronomical Observatory of Japan (dti02–1998).

References

Allan, R.R. (1969) Evolution of Mimas-Tethys Commensurability, *Astron. J.*, **74**, 497–506.

Brouwer, D. and Clemence, G.M. (1961) *Methods of Celestial Mechanics*, Academic Press, New York.

Brouwer, D. and van Woerkom, A.J.J. (1950) The secular variations of the orbital elements of the principal planets, *Astron. Pap. Amer. Ephemeris. Naut. Alm.*, **13**, 2, 81–107.

Chambers, J.E. and Wetherill, G.W. (1999) Making the terrestrial planets: N-body integrations of planetary embryos in three dimensions, *Icarus*, in press.

Chambers, J.E., Wetherill, G.W., and Boss, A.P. (1996) The stability of multi-planet systems, *Icarus*, **119**, 261–268.

Danby, J.M.A. (1992) *Fundamentals of Celestial Mechanics (second edition, third printing)*, Willmann-Bell Inc., Richmond, Virginia.

Elliot, J.L., French, R.G., Meech, K.J., and Elias, J.H. (1984) Structure of the uranian rings. I. squre-well model and particle-size constraints, *Astron. J.*, **89**, 1587–1603.

Esposito, L.W., Brahic, A., Burns, J.A., and Marouf, E.A. (1991) Particle properties and processes in Uranus's rings, in *Uranus*, The University of Arizona Press, Tucson & London, 410–468.

French, R.G., Nicholson, P.D., Porco, C.C., and Marouf, E.A. (1991) Dynamics and structure of the uranian rings, in *Uranus*, The University of Arizona Press, Tucson & London, 327–409.

Gladman, B. (1993) Dynamics of systems of two close planets, *Icarus*, **106**, 247–263.

Goldreich, P. and Porco, C.C. (1987) Shepherding of the Uranian rings. II. Dynamics, *Astron. J.*, **93**, 730–737.

Goldreich, P. and Tremaine, S. (1979a) Precession of the ϵ ring of Uranus, *Astron. J.*, **84**, 1638–1641.

Goldreich, P. and Tremaine, S. (1979b) Towards a theory for the uranian rings, *Nature*, **277**, 97–99.

Greenberg, R. (1973) The inclination-type resonance of Mimas and Tethys, *Mon. Not. R. Astron. Soc.*, **165**, 305–311.

Greenberg, R. (1975) On the Laplace relation among the satellites of Uranus, *Mon. Not. R. Astron. Soc.*, **173**, 121–129.

Greenberg, R. (1976) The Laplace relation and the masses of Uranus' satellite, *Icarus*, **29**, 427–433.

Greenberg, R. (1977) Orbit-orbit resonances in the solar system: varieties and similarities, *Vistas Astron.*, **41**, 209–239.

Hui, J., Horn, L.J., and Lane, A,L. (1991) Particle size of the Uranus delta ring's inner diffuse companion through comparison of RSS and PPS Voyager occultation data, *Icarus*, **93**, 347–353.

Ida, S. and Makino, J. (1992) N-body simulation of gravitational interaction between planetesimals and a protoplanet I. velocity distribution of planetesimals, *Icarus*, **96**, 107–120.

Innanen, K.A., Zheng, J.Q., Mikkola, S., and Valtonen, M.J. (1997) The Kozai mechanism and the stability of planetary orbits in binary star systems, *Astron. J.*, **113**, 1915–1919.

Innanen, K.A., Mikkola, S., and Wiegert, P. (1998) The Earth-Moon system and the dynamical stability of the inner solar system, *Astron. J.*, **116**, 2055–2057.

Ito, T. and Tanikawa, K. (1999) Stability and instability of the terrestrial protoplanet system and their possible roles in the final stage of planet formation, *Icarus*, in press.

Kokubo, E. and Ida, S. (1996) On runaway growth of planetesimals, *Icarus*, **123**, 180–191.

Kokubo, E. and Ida, S. (1998) Oligarchic growth of protoplanets, *Icarus*, **131**, 171–178.

Kokubo, E., Yoshinaga, K., and Makino, J. (1998) On a time-symmetric Hermite integrator for planetary N-body simulation, *Mon. Not. R. Astron. Soc.*, **297**, 1067–1072.

Kozai, Y. (1992) Shepherding satellites and dynamical structure of the rings of Uranus, *Publ. Astron. Soc. Japan*, **44**, 135–139.

Kozai, Y. (1993) Shepherding satellites and dynamical structure of the rings of Uranus. II, *Publ. Astron. Soc. Japan*, **45**, 263–267.

Levison, H.F., Lissauer, J.J., and Duncan, M.J. (1998) Modeling the diversity of outer planetary systems, *Astron. J.*, **116**, 1998–2014.

Makino, J. and Aarseth, S.J. (1992) On a Hermite integrator with Ahmad-Cohen scheme for gravitational many-body problems, *Publ. Astron. Soc. Japan*, **44**, 141–151.

Mizuno, H., Nakazawa, K., and Hayashi, C. (1978) Instability of gaseous envelope surrounding a planetary core and formation of giant planets, *Prog. Theor. Phys.*, **60**, 669–710.

Pauwels, T. (1983) Secular orbit-orbit resonance between two satellites with nonzero masses, *Celes. Mech.*, **30**, 229–247.

Peale, S.J. (1976) Orbital resonances in the solar system, *Annu. Rev. Astron. Astrophys.*, **14**, 215–246.

Pollack, J.B., Hubickyj, O., Bodenheimer, P., Lissauer, J.J., Podolak, M., and Greenzweig, Y. (1996) Formation of the giant planets by concurrent accretion of solid and gas, *Icarus*, **124**, 62–85.

Porco, C.C. and Goldreich, P. (1987) Shepherding of the Uranian rings. I. Kenematics, *Astron. J.*, **93**, 724–729.

Saha, P. and Tremaine, S. (1992) Symplectic integrators for solar system dynamics, *Astron. J.*, **104**, 1633–1640.

Saha, P. and Tremaine, S. (1994) Long-term planetary integrations with individual time steps, *Astron. J.*, **108**, 1962–1969.

Sinclair, A.T. (1972) On the origin of the commensurabilities amongst the satellites of Saturn, *Mon. Not. R. Astron. Soc.*, **160**, 169–187.

Sinclair, A.T. (1975) The orbital resonance amongst the Galilean satellites of Jupiter, *Mon. Not. R. Astron. Soc.*, **171**, 59–72.

Tajima, N. (1998) *Evolution and dynamical stability of the proto-giant-planet envelop*, PhD thesis, University of Tokyo, Tokyo.

Wisdom, J. and Holman, M. (1991) Symplectic maps for the N-body problem, *Astron. J.*, **102**, 1528–1538.

Formation of Protoplanets in the Region of Terrestrial Planets

KOUJI SHIIDSUKA

Department of Earth and Planetary Sciences, Faculty of Science
Tokyo Institute of Technology, Tokyo 152-8551, Japan
E-mail: kshiidsu@geo.titech.ac.jp

Abstract

Formation and evolution of protoplanets from a planetesimal swarm embedded in the protoplanetary solar nebula are investigated, by three-dimensional N-body simulation. Initially, the planetesimal swarm is distributed in the range of 0.5-1.5AU, in contrast to the local simulation by Kokubo & Ida (1998,1999), so that the formation of the system of multiple protoplanets which would evolve toward terrestrial planets is directly investigated. Finally, about 20 protoplanets are formed through runaway growth, whose masses almost agree with the isolation mass predicted by the earlier reference. The difference in semi-major axis between neighboring protoplanets are about 10 Hill's radius, which also agrees with the local simulation. We succeeded to scale the mass of protoplanets as a function of semi-major axis and time, which suggests that masses of protoplanets which have not yet reached the isolation mass would be distributed as $M \propto a^{-8}$.

1 INTRODUCTION

In the standard scenario of the planetary accretion, terrestrial planets are formed from small bodies called planetesimals whose initial masses are $\sim 10^{18}$g. The planetary accretion process have been mainly investigated by the statistical method (e.g. Greenberg et al. 1978; Wetherill & Stewart 1989; Weidenschilling et al. 1997) or N-body simulation (e.g. Kokubo & Ida 1996, 1998). According to them, there is a common stage called runaway growth where a massive planetesimal emerged from smaller planetesimals predominantly grows. Such a planetesimal is called a planetary embryo, or a protoplanet. In post-runaway stage, larger protoplanets grow more slowly than smaller ones, while the growth of protoplanets is still faster than that of planetesimals (Kokubo & Ida 1998). This type of growth is called oligarchic growth. In the oligarchic growth, the typical orbital separation is about $10R_{\rm H}$, where $R_{\rm H} \equiv (2M/3M_\odot)^{1/3}a$ is called Hill's radius ($M_\odot$ is the solar mass, M and a is the mass and semi-major axis of the protoplanet). Kokubo & Ida (1998) suggested if the oligarchic growth continues, a protoplanet with mass about $0.1M_\oplus$ is eventually formed at 1AU, which is much smaller than the terrestrial planets like the Earth and the Venus. This ultimate mass of the protoplanet is called the isolation mass. The protoplanet system thus formed is considered to become unstable in time scale longer than that for accretion and evolve toward the terrestrial planets through mutual collision. This final stage is studied by several authors (e.g. Chambers & Wetherill 1998, Ito & Tanikawa 1999).

However, Kokubo & Ida (1998) investigated the evolution of a planetary swarm distributed in narrow ring-like region around 1AU whose width is comparable to or narrower than the orbital separation of the protoplanet with the isolation mass. On the other hand, in the final stage toward the terrestrial planets, the instability of the protoplanet system would be caused by the gravitational interaction among three or more protoplanets (e.g. Chambers et al. 1996). Thus, it is necessary to directly investigate the masses and orbital separations of the protoplanets in more wider region which contains multiple protoplanets.

Weidenschilling et al. (1997) used multi-zoned simulation code to study the evolution of the planetesimal swarm in the region of terrestrial planets (0.5-1.5AU), and obtained the results

which is consistent with those suggested by Kokubo & Ida (1998). Their simulation is, however, based on the results of the local simulations. Thus, more direct and reliable approach is needed.

In this paper, we investigated the formation of the protoplanet system by three-dimensional N-body simulations. N-body simulation is more reliable method since orbits of all the particles are directly calculated. However, this method cannot start from sufficiently small planetesimals due to numerical difficulty. Instead, we estimated the mass of the protoplanet as a function of time, the semi-major axis, and initial mass of planetesimals. This would enable us to determine the evolution of protoplanet mass with arbitrary initial masses of planetesimals.

2 MODEL AND METHODS

2.1 Basic equations and force calculation

The equation of motion of i-th planetesimal is written as

$$\frac{\mathrm{d}\vec{v}_i}{\mathrm{d}t} = -\frac{GM_\odot}{|\vec{x}_i^3|}\vec{x}_i - \sum_{j\neq i}^{N} \frac{Gm_{\mathrm{p}j}(\vec{x}_i - \vec{x}_j)}{|\vec{x}_i - \vec{x}_j|^3} + \vec{f}_{\mathrm{gas},i}, \tag{1}$$

where $\vec{x}$ and $\vec{v}$ are the position and velocity of planetesimals respectively, G is the gravitational constant, m_{p} is the planetesimal mass, and $M_\odot$ is the solar mass. The r.h.s. of Eq.(1) represents the solar gravity, the mutual gravitational interaction with other planetesimals, and the (hydrodynamic) gas drag force from left to right. The gas drag force $\vec{f}_{\mathrm{gas}}$ takes the form (Adachi et al. 1976)

$$\vec{f}_{\mathrm{gas}} = -\frac{1}{2m_{\mathrm{p}}} C_{\mathrm{D}} \pi r_{\mathrm{p}}^2 \rho_{\mathrm{g}} |\vec{u}| \vec{u}, \tag{2}$$

where C_{D} is the gas drag coefficient, ρ_{g} is the mass density of the gas component of the solar nebula, and r_{p} is the radius of the planetesimal given by $r_{\mathrm{p}} = (3m_{\mathrm{p}}/4\pi\rho_{\mathrm{p}})^{1/3}$ (ρ_{p} is the internal density of planetesimals). We take $\rho_{\mathrm{g}} = 2\times10^{-9}(a/1\mathrm{AU})^{-11/4}\mathrm{g\,cm^{-3}}$ and $\rho_{\mathrm{p}} = 2\mathrm{g\,cm^{-3}}$, which are similar to the minimum-mass solar nebula model (Hayashi 1981).

We numerically integrate Eq.(1) by using the modified fourth order Hermite scheme (Kokubo et al. 1998) with the individual and hierarchical timestep (Makino 1991). We reduce the calculation cost by calculating the mutual gravity among planetesimals, which is the most expensive part, on special-purpose computer for N-body simulation, GRAPE-4 (Makino et al. 1997).

2.2 Accretion

For simplicity, we let two planetesimals always accrete when they contact. In the collision between larger planetesimal and smaller planetesimal, this assumption would be valid since in most cases, relative velocity between them is small compared with the escape velocity from larger one. However, in the protoplanet-dominated stage (see, section 3), collision between smaller planetesimals would lead to disruptive fragmentation and produce large amount of fragment bodies in real system. The existence of fragments would accelerate the growth of protoplanets since smaller bodies suffer strong gas drag (e.g. Wetherill & Stewart 1993). The effect of fragmentation should be investigated in the future work.

In order to save the computational time, we decrease the growth time scale by increasing a planetesimal's radius by amount of factor f, where f takes the value 5-10 in our simulations. As discussed in Kokubo & Ida (1996), moderate increment of radius makes no significant difference in the growth mode of planetesimals but only changes growth time scale. In our simulation, growth time scale is scaled as $\propto f^{-1}$ (Eq.(10).

2.3 Initial conditions

We set the surface mass density of planetesimals as $\Sigma_s = 10(a/1\mathrm{AU})^{-3/2}\mathrm{g\,cm^{-2}}$, which is about 50% larger than the minimum-mass solar nebula model. We distribute 4000 equal-mass planetesimals from 0.5 to 1.5 AU so that they satisfy the above surface mass density distribution. Thus, initial mass of a planetesimal is determined as 3.6×10^{24}g.

The initial distribution of eccentricity and inclination is given by the Rayleigh distribution (Ida & Makino 1992a), whose RMS eccentricity and inclination are determined by the equilibrium between increase due to gravitational stirring among planetesimals and damping due to gas drag (e.g. Kokubo & Ida 1999):

$$e_{\rm eq} = 2i_{\rm eq} \simeq 8.1 \times 10^{-3} \left(\frac{m_{\rm p}}{3.6\times10^{24}{\rm g}}\right)^{\frac{4}{15}} \left(\frac{a}{1{\rm AU}}\right)^{\frac{9}{20}} \left(\frac{\Sigma_{\rm s0}}{10{\rm g\,cm^{-2}}}\right)^{\frac{1}{5}} \left(\frac{\rho_{\rm g0}}{2\times10^{-9}{\rm g\,cm^{-3}}}\right)^{-\frac{1}{5}} \times \left(\frac{\rho_{\rm p}}{2{\rm g\,cm^{-3}}}\right)^{\frac{2}{15}} \left(\frac{C_{\rm D}}{10}\right)^{-\frac{1}{5}}, \tag{3}$$

where $\Sigma_{\rm s0}$ and $\rho_{\rm g0}$ are those at 1AU. In this study, we performed two simulations in which the value of f and $C_{\rm D}$ differ. In run 1, we choose $f = 10$ and $C_{\rm D} = 20$, and $f = 5$ and $C_{\rm D} = 10$ in run 2.

3 RESULTS

We found that the results of run 1 and 2 are qualitatively the same except that the time scale differs by amount of factor ~ 2 owing to our choice of f and $C_{\rm D}$. Thus, in this paper, we are concerned with the results of run 2. Snapshots of the systems on the a-e plane in run 2 is shown in Fig.1, where a is the semi-major axis and e is the eccentricity. The circles represent planetesimals and the radii of the circles are proportional to those of planetesimals. Due to larger surface density and shorter orbital period, larger planetesimals are initially formed at the innermost part of the particle disk, and the runaway growth seems to propagate outwards . Almost all the small planetesimals accrete to small number of larger planetesimals in 40000 years around 0.5AU, while there still remain small amount of planetesimals at 160000 years around 1.2-1.5AU. Hereafter, we simply call a larger planetesimal a protoplanet. In the snapshots, lines with the length of $5R_{\rm H}$ are drawn from the center of the protoplanets (with mass larger than 1/8 of the largest body in the snapshots) to both sides. After 160000 years, about 20 Mars-sized protoplanets are formed. The orbital separation is about $10R_{\rm H}$, which agrees with those obtained by Kokubo & Ida (1998,1999) and Weidenschilling et al. (1997). These results are also obtained by our gas-free simulations which we do not report in this paper.

To obtain the growth rate of protoplanets, we investigate the evolution of the random velocity of planetesimals, or equivalently, that of eccentricity and inclination. As shown in Fig.1, eccentricities (and inclinations) remain small for protoplanets owing to the dynamical friction from planetesimals. On the other hand, those of planetesimals increases monotonically with time despite of damping due to gas drag. The increase of eccentricity and inclination of planetesimal is responsible for the increase of the masses of the protoplanets as shown below. Figure 2 shows the evolution of RMS eccentricity and inclination of planetesimals in the region where $1.0\mathrm{AU} < a < 1.2\mathrm{AU}$, scaled by so-called the reduced Hill's radius of the largest body h_M, which is defined as $h_M = (M_{\rm L}/3M_\odot)^{1/3}$ where $M_{\rm L}$ is the mass of the largest body in that region. Note that h_M increases with time since $M_{\rm L}$ increases. In the figure, the scaled eccentricity and inclination remain almost constant after 4000 years, despite e, i, and $M_{\rm L}$ increase with time. Such a behavior is characteristic in the *protoplanet-dominated stage* (Ida & Makino 1993): when protoplanet mass exceeds some threshold $M_{\rm crit}$, e and i of planetesimals are determined by the equilibrium between damping of gas drag and gravitational stirring by a protoplanet, rather

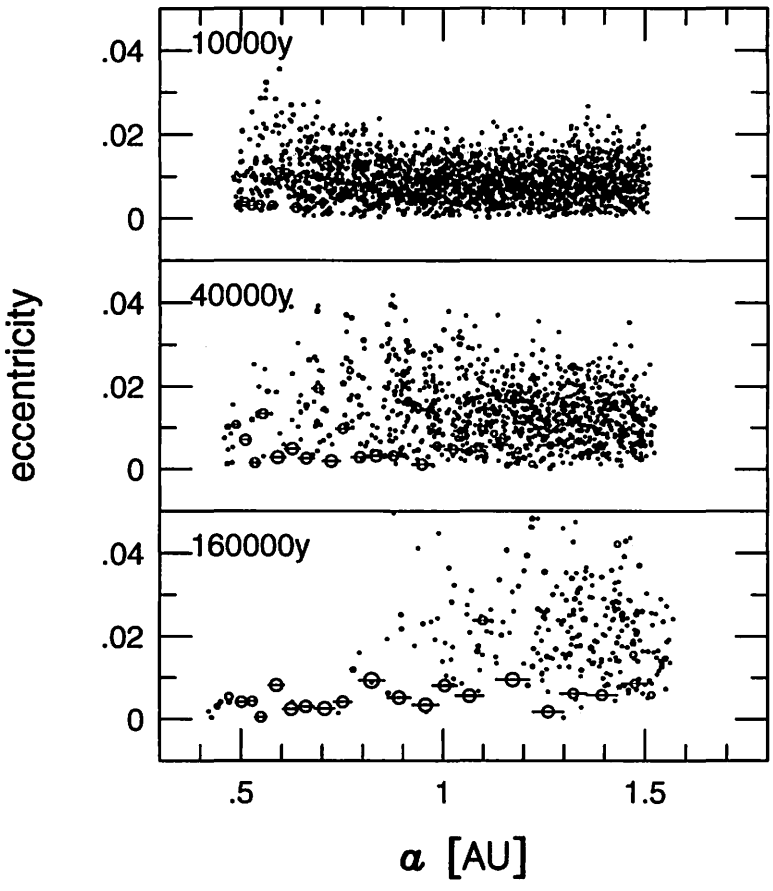


Figure 1: Time evolution of the planetesimal system on the a-e plane in run 2, where a is the semi-major axis and e is the eccentricity. The circles represent planetesimals and their radii are proportional to the radii of planetesimals. The system initially consists of 4000 equal mass planetesimals ($m_0 = 3.6 \times 10^{24}$g). In run 2, we used the five times larger radii than the realistic ones. The number of planetesimals are 2192 (10000 years), 1036 (40000 years), and 281 (160000 years). We draw lines from the center of the protoplanets to both sides, which have the length of $5R_H$. The protoplanets are selected if their masses are larger than 1/8 of the maximum mass in the system.

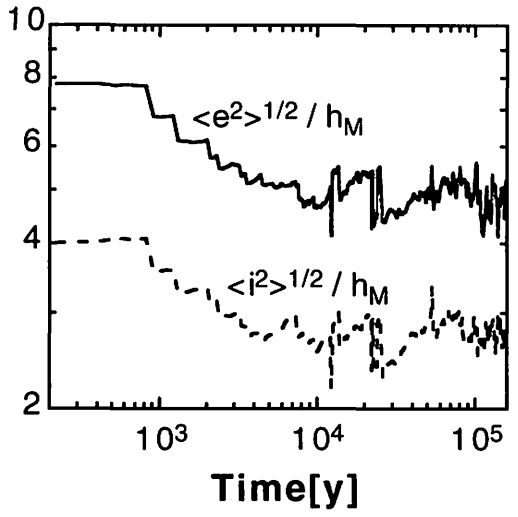


Figure 2: The evolution of the RMS eccentricity and inclination of planetesimals in the region where 1.0AU $< a <$ 1.2AU scaled by h_M. The scaling factor h_M is the reduced Hill's radius of the largest protoplanet with mass $M_{\rm L}$, defined by $h_M = (M_{\rm L}/3M_\odot)^{1/3}$.

than stirring among planetesimals. According to Ida & Makino (1993), $M_{\rm crit}$ is given by

$$\frac{M_{\rm crit}}{m_0} = 6.2\left(\frac{a}{1{\rm AU}}\right)^{\frac{3}{10}}\left(\frac{m_0}{3.6\times10^{24}{\rm g}}\right)^{-\frac{2}{5}}. \tag{4}$$

In our simulation, $M_{\rm crit}/m_0 = 8.0$ at 4000 year in the region, which well agrees with Eq. (4). Ida & Makino (1993) investigated a system with a single protoplanet and equal-mass planetesimals in its heated region defined as twice the feeding zone within which a planetesimal on an unperturbed orbit can collide with the protoplanet. In our calculation, however, other protoplanets exist with orbital separation of about $10R_{\rm H}$ thus we give width of the heated region to that extent, which is narrower than that given by Ida & Makino (1993). Then the equilibrium eccentricity of planetesimals is given by (for details, see Ida & Makino 1993).

$$e_{\rm eq} = 4.4\left(\frac{a}{1AU}\right)^{\frac{7}{10}}\left(\frac{m}{3.6\times10^{24}{\rm g}}\right)^{\frac{1}{15}}\left(\frac{C_{\rm D}}{10}\right)^{-\frac{1}{5}} h_M \tag{5}$$

Figure 3 shows the comparison of the scaled eccentricity obtained by the simulation and Eq.(5). Equation (5) almost agrees with the result of the simulation. By using this result, we can roughly estimate the mass of the protoplanets. The growth rate of the protoplanet is written as (e.g. Ida & Nakazawa 1989)

$$\frac{{\rm d}M}{{\rm d}t} \sim \int_{m_0}^{m_{\rm max}} {\rm d}m \frac{m\, n_{\rm s}(m)}{H}\pi r_{\rm p}^2\left(1+\frac{v_{\rm esc}^2}{v_{\rm rel}^2}\right) v_{\rm rel}, \tag{6}$$

where m_0 and $m_{\rm max}$ are the minimum and maximum mass of the system, H is the scale height of the planetesimal disk, $n_{\rm s}(m)$ is the surface number density of planetesimal with mass m, $r_{\rm p}$ is the radius of the protoplanet, $v_{\rm esc}$ is the escape velocity from the protoplanet, $v_{\rm rel}$ is relative velocity between the protoplanet and the planetesimals, and Ω is the Keplerian frequency. We assume $n_{\rm s}(m)$ is given by

$$n_{\rm s}(m){\rm d}m = 0.5\frac{\Sigma_{\rm s}}{m_0^2}\left(\frac{m}{m_0}\right)^{-5/2}{\rm d}m \quad (m_0 \le m \le m_{\rm max}) \tag{7}$$

(Kokubo & Ida 1996, 1999). Note that equation (7) would be appropriate in relatively early stage of accretion and invalid if the planetesimals are depleted owing to accretion to the protoplanets.

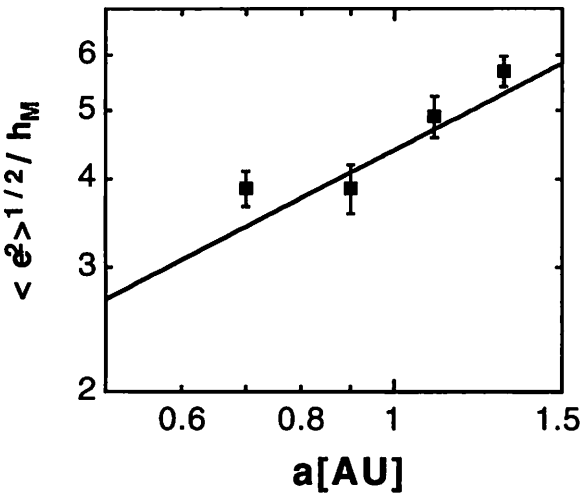


Figure 3: Filled squares with error bars denote the time averaged RMS eccentricity scaled by h_M in the region where 0.6-0.8AU, 0.8-1.0AU, 1.0-1.2AU, and 1.2-1.4AU, respectively. Error bars denote the 1σ deviation of time evolution. Time range for time averaging is chosen so that $M_{\rm L} > 50m_0$ and $N_{\rm rem} > 50$, where $N_{\rm rem}$ is remaining bodies in each region. Solid line expresses Eq.(5).

Now letting $H \simeq 2ia \sim ea$, $v_{\rm esc} \gg v_{\rm rel}$, and $v_{\rm rel} \sim ea\Omega$, Eq.(6) can be easily integrated and we obtain

$$\begin{cases} M = (M_{\rm crit}^{1/3} + Ft)^3 \quad ({\rm for} M < M_{\rm max}), \\ F = 3.6 \times 10^{-4} \left(\dfrac{\rho_p}{2{\rm g\,cm^{-3}}}\right)^{-\frac{1}{3}} \left(\dfrac{m}{3.6 \times 10^{24}{\rm g}}\right)^{-\frac{2}{15}} \left(\dfrac{C_{\rm D}}{10}\right)^{\frac{2}{5}} \left(\dfrac{f}{5}\right) \left(\dfrac{a}{1{\rm AU}}\right)^{-\frac{27}{10}}, \end{cases} \tag{8}$$

where we used $e_{\rm eq}$ given by Eq.(5) for $m = 3.6 \times 10^{24}$g, since mass dependence of $e_{\rm eq}$ is weak. In Eqs.(8), we take $M_{\rm crit}$ as the onset mass of the protoplanet-dominated stage given by Eq.(4), since that stage begins at early epoch in our simulation (i.e. in several thousand years). When protoplanet mass reaches the isolation mass $M_{\rm max}$, where

$$M_{\rm max} = 0.16 \left(\frac{b}{10R_{\rm H}}\right)^{\frac{3}{2}} \left(\frac{\Sigma_{\rm p0}}{10{\rm g\,cm^{-2}}}\right) \left(\frac{a}{1{\rm AU}}\right)^{\frac{3}{4}} M_\oplus, \tag{9}$$

we consider the protoplanet does not grow further since there is no planetesimals around it. The growth time scale of protoplanet $T_{\rm grow} \equiv M({\rm d}M/{\rm d}t)^{-1}$ is determined as that in the protoplanet-dominated stage since time scale to reach $M_{\rm crit}$ for largest planetesimal is shorter (Ida & Makino 1993, Inaba et al. 1999). Thus

$$\begin{aligned} T_{\rm grow} \sim \; & 9.3 \times 10^4 \left(\frac{M}{0.1M_\oplus}\right)^{\frac{1}{3}} \left(\frac{\rho_{\rm p}}{2{\rm g\,cm^{-3}}}\right)^{\frac{1}{3}} \left(\frac{\Sigma_{\rm p0}}{10{\rm g\,cm^{-2}}}\right)^{-1} \left(\frac{m_0}{3.6 \times 10^{24}{\rm g}}\right)^{\frac{2}{15}} \\ & \times \left(\frac{C_{\rm D}}{10}\right)^{-\frac{2}{5}} \left(\frac{f}{5}\right)^{-1} \left(\frac{a}{1{\rm AU}}\right)^{\frac{27}{10}} \quad [{\rm y}], \end{aligned} \tag{10}$$

which almost agree with the simulation. Figure 4 shows the comparison between results of run 2 with Eq.(8) and (9). Similar result is obtained for run 1, and the result of the simulations is well scaled by Eqs.(8) and (9)

4 DISCUSSION AND CONCLUSION

We investigated formation and evolution of protoplanets from a planetesimal swarm embedded in the solar nebula. The planetesimal swarm is distributed from 0.5 to 1.5AU, so that the difference of growth time scale of protoplanets can be clearly seen.

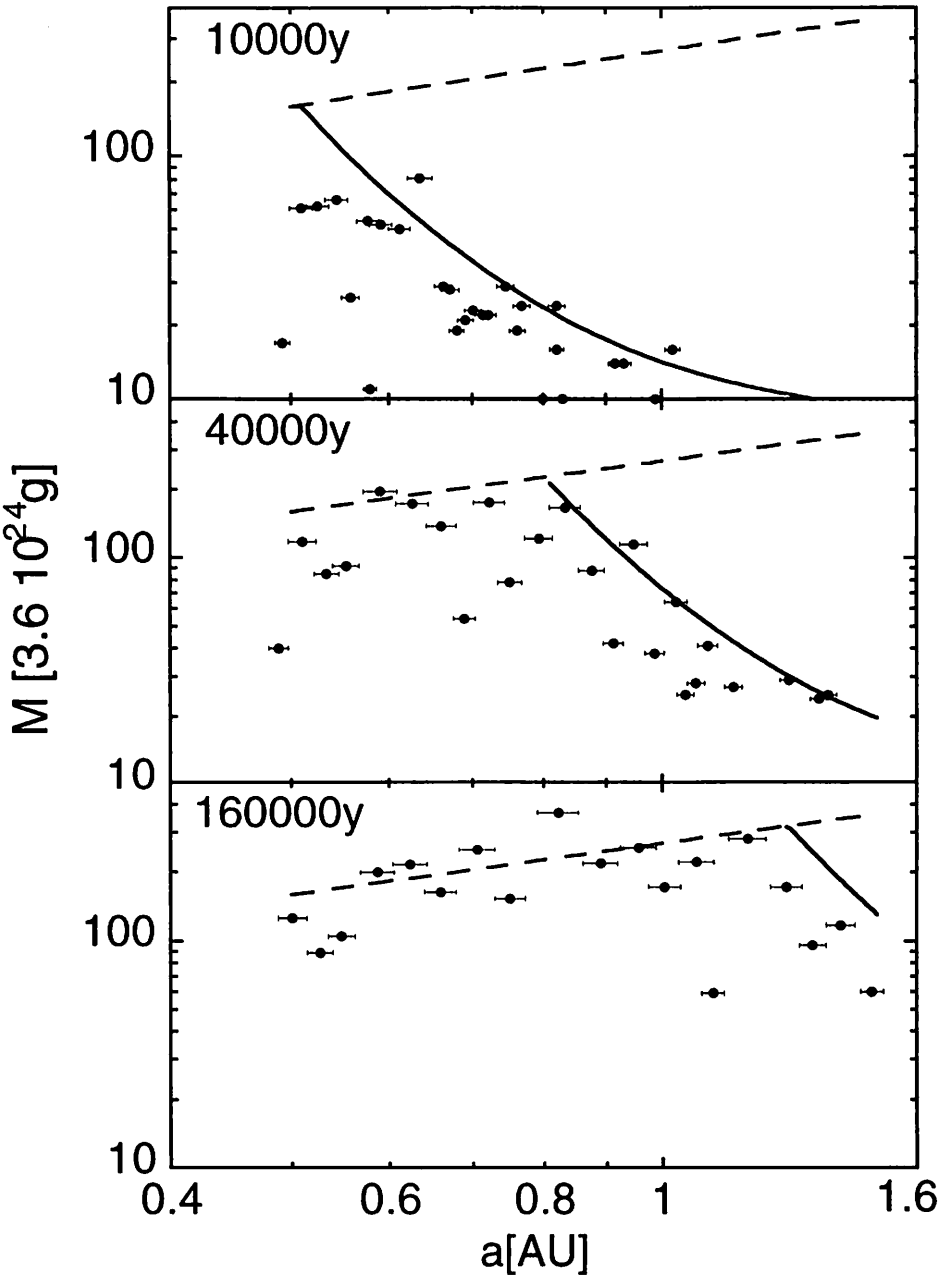


Figure 4: Comparison between the masses of the protoplanet in the N-body simulation (filled circles with bars) and those estimated by Eqs.(8) (solid lines) and (9) (dashed lines) at 10000 year, 40000 year, and 160000 year. Bars drawn from the center of circles are the same as Fig.1.

We directly found protoplanets which deplete surrounding planetesimals have masses predicted by Kokubo & Ida (1998), who found the oligarchic growth of protoplanets. Our results also agree with those obtained by Weidenschilling et al. (1997), who also investigated formation of the protoplanet system in the terrestrial planet region by using their own statistical method. However, Weidenschilling et al. (1997) did not gave quantitative argument about the evolution of the protoplanets. In contrast to Weidenschilling et al., we estimated protoplanet mass as function of semi-major axis and time by simple approximation, which well agrees with the simulation.

Equations (8) suggest that when $M_{\rm crit}$ is sufficiently smaller than M, it can be written as $M = F^3 t^3 \propto a^{-8.1} t^3$. However, the result of the simulation shows protoplanet mass distributes more gently, at most a^{-6} (Fig.4). This reflects our choice of the minimum mass in the simulation due to the computational limitation. In the protoplanetary solar nebula, it is considered that planetary accretion began with much smaller planetesimal (10^{18-20}g). Therefore the critical mass $M_{\rm crit}$ is much smaller than that of the simulation (Eq. (4)) and the distribution of the protoplanets would approach $M \propto a^{-8.1}$ as the protoplanets grow.

One may be afraid that we omit the somewhat important effect due to protoplanet-protoplanet interaction, since we reduce the accretion time scale while time scale of gravitational interaction is unchanged. Actually, protoplanets may migrate outwards substantially owing to orbital repulsion among them since outer protoplanet is less massive than the inner unless it reaches the isolation mass. We investigated this effect with simple formula by using the result of Hénon & Petit (1986) who investigated distant encounter between two bodies orbiting around the Sun, and found such a migration would not occur sensitively in the system considered in this paper, since masses of the protoplanets are relatively small. In a more massive disk or the region of jovian planets, however, this effect may play an important role on the evolution of the protoplanets. We will investigate this effect in detail in the forthcoming paper.

References

Adachi I., Hayashi C., Nakazawa K., 1976, Prog. Theor. Phys. 56, 1756
Chambers J. E., Wetherill G. W., Boss A., P., 1996, Icarus, 119, 261
Chambers J. E., Wetherill G. W., 1998, Icarus, 136, 304
Greenberg R., Wacker J. F., Hartmann W. K., Chapman C. R., 1978, Icarus, 35, 1
Hayashi C., 1981, Prog. Theor. Phys. Suppl., 70, 35
Hénon M., Petit J.-M., 1986, Celes. Mech. 38, 67
Ida S., Makino J., 1992a, Icarus, 96, 107
Ida S., Makino J., 1993, Icarus, 106, 210
Ida S., Nakazawa K., 1989, A&A, 224, 303
Inaba S., 1999, Ph. D. thesis, Tokyo Institute of Technology
Ito T., Tanikawa ., 1999, Accepted to Icarus
Kokubo E., Ida S., 1996, Icarus, 123, 180
Kokubo E., Ida S., 1998, Icarus, 131, 171
Kokubo E., Yoshinaga K., Makino J., 1998, MNRAS, 297, 1067
Kokubo E., Ida S., 1999, submitted to Icarus
Makino J., 1991, PASJ, 43, 859
Makino J., Taiji M., Ebisuzaki T., Sugimoto D., 1997, ApJ, 480, 432
Whetherill G. R., Stewart G. R., 1989, Icarus, 77, 330
Whetherill G. R., Stewart G. R., 1993, Icarus, 106, 190
Weidenschilling S. J., Spaute D., Davis D. R., Marzari F., Ohtsuki K., 1997, Icarus, 128, 429

Angular Momentum Transports in Dense, Self-Gravitating Particle Systems

Hiroshi Daisaka

Department of Earth and Planetary Sciences, Faculty of Science, Tokyo Institute of Technology
Tokyo 152-8551, Japan
Telephone: +81-3-5734-2243, Fax: +81-3-5734-3538
E-mail: hdaisaka@geo.titech.ac.jp

ABSTRACT

Angular momentum transports in dense, self-gravitating particle systems was investigated though local N-body simulations. In a Keplerian disk such as a planetary ring system, there exists originally velocity gradient and the interactions between particles tend to reduce the difference between their velocities. As a consequence, angular momentum transport occurs and the ring system evolves.

The transport arises through local, collisional, and gravitational transports. In a dense ring, transport due to collisional effect dominates because the mean free path is smaller than the particle physical radius. Recent local N-body simulations with large number of self-gravitating particles showed the formation of non-axisymmetric, wake-like structure (e.g., Salo 1995, Daisaka and Ida 1999) so gravitational transport would also be significant in the transport mechanisms. Furthermore, their simulations indicated that particles in the wake-like structure have coherent motion. Such particle motion would affect local and collisional transports.

Our numerical results show that the collisional effect in angular momentum transports is relatively reduced by the self-gravity and gravitational effect is the same order as local effect and is significant. The evidence of viscous instability is not found in $\tau \leq 0.6$ even if self-gravity is included .

Key words: angular momentum transport - local N-body simulation - planetary ring.

1 INTRODUCTION

By the observation of the Voyager, we have obtained detailed, much more informations about planetary rings, especially, microstructure in Saturn's B-ring, which could never see from grand-based observations (e.g., Esposito 1993). However, the origin of the rings itself and the structure have still remained as unsolved problems.

A ring system is composed of a lot of particles which rotate around a central planet. The system would evolve by exchanging momentum of particles through interactions such as mutual gravitational forces and direct collisions. This yields angular momentum transport and mass flow. Therefore it is essential for considering ring problems to understand the process of angular momentum transport.

It is considered that Saturn's B-ring is a dense particle system in which ring particles are closely packed (the most part of volume is occupied by ring particles) and mean free path may become the same order of particle physical size. Angular momentum transport in such a system was studied in relation to stability of ring. Theoretical studies by solving Boltzmann equation were done by several authors. Goldreich and Tremaine (1978) studied the transport of angular momentum due to movement of particles itself crossing a surface which is occurred in a dilute gas, but they neglected the effect that the particles have the finite size. Araki and Tremaine (1985) included the effect of particle size in the transports and they showed that this effect is very important in large optical depth (large surface density). Numerical works by performing local N-body simulation were done by Wisdom and Tremaine (1988), Salo (1991), and Richardson (1994). Their results showed good agreement with theoretical studies, especially, Araki and Tremaine (1985). However, in these studies the effect of self-gravity of particles was not taken into account.

Simulations with large number of self-gravitating particles showed that the formation of non-axisymmetric, wake-like structure naturally occurs due to gravitational instability (Salo 1995, Daisaka and Ida 1999). They also reported that the particles in the wake-like structure tend to move with a similar direction at a similar velocity, i.e., the particles move coherently and have a large systematic motion. In the numerical study by Richardson (1994), self-gravity was included but the effect of inhomogeneous spatial structure on the angular momentum transfer was not considered because he used only small number of particles ($N < 100$) for evaluation of angular momentum transport. However, it is expected that gravitational contribution in transport process due to non-axisymmetric structure becomes dominant effect.

On the other hand, angular momentum transport due to gravitational contribution in differentially rotating disk was studied in terms of generating mechanism of spiral structure in the galaxies (Lynden-Bell and Kalnajs 1972).

Recently, Takeda (1999) studied angular momentum transport by global N-body simulations in order to understand accretion time of proto-moon from a debris disk around proto-earth which is considered to be made in the framework of giant impact model. Global simulations show that diffusion of particles occurs and surface density is reduced quickly due to angular momentum transport. Indeed, system would evolve through such mass flow and that would reflect real situation. However, it is considered that density change would affects angular momentum transfer.

In this study, in order to understand basic transport mechanism in a particle system, we perform local N-body simulations. There we adopt periodic boundary conditions which can keep number of particles constant. Both mutual gravity between ring particles and direct collision are included. We consider the effect of non-axisymmetric structure on gravitational transport and the influence of systematic motion arising in the wake-like structure on local and collisional transports. As a preliminary result, we present the comparison between self-gravity and no self-gravity cases for local and collisional transports and the dependence of each transport on surface number density in this paper.

2 ANGULAR MOMENTUM TRANSPORTS

Interactions between ring particles in a differentially rotating disk give rise to angular momentum transport which make mass flow and make the system evolve. In a ring system as well as other Keplerian disk, circular velocity around a central planet decreases with distance from the planet according to Kepler's third law, i.e., a velocity gradient exists in the systematic flow velocity. Generally, interactions between particles affects to reduce the velocity difference. A radially inner particle revolve faster than an outer particle in Keplerian disk so that the inner gives its momentum to the outer particle. As a result, the outer particle obtain angular momentum and its orbit would move radially outward while the inner tends to fall to the central planet.

It is critical to understand what mechanism is dominant in angular momentum transport and know its rate for the arguments of the life time of a ring and the origin o f axisymmetric, record-like structures with a scale $\sim$ 10km in Saturn's ring discovered by Voyager and the stability of a ring. If a certain condition is satisfied, viscous instability could occur and a ring is divided into a lot of axisymmetric subrings (Lin and Bordenheimer 1981, Ward 1981, Lukkari 1981).

Angular momentum transports in a ring system could be classified as under three contributions:

(i) local (translational, streaming) transport.
This transport is caused by the movement of particles itself from one place to another across a shear surface. Particles moving to new place deliver momentum through collisions. Random motion of particles allows such transfer (Goldreich and Tremaine 1978, Greenberg 1988)

(ii) collisional transport.
Particles have a finite size so collision does not occur at a point. It is possible to transfer momentum by collisions which straddle the surface under consideration, i.e., without movement of particles itself. This effect could not be ignored in the case where mean free path is the same as particle physical radius (Araki and Tremaine 1985).

(iii) gravitational transport.
Ring particles have self-gravity so gravitational acceleration and deaccelaration of a particle by gravity from other particles are possible. This means that the exchange of momentum between particles arises through gravitational force. In the case where particles distribute homogeneously in a space, this transport should be insignificant since gravitational forces exerted on a particle by other particles would cancel out (in other words, acceleration and deacceleration due to mutual forces are canceled). But simulations with self-gravitating particles showed the formation of wake-like structure (Salo 1995, Daisaka and Ida 1999) so in a such case gravitational transport seems to be important process for the angular momentum transport. Note that collisional and gravitational transports are non-local transports.

Now we derive formulae of each transport from Boltzmann equation for local coordinate system, which is based on Takeda (1999). From a momentum equation of Boltzmann equation, we can obtain the relation of angular momentum and its flux as

$$\frac{\partial}{\partial t}[2\pi r\Sigma(r)\langle u_\theta\rangle] = -\frac{\partial}{\partial r}[F_{\rm L} + F_{\rm G} + F_{\rm C} + F_{\rm Conv}], \quad (1)$$

where

$$\begin{cases} \Sigma(r) &= \dfrac{1}{2\pi}\displaystyle\int_0^{2\pi} d\theta \int_{-\infty}^{\infty} dz\, \rho, \\ \langle u_\theta \rangle &= \dfrac{1}{2\pi\Sigma(r)}\displaystyle\int_0^{2\pi} d\theta \int_{-\infty}^{\infty} dz\, \rho u_\theta, \\ \rho &= \displaystyle\int m f(\boldsymbol{x}, \boldsymbol{v}, t)\, d\boldsymbol{v}, \\ \rho u_\alpha &= \displaystyle\int m f(\boldsymbol{x}, \boldsymbol{v}, t) v_\alpha\, d\boldsymbol{v}, \end{cases} \tag{2}$$

and $f(\boldsymbol{x}, \boldsymbol{v}, t)$ is a distribution function and m is mass of a particle. In above equation, cylindrical coordinate (r, θ, z) is chosen. $F_{\rm L}(r)$, $F_{\rm C}(r)$, $F_{\rm G}(r)$ and $F_{\rm conv}$ denotes local, collisional ,gravitational and convective flux (luminosity), respectively. Each flux in the above equations is represented as

$$\begin{cases} F_{\rm L}(r) &= r^2 \displaystyle\int_0^{2\pi} d\theta \int_{-\infty}^{\infty} dz\, P_{r\theta} \\ F_{\rm G}(r) &= \displaystyle\int dr' \int_0^{2\pi} d\theta \int_{-\infty}^{\infty} dz\, [r' \rho \frac{\partial \Phi}{\partial \theta}] \\ F_{\rm C}(r) &= -\displaystyle\int dr' \int_0^{2\pi} d\theta \int_{-\infty}^{\infty} dz\, [\frac{\partial}{\partial t}(r'^2 \rho u_\theta)]_{\rm coll} \\ F_{\rm conv}(r) &= r^2 \displaystyle\int_0^{2\pi} d\theta \int_{-\infty}^{\infty} dz\, u_r u_\theta \rho, \end{cases} \tag{3}$$

where Φ is the gravitational potential of a central planet and all particles and $P_{r\theta}$ is a pressure tensor defined as

$$P_{r\theta} = \int (v_r - u_r)(v_\theta - u_\theta) m f d\boldsymbol{v}. \tag{4}$$

$F_{\rm conv}$ is related with mass flux rate $\dot{M}_{\rm disk}$ and could be evaluated by $F_{\rm L}$, $F_{\rm C}$, and $F_{\rm G}$ (Takeda 1999). For later convenience, we define the variable as

$$H(r) = \frac{F(r)}{2\pi r^2} \tag{5}$$

for each flux and use $H(r)$ instead of $F(r)$ defined in Eq. (3).

In order to evaluate the angular momentum flux by using the result of N-body simulations, we substitute the distribution function which was represented by Dirac's δ-function as $f = \Sigma_i \delta(\boldsymbol{r} - \boldsymbol{r}_i)\delta(\boldsymbol{v} - \boldsymbol{v}_i)$ into Eq. (3). Moreover, by averaging in terms of r in the range $[r, r + \Delta r]$, and doing local approximation as

$$\begin{cases} x_i &= r_i - r_0, \\ y_i &= r_0(\theta_i - \Omega_0 t), \end{cases} \tag{6}$$

where we represent the coordinate of a reference point with "0", and we consider Cartesian coordinate at the reference point (see Fig. 1), we can get the formulae

$$\begin{cases} H_{\rm L} &= \dfrac{1}{L_x L_y} \Sigma_i m_i \dot{x}_i \dot{y}_{r,i}, \\ H_{\rm G} &= \dfrac{1}{L_x L_y} \Sigma_i \Sigma_{j, x_j - x_i > 0} G m_i m_j \\ &\qquad \times \dfrac{(x_j - x_i)(y_j - y_i)}{r_{ij}^3}, \\ H_{\rm C} &= \dfrac{1}{L_x L_y \Delta T} \Sigma_{coll} m_> \Delta \dot{y}_> (x_> - x_<), \end{cases} \tag{7}$$

where $\dot{y}_r = \dot{y} + 3\Omega_0 x_i/2$ is used and r_{ij} is relative distance, $r_{ij} = ((x_j - x_i)^2 + (y_j - y_i)^2 + (z_j - z_i)^2)^{1/2}$. "$>$" indicates a particle located radially outward at a collision. The "coll" means that collisions are counted until interval ΔT. $\Delta \dot{y}_>$ means the change of velocity by a collision. The expressions for local and collisional transports are the same as Wisdom and Tremaine (1988).

In a steady state, the angular momentum flux $H(r)$ (or $F(r)$) is related with the viscosity. Combining the equation transformed from Euler's equation of viscous fluid and Eq. (1) yields the relation between viscosity and the angular momentum flux as

$$H(r) = \frac{3}{2}\Omega(r)\Sigma(r)\langle \nu \rangle \tag{8}$$

where ν is kinematic viscosity. In this study, we will discuss about ν in stead of H (F).

3 NUMERICAL METHOD

In this section, we briefly explain local N-body method used in the present study as well as used in Daisaka and Ida (1999). This local method was first applied to dense ring systems by Wisdom and Tremaine (1988) and extended to include the effect of size distribution of ring particles and mutual gravity between particles (Salo 1991, 1992a, 1992b, 1995, Richardson 1994, Daisaka and Ida 1999). "Local" means that we consider a small square region (box) with width L_x and height L_y at a reference point with semimajor axis a_0 in the ring, which revolves along a circular orbit around a central planet with angular velocity $\Omega_0 = \sqrt{GM_s/a_0^3}$ of the reference point, and pursue the motion of particles only in the small region with periodic boundary conditions taking into account shearing motion of boxes (see Fig. 1). As seen in Fig. 1, eight copied boxes are considered to surround the concentrated box and radially inner and outer boxes are sliding with velocity $+3\Omega_0 L_x/2$ and $-3\Omega_0 L_x/2$ which arise from the difference of shear velocity of each reference points. When a particle goes out from the box at a boundary (for example, a particle with filled square in the central box), the corresponding particle (ghost particle) in the copied region comes in the box at the opposite boundary and we pursue the motion of this particle instead of outgoing particle. This procedure is the same with imposing periodic boundary condition. In this local method, number

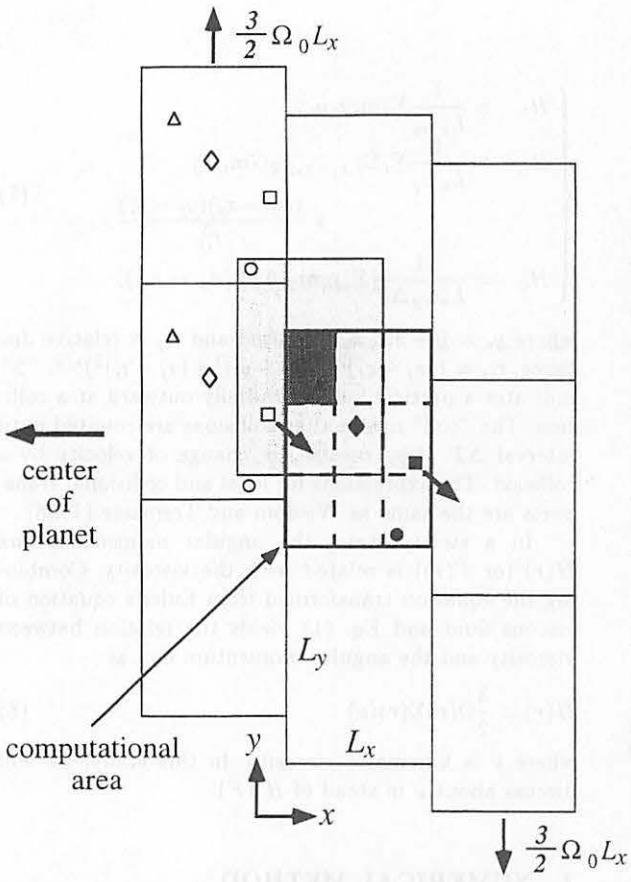


Figure 1. Schematic illustration of simulation region. Central box with solid line is computational area and motion of particles denoted by filled symboled are considered. Open symbols in boxes around the computational area denote ghost (copied) particles. The thick and thin shade regions denote a subregion and its virtual region which are used for evaluation of mutual gravitational forces in the subregion (detail is given in Daisaka and Ida 1999).

of particles is conserved as constant and it is convenient to study the dependence of equilibrium properties such as radial velocity dispersion and angular momentum flux on parameters which characterize ring systems such as optical depth.

Coordinate system we have chosen is as follow. We consider Cartesian coordinate at a reference point with semi major axis a_0 in ring, x-axis points radially outward and y and z-axes point to the direction of orbital motion and normal to the orbital plane, respectively (see Fig. 1). Since we ride on rotating system with angular velocity Ω_0 at the reference point and restrict the motion of particles in small region, motion of particles could be described by Hill's equation as (see, e.g., Hill 1878, Nakazawa and Ida 1988)

$$\begin{cases} \ddot{x}_i = 2\Omega_0 \dot{y}_i + 3\Omega_0^2 x_i + \sum_{j\neq i}^{N} \frac{Gm_j}{r_{ij}^3}(x_j - x_i), \\ \ddot{y}_i = -2\Omega_0 \dot{x}_i + \sum_{j\neq i}^{N} \frac{Gm_j}{r_{ij}^3}(y_j - y_i), \\ \ddot{z}_i = -\Omega_0^2 z_i + \sum_{j\neq i}^{N} \frac{Gm_j}{r_{ij}^3}(z_j - z_i), \end{cases} \quad (9)$$

where m_j and r_{ij} are mass of particle j and the relative distance between particle i and j, and M_s is mass of a central planet. This equation of motion is integrated by Hermite integrator, which is one of predictor-corrector integrators and is suitable to treat discontinue phenomenon such as jumps due to periodic boundary conditions and direct collisions because this integrator does not need data of past few step (Makino and Aarseth 1992).

The last terms on the right hand sides of Eqs. (9) denote the mutual gravitational force between particles and calculation cost of these terms is the most expensive in the integration because the cost is $O(N^2)$, where N is number of particles used in a simulation. Wisdom and Tremaine (1988) ignored self-gravity was ignored or only treated as approximate manner. But several authors showed that self-gravity plays an important role in the structure of rings (Salo 1992b, 1995, Richardson 1994, Daisaka and Ida 1999). In the present simulations, as well as Daisaka and Ida (1999), we take into account self-gravity of particles and perform direct sum of force from all particles by using a special purpose hardware for calculating gravitational force, HARP-2 (Makino et al. 1993). For efficient use of HARP-2, we also use subregion method for the evaluation of self-gravity (see Daisaka and Ida 1999).

In this study, we adopt a hard-sphere collision model in which velocity change of colliding particles after a collision is described by using the restitution coefficient of a ring particle ϵ. We assume that a ring particle is smooth, i.e., impact velocity only in normal direction is altered by collision. ϵ is generally a function of impact velocity as well as material property of colliding particles but precise knowledge on ϵ for a ice particle was not obtained in spite of many efforts (Bridges et al. 1984, Hatzes et al. 1988,Dilley 1993, Supulver et al. 1995). In this simulation, we will ϵ as a parameter.

Simulations are limited to the identical particle system, i.e., all particles have the same radius and mass. In this case, particle system is characterized by two parameters, dynamical optical depth τ and ratio $r_h/2r_p$ where r_h and r_p are particle physical radius and Hill's (tidal) radius defined by

$$r_h = (\frac{2m}{3M_s})^{1/3} a_0, \quad (10)$$

m is mass of a particle. The former is given by $\tau = N\pi r_p^2/L_x L_y$ and is equivalent to surface number (mass) density. The latter is a function of semi-major axis a_0 at a reference point and material density ρ and represents the effect of self-gravity. In the no self-gravity

case, $r_p/2r_p = 0$. Simulations are done with various set of these parameters.

After a few Keplerian time at the beginning of a simulation, particle system is relaxed and attains to a certain steady state in which heating due to gravitational scattering and direct collision and cooling due to inelastic collision are balanced. Such steady state does not depend on the choice of initial conditions but on above simulation parameters. Statistical values such as velocity dispersions are measured and are averaged after time when a equilibrium state is established.

4 RESULTS OF SIMULATIONS

In the beginning of this study, we investigate the dependence of viscosity which comes from each angular momentum transport on optical depth τ. We performed simulations varying optical depth τ with and without taking into self-gravity of particles for fixed parameters of restitution coefficient $\epsilon = 0.5$ and the ratio $r_h/2r_p = 0.82$ (we assume $a_0 = 10^8$m, typical semi-major axis of B-ring, and $\rho = 900\text{kg/m}^3$ for material density of ice). Optical depth is controlled by changing number of particles used in a simulation. Simulation region is chosen as square with the width of $L = 112$m (self-gravity case) and $L = 125$m (no self-gravity case), respectively. Numbers of particles used in simulations, for example, are $N = 400$ ($N = 500$) for $\tau = 0.1$ and $N = 2400$ ($N = 3000$) for $\tau = 0.6$ in the self-gravity (no self-gravity) cases, respectively. Simulations were done for $\tau \leq 0.6$ in self-gravity case, $\tau \leq 2.0$ in no self-gravity case.

First we show typical results obtained by simulations where both both self-gravity and inelastic, physical collision of particles are taken into account. Due to the gravitational instability, inhomogeneous, non-axisymmetric, and wake-like structure is formed spontaneously even if the initial condition of the simulations are set to avoid the instability. Figure 2 is spatial distributions of particles for $\tau = 0.2$ and $\tau = 0.6$ (self-gravity case) and $\tau = 0.6$ (no self-gravity case) in x-y and x-z plane and shows above result. In low τ cases, particles distribute homogeneously as seen in no self-gravity case but wake-like structure is clearly seen in large τ cases ($\tau = 0.6$ in Fig. 2). Typical scale of such structure is approximately given by the longest wavelength of axisymmetric gravitational instability in a thin disk as (Julian and Toomre 1966)

$$\lambda_{cr} = \frac{4\pi^2 G\Sigma}{\kappa^2}, \tag{11}$$

where Σ and κ are surface density and epicyclic frequency, respectively. As the formation of spatial structure, co-planar components in velocity dispersion v_x and v_y increase with large magnitude of fluctuation (which represents the amplitude of oscillation observed in the time evolution of velocity dispersions). This thing could be seen in Figure 3, which shows the dependence of velocity dispersions on τ. In no self-gravity case, velocity dispersions tend to have a value of $\sim r_p\Omega$ (it seems that velocity dispersions slightly decrease with increase of τ). On the other hand, in self-gravity case radial velocity dispersion attains surface escape velocity for small τ but for large τ it increases with increases of τ, satisfying the relation of $Q \sim 2$, Q is Toomre parameter for gravitational instability defined as (Toomre 1964)

$$Q = \frac{\kappa c_r}{3.36G\Sigma}, \tag{12}$$

where c_r denotes radial velocity dispersion. For non-axisymmetric perturbation, system would be unstable if $Q < 2$ (Griv 1998). This is the reason why it is considered that gravitational instability is responsible for the formation of spatial structure. The reason of increase in radial velocity dispersion is the development of systematic motion of particles forced by self-gravity in the wake-like structure. Such a systematic motion was not found in no self-gravity case. Thus, in the case where wake-like structure forms, the behavior of particles is altered compared with the no self-gravity case. This effect as well as the formation of spatial structure would affects to the angular momentum transfers.

First we compare local and collisional viscosities in the self-gravity case with these in the no self-gravity case. Figure 4 shows dimensionless viscosity and $\tau\nu$-relation as a function of τ in the self-gravity (with filled symbols) and no self-gravity cases (with open symbols). In no self-gravity case, both viscosities are increasing functions of τ for $\tau < 1$. ν_{local} is larger than ν_{coll} for $\tau \leq 0.6$ but vise versa for $\tau \geq 0.6$, ν_{coll}. This indicates that the contribution of collisional transport is important for large τ as pointed out by Araki and Tremane (1985) and Widsom and Tremaine (1988). On the other hand, in the self-gravity case, the situation is quite different. Both viscosities increase for $\tau \leq 0.4$ but take almost constant values for $\tau \geq 0.4$ (it seems that ν_{coll} slightly decreases). This behavior of ν comes from the fact that ν is normalized by velocity dispersion, which increases as the formation of spatial structure, i.e., with increase of τ. In $\tau\nu$-relation in Fig. 4(b), both curves of ν in the self-gravity case increase with τ rapidly than those in no self-gravity case. Furthermore, ν_{coll} becomes the same order of ν_{local} at $\tau = 0.6$ in the no self-gravity case but for $\tau \leq 0.6$, ν_{local} is always larger than ν_{coll}. This means that local contribution is still important for such large τ. This result would be related with systematic motion of particles which is dominant motion in the wake-like structure.

Next we show the significance of gravitational transport as well as local and collisional transports. Figure 5 shows dimensionless viscosity and $\tau\nu$-relation of gravitational contribution as a function of τ in addi-

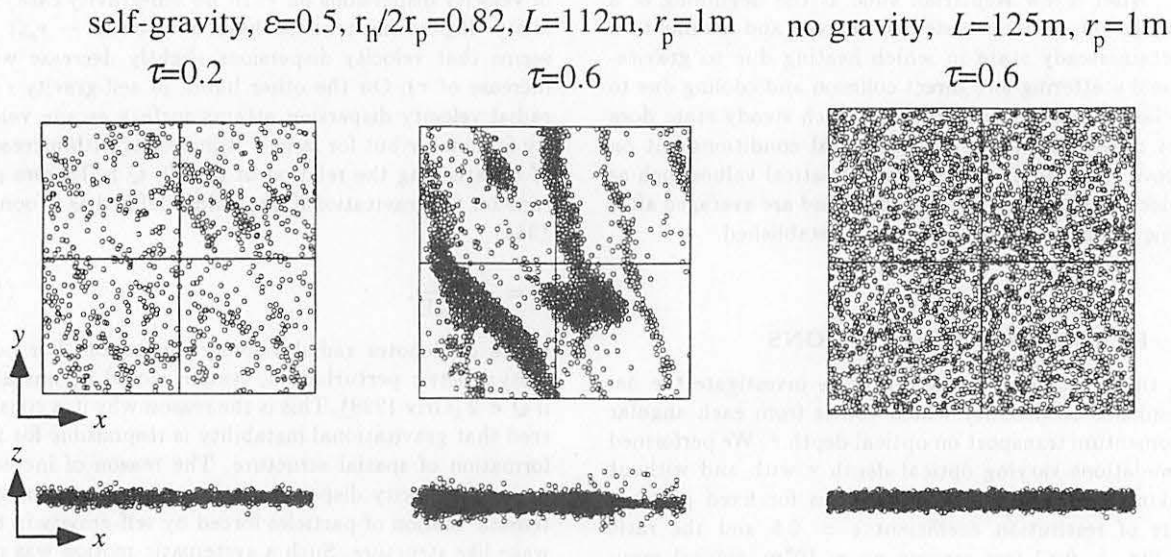


Figure 2. Particle distributions of face-on (upper panels) and edge-on views (lower panels) in self-gravity and no self-gravity cases. Homogeneous distribution is observed for $\tau = 2.0$ in no self-gravity case.

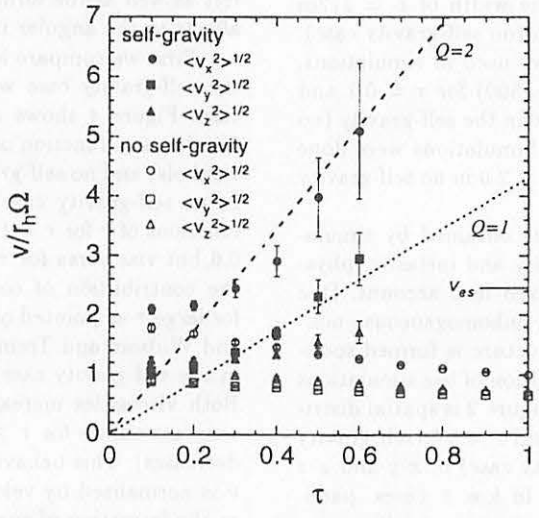


Figure 3. Velocity dispersions as a function of optical depth τ. Self-gravity and no self-gravity cases are represented by filled and open symbols, respectively. Error bar represents the magnitude of fluctuation of velocity dispersions. Broken and dotted lines are those obtained by putting $Q = 2$ and $Q = 1$, where Q is Toomre's parameter for gravitational instability. Each velocity is scaled by $r_h\Omega_0$.

tion to local and collisional contributions. For τ small enough not to induce gravitational instability and the formation of spatial structure, gravitational contribution is small. But for τ where wake structure could be clearly seen, ν_{grav} is larger than ν_{coll} and it becomes the same order of ν_{local} at $\tau \geq 0.5$. Figure 5(b) also shows that gravitational contribution is important in $\tau\nu$-relation. For $\tau \leq 0.6$, ν_{grav} normalized by velocity dispersion is an increasing function of τ since wake structure develops with increase of τ. $\tau\nu$-relations for all contributions increase with increase of τ, i.e., $\tau\nu$ of total viscosity does not satisfy the condition for viscous instability $\partial(\tau\nu)/\partial < 0$. Our simulations shows viscous instability does not occur even if gravitational contribution is taken into account for $\tau \leq 0.6$.

5 CONCLUSION AND DISCUSSION

In this study, we performed local N-body simulations and obtained each angular momentum flux of local, collisional, and gravitational contributions as viscosity.

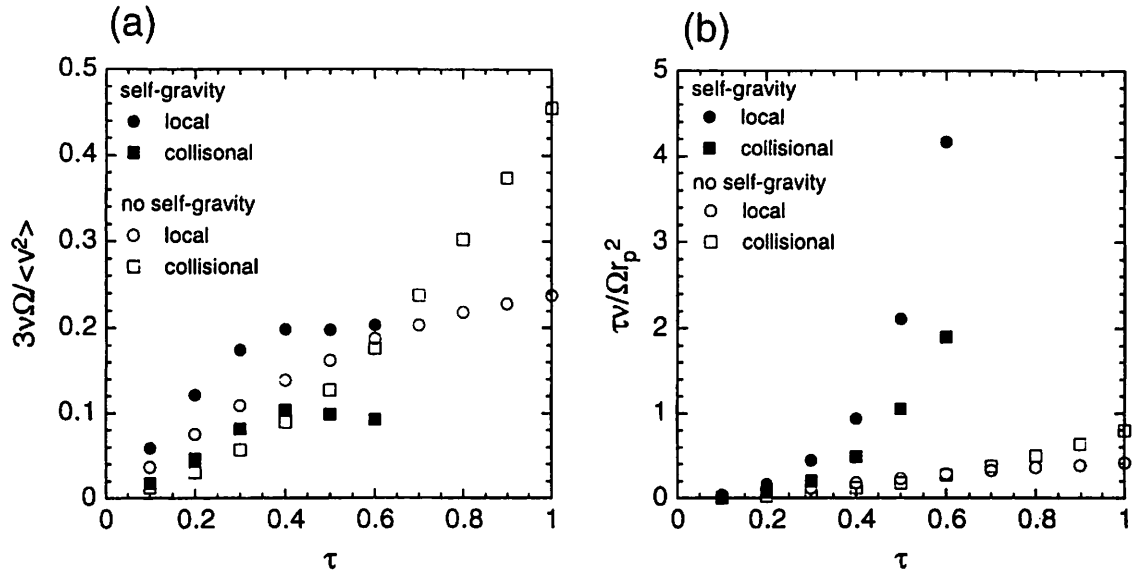


Figure 4. Comparison between self-gravity and no self-gravity cases of $\nu/\langle v^2\rangle$ (a) and $\tau\nu$ (b) as a function of τ for local and collisional contributions. In no self-gravity case, as well as Wisdom and Tremaine (1988), both local and collisional contributions are increasing functions of τ so that the viscous instability does not occur.

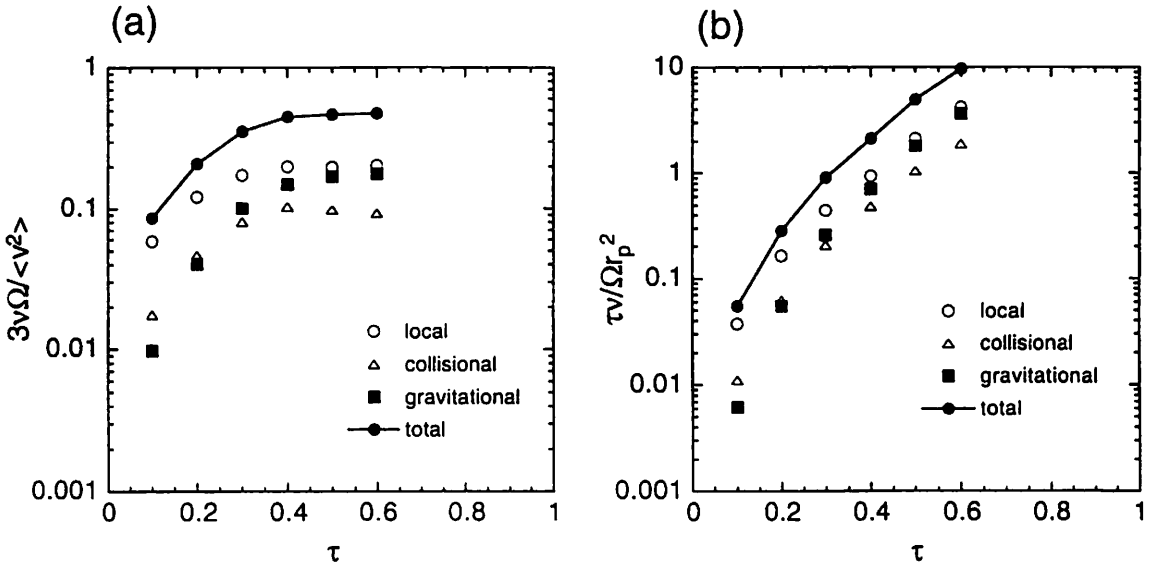


Figure 5. Dimensionless $\nu/\langle v^2\rangle$ (a) and $\tau\nu$ (b) in terms of optical depth τ for local, collisoinal, and gravitational contributions in self-gravity case. $\tau\nu$-relation for total is not decreasing function of τ so viscous instability does not occur in this case as well as no self-gravity case. Note that vertical axis is taken as logarithmic scale.

Our numerical results show that collisional effect in the angular momentum transports is relatively reduced in self-gravity case compared with that in no self-gravity case and gravitational effect is the same order as local effect and is considerably significant. The evidence of viscous instability was not found in $\tau \leq 0.6$ even if self-gravity is included . However it is insufficient for the argument of viscous instability because we study only small τ region. We must study the range of at least $\tau > 1$ for the argument of viscous instability. In this study, only optical depth is considered but the dependence on another ring's parameters should be studied. This will be done immediately as future work.

ACKNOWLEDGMENTS

The author thanks to H. Tanaka and S. Ida for useful discussions and comments. This manuscript is preparated with MN LaTeXstyle file v 1.4.

REFERENCES

ARAKI, S. AND S. TREMAINE, The Dynamics of Dense Particle Disks, *Icarus*, **65**, 83-109, 1986.

BINNEY, J. AND S. TREMAINE, *Galactic Dynamics*, 283pp., Princeton Univ. Press, Princeton. NJ, 1987.

BRIDGES, F. G., A. HATZES, AND D. N. C. LIN, Structure, stability and evolution of Saturn's rings, *Nature*, **309**, 333-335, 1984.

DAISAKA, H., AND S. IDA, Spatial Structure and Coherent Motion in Dense Planetary Rings Induced by Self-Gravitational Instability, *EPS*, subitted, 1999.

DILLEY, J. P., Energy Loss in Collisions of Icy Spheres: Loss Mechanism and Size-Mass Dependence, *Icarus*, **105**, 225-234, 1993.

ESPOSITO, L. W., Understanding planetary rings, *Annu. Rev. Earth Planet. Sci*, **21**, 487-523, 1993.

GOLDREICH, P. AND S. TREMAINE, The Velocity Dispersion in Saturn's Rings, *Icarus*, **34**, 227-239, 1978.

GOLDREICH, P. AND S. TREMAINE, The dynamics of planetary rings, *Ann. Rev. Astron. Astrophys.*, **20**, 249-283, 1982.

GREENBERG, R., Particle Properties and Large-Scale Structure of Planetary Rings: Rebound Characteristics and Viscosity, *Icarus*, **75**, 527-539, 1988.

GRIV, E., Local stability criterion for the Saturnian ring system, *Planet. Space Sci.*, **46**, 615-628, 1998.

HATZES, A. P., F. G. BRIDGES, AND D. N. C. LIN, Collision properties of ice spheres at low impact velocities, *Mon. Not. R. Astron. Soc.*, **231**, 1091-1115, 1988.

HILL, G. W., Researches in the Lunar Theory, *Amer. J. Math.*, **1**, 5-26, 129-147, 245-260, 1878.

JULIAN, W. H. AND A. TOOMRE, Non-axisymmetric responses of differentially rotating disks of starts, *Astrophys. J.*, **146**, 810-827, 1966.

LIN, D. N. C. AND P. BODENHEIMER, On the stability of Saturn's rings. *Astrophys. J. letter.* **248**, L83-L86. 1981.

LUKKARI, J., Collisional amplification of density fluctuations in Saturn's rings, *Nature*, **292**, 433-435, 1981.

LYNDEN-BELL, D. AND A. J. KALNAJS, On the generating mechanism of spiral structure, *Mon. Not. R. astr. Soc.*, **157**, 1-30, 1972.

MAKINO, J. AND S. J. AARSETH, On a Hermite integrator with Ahmad-Cohen Scheme for Gravitational Many-Body Problems, *Publ. Astron. Soc. Japan.*, **44**, 141-151, 1992.

MAKINO, J., E. KOKUBO, AND M. TAIJI, HARP: A Special-Purpose Computer for N-Body Problem, *Publ. Astron. Soc. Japan.*, **45**, 349-360, 1993.

NAKAZAWA, K. AND S. IDA, Hill's Approximation in the Three-Body Problem, *Prog. Theor. Phys. Suppl.*, **96**, 167-174, 1988.

RICHARDSON, D. C., Tree code simulations of planetary rings, *Mon. Not. R. Astron. Soc.*, **269**, 493-511, 1994.

SALO, H., Numerical Simulations of Dense Collisional Systems, *Icarus*, **90**, 254-270, 1991.

SALO, H., Numerical Simulations of Dense Collisional Systems. II. Extended Distribution of Particle Sizes, *Icarus*, **96**, 85-106, 1992a.

SALO, H., Gravitational wakes in Saturn's rings, *Nature*, **395**, 619-621, 1992b.

SALO, H., Simulations of Dense Planetary Rings. III. Self-Gravitating Identical Particles, *Icarus*, **117**, 287-312, 1995.

SUPULVER, K. D., F. G. BRIDGES, AND D. N. C. LIN, The Coefficient of Restitution of Ice Particles in Glancing Collisions: Experimental Results for Unfrosted Surfaces, *Icarus*, **113**, 188-199, 1995.

TAKEDA, T., in preperation.

TOOMRE, A., On the gravitational stability of a disk of stars, *Astrophys. J.*, **139**, 1217-1238, 1964.

WARD, W. R., On the radial structure of Saturn's ring, *Geophys. Res. Let.*, **8**, 641-643, 1981.

WISDOM, J. AND S. TREMAINE, Local simulations of planetary rings, *Astron. J.*, **95**, 925-940, 1988.

Accretional Process of the Moon from the Protolunar Disk

Takaaki Takeda
Earth and Planetary Science, Faculty of Science
Tokyo Institute of Technology, Ookayama, Meguro-ku, Tokyo 152-8551, Japan
E-mail:ttakeda@geo.titech.ac.jp

1 Introduction

The model of the origin of the Moon must satisfy the dynamical and geochemical characteristics of the present Earth/Moon system. Many models have been proposed in the past, but none of them could explain all the characteristics, such as large angular momentum of the system, small material density and iron depletion of the Moon. In past decades, the "Giant-Impact" scenario has become the most favored explanation of the lunar origin. This hypothesis proposes that the collision of a Mars-sized protoplanet with the early Earth splashed materials into the Earth's orbit to form the Moon (Hartmann and Davis 1975, Cameron and Ward 1976). This scenario is favored because of following two reasons:

1. It can potentially account for above major characteristics of the Earth/Moon system.
2. In the development in the theory of solar system formation, it is considered that large scale impacts between protoplanets may have occurred in the last stage of solar system evolution.

Several simulations showed that the plausible result of the impact is the formation of a circumterrestrial protolunar disk consisting of mantle materials of the impactor (e.g. Benz, Cameron and Merosh 1989; Cameron and Bentz 1991; Cameron 1997). After the impact with angular momentum similar to the present Earth/Moon system, disk materials (typically a few times of lunar mass) are distributed near or inside Roche limit, where the self-gravitating fluid body would be broken by tidal force. The radius of Roche limit is $a_R \sim 2.9R_\oplus$, with material density of the Moon ($\rho_M \sim 3.3\mathrm{g/cm^3}$). Since accretional process is restricted within the Roche limit (Ohtsuki 1993; Canup and Esposito 1995), these materials must spread out beyond the Roche limit to form large object(s).

Several researchers simulated the evolution of a protolunar disk by direct N-body simulations (Ida, Canup and Stewart 1997; Kokubo, Canup, and Ida 1999 ; Kokubo, Makino, and Ida 1999). Their result is that a single moon is formed from the protolunar disk on a time scale of $10^2 - 10^3$ orbital period, i.e. from a month to a year. In this process, disk materials fall to the Earth while some materials are transfered outward beyond the Roche limit in compensation. With these outward transfered materials, the moon is formed just outside Roche limit. The very rapid formation of the Moon may have many implications about geochemistry of the Moon (Abe et al. 1998), such as the depletion of volatile elements.

However, their simulations represented the disk by 1000-10000 particles due to CPU limit of N-body simulation. With a given total disk mass, the small particle number is compensated by the large size of each particle. In the case of $N \sim 10000$, the particles are of order of 100 km size. This large size of particles raises a question whether the results in N-body simulation of limited N is reliable. If diffusion process of protolunar disk or lunar accretion process have dependency on particle size, their results of rapid formation of the Moon may not be valid, since the details of characteristics of protolunar disk is still unknown. Furthermore, in a simple model neglecting self-gravitational interaction, it can be shown easily that a particle disk diffuses more rapidly if disk particles are larger and less numerous. We perform N-body simulations with several Ns (N = 1000-30000), with constant disk mass, and analyse the dependency of the evolution of the protolunar disk on N in detail. We conclude that the time scale of lunar formation does not depend on the particle size if they are sufficiently small ($N \gtrsim$ several thousands), and that the real formation of the Moon would be very rapid, from a month to a year.

2 Angular Momentum Transport

In a differentially rotating disk such as a protolunar disk, angular momentum is steadily transfered outward. Thus, inner part loses angular momentum and falls to the central body while outer part gains angular momentum and spreads outward (Lynden-Bell and Pringle 1974). We are interested in the mass

which spreads out beyond Roche Limit. The mass transport and angular momentum transport is related as follows.

Let $f(r, v, m)$ be the number density of disk particles in phase space (x, v, m). The density $f(r, v, m)$ satisfies the Boltzmann equation,

$$\frac{\partial f}{\partial t} + v_\alpha \frac{\partial f}{\partial x_\alpha} - \frac{\partial \Phi}{\partial x_\alpha}\frac{\partial f}{\partial v_\alpha} = \left(\frac{\partial f}{\partial t}\right)_c, \tag{1}$$

where $\Phi(x)$ and $(\partial f/\partial t)_c$ are gravitational potential and the rate of change of f due to collisions between particles respectively. Multiplying Eq. (1) by m and $mv_\beta(x)$ and integrating over velocity space v and mass m, we obtain the equation of continuity and equation of motion,

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial x_\alpha}(\rho u_\alpha) = \left(\frac{\partial \rho}{\partial t}\right)_c, \tag{2}$$

$$\frac{\partial}{\partial t}(\rho u_\alpha) + \frac{\partial}{\partial x_\beta}(P_{\alpha\beta} + \rho u_\alpha u_\beta) + \rho\frac{\partial \Phi}{\partial x_\alpha} = \left(\frac{\partial}{\partial t}\rho u_\alpha\right)_c, \tag{3}$$

where ρ, u, and $P_{\alpha\beta}$ are spatial density, averaged velocity vector, and pressure tensor defined by

$$\rho = \int g dv, \quad \rho u_\alpha = \int g v_\alpha dv, \tag{4}$$

$$P_{\alpha\beta} = \int g(v_\alpha - u_\alpha)(v_\beta - u_\beta)dv = \int g v_\alpha v_\beta dv - \rho u_\alpha u_\beta, \tag{5}$$

where g is defined as

$$g(x, v) = \int m f(x, v, m) dm. \tag{6}$$

In cylindrical coordinates, with equation of continuity and θ component of equation of motion, we have

$$\begin{aligned}\frac{\partial}{\partial t}(\rho r^2 u_\theta) + \frac{\partial}{\partial r}(r^2 u_r u_\theta) + r\frac{\partial u_\theta^2}{\partial \theta} + r^2\frac{\partial u_\theta u_z}{\partial z} &= -\left\{\frac{\partial}{\partial r}(r^2 P_{r\theta}) + r\frac{\partial P_{\theta\theta}}{\partial \theta} + r^2\frac{\partial P_{r\theta}}{\partial z}\right\} \\ &\quad + \left(\frac{\partial}{\partial t}\rho r^2 u_\theta\right)_c - \rho r\frac{\partial \Phi}{\partial \theta}.\end{aligned} \tag{7}$$

Averaging Eq.(7) over z and θ we have

$$\frac{\partial}{\partial t}(2\pi r\Sigma\langle r u_\theta\rangle) = -\frac{\partial}{\partial r}(2\pi r^2\overline{P_{r\theta}}) + 2\pi\overline{\left(\frac{\partial}{\partial t}\rho r^2 u_\theta\right)_c} - 2\pi r\Sigma\langle\frac{\partial \Phi}{\partial \theta}\rangle - \frac{\partial}{\partial r}(2\pi r\Sigma\langle u_r r u_\theta\rangle). \tag{8}$$

where $\Sigma(r)$, $\langle X\rangle$, and $\overline{X}$ are (averaged) surface density and averaging symbols defined by

$$\Sigma(r) = \frac{1}{2\pi}\int_0^{2\pi} d\theta \int_{-\infty}^{\infty} \rho dz, \tag{9}$$

$$\langle X\rangle = \frac{1}{2\pi\Sigma}\int_0^{2\pi} d\theta \int_{-\infty}^{\infty} \rho X dz, \quad \overline{X} = \frac{1}{2\pi}\int_0^{2\pi} d\theta \int_{-\infty}^{\infty} X dz. \tag{10}$$

The left hand side of Eq.(8) is the rate of the change of angular momentum density. Since total angular momentum is conserved, this term can be written as $-\partial F_{\rm AM}/\partial r$, where $F_{\rm AM}$ is outward radial angular momentum flux, or angular momentum luminosity. Then $F_{\rm AM}$ is obtained by integrating right hand side of Eq.(8).

$$F_{\rm AM} = F_{\rm local} + F_{\rm col} + F_{\rm grav} + F_{\rm conv}, \tag{11}$$

$$F_{\rm local} \equiv \int_0^{2\pi} r d\theta \int_{-\infty}^{\infty} dz\, r P_{r\theta}(r, \theta, z), \tag{12}$$

$$F_{\rm col} \equiv -\int_{r_{\rm min}}^{r} dr' \int_0^{2\pi} r' d\theta \int_{-\infty}^{\infty} dz \left(\frac{\partial}{\partial t}\rho r' u_\theta\right)_c, \tag{13}$$

$$F_{\rm grav} \equiv \int_{r_{\rm min}}^{r} dr' \int_0^{2\pi} r' d\theta \int_{-\infty}^{\infty} dz\, \rho \frac{\partial \Phi}{\partial \theta}, \tag{14}$$

$$F_{\rm conv} \equiv \int_0^{2\pi} r d\theta \int_{-\infty}^{\infty} dz \rho u_r r u_\theta, \tag{15}$$

where $r_{\rm min}$ is taken as inner boundary of the disk, so that $F_{\rm AM}$ is 0 outside the edge of the disk.

$F_{\rm local}$ is angular momentum transport due to pressure tensor, which comes from the difference between the velocity of each particle $\boldsymbol{v}$ and averaged velocity $\boldsymbol{u}$. $F_{\rm col}$ is angular momentum transport due to collisions between particles. $F_{\rm grav}$ is angular momentum transport due to mutual gravity. Spiral structure (see below) which develops in a protolunar disk transfers angular momentum outward, since outer part of an arm is accelerated being dragged by the gravity of inner part which rotates faster, while inner part loses angular momentum decelerated by outer part. Finally, $F_{\rm conv}$ is convective (or advective) angular momentum transport due to averaged velocity $\boldsymbol{u}$.

Note that whole angular momentum flux due to particles' movement is $F_{\rm local} + F_{\rm conv}$. However, distribution between $F_{\rm local}$ and $F_{\rm conv}$ depends on the definition of averaged velocity $\boldsymbol{u}$. Strictly speaking, defining $\boldsymbol{u}$ as Eq.(4), pressure tensor is always 0, since plural particles can not occupy the same place, so that $g(\boldsymbol{v}-\boldsymbol{u})$ in Eq.(5) is always 0. Thus, $F_{\rm local}$ is always 0 and $F_{\rm conv}$ bears whole angular momentum flux due to particle' movement. To make sense of "averaged" velocity $\boldsymbol{u}$ and pressure tensor $P_{\alpha\beta}$, we must integrate Eqs.(4) and (5) over finite small volume and divide them by the volume. To remove temporal fluctuations, above "finite small volume" must contain sufficient number of particles. Since we are interested in radial mass flux, we average them over θ- and z-direction so that $\boldsymbol{u}$ is a function of r only. As for radial direction, we integrate over sufficiently small width $[r, r+r_0]$, where r_0 is small enough compared to r so that $\boldsymbol{u}$ does not depend on r_0, but large enough to contain sufficient number of particles. With above particular definitions, radial mass flux and radial angular momentum flux can be correlated as follows.

Introducing specific angular momentum $h(r) \equiv r u_r$, $F_{\rm conv}$ is $2\pi r u_r h \Sigma$, and we can write Eq.(8) as

$$\frac{\partial}{\partial t}(2\pi r \Sigma h) = -\frac{\partial}{\partial r}(F_{\rm local} + F_{\rm col} + F_{\rm grav} + F_{\rm conv}). \tag{16}$$

With the equation of continuity, eq.(16) is reduced to

$$2\pi r \Sigma u_r \frac{dh}{dr} = -\frac{\partial}{\partial r}(F_{\rm local} + F_{\rm col} + F_{\rm grav}). \tag{17}$$

Then, outward mass flux $\dot{M}_{\rm disk} \equiv 2\pi r \Sigma u_r$ is expressed by three angular momentum fluxes $F_{\rm local}$, $F_{\rm col}$ and $F_{\rm grav}$ as follows.

$$\dot{M}_{\rm disk} = -\frac{1}{\frac{dh}{dr}} \frac{\partial}{\partial r}(F_{\rm local} + F_{\rm col} + F_{\rm grav}). \tag{18}$$

Thus, mass flux is determined by $F_{\rm local}$, $F_{\rm col}$, and $F_{\rm grav}$. Note that $F_{\rm conv} = h\dot{M}_{\rm disk}$ and it is rather the result of above three fluxes. In the analogy of viscous accretion disk model, $F_{\rm local} + F_{\rm col} + F_{\rm grav}$ can be regarded as "effective" viscous angular momentum flux due to "effective" viscosity (Lynden-Bell and Pringle 1974). Above argument is valid if specific angular momentum h is a function of r only (and not depends on θ or t), as is the case of a protolunar disk whose particles settle to nearly circular orbit quickly.

As is shown in Fig.1 below, most particles are composing aggregates or arms within Roche limit, and free particles are rather rare. This indicates that $F_{\rm grav}$ is rather expressed as the interactions between spiral arms as a whole, than the summation of interactions between free particles. In this case, $F_{\rm grav}$ must depends mainly on the density structure, and not on the size of particles. On the other hand, $F_{\rm local}$ and $F_{\rm col}$ may depend on size of particles, through the change of optical depth, collision frequency, and mean free path or radial distance between the center of particles at collision (e.g. Goldreich Tremaine 1978; Araki and Tremaine 1986). The theory of angular momentum transport in a self-gravitating particle disk is not well established. However, if self gravity is neglected, $F_{\rm local} + F_{\rm col}$ is approximately proportional to $N^{-1/3}$, if optical depth $\tau \gtrsim 1$. Small angular momentum flux means longer diffusion time of protolunar disk, and longer lunar formation time. Thus, if $F_{\rm local}$ and $F_{\rm col}$ is far greater than $F_{\rm grav}$ in simulation, the time scale of lunar formation may depends on the number of particles N, and the result of simulations with limited N may not be reliable.

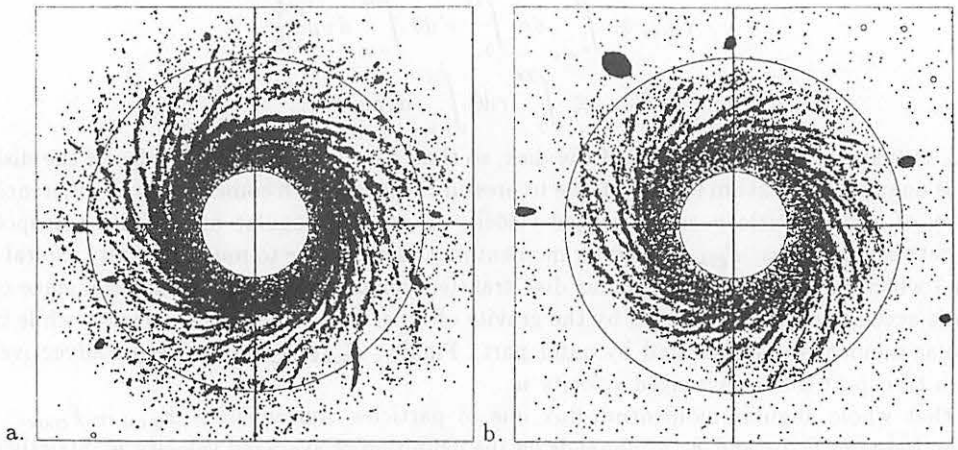


Figure 1: a. Snapshot after 6 Kepler time, in the case of $N = 30000$. Kepler time here means the Kepler time at Roche limit, which is about 7 hours. The two circles at center shows the Roche limit and the radius of the Earth. Spiral pattern dominates over the entire disk. No large aggregate grows yet. b. Snapshot after 24 Kepler time. Large aggregates grow just outside Roche limit. They continue to grow capturing particles diffused out.

3 Results

In our N-body simulation, we calculate the orbit of each particle by fourth-order Hermite scheme (which is a sort of predictor-corrector scheme). we assumed that particles are spherical and their normal coefficient of restitute is 0.1 and tangential one is 1. We started the simulations with following initial conditions:

1. The total mass of the disk is $0.03M_\oplus \sim 2.5M_{\rm L}$.
2. Surface density distribution is $\Sigma(r) \propto r^{-3}$ for $0.4R_{\rm Roche} < r < 1.2R_{\rm Roche}$.
3. Particles are distributed randomly in azimuthal direction.
4. All particles have the same size. The highest optical depth is obtained in this case.

Conditions 1. and 2. may be consistent with a plotolunar disk obtained by the impact simulations, but the particular choice is without any particular reason. Outer boundary at $1.2R_{\rm Roche}$ is chosen since our interest is the evolution of the disk within Roche limit. Initial eccentricities and inclinations are not important since they settle to quasi-equilibrium values quickly. We simulated the evolution of several disks with the same initial density distribution but different number of particles N. In any case, as random velocities of particles are damped by inelastic collisions, spiral structure develops in the entire disk (Fig.1a), due to gravitational instability (e.g. Toomre 1964). This spiral structure is rather material wave than density wave, and the arms are sheared apart and disappear in a time scale of period of a rotation, but they are constantly reformed and spiral structure remains. As angular momentum is transfered outward, the entire disk spreads out and large aggregates grow just outside Roche limit (Fig.1b). These aggregates grow capturing the particles which diffuse outside Roche limit. In the case of sufficiently large $N \gtrsim 10000$, a single large aggregate tends to grow, rather than several comparable aggregates.

We show the increase of the total mass outside Roche limit in Fig.2 We also show the total mass outside $0.9, 0.8$, and $0.7R_{\rm Roche}$ in this figure. The gradient of these lines in Fig.2 are the outward mass fluxes. In the case of larger N, the outward mass fluxes are smaller. Especially, mass fluxes through $0.7R_{\rm Roche}$ have large dependency on N. However, they are not so strongly effected by N in outer region. The mass fluxes through $1R_{\rm Roche}$ are nearly the same in $N > 3000$ cases. In the similar simulation which neglects self-gravity, the $N = 3000$ disk diffuses about two times faster than the $N = 30000$ disk.

We show the radial distribution of $F_{\rm local}$, $F_{\rm col}$ and $F_{\rm grav}$ in Fig.3. This figure shows that collisional angular momentum transport occupies rather large part of total angular momentum flux $F_{\rm local} + F_{\rm col} + F_{\rm grav}$. The fact that N-dependent component $F_{\rm col}$ occupies large part of total angular momentum is consistent with the fact that diffusion is faster in the case of smaller N (See Fig.2), since $F_{\rm col}$ is larger in the case of smaller N. Collisional momentum transport mainly dominates over inner region where

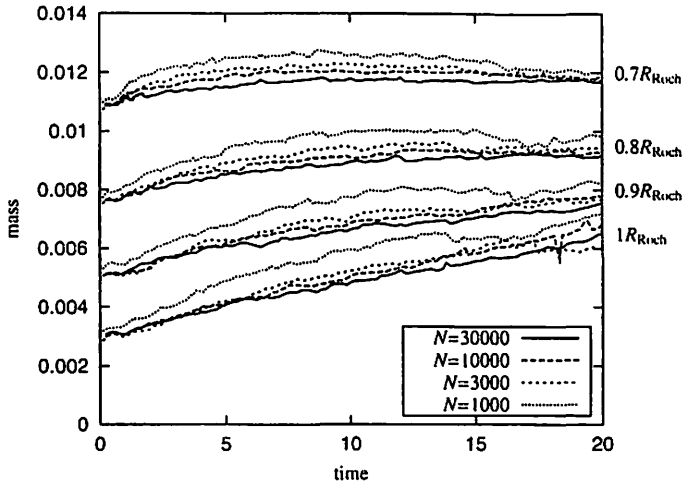


Figure 2: Mass outside 0.7,0.8,0.9 and 1 $R_{\rm Roche}$. The unit of mass is $M_\oplus$ and the unit of time is Kepler time at $1R_{\rm Roche}$. The lines which starts at 0.03 $M_\oplus$ show mass outside $1R_{\rm Roche}$. The gradient of each line shows the rate with which materials diffuse outward beyond the Roche radius (or 0.7, 0.8 and $0.9R_{\rm Roche}$). Inward mass flux begins at 0.7 or 0.8 $R_{\rm Roche}$ after a while in small N cases, since inner region thinned out rapidly.

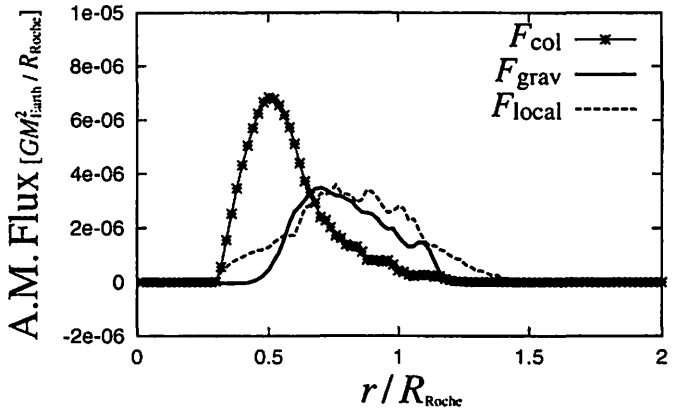


Figure 3: $F_{\rm local}$,$F_{\rm col}$ and $F_{\rm grav}$ in the case of $N = 30000$. They are averaged from 4 Kepler time to 6 Kepler time. The positive inclination means that there is inward mass flux and vice versa (See Eq.(18)). They are 0 inside the Earth radius since there is no particles there. The Earth radius is about $0.34R_{\rm Roche}$.

$r \lesssim 0.6R_{\rm Roche}$. In this region spiral structure is less apparent (Fig.1), since strong tidal force prevents the development of spiral structure near the Earth. The disk is nearly azimuthally uniform there, so that $F_{\rm grav}$ which comes from azimuthal asymmetry is 0. Thus, in this inner region, the evolution of radial mass distribution depends on particle number N strongly, which is consistent with Fig.2.

However, since mass flux through Roche limit is essential to lunar formation, outer region is more important. In outer region where the clearer spiral structure appears, local and gravitational angular momentum dominate over $F_{\rm col}$. This is true for all the cases in our simulation ($1000 < N < 30000$). In this range $F_{\rm grav}$ remains in a similar value and it occupies about half to one third of the total angular momentum transport. This indicates that the diffusion rate of this region, where spiral structure develops clearly, is not strongly dependent on N. Therefore, even if the particle size is much smaller in an actual protolunar disk, the time scale of diffusion of protolunar disk can not be longer than two or three times of the that in N-body simulation in limited N ($\sim$ 1 month).

4 Conclusion and Discussion

We find that the time scale of lunar formation is the time scale that angular momentum is transfered by self-gravity within a factor of 2-3, in the case of $N \gtrsim 3000$. We can roughly estimate $F_{\rm grav}$ as following.

The angular momentum flux due to self-gravitating spiral density structure is evaluated as follows (Lynden-Bell and Kalnajs 1972).

$$F_{\rm grav} = \frac{\pi^2 G r}{k^2} m H^2, \tag{19}$$

where k and H are radial wave number of spiral pattern and the density amplitude of the pattern. Liner analysis gives the wave number unstable to axisymmetric perturbation (e.g. Toomre 1984; Binney and Tremaine 1987). The critical wave number is,

$$k_{\rm c} \equiv \frac{\kappa^2}{\pi G \Sigma}, \tag{20}$$

where κ is epicyclic frequency. In Keplerian disk, it is equal to angular velocity of rotation Ω. This wave number is consistent with N-body simulations. Since the amplitude of density pattern $H \sim \Sigma$, as seen in Fig.1, the gravitational angular momentum flux is (e.g. Kokubo, Makino and Ida 1999; Takeda et al. 1999),

$$F_{\rm grav} \sim \frac{\pi^3 G}{M_\oplus} r^5 \Sigma^3. \tag{21}$$

From Eq.(21), we roughly evaluate the time scale that particles of mass $1 M_{\rm L}$ diffuses out beyond Roche limit, which is also the time scale of lunar formation. We have typically (Takeda et al. 1999),

$$\tau_{\rm L} \sim 2 \left(\frac{M_{\rm disk}}{2 M_{\rm L}} \right)^{-3} \quad \text{month}. \tag{22}$$

This estimate is consistent with the results of N-body simulations.

Even if particle size is infinitely small, total angular momentum flux can not be smaller than $F_{\rm grav}$. Therefore limited number of N affect the time scale of lunar formation by factor 2 at most if $N \gtrsim 3000$. We conclude that the result of N-body simulation is reliable at least with respect to the time scale of lunar formation, and that the Moon is formed very rapidly in a time scale of a few months. However, considering a less massive protolunar disk, the radial wave length of spiral structure becomes smaller (Eq.(20)). With a given N, the size of particles is proportional to $M_{\rm disk}^{1/3}$, while the wave length is proportional to $M_{\rm disk}$ through Σ. Therefore we may need larger N to express the spiral structure and to gain angular momentum flux due to the spiral structure properly in less massive case.

references

Abe, Y. , K.J. Zahnle, and A. Hashimoto 1998. Elemental Fractionation during Rapid Accretion of the Moon Triggered by a Giant Impact. In *Origin of the Earth and Moon,* pp.1. LPI contribution No.957, Lunar and Planetary Institute, Houston.

Araki, S. , and S. Tremain 1986. The Dynamics of Dense Particle Disks. *Icarus* **65**, 83-109

Benz, W. , A.G.W. Cameron, and H.J. Melosh 1989. The Origin of the Moon and the Single-Impact Hypothesis III. *Icarus* **81**, 113-131.

Binney, J. , and S. Tremaine 1987. *Galactic Dynamics,* Princeton Univ. Press.

Cameron, A.G.W. , and W.R. Ward 1976. *The Origin of the Moon. Proc. Lunar Planet. Sci. Conf. 7th,* 120-122

Cameron, A.G.W. , and W. Bentz 1991. The Origin of the Moon and the Single Impact Hypothesis IV. *Icarus* **92**, 204-216.

Cameron, A.G. 1997. The Origin of the Moon and the Single Impact Hypothesis V. *Icarus* **126**, 126-137.

Goldreich,P. and S.Tremaine 1978. The Velocity Dispersion in Saturn's Rings. Icarus 34,227-239.

Hartmann,W.K., and D.R.Davis 1975. Satellite-Sized Planetesimals and Lunar Origin. *Icarus* **24**, 504-515.

Ida S., Canup R.M. and Stewart G.R. 1997. Lunar accretion from an impact-generated disk. *Nature* **389**, 353-357.

Kokubo E., R.M. Canup, and S. Ida 1999. Lunar accretion from an impact-generated disk. In *Origin of the Earth and Moon,* eds. R.M. Canup and K. Righter, Univ. of Arizona Press, in press.

Kokubo E., J. Makino and S. Ida 1999. Evolution of a circumterrestrial disk and formation of a single moon. in preparation.

Lynden-Bell, D. , and A.J. Kalnajs 1972. On the Generating Mechanism of Spiral Structure *Mon. Not. Roy. Astron. Soc.* **167**, 1-30.

Lynden-Bell, D. and J.E. Pringle 1974. The evolution of viscous discs and the origin of the nebular variables. *Mon. Not. Roy. Astron. Soc.* **168**, 603-637.

Ohtsuki, K. 1993. Capture Probability of colliding planetesimals: Dynamical constraints on the accretion of planets, satellites, and ring particles. *Icarus* **106**, 228-246.

Takeda, T., E. Kokubo, and S. Ida 1999. in preparation

Toomre, A. 1964. On the Gravitational Stability of a Disk of Stars. *Astrophys. J.* **146**, 1217-1238

Titius-Bode's Law in a Simulated Planetary System of Two-Dimension

Toshio KAWAI, Chitose Institute of Science and Technology,
Chitose, Hokkaido 066-8655 Japan
Yukio YAMAMOTO, The Institute of Space and Astronautical Science,
Sagamihara, Kanagawa 229-0022 Japan

Abstract

We have explained the arrangement of our planets in accordance with Titius-Bode's law, by a mechanical model of coagulation and scattering of swarm of particles under mutual gravity. The model is simple and still realistic except for the assumption that all orbits are on a plane.

1 Introduction

We present a dynamical model, which reproduces the evolution of a planetary system, and specifically Titius-Bode's law.

It is generally agreed that our Planetary System has evolved from a swarm of small particles in the process of coagulation and scattering by mutual gravity. Studies have been accumulated on statistical and hydrodynamic behavior of the swarm, or on fundamental analysis of 2-3 body scattering.[1][2]

A special purpose computer GRAPE has been developed for such problems. This project has grown successfully to aim at the speed of PETA FLOPS now.[3][4] But until today the questions still remain: why do we have ten planets, or why are they arranged in a way as Titius and Bode correlated?

We utilize the notion and technique of modern orbital elements to formulate this problem. The elements change drastically when the particle collides or is scattered very much. Such drastic event turns out to occur $2 \sim 3$ times throughout the life of the particle. Except for such rare instantaneous events, the orbital elements stay essentially unchanged. Discovery of this property lead to a new algorithm, which saves computer time and gives insight into the physical process of planet formation. The model has succeeded to show clearly and quantitatively that our Planetary System evolved through accumulation of rather few drastic events of collisions and large scatterings.

2 The model and analysis

We assume that N small particles of equal mass existed on two-dimensional space initially, $m_i = pM_\odot/N$, where $M_\odot$ is the mass of the sun, $pM_\odot$ is the initial mass of all particles and p is a parameter. They are distributed randomly on a plane ranging from Mercury to Pluto (the distribution is specified later). Particles stay on an orbit except at collision and scattering. Such events are assumed to occur when a pair of orbits crosses. We assume that interactions among particles whose orbits do not cross can be ignored. When particles collide, they coagulate. The computation continues until there are not any orbits that cross.

Unit We employ three units:

mass: the mass of the sun, M $(1.9891 \times 10^{30} kg)$,

length: geometric mean of the orbital radius,

$$R \equiv \sqrt{R_{Mercury} \cdot R_{Pluto}} = 5.8526 \times 10^{11} m = 3.92\mathrm{AU} \quad (1)$$

time: $T = (R^3/GM)^{1/2} = 3.8873 \times 10^7 s = 1.232 year$.

For units of other quantities which appear in this paper, see Appendix A. Note that $r = 1$ approximately corresponds to the location of Uranus (5.2 AU).

In this unit system, $G = 1$ and $M_\odot = 1$. G and $M_\odot$ ($\equiv M$) have value 1, but they have dimensions as well as r, t, m and all quantities such as $h, v, a, b, \rho, \bar{f}$ etc. They are not non-dimensionalized.

Orbital elements We employ modern orbital elements ϵ (Lenz vector, eccentricity) and h (angular momentum/mass):

$$\begin{aligned} \epsilon &= \frac{v \times h}{GM} - e, \\ h &= r \times v, \end{aligned} \tag{2}$$

where $e = r/r$ is the unit vector directing to the particle from the origin (the sun). In our two-dimensional case, a set of three elements $(\epsilon_x, \epsilon_y, h_z)$ defines an orbit.

Condition of orbit crossing From the well known equation of orbit

$$r_i = \frac{h_i^2}{1 + \epsilon_i \cdot e_i} e_i, \quad r_j = \frac{h_j^2}{1 + \epsilon_j \cdot e_j} e_j. \tag{3}$$

Crossing occurs at e_0 if e_0 satisfies

$$r_i(e_0) = r_j(e_0), \quad \text{or} \quad \frac{1}{h_i^2} - \frac{1}{h_j^2} + \left(\frac{\epsilon_i}{h_i^2} - \frac{\epsilon_j}{h_j^2} \right) e_0 = 0. \tag{4}$$

Defining

$$a = \left(\frac{\epsilon_i}{h_i^2} - \frac{\epsilon_j}{h_j^2} \right), \quad b = \left(\frac{1}{h_i^2} - \frac{1}{h_j^2} \right), \tag{5}$$

eq.(4) is written

$$b + a \cdot e_0 = 0. \tag{6}$$

This equation has real roots when $|b| < |a|$ and

$$e_0 = \begin{pmatrix} e_{0x} \\ e_{0y} \end{pmatrix} = \frac{-ba + s\sqrt{a^2 - b^2} k \times a}{a^2}, \tag{7}$$

where k is the unit vector $(0, 0, 1)$. We have two cross points, which will be designated by a third index s, which takes two values $+1$ and -1. A unit vector e_0 pointing the cross point is designated by three suffixes i, j, s.

Velocity of particles at the point r Vector product of ϵ in eq.(2) with h yields

$$v = h \times (\epsilon + e)/h^2 GM. \tag{8}$$

Transformation upon collision When particles collide and thus coagulate at a cross point $P(i, j, s)$, new mass and orbital elements are generated by

$$m' = m_i + m_j, \tag{9}$$

$$m'h' = m_i h_i + m_j h_j, \tag{10}$$

$$m'v' = m_i v_i + m_j v_j. \tag{11}$$

The last equation, together with eq.(8), leads to an equation for ϵ':

$$m'(\epsilon' + e_0)/h' = \sum_{\alpha=i,j} m_\alpha (\epsilon_\alpha + e_0)/h_\alpha. \tag{12}$$

Timing of particle crossing Let the particles i, j pass the cross point at times τ_i and τ_j. We define the time difference $\tau = \tau_i - \tau_j$. Once in period T_j, the particle j is at the cross point. At this moment, probability of τ_i to be in interval $(\tau_i, \tau_i + \Delta\tau)$ is assumed to be $\Delta\tau/T_i$. Therefore the probability of event per time that τ is in $(\tau, \tau + \Delta\tau)$ is $\frac{\Delta\tau}{T_i T_j}$ (probability/time) which is denoted by $\tilde{p}(i, j, s)$. Here the period of rotation is $T_i = 2\pi h_i^3 (1 - \epsilon_i^2)^{(-3/2)}$.

Minimum distance of particles at an encounter We consider two particles near the cross point r_0. Let the particle i pass the cross point at the time $t = \tau_i$.

The motion is approximated by

$$r_i = v_i(t - \tau_i) + r_0, \quad r_j = v_j(t - \tau_j) + r_0. \tag{13}$$

The distance of two particles is

$$\begin{aligned} r_{ij} &= r_i - r_j \\ &= (v_i - v_j)t - (v_i\tau_i - v_j\tau_j) \\ &= v_{ij}t - C, \end{aligned} \tag{14}$$

where

$$v_{ij} = v_i - v_j, \quad C = v_i\tau_i - v_j\tau_j. \tag{15}$$

The distance $|r_{ij}|$ is minimum when

$$\frac{\partial r_{ij}^2}{\partial t} = 2v_{ij}^2 t - 2v_{ij} \cdot C = 0, \tag{16}$$

or

$$t_{min} = \frac{v_{ij} \cdot C}{v_{ij}^2}. \tag{17}$$

Substituting this for eq.(14), minimum of $r_{ij} = r_m$ is derived as follows.

$$\begin{aligned} r_m &= v_{ij}\frac{v_{ij} \cdot C}{v_{ij}^2} - C\frac{v_{ij} \cdot v_{ij}}{v_{ij}^2} \\ &= \frac{v_{ij} \times (v_{ij} \times C)}{v_{ij}^2} \\ &= -\frac{v_{ij} \times (v_i \times v_j)}{v_{ij}^2}(\tau_i - \tau_j) = -v_c\tau \end{aligned} \tag{18}$$

where

$$\tau = \tau_i - \tau_j, \quad v_c = v_{ij} \times (v_i \times v_j)/v_{ij}^2. \tag{19}$$

Probability of collision rate If the masses m_i and m_j are so small that the two orbits fly on straight lines as eq.(13), then collision occurs when $|r_m| < d \equiv (d_i + d_j)/2$, where

$$d_i = [6m_i/\pi\rho]^{1/3}. \tag{20}$$

It must be noted that ρ is the density of planet (a parameter, tentatively assumed to be that of water: $10^3 kg/m^3) = 10^8$ in our unit system. Using eq.(18), the condition of collision in terms of τ is

$$|\tau| < d/|v_c| \equiv \tau_c. \tag{21}$$

If we consider the effect of orbit deformation due to mutual gravity, this equation is revised as (see Appendix B)

$$\tau_c = \frac{d}{|v_c|}\left(1+\frac{2G(m_i+m_j)}{dv_{ij}^2}\right)^{\frac{1}{2}}. \tag{22}$$

This reduces to eq.(21) when $m \to 0$.

The probability of collision per time is

$$\bar{p}_c(i,j,s) = \frac{2\tau_c(i,j,s)}{T_iT_j}. \tag{23}$$

In other words, particles of the pair collide once in $T_iT_j/2\tau_c$.

Analysis of scattering The impulse on i from j is $\bar{\boldsymbol{f}}_{ij} = -\int Gm_im_j\boldsymbol{r}_{ij}/r_{ij}^3dt$. This is readily integrated (see Appendix B)

$$\bar{\boldsymbol{f}}_{ij} = \frac{2Gm_im_j}{|\boldsymbol{v}_i\times\boldsymbol{v}_j|\tau}\hat{\boldsymbol{v}}_c = \frac{2Gm_im_j}{v_{ij}v_c\tau}\boldsymbol{v}_c \tag{24}$$

$\hat{\boldsymbol{v}}_c$ is the unit vector directing $\boldsymbol{v}_c$. From the impulse, $\Delta\boldsymbol{v}_i$, $\Delta\boldsymbol{v}_j$ and hence Δh, $\Delta\epsilon$ are derived:

$$\Delta\boldsymbol{v}_i = \bar{\boldsymbol{f}}_{ij}/m_i. \tag{25}$$

Change of angular momentum $\boldsymbol{h}$ and eccentricity $\Delta\epsilon$ are

$$\Delta\boldsymbol{h}_i = \boldsymbol{r}_0\times\Delta\boldsymbol{v}_i, \tag{26}$$

$$\Delta\epsilon_i = \frac{\Delta\boldsymbol{v}_i\times\boldsymbol{h}_i + \boldsymbol{v}_i\times\Delta\boldsymbol{h}_i}{GM}. \tag{27}$$

Similarly,

$$\Delta\boldsymbol{v}_j = -\bar{\boldsymbol{f}}_{ij}/m_j \tag{28}$$

$$\Delta\boldsymbol{h}_j = \boldsymbol{r}_0\times\Delta\boldsymbol{v}_j \tag{29}$$

$$\Delta\epsilon_j = \frac{\Delta\boldsymbol{v}_j\times\boldsymbol{h}_j + \boldsymbol{v}_j\times\Delta\boldsymbol{h}_j}{GM}. \tag{30}$$

Particles i,j undergo many scattering until collision occurs. The condition of collision is given by eq.(21). Let τ play "random walk" in a range $[-T/2, T/2]$. As a consequence of random walk, particles are scattered many times. It is well known that the magnitude of the random impulse until a collision is represented by one impulse $\bar{F}[impulse/collision]$ which is the "root mean square" of $\bar{\boldsymbol{f}}$.

$$\bar{F}^2 = \frac{2\int_{\tau_c}^{T}\bar{\boldsymbol{f}}^2d\tau}{2\int_0^{\tau_c}d\tau} = \frac{\left(\frac{2m_im_j}{v_{ij}v_c}\right)^2\frac{1}{\tau_c}}{\tau_c} = \left(\frac{2m_im_j}{v_{ij}v_c\tau_c}\right)^2. \tag{31}$$

In summary this represents effects of many scattering until collision. Here $1/T \to 0$ is used because $T \gg \tau_c$. T has no effect on the result.

Selection of a pair (i,j,s) and time interval Δt $\sum \tilde{p}_c(i,j,s)$ is the collision rate per unit time. Now we choose Δt by

$$\sum \tilde{p}_c \Delta t = \frac{1}{2}. \tag{32}$$

In the interval Δt, we assume that one collision or one scattering occurs on one pair. The probability of the collision and scattering is 1/2 each.

The choice of the pair (i,j) and cross point s is made randomly with the probability proportional to $\tilde{p}_c(i,j,s)$.

Time sequence of the evolution is traced until all orbits do not cross, and nothing changes any more.

Initial condition Initial orbital elements are randomly determined in the range

$$-\alpha < \log h < \alpha \tag{33}$$

$$|\epsilon| < \beta \tag{34}$$

where α, β are parameters. For the choice of initial mass distribution, see Appendix C.

The parameter $\alpha = 1.156$ corresponds to the distance from Mercury to Pluto.

3 Algorithm

The algorithm is simple as outlined below.

input input 6 parameters of the model: α, β, γ, N_0, p, ρ.

initialize state table generate m_i, h_i, ϵ_i $(i = 1, ...N_0)$

initialize cross table check the crossing for all pairs of orbit. When a pair crosses, register i, j, $\tilde{p}_{ijs}$ successively $(k = 1, 2, ...K)$ where K is the total number of cross points.

until $K = 0$, repeat $\Delta t = \frac{1}{2}\sum \tilde{p}_{ijs}$, select i, j, s with probability proportional to $\tilde{p}_{ijs}$.

Do either collision or scattering transformation (probability 1/2 each)

update t, $t = t + \Delta t$

update state table m_i, h_i, ϵ_i $(i = 1, ...N)$

update cross table $i, j, \tilde{p}_{ijs}$ $(k = 1, ...K)$

4 Results and discussions

Our model has 6 parameters. Their meaning and typical values are:

Table 1: Parameters of this model and standard values

N_0	300	number of initial particles
α	1.156	range of initial distribution
β	0.3	range of initial eccentricity
p	0.001	initial total mass of planets $/M_\odot$
ρ	$10^3[kg/m^3]$	density of planet
γ	2	initial distribution of particles

Figure 1 shows "evolution diagram" at this typical values except for $N_0 = 150$. The horizontal axis stands for $\log t$ and the transverse axis is $\log h = \frac{1}{2}\log a$. At each event, circles whose diameter is proportional to the mass are plotted on (t, h_i).

As time goes on, Mercury is born at first followed by Venus, Earth.... After the final event, the system is completed. It takes about 10^8 years. Final number of particles N_F is 8 in this example, and the radius of revolution ($\propto h^2$) is nearly equally spaced in log scale, namely the geometric series.

Fig.1 shows a result of a case, which depends on the set of random numbers. Using the same set of parameters, we obtain different history and different final number of planets, because the event of collision and scattering is stochastic. In view of chaotic behavior in the dynamical system, this stochastic property appears even in exact Newton dynamics. Furthermore, the initial condition is accidental. So we cannot readily conclude that our stochastic model has inherent defect. Oppositely, if we find many planetary systems in the future, the variation may be observed.

In any case, we calculated 100 cases of different set of random numbers and obtained $< N_F >$ and $\sigma_{N_F} = \sqrt{< (N_{Fi} - < N_F >)^2 >}$. As the result, in the standard condition $N_F = 7.2 \pm 2.1$. In the same condition, from the calculation of 1000 cases, $N_F = 7.1 \pm 2.1$ is obtained. We can conclude that 100 cases is enough. Fig.2 is the histogram of N_F obtained from this 1000 cases.

We investigate the effects of the six parameters on N_F and σ_{N_F}. We change one parameter each from the standard condition. α is changed in Fig.3. N_F increase linearly to α. That is, N_F increases by 1 every $\Delta\alpha = 0.154$. The radius of revolution forms geometric series which has a common ratio of 1.85. Fig.4 shows the actual configuration of planets. The solid line stands for the geometric series whose common ratio is 1.85 and the dotted line is Bode's law.

On Figs.5,6,7,8 and 9, N_0, β, ρ, γ and p are changed from the standard condition. N_F is independent of these parameters to the first approximation. We can conclude that N_F/α is approximately a universal constant.

Fig.10 shows the process of collision and scattering on the revolution diagram in the standard condition. The particle after collision is expressed by $\circ$ and the particle scattered away is shown by $\uparrow$. The graphs 11 $\sim$ 14 show the orbits every $t = 10^n$ $(n = 3, 5, 7, 9)$.

Fig.15 is the final state of the standard condition where the effect of scattering is removed. In the process which has collision only, orbits get less eccentricity and less chance of crossing, leaving more particles in the final state. Scattering is so important to explain our solar system of 10 planets.

Fig.16 shows the process of collision and scattering. A point in this figure represents the mass and location (h) of a particle. Two points before a collision are connected to a point representing the new particle. We can see the fate of all initial particles. Some remain in a part of a planet, and some are scattered out of the planetary system. We note that the masses of particles scattered away are light, and the final particles, namely planets, have gathered the particles that are initiallyin their neighborhood.

Fig.17 is the graph of total mass of particles vs time. Total mass is reduced by scattering. In this case 1/3 of initial mass is scattered away.

In all these cases, we have assumed that initial mass of particles are equal ($= m_0$).

We calculate a case where initial mass of particles has distribution

$$m = Ae^{-(m-m_0)^2/2m_0^2}. \tag{35}$$

The statistics of the final number of stars out of 100 cases was

$$N_F = 6.95 \pm 2.0. \tag{36}$$

Therefore, N_F neither depends on initial mass distribution of particles.

Summing up these results, $< N_F >$ does not change so much even if parameters are drastically changed around the "standard case" shown in Table 1.

So far we have discussed number of planets N_F as a representative quantity. We now show the location of the planets and confirm the geometric or exponential law. For this purpose we select 100 cases which yield $N_F = 7$, and calculated

$$\Delta_i = \log h_{i+1} - \log h_i \quad (i = 1, 2, ...6). \tag{37}$$

Its average $< \Delta_i >$ and standard deviation is shown in Fig.18 for the standard condition and other cases with different initial mass distribution. $< \Delta_i >$ is found to be nearly constant with some systematic deviation.

5 Conclusion

We have shown that the essential features of our planetary system can be explained by a simple, mechanical, two-dimensional model. From the initial standard condition, about 7 planets evolve in 10^8 years in the space ranging from Mercury to Pluto. The number of final planets is approximately 7 for all parameters of the model for a wide range of variation. The location of the planets obeys Bode's law $0.4 + 0.32^n$ (AU) $\sim 1.85^{(n-1)}$ (n :planet number. $n_{earth} = 1$). The evolution is simulated easily: formulation is simple and computer time is only several seconds by a present personal computer.

The evolution could be explained by tracing a thousand particles by Newton's law for 10^8 years. Even today, the lack of computing power is fatal for such an approach. The difficulty is common for all frontiers of computational sciences such as protein dynamics, where we can simulate the molecules for a period of time too short to obtain valuable practical information.

Among the computation-intensive problems, the planetary system has a peculiar property that the essence of the phenomena is successive events (of collision and large scattering). Other than such instantaneous events, the particles turn around the sun "without change". Of course they move around in the ordinary phase space, but they almost stay at fixed points in the 5 dimensional parameter space (the space of orbital elements $\epsilon, \boldsymbol{h}$). Successive transformation in the orbital space occurs 1000 times, and the system reaches the final state.

Even though we need 10^4 floating point operations for one transformation, computer can compute one case in a few seconds, and generate an evolution diagram. This is an exceptionally lucky case for a computational physicist

We would like to acknowledge many persons for valuable discussions. To name just a few they are:

Yoshitsugu Ohno (Physics, Illinois University), Daiichiro Sugimoto and Eiichiro Kokubo (The University of Tokyo), Hiroshi Kinoshita and Shoken Miyama (National Astronomical Observatory), Mikio Shimizu (The Institute of Space and Astronautical Science), Nobuyoshi Wakabayashi (Keio University).

Appendix

A Units and dimensions

Units and dimensions employed in this paper are as follows.

$M_\odot$	mass of the sun	1	$[M]$
m	mass of initial particles	$1.9891 \times 10^{30} kg$	$[M]$
v	velocity	$1.5055 \times 10^{4} m/s$	$[RT^{-1}]$
ρ	density	$0.9921 \times 10^{-5} kg/m^3$	$[MR^{-3}]$
ω	angular velocity	$2.5724 \times 10^{-8} s^{-1}$	$[T^{-1}]$
h	angular momentum/mass	$8.8155 \times 10^{15} m^2/s$	$[R^2T^{-1}]$
ϵ	eccentricity		$[1]$
a, b	auxiliary quantities	$1.2880 \times 10^{-32} s^2/m^4$	$[R^{-4}T^2]$
d	diameter of particle	$5.8526 \times 10^{11} m$	$[R]$
$\tilde{p}$	probability/time	$2.5724 \times 10^{-8} s^{-1}$	$[T^{-1}]$
$\bar{f}$	impulse	$1.9947 \times 10^{34} kg \cdot m/s$	$[MRT^{-1}]$
G	constant of gravity	1	$[M^{-1}R^3T^{-2}]$

It must be noted that all quantities have dimensions. They are not non-dimensionalized. However, $G = 1$ and $M_\odot = 1$ wherever they appear in this paper.

B The argument when orbits are deformed by attraction

We consider two particles attracting by gravity

$$m_i \ddot{\boldsymbol{r}}_i = -\frac{G m_i m_j}{|\boldsymbol{r}_{ij}|^3}(\boldsymbol{r}_i - \boldsymbol{r}_j), \quad m_j \ddot{\boldsymbol{r}}_j = -\frac{G m_i m_j}{|\boldsymbol{r}_{ij}|^3}(\boldsymbol{r}_j - \boldsymbol{r}_i). \tag{38}$$

Dividing both sides by m_i, m_j and subtracting, we have

$$\ddot{\boldsymbol{r}}_{ij} = -\frac{G(m_i + m_j)}{|\boldsymbol{r}_{ij}|^3}\boldsymbol{r}_{ij}. \tag{39}$$

Abbreviating $\boldsymbol{r}_{ij}$ and $m_i + m_j$ by $\boldsymbol{r}$ and m, we have two conservation laws.

$$\frac{1}{2}\dot{\boldsymbol{r}}^2 - \frac{Gm}{r} = E_0 \tag{40}$$

$$\boldsymbol{r} \times \dot{\boldsymbol{r}} = \boldsymbol{L}_0. \tag{41}$$

Let us apply these laws to our case at $t = -\infty$ and $t = t_{min}$ where $|\boldsymbol{r}| = x$ and $|\boldsymbol{v}| = U$ (unknown).

$$E_0 = \frac{1}{2}v_{ij}^2 = \frac{1}{2}U^2 - \frac{Gm}{x} \tag{42}$$

$$\boldsymbol{L}_0 = (0, 0, -r_m v_{ij}) = (0, 0, -xU). \tag{43}$$

Eliminating U from the two equalities, we obtain

$$x = -\frac{Gm}{v_{ij}^2} + \sqrt{\left(\frac{Gm}{v_{ij}^2}\right) + r_m^2}. \tag{44}$$

The condition of collision, $x < d$, leads to

$$d\left(1 + \frac{2Gm}{dv^2}\right)^{\frac{1}{2}} > r_m = \tau v_c. \tag{45}$$

The critical value of τ is

$$\tau_c = \frac{d}{v_c}\left(1+\frac{2Gm}{dv^2}\right)^{\frac{1}{2}}. \tag{46}$$

Now let us consider the impulse on i from j

$$\bar{f} = -Gm_im_j\int\frac{|r_i-r_j|}{r_{ij}^3}dt. \tag{47}$$

As in Fig.19, let us take $\hat{v}_{ij}$ as x axis.

Let us note that $v_{ij}\cdot r_m = 0$, and $r_m = v_c\tau$. The motion is described by

$$r = (vt, r_m, 0), \tag{48}$$

and

$$\begin{aligned}\bar{f} &= -\int_{-\infty}^{\infty}Gm_im_j\frac{(vt,rm,0)}{(v^2t^2+r_m^2)^{\frac{3}{2}}}dt\\ &= -Gm_im_jr_m\int_{-\infty}^{\infty}\frac{dt}{(v^2t^2+r_m^2)^{\frac{3}{2}}}\\ &= -\frac{Gm_im_j}{vr_m}(0,1,0) = -\frac{Gm_im_j}{r_m^2}\frac{2r_m}{v}(0,1,0).\end{aligned} \tag{49}$$

The magnitude can be interpreted as the force (at r_{min}) lasts for the pass time $\Delta t = 2r_m/v$. From this observation, we can well-approximate the impulse when orbit deformation is taken into account

$$\frac{2Gm_im_j}{xU} = \frac{2Gm_im_j}{r_mv}. \tag{50}$$

The expression of the magnitude of the impulse is common in terms of bv, although the direction of the impulse is no more perpendicular to v_{ij}. We will ignore this scattering angle.

Then in general the impulse is expressed as

$$\bar{f} = \frac{2Gm_im_j}{r_mv_{ij}}\hat{v}_c = \frac{2G_im_j}{v_c\tau v_{ij}}\hat{v}_c, \tag{51}$$

(note $r_m = v_c\tau$). Now

$$\begin{aligned}v_c &= \frac{|v_{ij}\times(v_i\times v_j)|}{v_{ij}^2}\\ &= \frac{\hat{v}\times k|v_i\times v_j|}{v_{ij}} = \frac{|v_i\times v_j|}{v_{ij}},\\ \hat{v}_c &= \hat{v}_{ij}\times\hat{k},\end{aligned} \tag{52}$$

thus

$$\bar{f} = \frac{2Gm_im_j}{\tau|v_i\times v_j|}\hat{v_{ij}}\times\hat{k}. \tag{53}$$

C r^m distribution

Let the probability of a particle be in $(r, r+dr)$ be

$$n(r)2\pi r dr \tag{54}$$

where r is the distance from the origin. We assume particles are distributed from $r = r_0$ to $r = r_1$. Then

$$\int_{r_0}^{r_1} n(r)2\pi r dr = 1. \tag{55}$$

If we use x which has uniform distribution in range $[0, 1]$, r' that satisfy

$$\int_{a}^{r'} n(r)2\pi r dr = x \tag{56}$$

has $n(r)$ distribution.

Now if distribution of r' is uniform in log scale, from

$$\int^{r'} n(r)2\pi r dr \propto \log h = \frac{1}{2}\log r', \tag{57}$$

we can conclude

$$n(r) \propto \frac{1}{r^2}. \tag{58}$$

When we choose $n(r) = Ar^m$ distribution, using

$$A\int_{r_0}^{r_1} 2\pi r^{m+1} dr = \left[\frac{2\pi A}{m+2} r^{m+2}\right]_{r_0}^{r_1} = 1, \tag{59}$$

we get the value of A as

$$\frac{1}{A} = \frac{2\pi}{m+2}(r_1^{m+2} - r_0^{m+2}). \tag{60}$$

So from eq.(57), random number r which obeys r^m distribution is made by following equation:

$$\sqrt{r} = h = [x(r_1^{m+2} - r_0^{m+2}) + r_0^{m+2}]^{\frac{1}{2(m+2)}} \tag{61}$$

where $m \neq -2$.

References

[1] Wetherill G. W. and G. R. Stewart: Accumulation of a swarm of small planetesimals, *Icarus*, **77**, 330-367 (1989).

[2] T. Takeuchi, S. M. Miyama, and D. N. E. Lin: Gap formation in protoplanetary discs, *Astrophysical J.*, **460**, 832-847 (1996).

[3] Sugimoto D., et al: Astronomy with Peta Flops dedicated machine, *Trans. Astronomical Soc. Jpn.*, (Spring 1996).

[4] E. Kokubo, S. Ida: Orbital evolution of protoplanets embedded in a swarm of planetessimals, *Icarus*, **114**, 147-257 (1995).

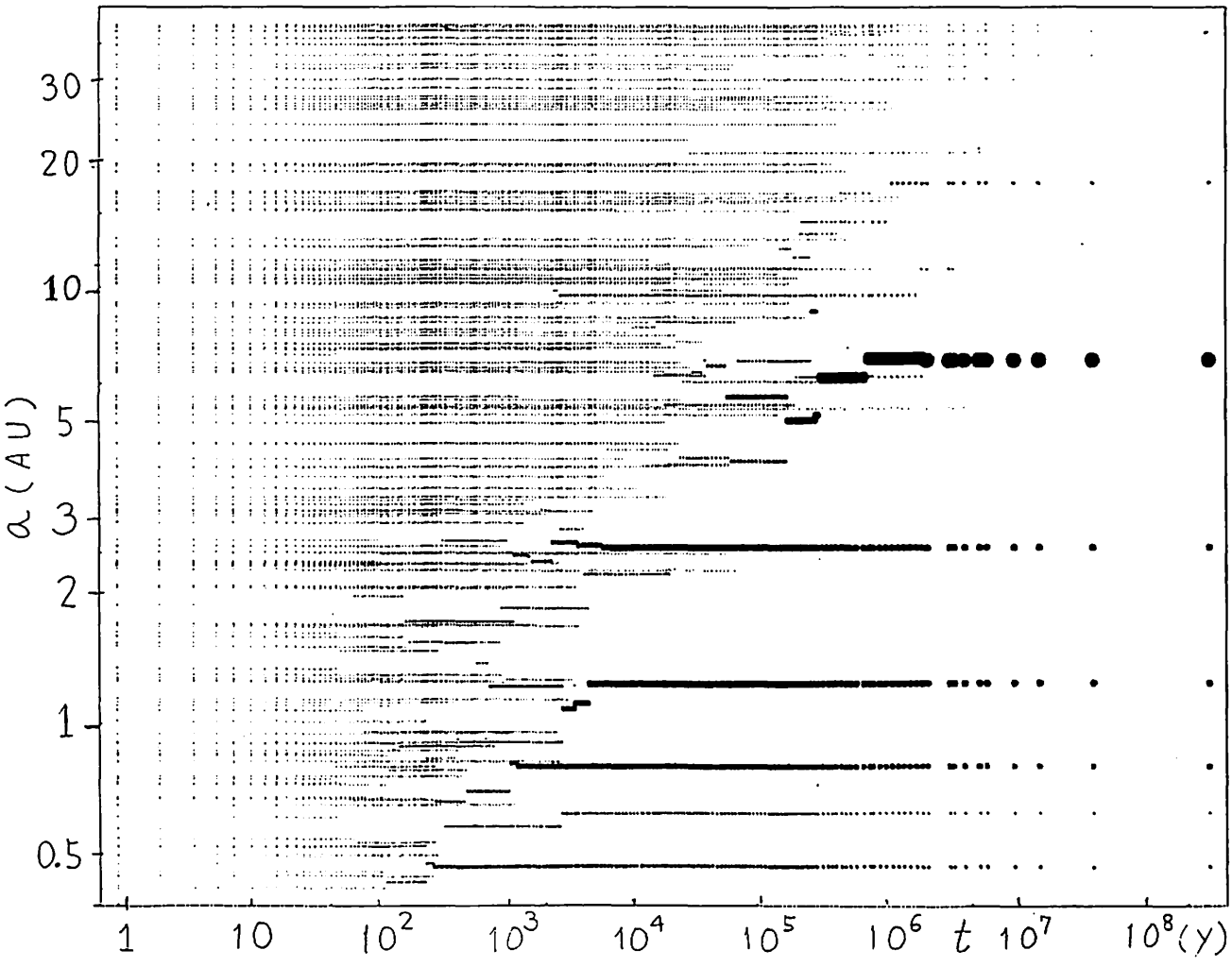


Fig.1 Evolution diagram. A dot shows time and location of a particle. The radius of the dot is proportional to the mass. The dots are plotted when a collision or a scattering occurs in the planetary system.

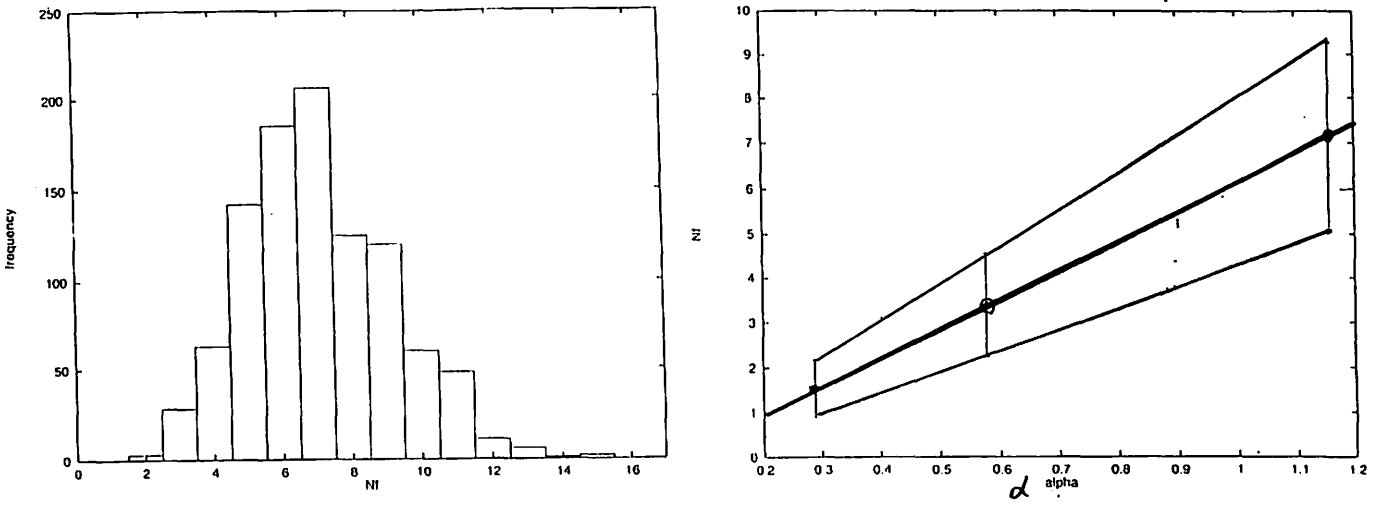


Fig.2 Histogram of N_F (final number of planets) obtained by 1000 cases of the standard condition.

Fig.3 N_F (number of final planets) vs range of initial mass distribution α ($\equiv \frac{1}{4}\log_e \frac{a_{max}}{a_{min}}$).

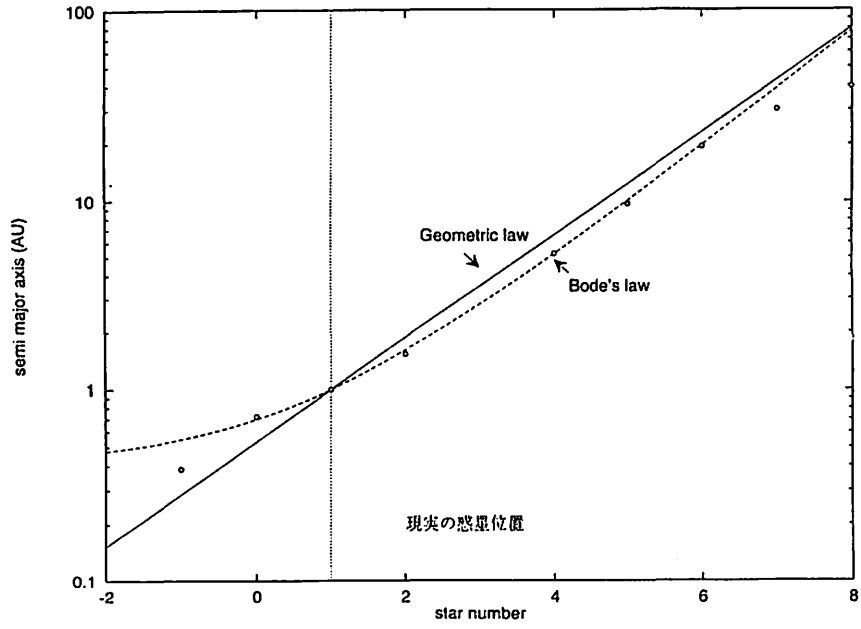


Fig.4 Actual location of our planets, Bode's law and exponential law.

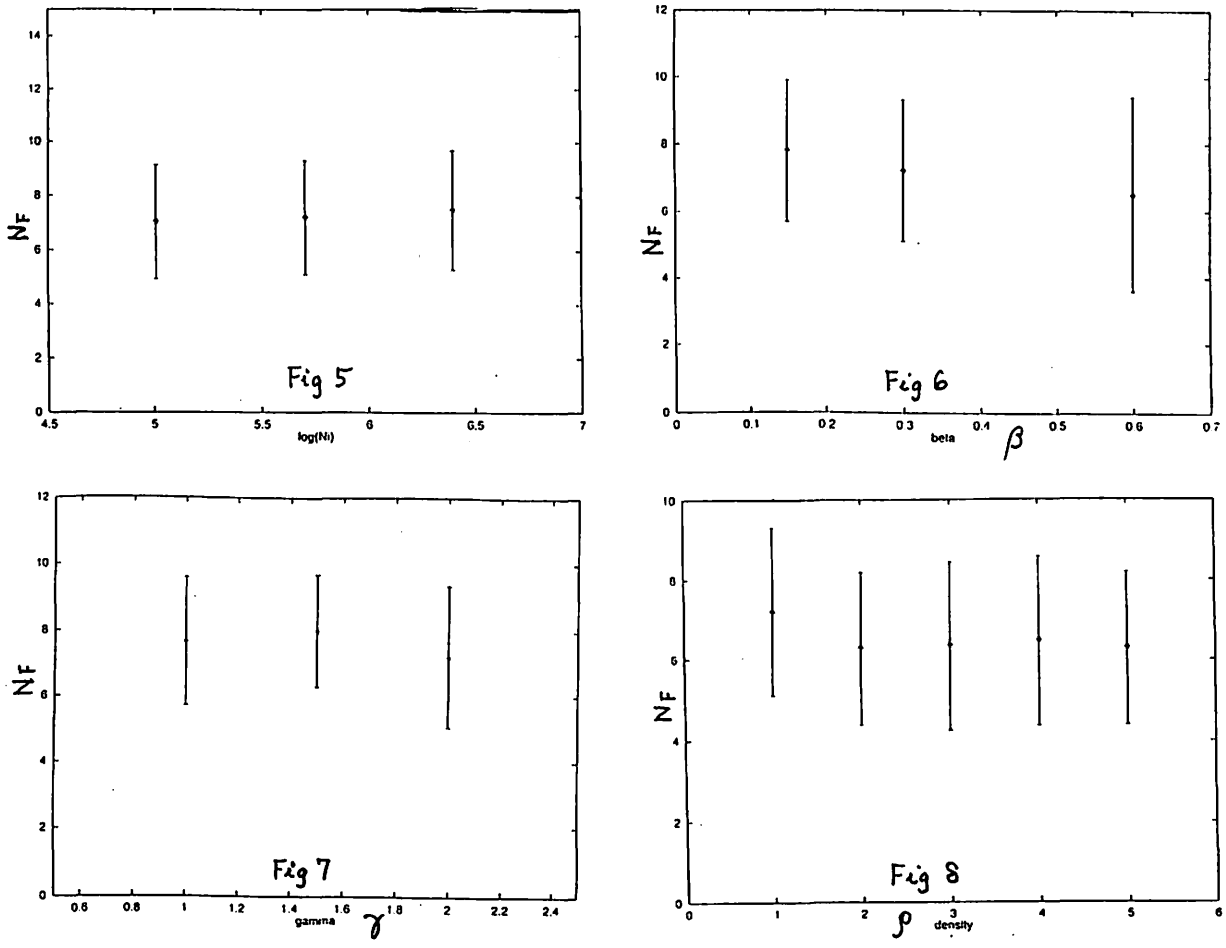


Fig.5 N_F vs number of initial particles. Other parameters are standard.

Fig.6 N_F vs β, the upper limit of initial eccentricity $|\epsilon| < \beta$. Other parameters are standard.

Fig.7 N_F vs ρ, the density of the planet in $10^3 kg/m^3$. Other parameters are standard.

Fig.8 N_F vs γ, the exponent of initial mass distribution $r^{-\gamma} 2\pi r dr$. Other parameters are standard.

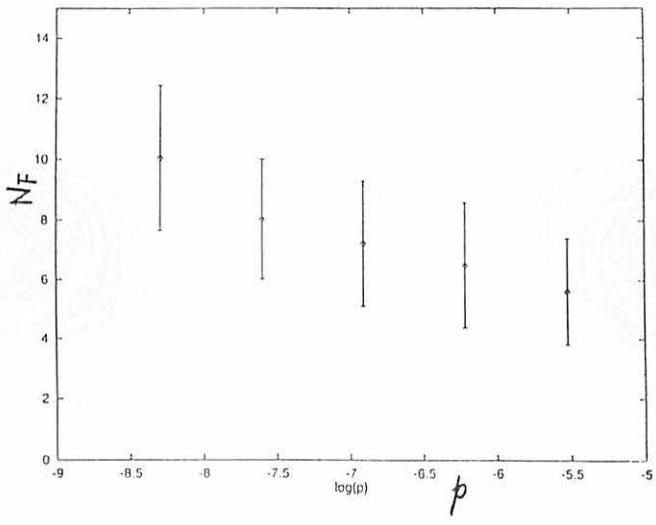


Fig.9 N_F vs p, the initial total mass of the particles normalized by $M_\odot$. Other parameters are standard.

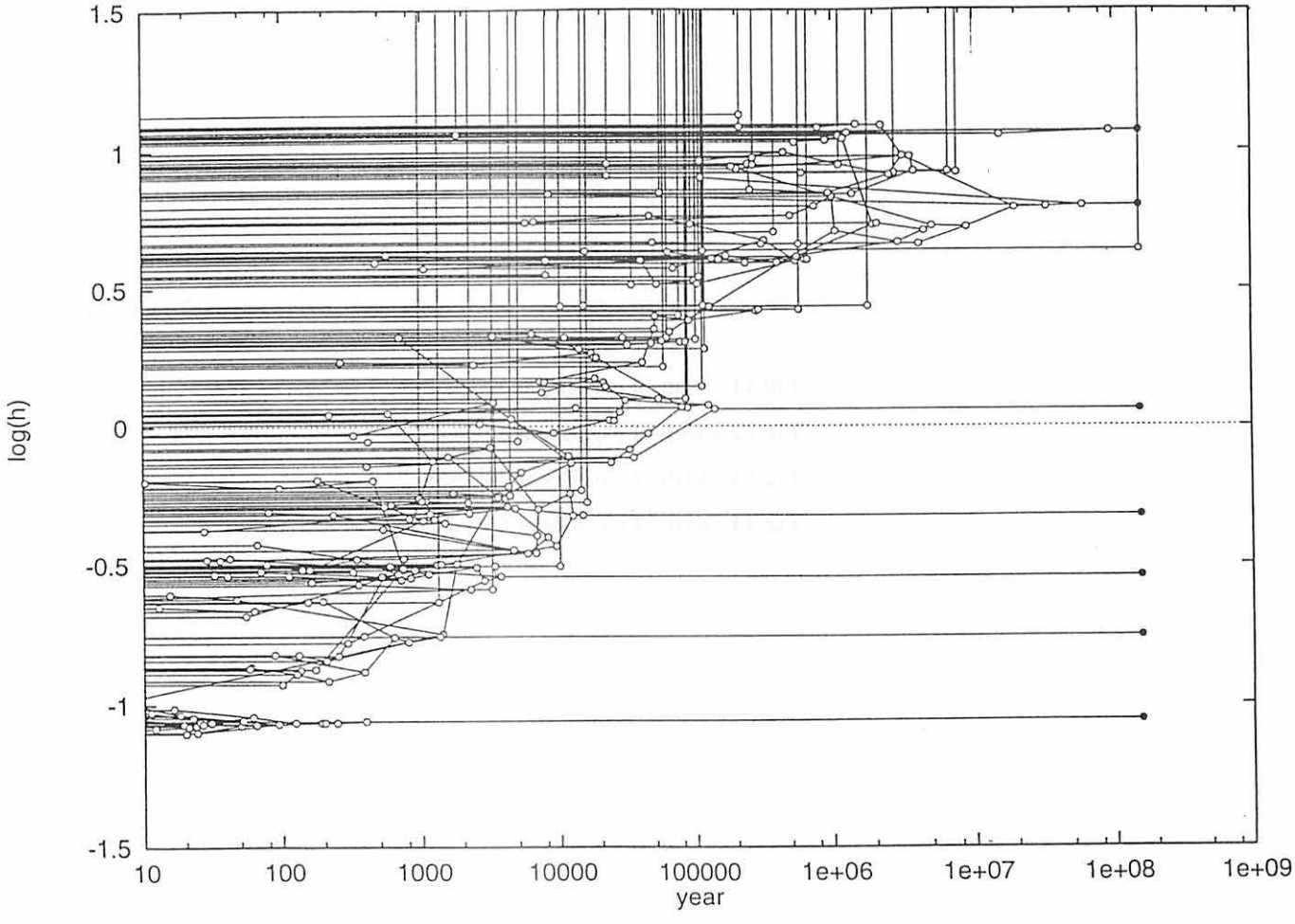


Fig.10 Details of the evolution diagram. Standard case. White circles indicate the outcome of collision. A vertical line indicate the escape of the particle by scattering.

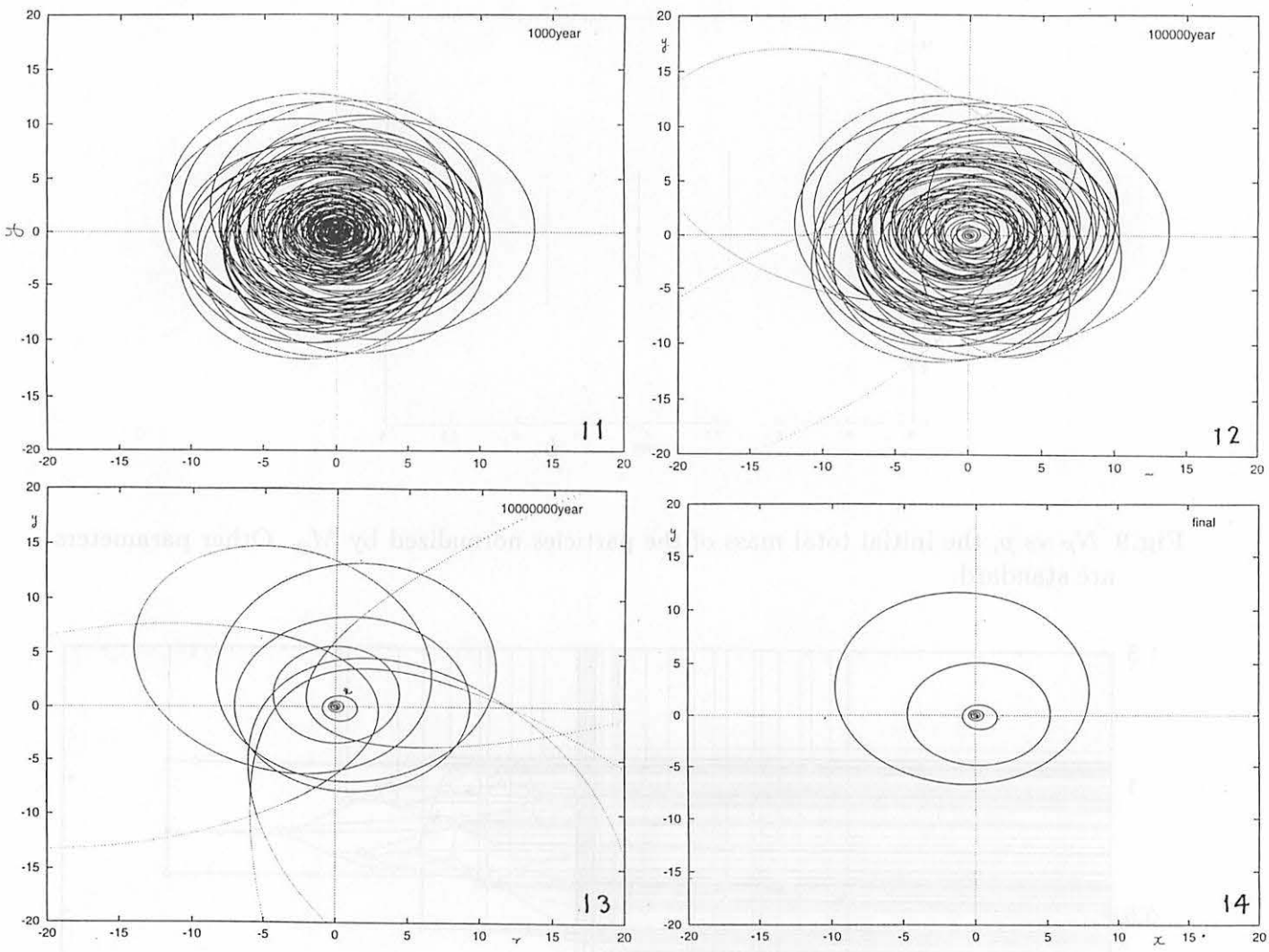


Fig.11 Orbits of particles at 10^3 year.

Fig.12 Orbits of particles at 10^5 year.

Fig.13 Orbits of particles at 10^7 year.

Fig.14 Orbits of particles at final state.

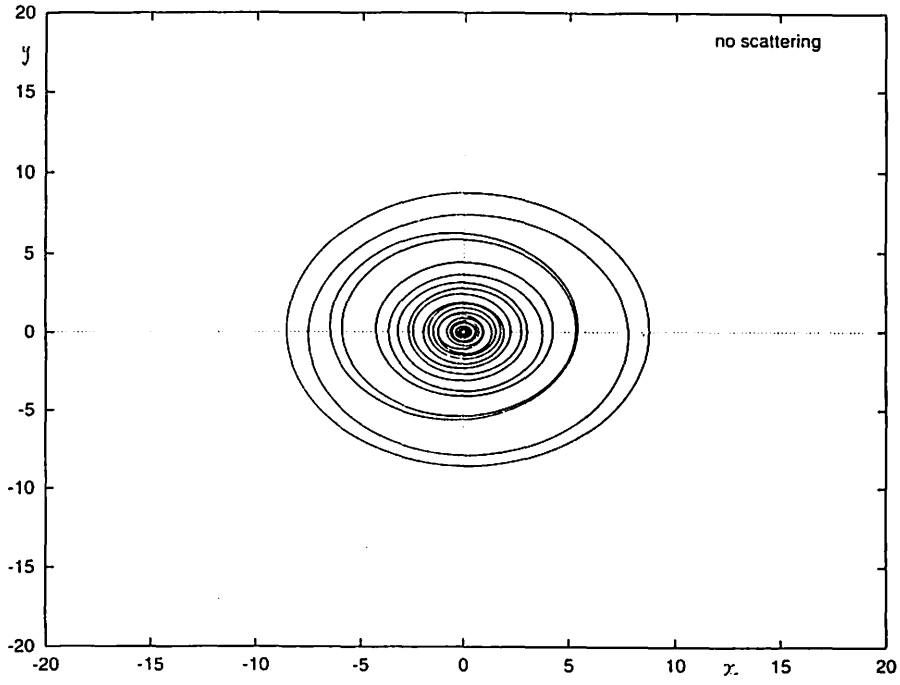


Fig.15 Orbits of particles at the final state for a model without scattering (parameters are standard).

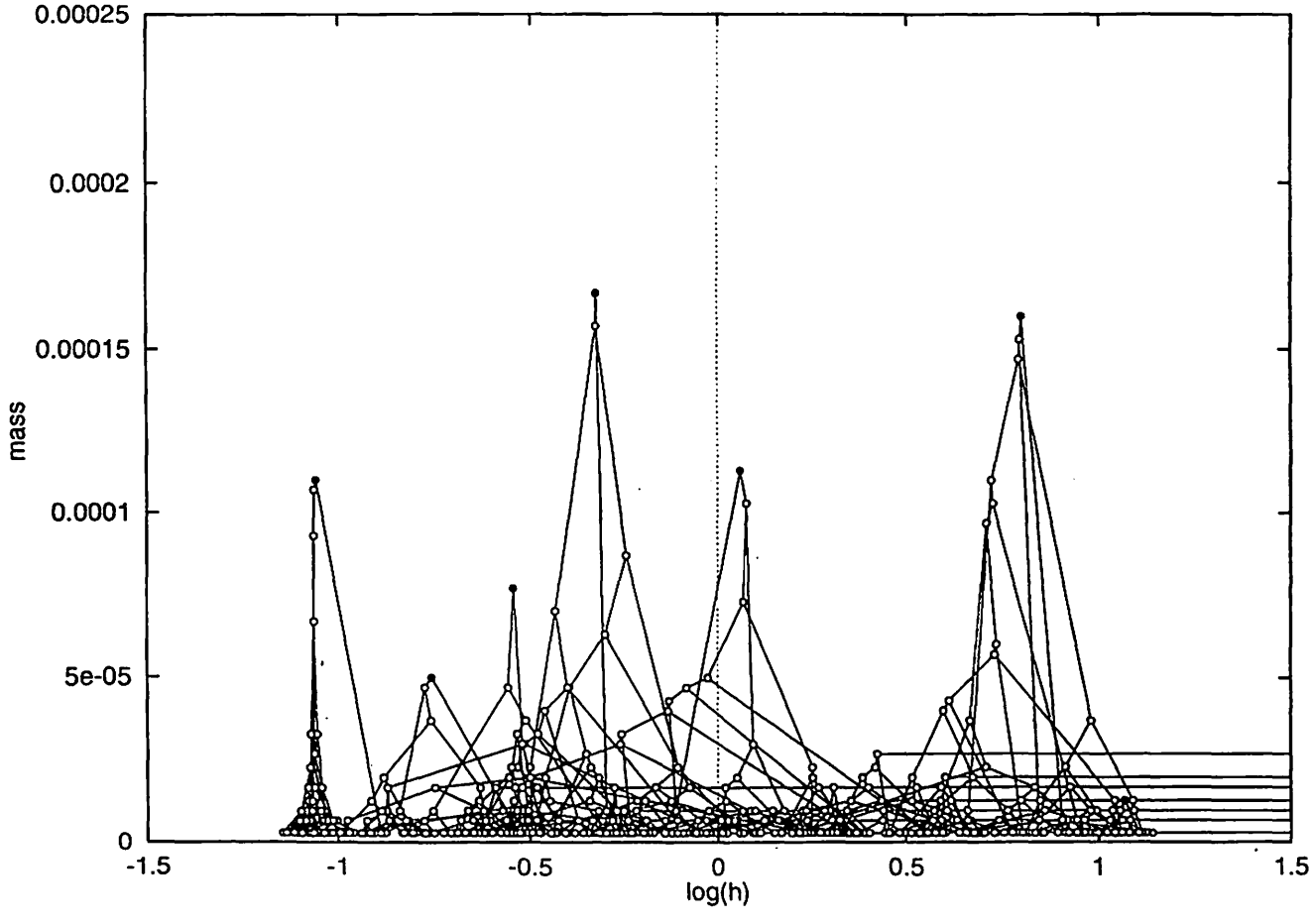


Fig.16 Coagulation map. A particle of mass m generated at h is plotted with lines pointing to two original particles. Horizontal lines show particles that are scattered out of the system.

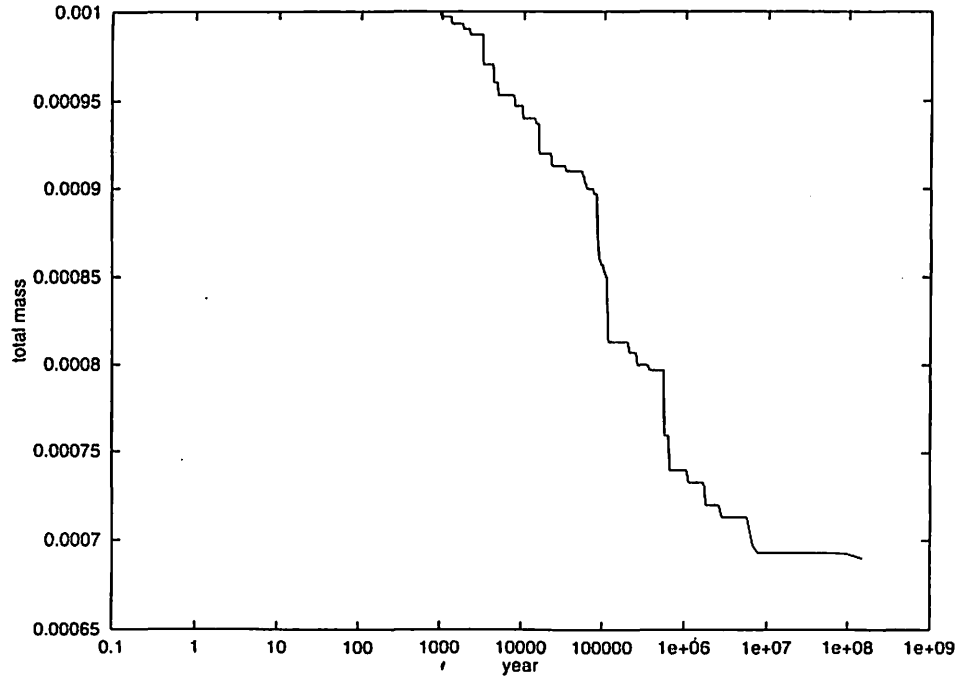


Fig.17 Total mass decrease due to scattering. Parameters are standard.

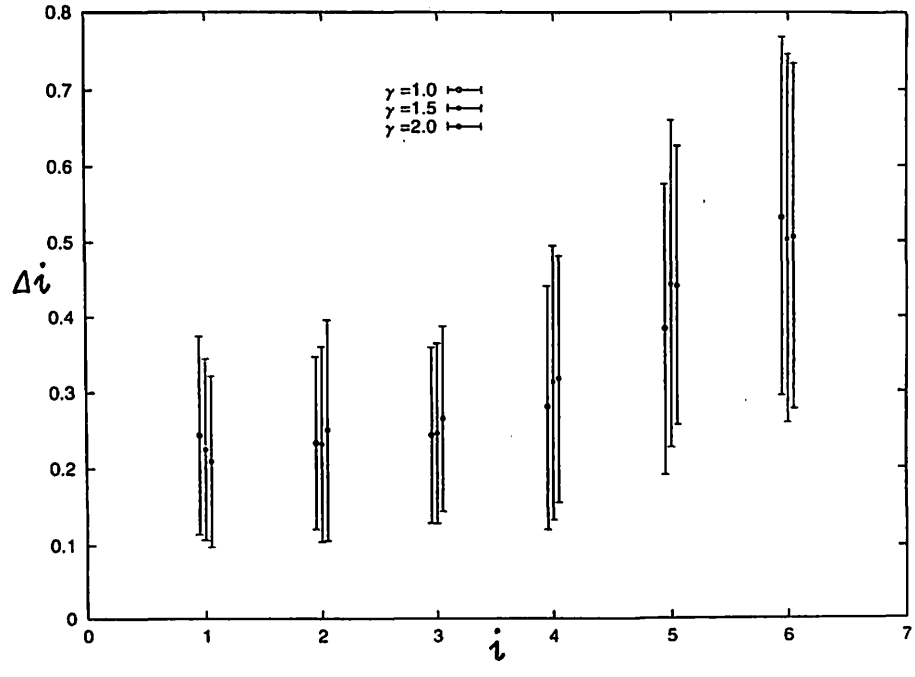


Fig.18 Interval of planets vs gap number.

Structure of Mean Motion Resonances in Outer Solar System

Tetsuharu Fuse*,
Hiroshi, Kinoshita** and Hiroshi, Nakai**
*Department of Astronomical Science, Graduate University for Advanced Studies, Mitaka, Tokyo 181-8588
** National Astronomical Observatory, Mitaka, Tokyo 181-8588
* E-mail tetsu.fuse@nao.ac.jp

Abstract

Recently, about 130 celestial bodies are discovered outside Neptune. They are called *Edgeworth-Kuiper Belt Objects* (EKBOs), and are seemed to be remnants of planetesimals. The characteristics of the EKBOs is that many of them are in mean motion resonances with Neptune. We investigate the mean motion resonances with two semi-analytical methods and numerical integration methods. We discuss the dynamical structure of each resonance, and show a role of Uranus in the resonances. We also describe effect of the ν_8 secular resonance on the structure.

1 背景

太陽系の広さは，新しい惑星が発見されるにつれて大きくなってきたといえる。1950 年前後になると，Edgeworth と Kuiper が最遠の冥王星周辺に関する新しい解釈を発表した:「太陽系外縁部には惑星になりきれなかった微惑星が存在しているに違いない」。このような海王星軌道の外側の領域を Edgeworth-Kuiper Belt (EKB), さらに EKB に存在する天体を Edgeworth-Kuiper Belt Objects (EKBOs) と呼ぶ。最近の数値シミュレーションの結果から，EKBOs は短周期彗星の起源であるとも考えられている。

CCD カメラなどの観測装置が整い，EKBOs のサーベイ観測が本格的に行われていた中，1992 年 8 月ハワイ大学の 2.2m 望遠鏡により最初の EKBO が発見された。その後から現在までの発見総数は約 130 にのぼる。発見された EKBOs は，明るさから直径数百 km と見積もられ，またその多くは海王星と平均運動共鳴 (以後レゾナンスという) にある。レゾナンスは，天体間の公転周期の比が整数で表される[1]ときに起こる得る共鳴現象であり，臨界引数と呼ばれるパラメータにより特徴づけられる[2]。

[1]たとえば，海王星と EKBOs の公転周期の比が 2 対 3 のとき，"2:3 レゾナンス" とする。
[2]レゾナンスにある天体の臨界引数は，ある角度の周りを秤動する。

2 目的と手法

本研究の目的は,「3 次元モデルを採用した半解析的手法と多くのテスト天体を用いた数値シミュレーションにより EKB 内のレゾナンス構造について詳細かつ大局的に理解すること」である。2 つのそれぞれの手法に関して，次のような開発を行った:

- 半解析的手法[3]

  離心率を持った海王星の軌道面に対して， EKBOs に軌道傾斜角を持たせた 3 次元モデルを採用する。このモデルを表すハミルトニアンにおいて，数値的に平均化するパラメータの違いにより，*Model A* と *Model E* の 2 種類の方法がある:

  * *Model A*: 短周期項を平均化し，さらに長周期項を固定することで，中周期項である天体の臨界引数と軌道長半径の変化を見ることができる。この方法では臨界引数の振る舞いだけでなく，レゾナンスの領域を決定することができる。
  * *Model E*: 天体が *Model A* により求めたレゾナンス状態にあると仮定し，短周期・中周期項を平均化することで長周期項である天体の離心率と近日点の変化を見ることができる。これにより，海王星との永年共鳴 (ν_8 永年共鳴) の存在を調べることができる。天体が ν_8 永年共鳴にあると，離心率が大きく変化する可能性がある。

- 数値シミュレーション法

  半解析的手法による結果を検証し，さらにレゾナンスの構造をより詳しく明らかにするために必要な高速数値計算アルゴリズムを開発した。このアルゴリズムは太陽系内天体の軌道計算に広く用いられている Mixed Variable Symplectic 積分法を改良し，太陽から遠く離れた数多くの天体を高速に数値積分することを可能にした (Fuse *et al.* 1998)。

以上の 2 つの半解析的手法と数値シミュレーション法を用いて，30 ～ 50 AU の間にある 1:2, 2:3, 5:6, 3:4, 3:5 の 5 つのレゾナンスについて解析を行った。

[3]小惑星と木星のレゾナンス (摂動天体が外側) の解析に用いられていた手法 (Yoshikawa 1990) を海王星と EKBOs のレゾナンス (摂動天体が内側) 用に改良した。なお木星・土星・天王星による摂動は，リング近似として考慮する。

3 レゾナンスの構造

3.1 半解析的手法

Model A と *Model E* の 2 つの半解析的手法により得られた結果について，対応する木星と小惑星のレゾナンス (Yoshikawa 1989,1990,1991) と比較しながらまとめる:

- *Model A* による結果

  - 1:2 レゾナンスにある EKBOs の臨界引数は，180 度とさらに離心率に依存する 2 つの角度の計 3 つの角度の周りを秤動する可能性がある。1:2 レゾナンスには，1 個の EKBOs が発見されている。

    一方，2:1 レゾナンスには小惑星はほとんど存在しておらず (ギャップと呼ばれる)，臨界引数は 0 度の周りのみを秤動する。

  - 2:3 レゾナンスにある EKBOs の臨界引数は，180 度の周りを秤動する。2:3 レゾナンスには，21 個の EKBOs が発見されている。

    3:2 レゾナンスに少数存在する小惑星はヒルダ群と呼ばれ，その臨界引数は 0 度の周りを秤動する。

  - 5:6 レゾナンスにある EKBOs の臨界引数も 180 度の周りを秤動する。

    6:5 レゾナンスにある小惑星は発見されていない。臨界引数は 0 度の周りを秤動する。

  - 3:4 レゾナンスにある EKBOs の臨界引数も同様に 180 度の周りを秤動する。3:4 レゾナンスには，1 個の EKBOs が発見されている。

    4:3 レゾナンスには数個の小惑星が存在し，チューレ群と呼ばれる。臨界引数は 0 度の周りを秤動する。

  - 3:5 レゾナンスにある EKBOs も臨界引数は 180 度の周りを秤動する。3:5 レゾナンスには，1 個の EKBOs が発見されている。

    数値計算によると 5:3 レゾナンスにある小惑星は安定とされるが，いまだ発見されていない。臨界引数は 180 度の周りを秤動する。

  - これらの結果から各レゾナンスの領域を求め,「軌道長半径－離心率平面」上に表した (図 1 のダイヤモンドの形をした領域内)。2 つの曲線は，それぞれ天王星・海王星軌道と交差する軌道[4]を意味する。

[4]天王星軌道と交差する軌道 ≡ 近日点距離が 19.2 AU，海王星軌道と交差する軌道 ≡ 近日点距離が 30.1 AU。

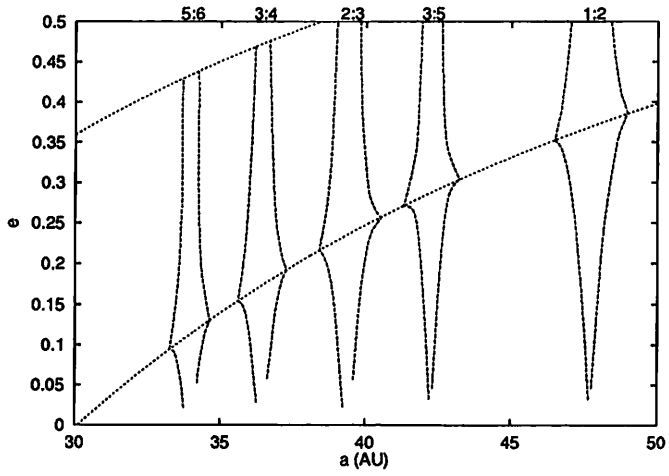


Fig. 1: The structure of the five mean motion resonances by *Model A*.

- *Model E* による結果

  - 5 つのすべてのレゾナンスにおいて，特定の条件のもとでは ν_8 永年共鳴が起こり得るが，全体的には離心率の変化は小さい。またこのような離心率の変化は，軌道傾斜角にも大きく依存する。

    一方，2:1 レゾナンスにある小惑星は，木星による永年共鳴のため離心率が非常に大きく変化する可能性がある。また 3:2，4:3 レゾナンスにある小惑星の離心率の変化は小さい。

3.2 数値シミュレーション

3.2.1 5 つのレゾナンスについて

5 つの 1:2，2:3，5:6，3:4，3:5 レゾナンスにおいた総計 20 数万体のテスト天体について 500 万年間の数値シミュレーションを行った。ここでは，木星・土星・天王星・海王星を摂動天体として考慮している[5]。テスト天体の軌道長半径と離心率の初期値空間は図 2 のように広範囲にとり，軌道傾斜角は 1 度とした。

それぞれの計算結果を「時間平均した軌道長半径と離心率の分布」として図 3 に示す。図中の曲線は，近日点距離が 25.4 AU の軌道を意味する。これは，各レゾナンスの”時間平均した離心率”の最大値を平均した値に相当する。またこの図では，今まで行われてきた初期値空間図による評価ではわからなかった，各レゾナンスの詳細な構造が見える。一方，レゾナンスにある天体はそれぞれレゾナンス領域内で直線状に並んでいる。

以下に各レゾナンスについての結果を示す:

[5]惑星の作用圏に入った天体は，その時点で計算を止める。

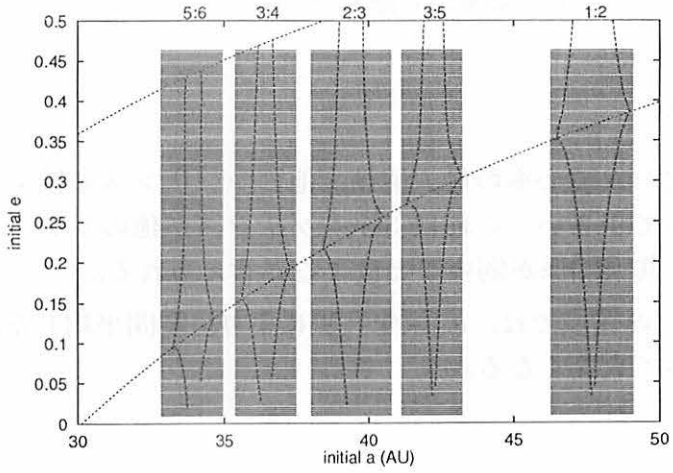


Fig. 2: The initial (a, e) of EKBOs.

- 1:2 レゾナンス
  - 臨界引数の振る舞いは，半解析的手法による結果に非常によく一致する (他の 4 つのレゾナンスについても同様)。
  - 離心率の大きく安定な天体は，他のレゾナンスに比べて少ない。*Model E* によると，このような天体は ν_8 永年共鳴によりさらに離心率が大きく変化し，レゾナンスから出てしまうためと考えられる。
  - レゾナンスになく離心率が小さい天体も，ν_8 永年共鳴により離心率が大きく変化し，その結果としてレゾナンスに入る天体もある。
- 2:3 レゾナンス
  - 時間平均した離心率の最大値は，5 つのレゾナンスの平均値に比べて最も大きく，レゾナンスが安定であることを示す。
  - 40 AU ～ の領域では，ν_8 永年共鳴により時間平均した離心率が他の領域に比べて大きくなる。
- 5:6 レゾナンス
  - レゾナンスにない天体は，離心率が非常に小さい場合のみ安定である。
  - レゾナンスにある天体は，離心率がある程度大きいものは ν_8 永年共鳴により離心率がさらに大きく変化し，レゾナンスから出てしまう。
  - 5:6 レゾナンスより太陽から遠い所にある 4:5 レゾナンスに入る天体も多い。
- 3:4 レゾナンス
  - 36 AU 付近の離心率の小さな天体は，ν_8 永年共鳴により時間平均した離心率が大きくなる。

- 3:4 レゾナンスより太陽から離れた所にある 5:7 レゾナンスに入る天体も多い。

- 3:5 レゾナンス
  - 時間平均した離心率の最大値は，他のレゾナンスの結果から推測される値に比べて小さい。これは 3:5 レゾナンスが他のレゾナンスよりも次数が高く，共鳴関係が弱いことによると考えられる。
  - ～ 42 AU の領域では，ν_8 永年共鳴により，時間平均した離心率が他の領域に比べて大きくなる。

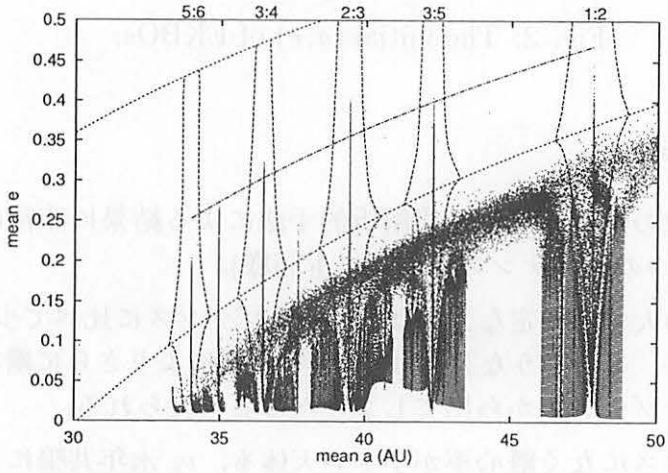


Fig. 3: The mean values of (a, e) for 5 Myr.

全体を通して明らかになったことは，次のようである:

- ν_8 永年共鳴がない場合，レゾナンスにある天体が持つ最大の離心率は天王星軌道との距離により決まる。
- ν_8 永年共鳴は，各レゾナンスの局所的な構造を決める大きな役割をしている。
- *Model A* により求めたレゾナンス領域のはるか外側から，レゾナンスに飛び込む天体が見られる。この現象は，天王星の存在による。

References

Fuse,T., Kinoshita,H., and Nakai,H., (1998) 第 30 回天体力学研究会集録, 255-259.

Yoshikawa,M., (1989) *Astron.Astrophys.* **213**, 436-458.

Yoshikawa,M., (1990) *Icarus* **87**, 78-102.

Yoshikawa,M., (1991) *Icarus* **92**, 94-117.

カイパーベルトにおける共鳴と軌道の安定性
The relation between Neptune resonance and stability of the orbit in the Kuiper belt

中井宏、木下宙（国立天文台）
H. Nakai and H. Kinoshita
National Astronomical Observatory

Abstract

We investigate the relation between the regions of stable orbits and the locations of the Kozai resonance and ν_{18} resonance near the mean motion resonance with Neptune in the Kuiper belt. We study these resonances by a semi-analytical method and a numerical integration. Then we get the following results.

- In the 1:2 mean motion resonance with Neptune, the regions of the mean motion resonance and the Kozai resonance do not overlap each other.
- In the 2:3 mean motion resonance with Neptune, the stable orbits occupy large area in the eccentricity-inclination plane. The region of the mean motion resonance and Kozai resonance overlap each other, and the region of the secondary resonance appears. But the region of Kozai resonance and ν_{18} resonance do not overlap each other.
- In each 4:5 and 3:4 mean motion resonance, the stable orbits occupy small area in the eccentricity-inclination plane. The region of the Kozai resonance and ν_{18} resonance overlap each other, and some orbits are in the multiple resonance.
- Each single resonance with Neptune has the effect of increasing the least approach distance between the Kuiper belt object and Neptune.
- In the multiple resonance, the distance between the Kuiper belt object and Neptune is larger than that in a single resonance, consequently the multiple resonance acts as the mechanism of stability.

1. はじめに

カイパーベルト領域には約130個の小天体が見つかっている（1999年3月）。観測が少なく、軌道決定に離心率や平均運動共鳴の条件を仮定することもあり、不確実な軌道要素であるが、明らかに小天体は一様に分布していない。メインベルトにおける共鳴構造と同じように、カイパーベルトにおいても何らかの共鳴構造が考えられる。

半解析的手法を用いて海王星からの距離（$n:n+1$平均運動共鳴の領域）によって、離心率ー軌道傾斜角空間で、古在共鳴、ν_{18}共鳴になる領域を求め、共鳴現象によってカイパーベルト天体と海王星との最接近距離がどう変化するかを調べた。また、平均運動共鳴領域に離心率、軌道傾斜角が異なる多数のテスト天体を配置し、１億年の数値シミュレーションを行い、半解析手法の結果と比較し、安定な軌道と共鳴の関係を調べた。

２．海王星との最接近距離

カイパーベルト天体と海王星の最接近距離は軌道要素や共鳴の状態によって変化する。図１は

様々な共鳴によって最接近距離がどのように変化するかを図示したもので、図 1-a,b はそれぞれ、軌道長半径が35.0 AU（4:5 平均運動共鳴付近）、47.8 AU（1:2 平均運動共鳴付近）の場合である。

2-1　両天体が平均運動共鳴の関係にない場合

両天体が同一平面上にある場合の両天体の最接近距離はカイパーベルト天体の離心率と共に直線的に減少する（図 1-a,b 実線）。

2-2　両天体が平均運動共鳴の関係にある場合

カイパーベルト天体と海王星の平均経度をそれぞれ、λ、λ_N、近日点経度を ϖ とすると、臨界引数は $\sigma = (n+1)\lambda - n\lambda_N - \varpi$、($n$: 整数) と定義され、臨界引数が常に180度となるように λ、λ_N を決め、両天体の位置から、最接近距離を求めたものが図 1-a,b の点線である。最接近距離は離心率が小さい範囲では離心率と共に増加するが、離心率が大きくなるとカイパーベルト天体の近日点が海王星軌道の内側に入り、軌道が交差するために、逆に最接近距離は減少する。

2-3　同時に平均運動共鳴と古在共鳴の関係にある場合

カイパーベルト天体の軌道傾斜角を40度、近日点引数を90度に固定すると、古在共鳴の効果により、両天体の最接近距離は図 1-a,b 短鎖線のように変化する。最接近距離が大きくなるのは古在共鳴によって、遠日点が海王星の軌道面から離れた所に固定されるためである。

2-4　同時に平均運動共鳴、古在共鳴、ν_{18}共鳴の関係にある場合

上記の状態で海王星の軌道傾斜角を5度（ν_{18}共鳴の効果を大きくするため）とし、両天体の昇交点経度の差を180度に固定したときの最接近距離は図 1-a,b の長鎖線となる。ν_{18}共鳴の効果はカイパーベルト天体の軌道傾斜角が変化したことと同じような働きをする。

一つの軌道で平均運動共鳴、古在共鳴、ν_{18}共鳴等、同時に複数の共鳴状態になると、海王星との最接近距離は非常に大きくなり、両天体は大接近しなくなる。しかし、軌道要素が一定でなく、秤動の振幅が0でないために、実際の小天体と海王星の最接近距離は図の値より小さくなる。

3．古在共鳴、ν_{18}共鳴

近日点引数が90度または270度の回りを秤動する状態を古在共鳴といい、小天体と海王星の軌道面が同じ周期で回転する状態、すなわち、小天体と海王星の昇交点経度の差が秤動する状態をν_{18}共鳴という。

半解析的手法（Kozai,1983）によって、古在共鳴の領域を求める。カイパーベルト天体に及ぼす惑星の摂動は $R = m_N/\Delta$ で表される。但し、m_Nは摂動天体（海王星）の質量、Δは相互距離である。カイパーベルト天体の日心距離、近日点引数、真近点離角をそれぞれ、r、g、fとする。摂動天体の軌道面に対するカイパーベルト天体の軌道傾斜角をiとし、昇交点から測った摂動天体の平均経度を λ_Nとすると

$$\Delta^2 = r^2 + r_N^2 - 2rr_N \cos S,$$

$$\cos S = \cos^2(i/2)\cos(f+g-\lambda_N) + \sin^2(i/2)\cos(f+g+\lambda_N)$$

となる。$\Theta = \sqrt{1-e^2}\cos i$ を一定とし、数値的にRを求め、平均操作により、短周期項を消去すると、系の自由度は1になり、等ポテンシャルカーブから離心率と近日点引数の関係が分かる。

図 2-a は軌道長半径 39.4 AU (2:3 平均運動共鳴)、離心率 0.32、軌道傾斜角 42 度（$\Theta = 0.704$）のテスト天体の半解析的手法で求めた等ポテンシャルカーブである。点は上記の値を初期値とし、数値積分によって求めた離心率と近日点引数の値である。太線は近日点引数が 90 度の回りを秤動する限界で、近日点引数が 90 度のとき、離心率が $e_{min} \le e \le e_{max}$を満足すれば、古在共鳴の状態になる。すなわち、一つのΘが決まれば、軌道が古在共鳴になる離心率の範囲$e_{min} \le e \le e_{max}$

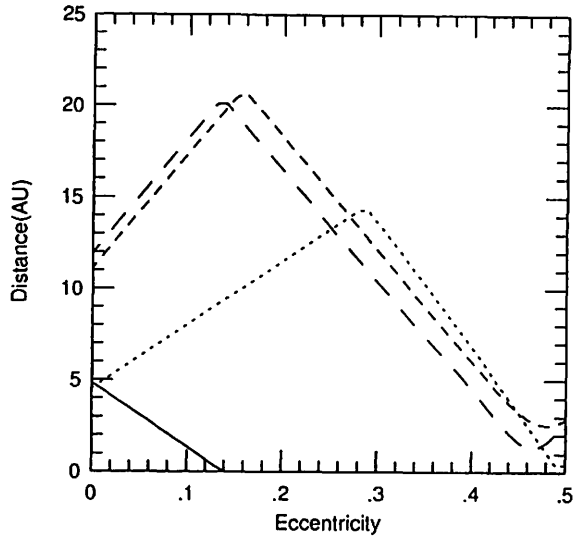


Fig. 1-a. a=35.0AU (4:5 resonance)

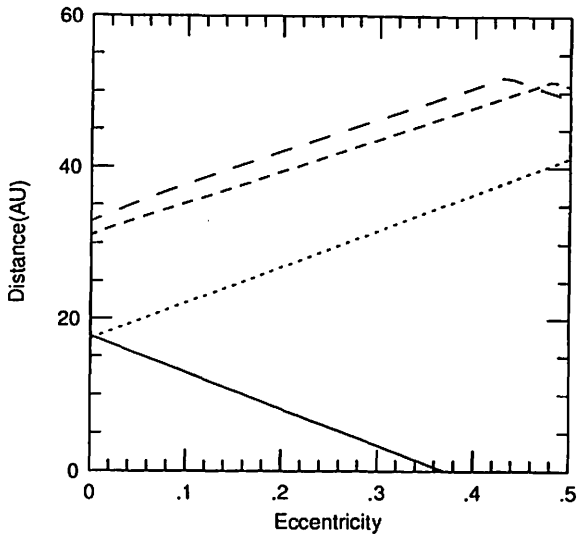


Fig. 1-b. a=47.8AU (1:2 resonance)

Figure 1. The least approach distance between Neptune and the Kuiper belt object. The solid line, the dotted line, the short dash line and the long dash line show non-resonant case, mean motion resonant case, double resonant (the mean motion and the Kozai resonant) case, triple resonant (the mean motion, the Kozai resonant, and the ν_{18} resonant) case, respectively.

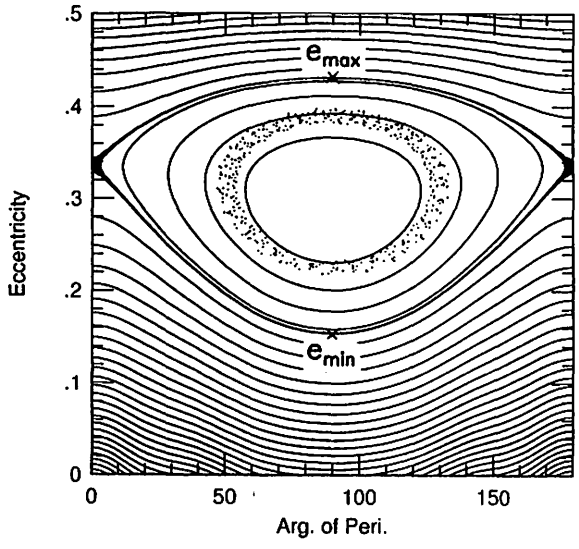


Figure 2-a. Equi-potential value curves for 2:3 mean motion resonance with Neptune for $\Theta = 0.704$. The thick lines denote the border of the Kozai resonance.

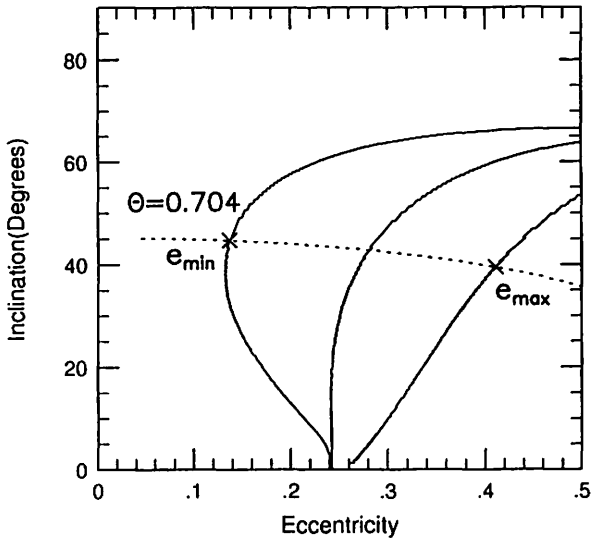


Figure 2-b. The picture shows the Kozai resonance. The center line denoting its location and the side lines its width. The dotted line shows the case of $\Theta = 0.704$.

が求まる。図 2-b は軌道長半径（$a = 39.4$ AU）を固定し、Θを変化させながら古在共鳴になる領域を離心率ー軌道傾斜角面に求めたものである。海王星との平均運動共鳴に相当する軌道長半径で、古在共鳴になる領域を求めれば、その領域は平均運動共鳴と古在共鳴両方の共鳴状態の領域になる。

同様な方法で、カイパーベルト天体と海王星の昇交点経度の差と軌道傾斜角の関係も等ポテンシャルカーブから求めることが出来、離心率ー軌道傾斜角面で、ν_{18}共鳴になる領域を知ることが出来る（Nakai and Kinoshita,1985)。

4．数値シミュレーション

カイパーベルト領域の平均運動共鳴付近に離心率（$0.0 \leq e \leq 0.5$）と軌道傾斜角（$0 \leq i \leq 90°$）が異なる多数のテスト天体を配置し、1.1億年の数値積分を行った。摂動天体は木星、土星、天王星、海王星の４天体で、積分法は対称型線形多段法（Quinlan and Tremaine,1990）を用いた。カイパーベルト天体の質量を 0 とし、離心率と軌道傾斜角以外のテスト天体の初期値は $\omega = 90°$、$\Omega = \Omega_N + 180°$、$l = ((\lambda_N - \varpi)n + 180° + 360°m)/(n+1)$ とした。但し、ω は近日点引数、Ω は昇交点経度、l は平均近点離角、λ は平均経度、m、nは整数、添え字Nは海王星を表す。数値積分で求まった $4:5, 3:4, 2:3, 1:2$ 平均運動共鳴付近の軌道の特徴をそれぞれ、図 3-a,b,c,d に示す。積分期間中、離心率や軌道傾斜角は一定ではないので、初期の離心率と軌道傾斜角の値の位置に、軌道の特徴を一連の色と記号で示している。黒色は 1.1億年間の軌道の状態を表し、黒丸は軌道が安定、黒 □ は短期間ではあるが軌道が乱れ、黒 × は軌道が不安定であることを表す。緑、赤、青緑、赤紫、黄の各色は共鳴の種類を表し、それぞれ、平均運動共鳴、古在共鳴、ν_{18}共鳴、ν_8共鳴、第２次共鳴に対応している。記号は 1.1億年間の共鳴の状態を表し、丸は安定な共鳴状態、□ は短期間共鳴状態が崩れ、× は共鳴状態にないことを表す。例えば、緑丸は 1.1億年間平均運動共鳴状態にあり、青緑の □ は ν_{18}共鳴状態が短期間崩れたことを、緑と赤の× は、それぞれ、平均運動共鳴、古在共鳴でないことを表す。赤色の、⊘、⊘、⊗は、それぞれ、近日点引数が順行、逆行、および、順行逆行を繰り返す事を表し、⊡、⊡、⊠ は近日点引数の順行、逆行、および、順行逆行の繰り返しが短期間ではあるが崩れた事を表す。赤色の ■ は近日点引数が0° または180° の回りを秤動することを表し、青緑色の ■ はカイパーベルト天体と海王星の昇交点経度の差が 0° の回りを秤動することを表す。

5．軌道の安定性と共鳴

5-1　1:2 平均運動共鳴付近

離心率が 0.1 以上で、軌道傾斜角 20 度以下の安定な軌道は全て平均運動共鳴の状態にあり、近日点引数は順行している。古在共鳴の領域は軌道傾斜角が 50 度以上で現れ、この領域では全ての軌道が平均運動共鳴の状態ではない。軌道傾斜角が 0 度付近では、部分的にν_{18}共鳴になる軌道があり、また、離心率が 0 付近では ν_8共鳴になる軌道もある。海王星から離れているため、離心率が小さければ、平均運動共鳴か古在共鳴であれば、海王星との最接近距離は 10 AU程度以上になり、軌道は乱れない。半解析的手法では、古在共鳴領域を求めるために、臨界引数は 180 度の回り 90 度の振幅で秤動すると仮定して短周期成分を平均して消去している。しかし、実際の数値積分の結果によると、古在共鳴の領域では、平均運動共鳴が崩れ、臨界引数の振幅は 180 度を超えている。臨界引数の振幅が大きい場合の古在共鳴、ν_{18}共鳴、ν_8共鳴領域については検討する必要がある。また、臨界引数は振幅が小さい場合は90度や 270 度の回りを秤動し、振幅が大きくなる

と 180度の回りの秤動に変わる。このような変化は、他の平均運動共鳴では見られない、1:2平均運動共鳴の特徴である。

5-2　2:3平均運動共鳴付近

離心率ー軌道傾斜角（$e-i$）面で軌道が安定な領域は他の平均運動共鳴付近に比べ非常に広い。特に、軌道傾斜角が 60度付近では離心率が 0.5 でも軌道は安定である。安定な軌道はほとんどの場合平均運動共鳴にあり、軌道傾斜角が小さく離心率が大きい範囲では同時に古在共鳴の状態になっている。e が 0.3、iが 20度では近日点引数がカイパーベルト天体と海王星の昇交点経度の差の周期に一致する第２次共鳴の状態になっている。（冥王星も第２次共鳴の状態になっている。）ν_{18}共鳴は軌道傾斜角が小さい軌道で現れるが、古在共鳴の領域とは別の場所である。半解析的手法で求めた共鳴領域と数値積分の結果は良く一致している。

5-3　3:4平均運動共鳴付近

離心率と軌道傾斜角の小さい領域で安定な軌道がある。また、離心率が 0.2 で軌道傾斜角が10－30度の場所で古在共鳴、第２次共鳴になる軌道がある。その場所は半解析的手法による古在共鳴の離心率の中心の場所に相当する。小さい離心率で軌道傾斜角が 10－20度付近にν_{18}共鳴の領域があるが、古在共鳴とは重ならない。

5-4　4:5平均運動共鳴付近

ν_{18}共鳴の中心の軌道傾斜角が大きくなり、また、その幅も大きくなる。古在共鳴の中心の離心率は小さい方に移動し、その幅も小さくなる。平均運動共鳴、古在共鳴、ν_{18}共鳴の領域が重なり、３種類の共鳴状態になる軌道が現れる。安定な軌道の領域は狭くなり、複数の共鳴状態になる割合は大きくなる。

5:6平均運動共鳴では安定な軌道になる領域は非常に狭くなり、海王星に近づくと安定な軌道は存在出来なくなる。

6．　まとめ

海王星に近い平均運動共鳴になるにつれて、古在共鳴の中心の離心率は小さくなり、古在共鳴の範囲も小さくなる。逆に ν_{18}共鳴の中心の軌道傾斜角は大きくなりν_{18}共鳴の範囲も大きくなる。海王星に近い平均運動共鳴付近では一つの軌道で同時に平均運動共鳴、古在共鳴、ν_{18}共鳴の状態となる軌道の割合が高くなる。数値積分による結果は1:2平均運動共鳴付近以外では半解析的手法の結果と定性的に良く一致する。海王星により近い平均運動共鳴では、軌道が安定な領域が狭くなり、また、隣接する平均運動共鳴領域が接近して存在するために、9:10平均運動共鳴（32.3 AU）のように海王星に非常に近い場合は1.1億年間安定な軌道は存在できない。

共鳴の周縁域では軌道要素が大きく変化するため、軌道はカオス的になる。特に、古在共鳴が崩れる最大離心率（図 2-a,b 中の e_{max}）以上では殆どの軌道は不安定である。

共鳴の中心付近ではカイパーベルト小天体と海王星の間の最接近距離が大きくなり軌道の安定性が保たれている。共鳴現象はカイパーベルト領域の天体の軌道安定化機構として働き、共鳴構造の形成に重要な役割を果たしている。

7．　おわりに

今回は $n:n+1$ 平均運動共鳴（$n=1,2,3,4\cdots$）という特殊な場合の検討であった。半解析手法によって　ν_8共鳴の領域を調べ、古在共鳴、ν_{18}共鳴、ν_8共鳴等について、カイパーベルト全域での共鳴構造を明らかにする。そして、メインベルト領域における構造との類似点、相違点を明

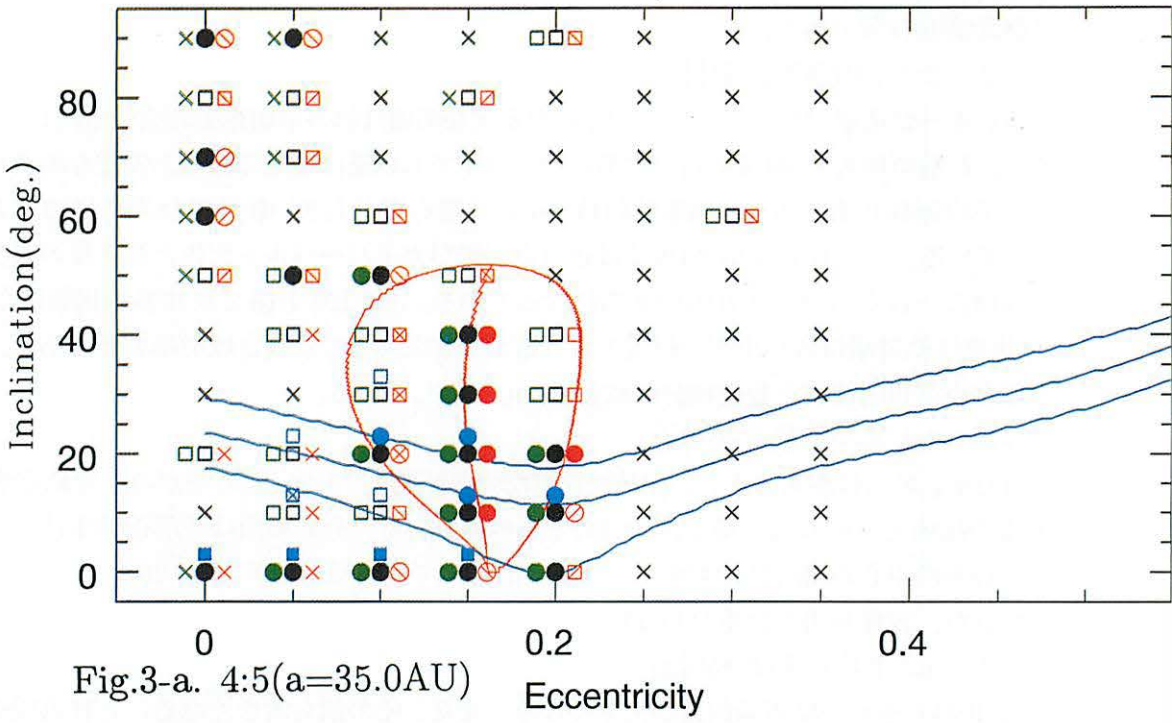


Fig.3-a. 4:5(a=35.0AU)

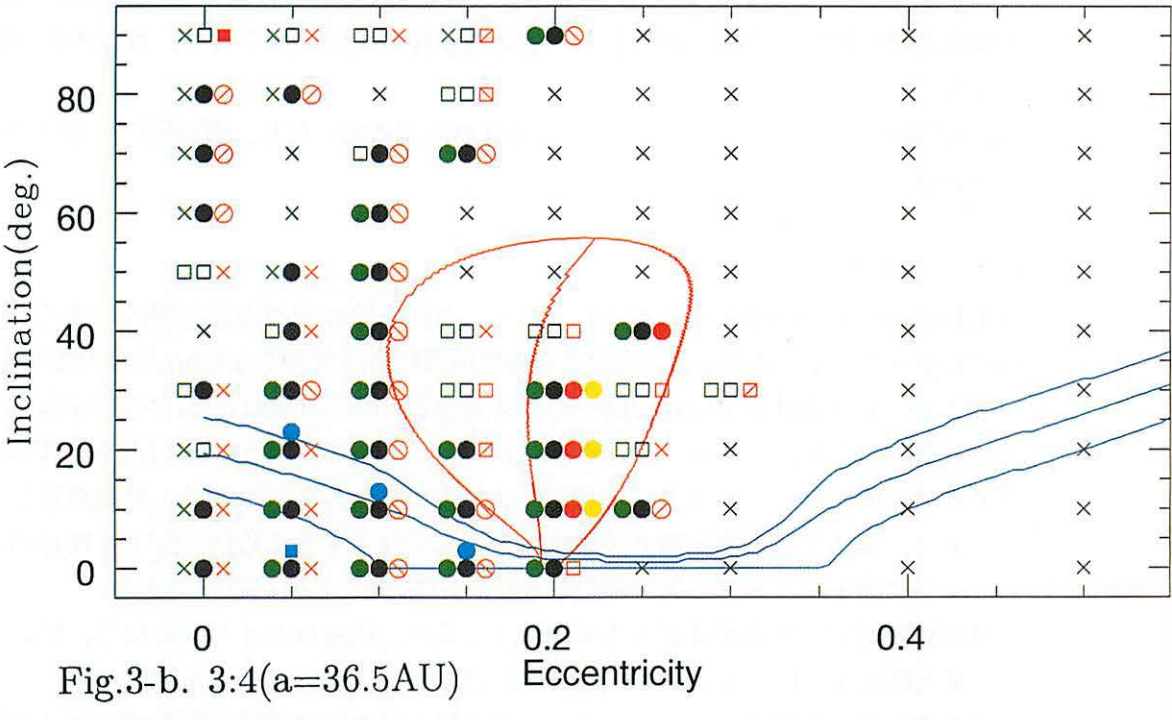


Fig.3-b. 3:4(a=36.5AU)

Figure 3. The phase space structures near mean motion resonances with Neptune. The colors and symbols in figures denote the features of orbit from numerical integration of 110 million years. The black filled circles, black open squares and black crosses indicate the stable orbits, unstable orbits for a short time and unstable orbits, respectively. Each color, that is, green, red, cyan, magenta and yellow is associated with the mean motion resonance, the Kozai resonance, the ν_{18} resonance, the ν_8 resonance, the secondary resonance, respectively. The filled circles show the orbits which stay in the resonance for 110 million years, the open squares denote the orbits which are temporarily captured into the resonance, and the crosses denote the orbits

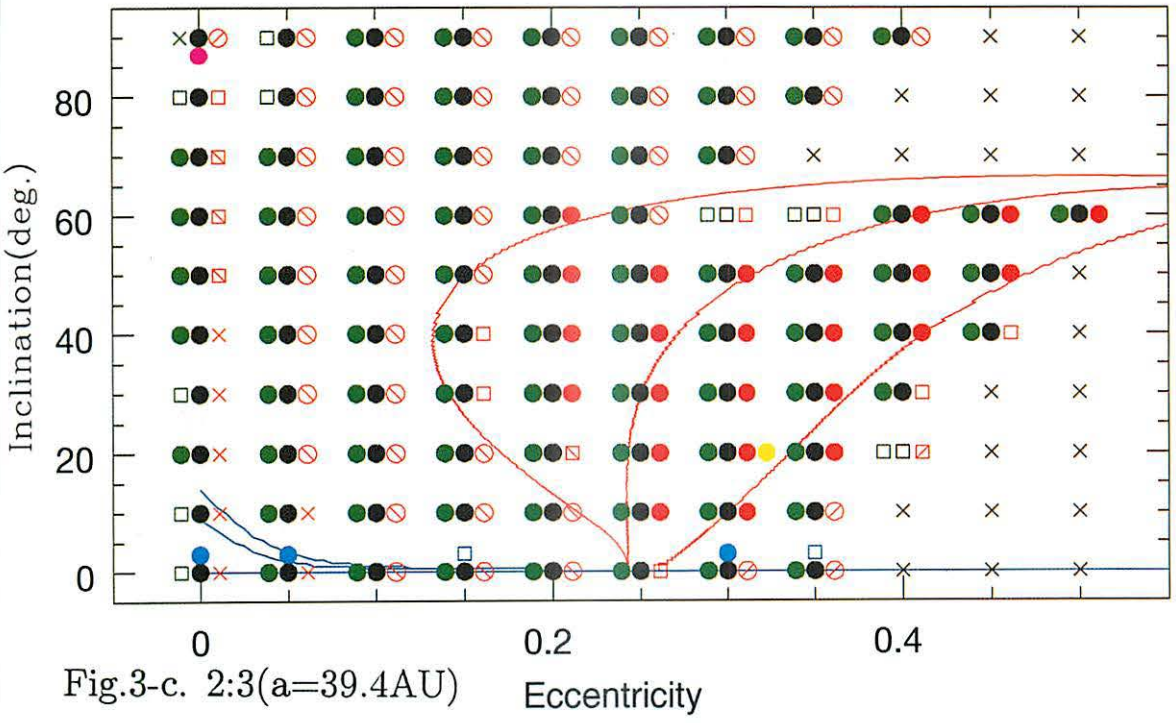


Fig.3-c. 2:3(a=39.4AU)

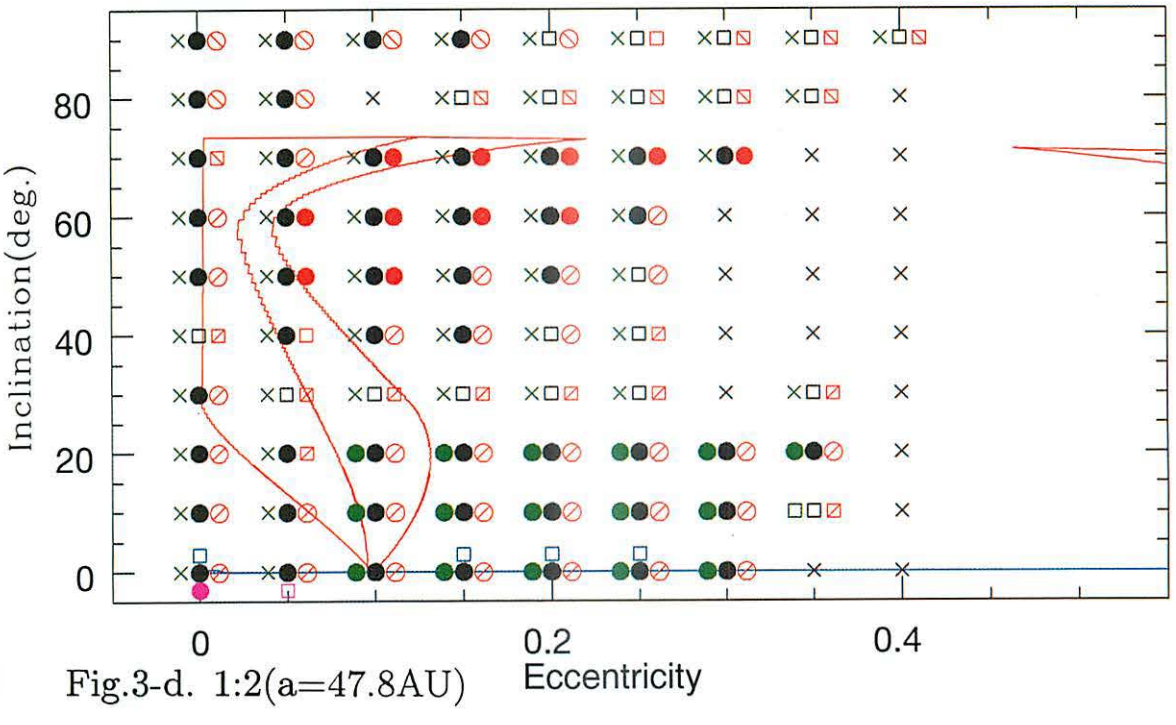


Fig.3-d. 1:2(a=47.8AU)

Figure 3.(continued)

which are not in the resonant state. For example, the filled green circles indicate the orbits stay in the state of the mean motion resonance with Neptune for 110 million years. The red symbols($\oslash, \otimes$) represent the motions of the argument of perihelion are direct and retrograde, respectively. The filled red square symbol represents ω librates around 0° or 180°, and the filled cyan squares represent $\Omega - \Omega_N$ librates around 0°. The red lines and cyan lines, respectively, show the locations of the Kozai resonance and the ν_{18} resonance by a method of semi-analytical secular perturbations.

らかにすることが今後の課題である。また、1:2平均運動共鳴で数値積分と半解析的手法で求めた古在共鳴の領域が異なっていた。この理由についても検討する必要がある。

8. 参考文献

Duncan M.J., Levison H.F. and Budd S.M. : 1995, The Dynamical Structure of the Kuiper Belt, *Astron. J.*,**110**, pp.3073-3081.

Knežević Z., Milani A., Farinella P., Froeschle Ch., and Froeschle Cl. : 1991, Secular Resonances from 2 to 50 AU, *ICARUS*,**93**, pp.316-330.

Kozai Y. : 1962, Secular Perturbations of Asteroids with High Inclination and Eccentricity, *Astron. J.*,**67**, pp.591-598.

Kozai Y. : 1983, Secular Perturbations of Asteroids with Commensurable Mean Motions, Celestial Mechanics of the International Symposium G. LEMAITRE, Louvain-la-Neuve, Belgique, pp.1-20.

Malhotra R. : 1996, The Phase Space Structure Near Neptune Resonances in the Kuiper Belt, *Astron. J.*,**111**, pp.504-516.

Morbidelli A., Thomas F. and Moons M. : 1995, The Resonant Structure of the Kuiper Belt and the Dynamics of the First Five Trans-Neptunian Objects, *Icarus*, **118**, pp.322-340.

MPEC : Minor Planet Electronic Circulars, Minor Planet Center, Smithsonian Astrophysical Observatory.

Nakai H. and Kinoshita H. : 1985, Secular Perturbations of Asteroids in Secular Resonance, *Celest. Mech.*,**36**, pp.391-407.

Nakai H. and Kinoshita H. : 1996, Orbits of Kuiper Belt Objects, Proceedings of the 28th Symposium on Celestial Mechanics in Tokyo, pp.143-152(in Japanese).

Nakai H. and Kinoshita H. : 1997, Kozai Resonance and ν_{18} Secular Resonance in the 2:3 Mean motion Resonance in Kuiper Belt Region, Proceedings of the 29th Symposium on Celestial Mechanics in Tokyo, pp.279-289.

Quinlan G D, and Tremaine S. : 1990, Symmetric Multistep Methods for the Numerical Integration of Planetary Orbits, *Astron. J.*,**100**, pp.1694-1700.

Analytical Solution of the Kozai Resonance and its Application

Hiroshi Kinoshita and Hiroshi Nakai
National Astronomical Observatory
2-21-1 Osawa, Mitaka, Tokyo, Japan
e-mail: Kinoshita@nao.ac.jp

Abstract

When Kozai(1962) studied the secular resonance of asteroids, he found the so-called Kozai resonance and expressed the analytical solution with use of Weierstrass $\wp$. Here we give the analytical solution of the eccentricity, the inclination and the argument of pericenter with use of the Jacobi elliptic functions, which are more familiar than the Weierstrass $\wp$. Then we derive the Fourier expansion of the longitude of node and the mean anomaly. The analytical expressions obtained here can be used for any value of the eccentricity and the inclination. Finally we applied these analytical expressions to several dynamical systems: Nereid, that is a highly eccentric satellite of Neptune, and newly discovered retrograde satellites of Uranus.

Key words: Kozai resonance - secular perturbation - Nereid

1 Introduction

When Kozai(1962) investigated the secular resonance of asteroids with high inclination and eccentricity, he found the argument of the perihelion of an asteroid does librate and even if the initial eccentricity is very small, the eccentricity does become very large. This phenomenon is recently called the Kozai mechanism or the Kozai resonance, since the mean motion of the perihelion longitude is equal to that of the longitude of node. The Kozai mechanism appears in the similar dynamical systems as the asteroidal case. Kinoshita and Nakai(1991) investigated the fictitious satellites located in the region where the solar perturbation is dominant and found that they become highly eccentric and have a possibility to collide with Uranus due to the Kozai mechanism. Wiegert and Holman(1997) investigated the long-term orbital stability of small bodies near the central binary of the Alpha Centauri system and found the region whose stability is determined by the Kozai mechanism. Holman, Touma, and Tremaine (1997) discussed the planet recently discovered orbiting the star 16 Cyg B, whose eccentricity is 0.67 and proposed its high eccentricity is caused by the Kozai mechanism from the distant companion star, 16 Cyg A. Innanen et al.(1997) explored planetary orbits in binary systems, where the planetary orbit has high inclination relative to the orbital plane of the binary and found the Kozai mechanism is suppressed when the mutual perturbation among the planets are strong enough compared with the tidal force due to the companion star. Thomas and Morbidelli(1996) discussed the Kozai mechanism, where the disturbing body is inside of the disturbed body.

Kozai(1962) obtained the analytical solution of the Kozai resonance by an elliptic function of Weierstrass $\wp$. Since Weierstrass $\wp$ is not familiar, Kinoshita and Nakai (1991) expressed the solution with use of Jacobi elliptic functions. The analytical solution in the

paper mentioned above, however, turned out to be only applicable to the case where the argument of the perihelion circulates. In section 2 we derive the equations of motions from the averaged Hamiltonian and in section 3 we give the analytical solutions for the case where the argument of perihelion circulates and the case where it librates. In section 4 avoiding the use of third kind of elliptic integrals we give the Fourier expansions of the longitude of node and the mean anomaly. In section 5 we applied the analytic solutions to the newly discovered satellites of Uranus and Nereid, which is a satellite of Neptune and whose eccentricity is 0.67 at present. We numerically integrated the orbit of Nereid and compared our analytic solutions with the numerical integrated orbit and the analytic solution for Nereid by Mignard(1975)

2 Equations of Motion

In this paper we discuss only the system where the disturbing body is orbiting outside of the disturbed body. Kozai(1962) discussed a massless asteroid disturbed by the outer giant planets. The Hamiltonian for this system in terms of Delaunay elements $(L, G, H, l, \omega, \Omega)$ is

$$F = F(L, G, H, l, \omega, \Omega, \lambda'), \tag{1}$$

where λ' is the longitude of the disturbing planet. The explicit expression of F is given in Kozai(1962). After the elimination of short periodic terms the new Hamiltonian F^* does not include l and λ' and also h since the disturbing potential becomes axial symmetric. Since the new L and H (the z-component angular momentum) are constant, we have the following relation between the eccentricity and the inclination

$$\frac{H}{L} = \sqrt{1-e^2}\cos I = \text{constant} = \sqrt{h}. \tag{2}$$

Now we give the explicit expression the new Hamiltonian(Kozai 1962):

$$\begin{aligned} F^* &= \gamma\Big((2+3e^2)(3\cos^2 I - 1) + 15e^2 \sin^2 I \cos 2\omega\Big) \\ &\quad + O((a/a_d)^2), \end{aligned} \tag{3}$$

$$\gamma = \frac{1}{16}\frac{m_d}{m_d + m_c}\frac{n_d^2}{(1-e_d^2)^{3/2}}a^2 \tag{4}$$

where m_d, n_d, a_d, and e_d are the mass, the mean motion, the semi-major axis, and the eccentricity of a disturbing body and m_c is the mass of the central body. For the asteroidal case the central body is the Sun and the disturbing bodies are the outer planets. Since the number of freedom of the system after the elimination of short periodic terms is one, we can draw equi-F^* curves in $e-\omega$ plane and can know from the $e-\omega$ diagram the global behaivor of the system without solving the differentail equations (see the figures 2-5 of Kozai's paper (1962)). The topology of the $e-\omega$ diagram depends on the value of h. When $h > 0.6$, the argument of pericenter circulates and the librational region does not exist. When $h < 0.6$, the region, where the argument of pericenter librates, does appear (see the figures 2-5 of Kozai's paper (1962)).

From the Hamiltonian (3) the equations of motion take the form

$$\frac{d\eta}{dt^*} = -\frac{15}{8}e^2 \sin^2 I \sin 2\omega \tag{5}$$

$$\frac{d\omega}{dt^*} = -\frac{3}{8\eta}\left(\eta^2 - 5\cos^2 I + 5(\cos^2 I - \eta^2)\cos 2\omega\right) \tag{6}$$

with the constraint $\eta \cos I = \text{constant} = \sqrt{h}$, where

$$\eta = \sqrt{1-e^2}, \tag{7}$$

$$t^* = \gamma^* t, \gamma^* = \frac{m_d}{m_d + m_c}(1-e_d^2)^{-3/2}\frac{n_d^2}{n}. \tag{8}$$

Even though these equations do not seem to be integrable, the solution can be obtained with use of an elliptic function. Kozai(1962) expressed the solution by an elliptic function of Weierstrass $\wp$. Kinoshita and Nakai(1991) gave the analytical expression with use of Jacobian elliptic functions, which are more familiar and easier in handling than Weierstrassian elliptic functions. This analytical expression, however, turned out to be only applicable to the case where the argument of the perihelion circulates. So here we give the analytical solution for both cases, the librational case and the circulational case.

3 Analytical Solution of e, I, and ω

3.1 Circulational case of ω

First we discuss the case where the argument of pericenter circulates. The energy integral(3) is

$$(5-3x)(3h/x-1) + 15(1-x)(1-h/x)\cos 2\omega = C, \tag{9}$$

where $x = \eta^2$. We define $x_0 = 1 - e_0^2$ and $I = I_0$ at the argument of perihelion $\omega = 0$ and e_0 is the minimum value for the orbit with this initial value. The energy integral evaluated at $\omega = 0, x = x_0$ is

$$C = 10 - 12x_0 + 6h. \tag{10}$$

From (9) and (10), we have

$$\cos 2\omega = \frac{Q(x)}{5(1-x)(x-h)}, \tag{11}$$

where

$$Q(x) = -x^2 + (5(1+h) - 4x_0)x - 5h. \tag{12}$$

From (12) we have

$$\sin^2\omega = \frac{2(x_0 - x)x}{5(1-x)(x-h)}, \tag{13}$$

$$\cos^2\omega = \frac{y}{5(1-x)(x-h)}, \tag{14}$$

where

$$y = -3x^2 + (5 + 5h - 2x_0)x - 5h. \tag{15}$$

The equation of motion of x from (5) is

$$\frac{dx}{dt^*} = 2\eta\frac{d\eta}{dt} = -\frac{15}{4}\eta e^2 \sin^2 I \sin 2\omega = -\frac{15}{2}x^{-\frac{1}{2}}(1-x)(x-h)\sin\omega\cos\omega. \tag{16}$$

Substituting (13) and (14) into (16),we have

$$\frac{dx}{dt^*} = -\frac{3}{2}\sqrt{2(x_0 - x)y}. \tag{17}$$

We define $x_1, x_2 (x_1 < x_2)$ of the sotution of the quadratic equation $y = 0$:

$$x_1 + x_2 = \frac{1}{3}(5 + 5h - 2x_0), \tag{18}$$

$$x_1 x_2 = \frac{5}{3}h. \tag{19}$$

The differential equation for x becomes

$$\frac{dx}{dt^*} = -\frac{3\sqrt{6}}{2}\sqrt{(x - x_1)(x_0 - x)(x_2 - x)}. \tag{20}$$

Since the order of x inside of the square root is three, the solution of (20) is expressed in terms of Jacobian elliptic function:

$$x = x_0 + (x_1 - x_0)\mathrm{cn}^2\theta, \tag{21}$$

where

$$\theta = \frac{2K}{\pi}(\omega^* + \frac{\pi}{2}), \tag{22}$$

$$\omega^* = n_{\omega^*} t, \tag{23}$$

$$n_{\omega^*} = \frac{3\sqrt{6}\pi}{8K}\sqrt{x_2 - x_1}\gamma^*. \tag{24}$$

In the above expressions K is the complete elliptic integral of the first kind with the modulus k^2

$$k^2 = \frac{x_0 - x_1}{x_2 - x_1}. \tag{25}$$

In (22) ω^* is the angle variable corresponding to the argument of perihelion and n_{ω^*} is the mean motion of ω^*. The maximum eccentricity is obtained from (21) as

$$e_{\max} = \sqrt{1 - x_1} \tag{26}$$

The inclination corresponding to $e_{\max}$ is

$$I_{\min} = \cos^{-1}\sqrt{h/(1 - e_{\max}^2)} \text{ for } I_0 < 90° \tag{27}$$

or

$$I_{\max} = \cos^{-1}\sqrt{h/(1 - e_{\max}^2)} \text{ for } I_0 > 90° \tag{28}$$

3.2 Librational case of ω

In the libration region the eccentricity has two values for ω. Here we define the smaller eccentricity as $e_0 (x_0 = 1 - e_0^2)$ at $\omega = \pi/2, t = 0$ for the initial condition of the equations of motion. The energy integral at this point is

$$C = 30h/x_0 - 20 + 18x_0 - 24h. \tag{29}$$

From (9) and (29) we have

$$\cos 2\omega = \frac{Q(x)}{5(1-x)(x-h)}, \tag{30}$$

where

$$Q(x) = -x^2 + (-5(1+h) + 10h/x_0 + 6x_0)x - 5h. \tag{31}$$

From (30) we have

$$\sin^2\omega = \frac{2(x_2-x)x}{5(1-x)(x-h)}, \tag{32}$$

$$\cos^2\omega = \frac{3(x_0-x)(x-x_1)}{5(1-x)(x-h)} \tag{33}$$

where

$$x_1 = \frac{5h}{3x_0}, \tag{34}$$

$$x_2 = \frac{5}{2} + \frac{5}{2}h - \frac{3}{2}x_0 - \frac{3}{2}x_1 \tag{35}$$

Since $x_2 - x_0 = 5/2(1-x_0)(x_0-h)/x_0 > 0$,$x_2$ is larger than x_0. Finally the equation of motion for x becomes

$$\frac{dx}{dt^*} = -\frac{3\sqrt{6}}{2}\sqrt{(x-x_1)(x_0-x)(x_2-x)}. \tag{36}$$

Since the functional form of this equation (36) is the same as that of (20) for the circulational case except that the meaning of the constant x_1 and x_2 are different between the circulational and librational cases, the solution of the differential equation (36) is

$$x = x_0 + (x_1 - x_0)\mathrm{cn}^2\theta, \tag{37}$$

where

$$\theta = \frac{2K}{\pi}(\omega^* + \frac{\pi}{2}), \tag{38}$$

$$\omega^* = n_{\omega^*}t + \frac{\pi}{2}, \tag{39}$$

$$n_{\omega^*} = \frac{3\sqrt{6}}{8K}\sqrt{x_2 - x_1}\gamma^*. \tag{40}$$

4 Solution of Ω and l

The equation of the longitude of node is obtained from the Hamiltonian (3):

$$\frac{d\Omega}{dt} = \frac{3}{8}\gamma^*\frac{\sqrt{h}}{x}(3x - 5 + \frac{Q(x)}{x-h}). \tag{41}$$

Since γ does depend on one of Delaunay variable L (see (4)), the equation of the mean anomaly is

$$\frac{dl}{dt} = -\gamma\frac{\partial C}{\partial L} - \frac{\partial\gamma}{\partial L}C = -\frac{3}{8}\gamma^*(3h - x + \frac{Q(x)}{1-x}) - \frac{1}{4}C\gamma^*. \tag{42}$$

4.1 circulational case

By substituting (10) and (12) into (41) and (42),we have the equations of motion for the longitude of node and the mean anomaly:

$$\frac{d\Omega}{dt} = \frac{3}{4}\sqrt{h}\gamma^*\Big(1 - \frac{2(x_0 - h)}{x - h}\Big), \tag{43}$$

$$\frac{dl}{dt} = \frac{1}{4}\gamma^*\Big(-4 + 6x_0 - 3h - \frac{6(1 - x_0)}{1 - x}\Big). \tag{44}$$

Since the lefthand side of (43) and (44) are known functions of time, Ω and l are obtained simply by quadrature. The integrals, however, include third kind of elliptic integrals, which forms are not practically useful both in numerical evaluation and analytical treatment. Especially the important quantities, the mean motions of Ω and l do not explicitly appear in the forms. Therefore we give the Fourier expansions of Ω and l in the followings, whose derivation is given in the Appendix.

The longitude of node from Appendix 6.1 is

$$\Omega = \Omega^* + \sum_{m=1}^{\infty} b_m \sin 2m\omega^*, \tag{45}$$

$$\Omega^* = n_{\Omega^*} t + \Omega_0^*, \tag{46}$$

where Ω^* is the angle variable corresponding to the longitude of node and Ω_0 is one of the integral constants. The amplitude of periodic terms with arguemnts $2\omega^*$ is

$$b_m = -\frac{2(-q)^m}{m(1 - q^{2m})} \sinh m\frac{\pi c}{K}, \tag{47}$$

where

$$c = F(\xi, k'), \tag{48}$$

$$\sin\xi = \sqrt{\frac{x_2 - x_1}{x_2 - h}}, \tag{49}$$

$$k' = \sqrt{1 - k^2}. \tag{50}$$

In (48) $F(\xi, k')$ is the normal elliptic integral of the first kind and q in (47)is Jacob's nome, which is a function of k^2. The mean motion of the longitude of Ω is give by

$$n_{\Omega^*} = -\frac{3}{4}\sqrt{h}\gamma^*(-1 + 2\frac{x_0 - h}{x_2 - h}) - \Lambda_0(\xi.\, k) n_{\omega^*}, \tag{51}$$

where Λ_0 is Heuman's Lambda funtion:

$$\Lambda_0(\xi, k) = \frac{2}{\pi}(EF(\xi, k') + KE(\xi, k') - KF(\xi, k')). \tag{52}$$

Here E is the complete elliptic integral of the second kind and $F(\xi, k')$ is the normal elliptic integral of the second kind.

The Fourier expansion of the mean anomaly from Appendix 6.2 is given by

$$l = l^* + \sum_{m=1}^{\infty} f_m \sin 2m\omega^*, \tag{53}$$

$$l^* = n_{l^*} t + l_0^*. \tag{54}$$

The amplitude of a periodic term with argument $2m\omega$ is

$$f_m = -\frac{2q^m}{m(1-q^{2m})} \sinh m\frac{\pi c'}{K}, \tag{55}$$

where

$$c' = F(\zeta, k'), \tag{56}$$

$$\sin\zeta = \sqrt{\frac{(x_2-1)(1-x_0)}{(1-x_1)(x_2-x_0)}}. \tag{57}$$

The secular perturbatin in the mean anomaly is given by

$$n_{l^*} = \frac{1}{4}\gamma^*(-4+6x_0-3h) - \Lambda_0(\zeta, k')n_{\omega^*}. \tag{58}$$

When both e_0 and I_0 are small, the modulus k becomes small and when e_0 is large and h is small, the modulus k also becomes small. Here we give the analytical expressions for small k, which is useful for rough estimations of the mean motions of ω and Ω without using elliptic integrals.

The Fourier expansion of the argument of pericenter for small k is

$$b_m = -\sqrt{\frac{2h(x_0-h)}{3(x_1-h)(x_2-x_1)}}\frac{(-a)^m}{m} + O(k^2), \tag{59}$$

$$a = \frac{2}{\beta}(1-\frac{\beta}{2}-\sqrt{1-\beta}), \beta = \frac{x_0-x_1}{x_0-h}, \tag{60}$$

$$n_{\Omega^*} = -\frac{3}{4}\gamma^*\sqrt{h}(-1+2\sqrt{\frac{x_0-h}{x_1-h}}) + O(k^2). \tag{61}$$

The Fourier expansion of the mean anomaly for the case of small k is

$$f_m = -\sqrt{\frac{2(1-x_0)}{3(1-x_1)(x_2-x_1)}}\frac{a_1^m}{m} + O(k^2), \tag{62}$$

$$a_1 = \frac{2}{\beta_1}(1+\frac{\beta_1}{2}-\sqrt{1+\beta_1}), \tag{63}$$

$$\beta_1 = \frac{x_0-x_1}{1-x_0}. \tag{64}$$

$$n_{l^*} = \frac{1}{4}\gamma^*(-4+6x_0-3h-6\sqrt{\frac{1-x_0}{1-x_1}}) + O(k^2). \tag{65}$$

4.2 Librational case

By substituting (29) and (31) into (41) and (42), we have the equations of motion for the longitude of node and the mean anomaly for the librational case:

$$\frac{d\Omega}{dt} = \frac{3}{4}\sqrt{h}\gamma^*\Big(1-\frac{3(x_0-h)(x_1-h)}{h(x-h)}\Big), \tag{66}$$

$$\frac{dl}{dt} = \frac{1}{4}\gamma^*\Big(11 + 12h - 9x_0 - 9x_1 + \frac{9(1-x_0)(1-x_1)}{1-x}\Big). \tag{67}$$

In the derivation of (66) and (67), the relation (34) is used. Since the differences between (43) and (66) and between (44) and (67) are only the factor of $1/(x-h)$ and $1/(1-x)$,we can get the Fourier expansions of the longitude of node and the mean anomaly by the similar procedure as for the circulational case (section 4.1). In this section the same symbols as those in section 4.1 are used but the parameters x_1 and x_2 should be computed from (34) and (35) istead of (18) and (19) in the numerical evaluations .

The longitude of node is

$$\Omega = \Omega^* + \sum_{m=1}^{\infty} b_m \sin 2m\omega^*, \tag{68}$$

$$\Omega^* = n_{\Omega^*} t + \Omega_0^*. \tag{69}$$

The amplitude of a periodic term with argument $2m\omega^*$ is given by

$$b_m = -\frac{2(-q)^m}{m(1-q^{2m})} \sinh m\frac{\pi c}{K}, \tag{70}$$

$$c = F(\xi, k'), \tag{71}$$

$$\sin\xi = \sqrt{\frac{x_2 - x_1}{x_2 - h}}, \tag{72}$$

$$k' = \sqrt{1-k^2}. \tag{73}$$

The mean motion of the longitude of Ω^* is

$$n_{\Omega^*} = -\frac{3}{4}\sqrt{h}\gamma^*(-1 + 3\frac{(x_0-h)(x_1-h)}{h(x_2-h)}) - \Lambda_0(\xi,k)n_{\omega^*}. \tag{74}$$

The Fourier expansion of the mean anomaly is

$$l = l^* + \sum_{m=1}^{\infty} f_m \sin 2m\omega^*, \tag{75}$$

$$l^* = n_{l^*} t + l_0^*. \tag{76}$$

The periodic amplitude of argument $2m\omega^*$ is

$$f_m = \frac{2q^m}{m(1-q^{2m})} \sinh m\frac{\pi c'}{K}, \tag{77}$$

$$c' = F(\zeta, k'), \tag{78}$$

$$\sin\zeta = \sqrt{\frac{(x_2-1)(1-x_0)}{(1-x_1)(x_2-x_0)}}, \tag{79}$$

The secular perutrbation of the mean anomaly is

$$n_{l^*} = \frac{1}{4}\gamma^*(11 + 12h - 9x_0 - 9x_1) + \Lambda_0(\zeta,k')n_{\omega^*}. \tag{80}$$

The Fourier expansion of the longitude of node for small k is

$$b_m = -\sqrt{\frac{3(x_0-h)(x_1-h)}{2h(x_2-x_1)}}\frac{(-a)^m}{m} + O(k^2), \tag{81}$$

$$a = \frac{2}{\beta}(1-\frac{\beta}{2}-\sqrt{1-\beta}), \beta = \frac{x_0-x_1}{x_0-h}, \tag{82}$$

$$n_{\Omega^*} = -\frac{3}{4}\sqrt{h}\gamma^*(-1+3\sqrt{(x_0-h)(x_1-h)}) + O(k^2). \tag{83}$$

The Fourier expansion of the mean anomaly for the case of small k is

$$f_m = \sqrt{\frac{3(1-x_0)(1-x_1)}{2(x_2-x_1)}}\frac{a_1^m}{m} + O(k^2) \tag{84}$$

$$a_1 = \frac{2}{\beta_1}(1+\frac{\beta_1}{2}-\sqrt{1+\beta_1}), \tag{85}$$

$$\beta_1 = \frac{x_0-x_1}{1-x_0}. \tag{86}$$

$$n_{l^*} = \frac{1}{4}\gamma^*\Big(11+12h-9x_0-9x_1+9\sqrt{(1-x_0)(1-x_1)}\Big) + O(k^2) \tag{87}$$

5 Application

5.1 Retrograde Satellites of Uranus

The newly discovered retrograde satellites of Uranus (S/1997 U1 and S/1997 U2) are both located in the region where the solar perturbation is dominant. We numerically integrated the outer satellite S/1997 U2 over 10000 years with use of the initial elements that are given in IAUC 6869(1998 April 7):Epoch = 1998 July 6.0 TT, $a = 0.081643AU, e = .509386, I = 152.6686, \omega = 18.0055, \Omega = 255.8085$, time of pericenter passage = 1997 June 16.0390 TT. Figure 1 shows the time variations of e, I, ω, Ω of the first 2700 years. In order to evaluate the analytical solutions, we need three parameters, e_0, I_0 and t_0 when the argument of pericenter is zero. These three parameters could be determined from the initial elements at $t = 0$ by an iterative method. Instead of this rigorous procedure, we determine e_0 and I_0 from the numerically integrated orbit after removing short periodic variations by running means: $e_0 = 0.512$ and $I_0 = 151.1$. Since the argument of pericenter circulates, x_1 and x_2 are calculated from (18) and (19): $x_1 = 0.6365, x_2 = 1.4808$. The maximum eccentricity and the minimum inclination of the the analytical solution are $e_{\max} = 0.603, I_{\max} = 160.5$, which are computed from (26) and (28). The corresponding values of the numerically integrated orbit are are $e_{\max} = 0.605, I_{\max} = 160.8$, which are in good agreement with those of the analytical solution. The circulation periods of ω and Ω of the numerically integrated orbit are $P_{\omega^*} = 1350$ years and $P_{\Omega^*} = 1770$ years. On the other hand the periods of analytical solution, which are computed from (40) and (51), are $P_{\omega^*} = 1220$ years and $P_{\Omega^*} = 1780$, whose relative accuracies with respect to the numerically integrated orbit are 0.1 and 0.01,respectively. The order of the disagreement in the secular perurbation is the ratio between the orbital periods of 1997 U2 and Uranus, $m = P(1997\ U2)/P(\text{Uranus})$=0.04.

	O	C		(O-C)/O	
	N.I.	Mignard	K&N	Mignard	K&N
P_Ω	17890y	15000y	17980y	0.162	0.004
P_ω	13600y	13400y	13670y	0.015	0.005

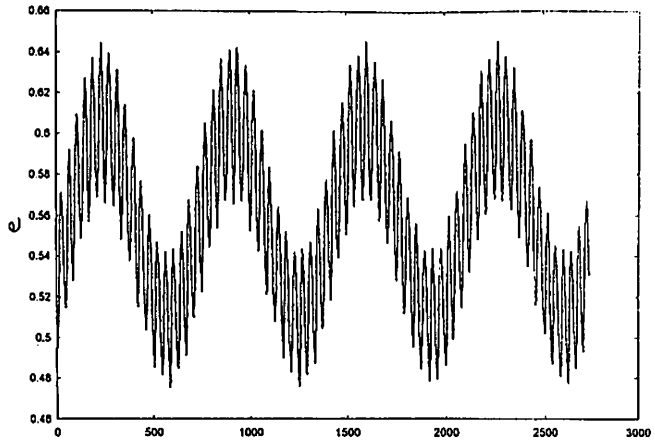
Table 1: secular perturbation of Nereid

	O	C		(O-C)/O	
	N.I.	Mignard	K&N	Mignard	K&N
Longitude of Node					
2ω	19.°795	19.°7	19.°921	0.005	0.006
4ω	3.579	3.3	3.463	0.078	0.032
6ω	0.795	0.7	0.803	0.227	0.11
Argument of Pericenter					
2ω	19.°326	19.°25	19.°443	0.004	0.006
4ω	3.573	3.23	3.400	0.096	0.006
6ω	0.902	0.725	0.785	0.196	0.129
Eccentricity					
2ω	0.00546	0.0060	0.00548	0.099	0.004
Inclination					
2ω	3.°123	3.°09	3.°137	0.106	0.004

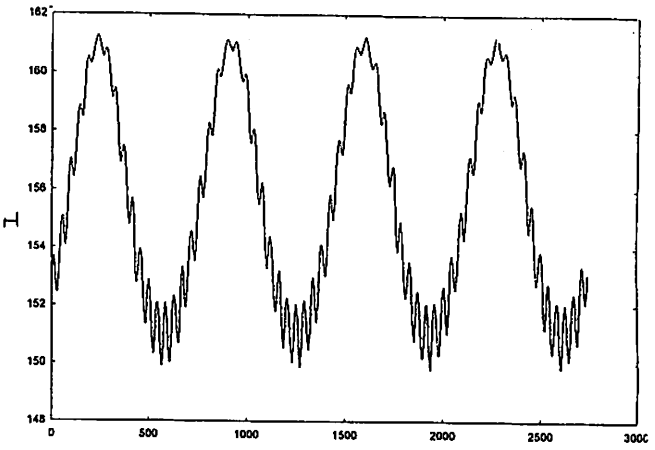
Table 2: Periodic Perturbations of Nereid

5.2 Nereid

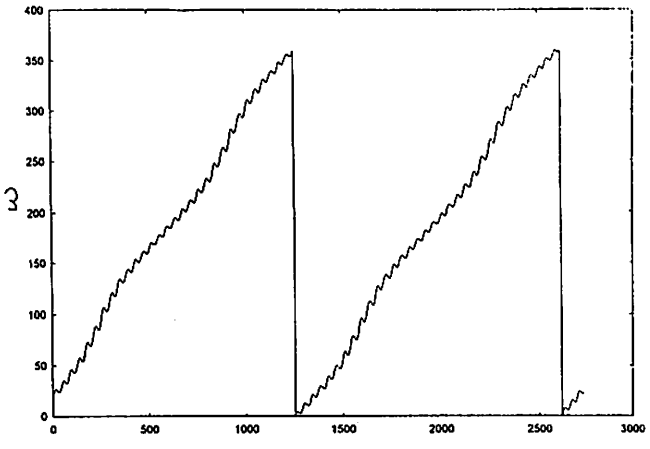
Nereid is the remotest satellites of Neptune and the eccentricity at present is about 0.76, which is the highest in the natural satellites in the solar system. The sun is also the dominant disturber, whose order is m^2 ($m = P(\text{Nereid})/P(\text{Neptune})=1/162$). The next larger perturbation comes from the inner satellite, Triton, whose order is about 1/1000 of the Sun's perturbation. Mignard(1975) constructed the analytical theory of Nereid up to m^3 with use of Von-Zeipel method. In his theory the eccentricity can take any value but the inclination is assumed to be small and the secular and long-periodic perturbations are obtained up to fourth order of the inclination. On the other hand in the theory of this paper both the eccentricity and the inclination can take any value but the order of perturbations is second order with respect to m. Mignard (1981) determine the mean elements of Nereid from the 44 known observations which were available at that time and gave the longperiodic ephemerides of Nereid. We numerically integrated the orbit of Nereid over 100,000 years with use of the osculating elements, which are given by Mignard (1981). Then similarly in the case of S/1997 U2 we evaluated our analytical solution and compared it with the numerically integrated orbit and the theory of Mignard (1981).



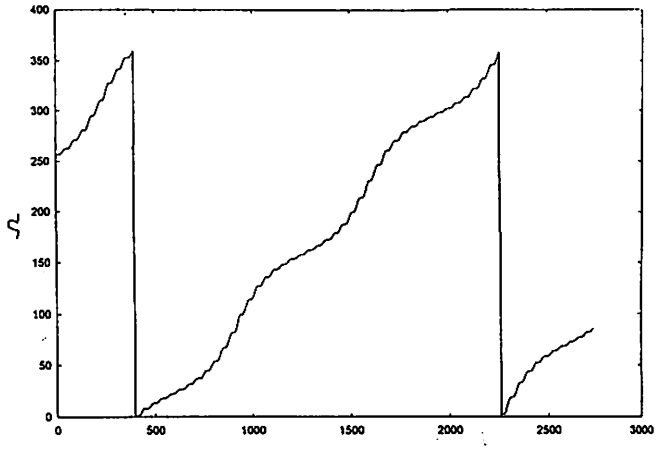
Eccentricity



Inclination



Argument of Pericenter



Longitude of Node

Figure 1: Orbital Elements of 1997 U2

The comparison for the secular perturbations are summarized in Table 1 and those for longperiodic perturbations in Table2,where the amplitudes of arguments $2\omega, 4\omega, 6\omega$ are given for four orbital elements, e, I, ω, Ω. In these tables 'O' stands for the numerically integrated orbit and 'C' stands for Mignard's theory (1981) or the theory in this paper. The last two columns of the tables give the relative accuracies of two theories with respect to the numerically integrated orbit. As for the secular perturbation the solution in this paper is much better than Mignar's solution. As for the long periodic perturbations the solution of this paper is in better agreement than Mignard's solution except the amplitudes of 2ω for ω and Ω. Since the disagreement in these two terms is of order m, we can expect the differential correction of order m^3 to the solution in this paper will give much better solution.

6 Appendix

6.1 Fourier Expansion of Longitude of Node

From (21) we have

$$x - h = (x_0 - h)(1 - \beta \mathrm{cn}^2\theta) = (x_0 - h)(1 - \beta)(1 + \beta^* \mathrm{sn}^2\theta) \tag{88}$$

where

$$\beta = \frac{x_0 - x_1}{x_0 - h}, \beta^* = \frac{\beta}{1 - \beta} \text{ and } \beta > 0, \beta^* > 0. \tag{89}$$

We define a constant c such that

$$k^2 \mathrm{sn}^2 ic = -\beta^*, i = \sqrt{-1} \tag{90}$$

With use of the formular of Jacobian elliptic functions with imaginary arguments (Byrd and Friedman 1971, p.24)

$$\mathrm{sn}(ic, k) = i\,\mathrm{tn}(c, k'), k' = \sqrt{1 - k^2}, \tag{91}$$

we obtain the expression of c in real number

$$c = \mathrm{tn}^{-1}(\sqrt{\beta^*}/k, k') = F(\xi, k'), \tag{92}$$

$$\sin\xi = \sqrt{\frac{\beta^*}{k^2 + \beta^*}} = \sqrt{\frac{x_2 - x_1}{x_2 - h}}, \tag{93}$$

where F is the normal elliptic integral of the first kind. From (90) we have

$$\mathrm{cn}^2 ic = 1 - \mathrm{sn}^2 ic = 1 + \frac{\beta^*}{k^2}, \tag{94}$$

$$\mathrm{dn}^2 ic = 1 - k^2 \mathrm{sn}^2 ic = 1 + \beta^*. \tag{95}$$

With use of (90), (94), (95), the second term of the right-hand side of equation (43) is transformed in the following expression:

$$\begin{aligned} \frac{x_0 - h}{x - h} &= \frac{1}{1 - \beta} \frac{1}{1 + \beta^* \mathrm{sn}^2\theta} = \frac{1}{1 - \beta}\left(1 - \frac{\beta^* \mathrm{sn}^2\theta}{1 + \beta^* \mathrm{sn}^2\theta}\right) \\ &= \frac{1}{1 - \beta}\left(1 - p \frac{k^2 \mathrm{sn}\, ic\, \mathrm{cn}\, ic\, \mathrm{dn}\, ic\, \mathrm{sn}^2\theta}{1 - k^2 \mathrm{sn}^2 ic\, \mathrm{sn}^2\theta}\right) \end{aligned} \tag{96}$$

$$p = \frac{\beta^*}{i\sqrt{\beta^*(1+\beta^*)(k^2+\beta^*)}} \tag{97}$$

The second term of (96) can be expressed in terms of Jacobian Zeta functions (Whittaker and Watson 1965, p.522)

$$\frac{k^2 \mathrm{sn}\, ic\, \mathrm{cn}\, ic\, \mathrm{dn}\, ic\, \mathrm{sn}^2\theta}{1 - k^2 \mathrm{sn}^2 ic\, \mathrm{sn}^2\theta} = \frac{1}{2}(Z(\theta - ic) - Z(\theta + ic) + 2Z(ic)). \tag{98}$$

The Fourier expansion of $Z(u)$ (Byrd and Friedman 1971,p.301) is

$$Z(u) = \frac{2\pi}{K} \sum_{m=1}^{\infty} \frac{q^m}{1 - q^{2m}} \sin \frac{m\pi}{K} u. \tag{99}$$

where K is the complete elliptic integral of the first kind and q is Jacobi's nome, and they are fuctions of k^2. By applying the expansion (99) to (98), the Fourier expansion of (97) is

$$\frac{x_0 - h}{x - h} = a_0 + \sum_{m=1}^{\infty} a_m \cos m \frac{\pi}{K}\theta, \tag{100}$$

where

$$a_0 = \frac{1}{1-\beta}(1 - \frac{1}{i}\sqrt{\frac{\beta^*}{(1+\beta^*)(k^2+\beta^*)}} Z(ic)), \tag{101}$$

$$a_m = \frac{2\pi}{K(1-\beta)}\sqrt{\frac{\beta^*}{(1+\beta^*)(k^2+\beta^*)}} \frac{q^m}{1-q^{2m}} \sinh m\frac{\pi c}{K}. \tag{102}$$

$Z(ic)$ can be expressed in term of Heuman's Lambda function Λ_0 as

$$Z(ic) = \frac{1}{i}(-\sqrt{\frac{\beta^*(1+\beta^*)}{k^2+\beta^*}} + \frac{\pi}{2K}\Lambda_0(\xi, k)) \tag{103}$$

(see Byrd and Friedman 1971,pp.318-319),

$$\Lambda_0(\xi, k) = \frac{2}{\pi}(EF(\xi, k') + KE(\xi, k') - KF(\xi, k')), \tag{104}$$

where E and $E(\xi, k')$ are the complete and normal elliptic integral of the second kind, respectively. Substituting (103) into (101) we have

$$a_0 = \frac{1}{1-\beta}\Big(\frac{k^2}{k^2+\beta^*} + \frac{\pi}{2K}\sqrt{\frac{\beta^*}{(1+\beta^*)(k^2+\beta^*)}}\Lambda_0(\xi, k)\Big). \tag{105}$$

Now the equation of Ω takes the following form with use of (43):

$$\frac{d\Omega}{dt} = \frac{3}{4}\gamma^*\sqrt{h}(1 - 2a_0 - 2\sum_{m=1}^{\infty}(-1)^m a_m \cos 2m\omega^*) \tag{106}$$

and is integrated as

$$\Omega = \Omega^* + \sum_{m=1}^{\infty} b_m \sin 2m\omega^*, \tag{107}$$

where

$$\Omega^* = n_{\Omega^*} t + \Omega_0^*, \tag{108}$$

$$n_{\Omega^*} = \frac{3}{4}\gamma^*\sqrt{h}(1-2a_0), \tag{109}$$

$$b_m = -\frac{3}{4}\gamma^*\sqrt{h}\frac{(-1)^m a_m}{m n_{\omega^*}}. \tag{110}$$

The variable Ω^* is the angle variable corresponding to the longitude of node and Ω_0^* is one of integral constants. With use of (18) and (19) we have

$$\frac{1}{1-\beta}\sqrt{\frac{\beta^*}{(1+\beta^*)(k^2+\beta^*)}} = \sqrt{\frac{x_0-h}{(x_1-h)(x_2-h)}}\sqrt{x_2-x_1} = \sqrt{\frac{3}{2h}}\sqrt{x_2-x_1}. \tag{111}$$

and with help of the expression of mean motion of perihelion n_{ω^*}(24),we finally get the expression of the perturbed mean motion of Ω and the amplitudes of arguments of $2m\omega^*$:

$$n_{\Omega^*} = \frac{3}{4}\sqrt{h}\gamma^*(1-2\frac{x_0-h}{x_2-h}) - \Lambda_0(\xi,k)n_{\omega^*}, \tag{112}$$

$$b_m = -\frac{2(-q)^m}{m(1-q^{2m})}\sinh m\frac{\pi c}{K}. \tag{113}$$

6.2 Fourier Expansion of Mean Anomaly

From(21) we have

$$1-x = (1-x_0)(1+\beta_1 \mathrm{cn}^2\theta) = (1-x_0)(1+\beta_1)(1-\beta_1^*\mathrm{sn}^2\theta), \tag{114}$$

where

$$\beta_1 = \frac{x_0-x_1}{1-x_0}, \beta_1^* = \frac{\beta_1}{1+\beta_1} = \frac{x_0-x_1}{1-x_1} > k^2. \tag{115}$$

and then we get from (44) the equation for the mean anomaly:

$$\frac{dl}{dt} = \frac{1}{4}\gamma^*\Big(-4+6x_0-3h-\frac{6}{(1+\beta_1)(1-\beta_1^*\mathrm{sn}^2\theta)}\Big). \tag{116}$$

From the formula (434.01) of the elliptic integrals of the third kind (Byrd and Friedman 1971, p.235)

$$\int_0^\theta \frac{d\theta}{1-\beta_1^*\mathrm{sn}^2\theta} = \frac{\sqrt{\beta_1^*}\pi(\theta\Lambda_0(\zeta,k)-\Omega_6)}{2K\sqrt{(\beta_1^*-k^2)(1-\beta_1^*)}}, \tag{117}$$

where

$$\Omega_6 = \frac{K}{i\pi}\ln\frac{\theta_3(v-iw)}{\theta_3(v+iw)}, v = \frac{\pi}{2K}\theta. \tag{118}$$

and

$$\Lambda_0(\zeta,k) = \frac{2}{\pi}(EF(\zeta,k') + KE(\zeta,k') - KF(\zeta,k')), \tag{119}$$

which is Heuman's Lambda function. In the expressions of (118) and (119), the parameters w and ζ are

$$w = \frac{\pi}{2K}F(\zeta,k'), \tag{120}$$

and

$$\sin\zeta = \sqrt{\frac{\beta_1^*-k^2}{\beta_1^*k'^2}} = \sqrt{\frac{(x_2-1)(1-x_0)}{(1-x_1)(x_2-x_0)}}. \tag{121}$$

For the Fourier expansion of Ω_6 we use the formula (1050.02) of the Theta functions (Byrd and Friedman 1971, p.316)

$$\ln \theta_3(v) = \ln C_E + 2\sum_{m=1}^{\infty} \frac{(-1)^{m-1}q^m}{m(1-q^{2m})} \cos 2mv, \tag{122}$$

where C_E is Euler's number(0.577216). Substituting (122) into (118) we have the partial derivative of Ω_6 with respect to θ:

$$\frac{\partial \Omega_6}{\partial \theta} = -4\sum_{m=1}^{\infty} \frac{(-q)^m}{1-q^{2m}} \sinh 2mw \cos 2mv \tag{123}$$

then by differentiating (117) with respect to θ we get the following Fourier expansion:

$$\frac{1}{1-\beta_1^* \mathrm{sn}^2\theta} = c_0 + \sum_{m=1}^{\infty} c_m \cos 2mv, \tag{124}$$

$$c_0 = \frac{\pi}{2K}\sqrt{\frac{\beta_1^*}{(\beta_1^*-k^2)(1-\beta_1^*)}}\Lambda_0, \tag{125}$$

$$c_m = \frac{2\pi(-q)^m}{K(1-q^{2m})}\sqrt{\frac{\beta_1^*}{(\beta_1^*-k^2)(1-\beta_1^*)}} \sinh 2mw. \tag{126}$$

Finally substituting (124) into (116) we have the equation for the mean anomaly:

$$\frac{dl}{dt} = \frac{1}{4}\gamma^*\Big(-4+6x_0-3h-\frac{6}{1+\beta_1}(c_0+\sum_{m=1}^{\infty} c_m \cos 2mv)\Big), \tag{127}$$

where $v = \omega^* + \pi/2, \omega^* = n_{\omega^*}t$ (see (23)and (24)), and get the Fourier expansion of the mean anomaly in the following form:

$$l = l^* + \sum_{m-1}^{\infty} f_m \sin 2m\omega^*, \tag{128}$$

$$l^* = n_{l^*}t + l_0^*, \tag{129}$$

where

$$n_{l^*} = \frac{1}{4}\gamma^*(-4+6x_0-3h-\frac{6}{1+\beta_1}c_0), \tag{130}$$

$$f_m = -\frac{3}{4}\frac{\gamma^*}{1+\beta_1}\frac{(-1)^m c_m}{m n_{\omega^*}}. \tag{131}$$

The variable l^* is the angle variable and l_0^* is one of the integral constants. With help of (18) and (19), we have

$$\frac{1}{(1+\beta_1)\sqrt{x_2-x_1}}\sqrt{\frac{\beta_1^*}{(\beta_1^*-k^2)(1-\beta_1^*)}} = \sqrt{\frac{3}{2}} \tag{132}$$

and we obtain the following much simpler exppressions of n_{l^*} and f_m:

$$n_{l^*} = \frac{1}{4}\gamma^*(-4+6x_0-3h) - \Lambda_0 n_{\omega^*}, \tag{133}$$

$$f_m = -\frac{2q^m}{m(1-q^{2m})}\sinh 2mw. \tag{134}$$

References

Byrd, P. F. and Friedman, M. D.:1971, Handbook of Elliptic Integrals for Engineers and Scientists, *Springer-Verlag.*

Holman, M., Touma, J., and Tremaine, S.:1997, Chaotic Variations in the Eccentricity of the Planet Orbiting 16 Cygni B, *Nature*, **386**, 254-256.

Innanen, K.A.,Zheng, J.Q., Mikkola, S., and Valtonen, M.J.:1997, The Kozai Mechanism and the Stability of Planetary Orbits in Binary Star Systems, *Astron. J.*, **113**, 1915-1919.

Kinoshita, H. and Nalkai, H.:1991, Secular Perturbations of Fictitious Satellites of Uranus, *Celest. Mech.*, **52**, 293-303.

Kozai,Y:1962, Secular Perturbations of Asteroids with High Inclination and Eccentricity, *Astron. J.*, **67**, 591-598.

Mignard, F.:1975, Satellite à Forte Excentricité. Application à Néréide, *Astron. & Astrophys.*, **43**, 359-379.

Mignard, F.: 1981, The Mean Elements of Nereid, *Astron. J.*, **86**, 1728-1729.

Thomas, F. and Morbidelli,A.: 1996, The Kozai Resonance in the Outer Solar System and the Dynamics of Long-Period Comets, *Celest. Mech.*, **64**, 209-229.

Wiegert, P.A. and Holman, M.:1997, The Stability of Planets in the Alpha Centauri System, *Astron. J.*, **113**, 1445-1450.

AN ANALYTICAL THEORY ON A SATELLITE MOTION WITH HIGHLY ECCENTRIC ORBIT

ABDEL-NABY S. SAAD[1] and HIROSHI KINOSHITA[2]
1 *Dept. of Astron. Sci., School of Mathematical and Physical Science, The Graduate Univ. For Advanced Studies, Mitaka, Tokyo 181, Japan.*
E-mail : saad@pluto.mtk.nao.ac.jp

2 *National Astronomical Observatory, Mitaka, Tokyo 181, Japan.*
E-mail : Kinoshita@nao.ac.jp

(Received ..... ; Accepted in final form .....)

Abstract. In this paper , we offer an analytical theory for the motion of a satellite with highly eccentric orbit and zero inclination. The solution of the equations of motion is given using Lie transformations approach in Hori's version. The solar perturbations effects are taken into account. The disturbing function is developed in powers of the ratio of the semimajor axes of the satellite and the sun and put in a closed form with respect to the eccentricity. An application on the second satellite of Neptune is considered. The secular and periodic perturbations are obtained up to the fifth and fourth order respectively. The comparison with the numerical integration of the equations of motion gives an accuracy on the level of 300 m for the semi-major axis, 3×10^{-8} for the eccentricity and 0.004 arc second for the angular variables.

Keywords: Satellite theory, Hori's perturbation method, Nereid.

1. Inroduction

In the present paper, we construct an analytical theory for the motion of a satellite with high eccentricity and zero inclination using Hori's method (1966). The perturbing forces considerd are only due to the sun. For the verification of the theory, an application to the second Neptunian satellite is performed. The results of the analytical solution are compared with those of the numerical integration of the equations of motion. The disturbing function is developed in powers of the ratio between the smei-major axis of the satellite and the semi-major axis of the sun, and put in a closed form with respect to the eccentricity. A similar theory on the lunar motion was investigated by Hori(1963) based on Von Zeipel's method.

Hori's method has many advantages over the usual Von Zeipel's method: (i) it is based on Poisson brackets, (ii) all the formulae are canonically invariant, and (iii) the perturbations in any element are given in an explicit and simple form. The method introduced by Deprit(1969) based

on Lie approach is equivalent to that by Hori(Campbell and Jefferys, 1970; Henrard and Roels, 1974).

Nereid has little been studied analyticaly as well as observationally since its discovery by Kuiper in 1949. Perhaps, because of its unusually large eccentricity and faintness. Mignard (1975, 1981) introduced a second order theory on Nereid using Von Zeipel's method. He determined the mean elements of Nereid from 44 Earth-based observations (van Briesbroek, 1951,1957; Rose, 1974). Segerman and Richardson (1997) offered an analytical theory for the orbit of Nereid. In their theory the perturbations (due to the sun, Triton and the oblateness of Neptune) are treated as a first-order effects, and a first order transformation using Deprit's method is obtained.

The orbital determination of Nereid is carried out by Rose(1974) using Keplerian elements and by Veillet(1882, 1988) using Mignard's theory. Although these models are adequate for orbital fit to the Earth-based observations, they are not sufficiently complete nor precise enough to support Voyager navigation(Jacobson, 1990, 1991). For a better understanding the motion of Nereid we extended our theory up to fifth order for secular perturbations and up to fourth order for periodic. In Section 2 we give the equations of motion and the desturbing function. In Section 3 we introduce briefly the method of solution, while Section 4 is devoted to the analytical results for the short-periodic, long periodic and secular terms respectively. In Section 5 we offer a comparison between the numerical and the analytical results, then our conclusions.

2. Formulations

The equations of motion of a satellite disturbed by the sun are given in Delaunay's elements as follows:

$$\frac{d(L,G,H,K)}{dt} = \frac{\partial F}{\partial(l,g,h,k)}, \tag{2.1}$$

$$\frac{d(l,g,h,k)}{dt} = -\frac{\partial F}{\partial(L,G,H,K)}.$$

with the Hamiltonian

$$F = \frac{\mu^2}{2L^2} - \nu K + R \tag{2.2}$$

where $\mu = n^2a^3$ (n and a define the mean motion and the semimajor-axis of the satellite respectively), ν is the mean motion of the sun, k (the mean longitude of the sun) is given by

$$k = \nu t + const.$$

K is a conjucate momentum of k which is introduced for the Hamiltonian to be independent of time (Brouwer and Clemence, 1961).

R: is the disturbing function due to the sun

$$R = R_1 + R_2 + ...$$

$$R = \nu^2 a^2 (\frac{r}{a})^2 (\frac{a_\odot}{r_\odot})^3 \left\{ (-\frac{1}{8} + \frac{3}{8}\cos^2(i)) + \frac{3}{16}(1+\cos(i))^2 \cos(2f - 2f_\odot + 2g + 2h) + \right.$$

$$\left. + \frac{3}{8}(1-\cos^2(i))[\cos(2f+2g)+\cos(2f_\odot-2h)] + \frac{3}{16}(1-\cos(i))^2 \cos(2f+2f_\odot+2g-2h) \right\}. \quad (2.3)$$

Here, R_2 is neglected by considering the ratio $a/a_\odot$ equal to zero. Delaunay elements are given by:

$$L = \sqrt{\mu a}, \quad l$$

$$G = L\sqrt{1-e^2}, \quad g$$

$$H = G\cos(i), \quad h$$

$$\cos(i) = \frac{H}{G}$$

Considering the case of zero inclination and the circular motiom of the sun, the Hamiltonian is reduced to the following form

$$F = \frac{\mu^2}{2L^2} - \nu K + \frac{1}{4}\nu^2 a^2 \left[\frac{r^2}{a^2} + 3\frac{r^2}{a^2}\cos(2f + 2g + 2h - 2k) \right]. \quad (2.4)$$

Following Hori's approach(1963), we consider the canonical transformation of the variables

$$x_1, x_2, x_3, x_4 \; ; \; y_1, y_2, y_3, y_4$$

wher, $x_1=L$, $x_2=G$, $x_3=H\text{-}G$, $x_4=K+G$, and

$$y_1=l, \; y_2=g+h\text{-}k, \; y_3=h, \; y_4=k.$$

The new Hamiltonian does not depend on y_3 or y_4, consequently x_3 and x_4 are constants, then we have the new equations of motion

$$\frac{dx_1}{dt} = \frac{\partial F}{\partial y_1}, \; \frac{dy_1}{dt} = -\frac{\partial F}{\partial x_1}, \tag{2.5}$$

$$\frac{dx_2}{dt} = \frac{\partial F}{\partial y_2}, \; \frac{dy_2}{dt} = -\frac{\partial F}{\partial x_2}.$$

and the associated Hamiltonian

$$F = \frac{\mu^2}{2L^2} + \nu x_2 + \frac{1}{4}\nu^2 a^2 \left[\frac{r^2}{a^2} + 3\frac{r^2}{a^2}\cos(2f + 2y_2)\right]. \tag{2.6}$$

$$F = F_0 + F_1 + F_2 + \ldots$$

where

$$F_0 = \frac{\mu^2}{2{x_1}^2}, \qquad F_1 = \nu x_2, \tag{2.7}$$

$$F_2 = \frac{1}{4}\nu^2 a^2 \left[\frac{r^2}{a^2} + 3\frac{r^2}{a^2}\cos(2f + 2y_2)\right]. \tag{2.8}$$

3. Hori's Perturbation Method

In this section we introduce briefly Hori's perturbation method(1966). To solve the dynamical system equations (2.5) we consider the canonical transformation

$$x_j, y_j \longrightarrow x_j', y_j', \; j=1,2$$

The perturbations of any quantity x_j, and y_j can be given by the series

$$x_j = x_j' + \left\{x_j', S\right\} + \frac{1}{2}\left\{\left\{x_j', S\right\}, S\right\} + \ldots$$

$$y_j = y_j^{'} - \left\{y_j^{'}, S\right\} + \frac{1}{2}\left\{\left\{y_j^{'}, S\right\}, S\right\} + ... \tag{3.1}$$

where, the braces are the Poissn brackets,

$$\epsilon S(x_j^{'}, y_j^{'}) = \sum_{k=1} S_k(x_j^{'}, y_j^{'})$$

is the determining function of the new variables, S_k has a factor ϵ^k, where ϵ^k is a small parameter.
Any function of x_j, y_j can be expressed on the form

$$f(x,y) = f(x',y') + \{f,S\} + \frac{1}{2}\{\{f,S\},S\} + ... \tag{3.2}$$

The new canonical equations take the form

$$\frac{dx_1^{'}}{dt} = \frac{\partial F^*}{\partial y_1^{'}}, \frac{dy_1^{'}}{dt} = -\frac{\partial F^*}{\partial x_1^{'}}, \tag{3.3}$$

$$\frac{dx_2^{'}}{dt} = \frac{\partial F^*}{\partial y_2^{'}}, \frac{dy_2^{'}}{dt} = -\frac{\partial F^*}{\partial x_2^{'}}.$$

where F^* is the new Hamiltonian. Since F is free from the time, we have the energy integral:

$$\sum_{k=0} F_k^*(x_1^{'}, x_2^{'}, y_1^{'}, y_2^{'}) = \sum_{k=0} F_k(x_1, x_2, y_1, y_2). \tag{3.4}$$

F_k has the factor ϵ^k.

In order to obtain the determining function S_k and the new Hamiltonian F_k^*, a new parameter t^* is introduced to get what called an auxilary equation

$$\frac{dx_j^{'}}{dt^*} = \frac{\partial F_0}{\partial y_j^{'}}, \frac{dy_j^{'}}{dt^*} = -\frac{\partial F_0}{\partial x_j^{'}}. \tag{3.5}$$

The averaging with respect to the parameter t^* eliminates this parameter from the new Hamiltonian F^*. Equations (3.1), (3.4) and (3.5) are used together to get the determining function and the new Hamiltonian by the following algorithm.

0-order

$$F_0^* = F_0,$$

first-order

$$F_1^* = F_{1s},$$

$$S_1 = \int F_{1p} dt^*,$$

second-order

$$F_2^* = F_{2s} + \frac{1}{2}\{F_1 + F_1^*, S_1\}_s,$$

$$S_2 = \int \left(F_{2p} + \frac{1}{2}\{F_1 + F_1^*, S_1\}_p\right) dt^*,$$

third-order

$$F_3^* = F_{3s} + \frac{1}{12}\{\{F_{1p}, S_1\}, S_1\}_s + \frac{1}{2}\{F_2 + F_2^*, S_1\}_s + \frac{1}{2}\{F_1 + F_1^*, S_2\}_s,$$

$$S_3 = \int \left(F_{3p} + \frac{1}{12}\{\{F_{1p}, S_1\}, S_1\}_p + \frac{1}{2}\{F_2 + F_2^*, S_1\}_p + \frac{1}{2}\{F_1 + F_1^*, S_2\}_p\right) dt^*.$$

and so on.

The subscripts s and p stand for the secular and periodic parts respectively. This process can be repeated to another set of variables x_j'', y_j'' with introducing a new parameter t^{**}. The new determining function S_k^* and the new Hamiltonian F_k^{**} can be obtained using the algorithm

0-order

$$F_0^{**} = F_0^*,$$

first-order

$$F_1^{**} = F_1^*,$$

Second-order

$$F_2^{**} = F_{2s}^*,$$

$$S_1^* = \int F_{2p}^* dt^{**},$$

third-order

$$F_3^{**} = F_{3s}^* + \frac{1}{2}\{F_2^* + F_2^{**}, S_1^*\}_s,$$

$$S_2 = \int \left(F_{3p}^* + \frac{1}{2}\{F_2^* + F_2^{**}, S_1^*\}_p\right) dt^{**},$$

and so on.

4. Results

4.1. SHORT PERIODIC TERMS

To remove the short periodic terms(y_1), we average on the mean anomaly $l \equiv y_1$ by considering the canonical transformation

$$(x_1, x_2, y_1, y_2) \longrightarrow (x_1', x_2', y_1', y_2')$$

Then, F will be

$$F(x_1, x_2, y_1, y_2) \longrightarrow F^*(x_1', x_2', \cdot, y_2')$$

with the determining function $S \equiv S(x_1', x_2', y_1', y_2')$, where the new Hamiltonian and the determining function are

$$F^* = F_0^* + F_1^* + F_2^* + F_3^* + F_4^* + F_5^*,$$

$$S = S_1 + S_2 + S_3 + S_4.$$

The averaging is performed such that the new Hamiltonian F^* does not include y_1. Applying the first algorithm in section 3, the new Hamiltonian and the determining function have the forms

$$F_0^* = \frac{\mu^2}{2x_1'^2}, \tag{4.1.1}$$

$$F_1^* = \nu x_2'^2, \tag{4.1.2}$$

$$S_1 = 0,$$

$$F_2^* = \frac{1}{4}\nu'^2 a'^2 \left(\frac{3}{2}e'^2 + \frac{15}{2}e'^2\cos(2y_2')\right), \tag{4.1.3}$$

It is convenient to express the formulae in terms of the eccentric anomaly u instead of the mean anomaly $l'(\equiv y_1')$

$$l' = u' - e'\sin(u').$$

The dashed orbital elements a', e', n', and η' obtained after the elimination of short periodic terms are computed from

$$a' = \frac{x_1'^2}{\mu}, e' = \sqrt{1 - (\frac{x_2'}{x_1'})^2}, n' = \frac{\mu^2}{x_1'^3}, \eta' = \frac{x_2'}{x_1'}.$$

Hereafter in this section for simplicity the superscript dash attached to a, e, n, η, u and y_2 are omitted, which will not cause confusion.

$$S_2 = \frac{1}{4}\frac{\nu^2 a^2}{n}\left(A_1 + B_1\cos(2y_2) + C_1\sin(2y_2)\right), \tag{4.1.4}$$

$$F_3^* = 0$$

$$S_3 = \frac{1}{4}\frac{\nu^3 a^2}{n^2}\left(-2B_2\sin(2y_2) + 2C_2\cos(2y_2)\right), \tag{4.1.5}$$

$$F_4^* = \frac{1}{16}\frac{\nu^4 a^2}{n^2}\left\{-\frac{49}{4} + \frac{873}{4}e^2 - \frac{4347}{32}e^4 + (\frac{333}{4}e^2 - \frac{237}{8}e^4)*\right.$$

$$\left.*\cos(2y_2) + \frac{615}{32}e^4\cos(4y_2)\right\}, \tag{4.1.6}$$

$$S_4 = \frac{1}{16}\frac{\nu^4 a^2}{n^3}\{[S_4]_0 + [S_4]_{2c}\cos(2y_2) + [S_4]_{2s}\sin(2y_2) + \\ + [S_4]_{4c}\cos(4y_2) + [S_4]_{4s}\sin(4y_2)\}, \tag{4.1.7}$$

$$F_5^* = \frac{1}{16}\frac{\nu^5 a^2}{n^3}\eta\left[-\frac{97}{2} + \frac{2335}{4}e^2 - \frac{1545}{8}e^4 + e^2(101 - 17e^2)\cos(2y_2)\right]. \tag{4.1.8}$$

Where

$$A_1 = (-2e + \frac{3}{4}e^3)\sin(u) + \frac{3}{4}e^2\sin(2u) - \frac{1}{12}e^3\sin(3u),$$

$$B_1 = (\frac{-15}{2}e + \frac{15}{4}e^3)\sin(u) + (\frac{3}{2} + \frac{3}{4}e^2)\sin(2u) + (-\frac{1}{2}e + \frac{1}{4}e^3)\sin(3u),$$

$$C_1 = \eta\left(\frac{-15}{4}e^2 - \frac{15}{2}e\cos(u) + (\frac{3}{2} + \frac{3}{2}e^2)\cos(2u) - \frac{1}{2}e\cos(3u)\right),$$

$$B_2 = \frac{33}{8}e^2 - \frac{27}{16}e^4 + (\frac{33}{4}e - \frac{27}{8}e^3)\cos(u) + (\frac{-3}{4} - \frac{19}{8}e^2 + e^4)\cos(2u) + \\ + (\frac{15}{2}e + \frac{1}{24}e^3\cos(3u) + (\frac{-1}{16}e^2 + \frac{1}{32}e^4)\cos(4u),$$

$$C_2 = \eta\left[(\frac{-33}{4}e + 3e^3)\sin(u) + (\frac{3}{4} + \frac{11}{4}e^2)\sin(2u) + (\frac{-15}{12}e - \frac{1}{4}e^3) * \right. \\ \left. * \sin(3u) + \frac{1}{16}e^2\sin(4u)\right],$$

$$[S_4]_0 = (-\frac{127}{4}e - \frac{857}{4}e^3 + \frac{8383}{96}e^5)\sin(u) + (-\frac{427}{16}e^2 + \frac{7739}{96}e^4)\sin(2u), \\ + (-\frac{19}{24}e^3 - \frac{547}{96}e^5)\sin(3u) - \frac{7}{384}e^4\sin(4u) + \frac{1}{24}e^5\sin(5u),$$

$$[S_4]_{2c} = (-61e - 87e^3 + \frac{143}{4}e^5)\sin(u) + (4 + \frac{21}{8}e^2 + \frac{209}{8}e^4)\sin(2u) +$$

$$+(\frac{7}{2}e-\frac{5}{2}e^3-\frac{21}{8}e^5)\sin(3u)+(-\frac{21}{16}e^2+\frac{25}{32}e^4)\sin(4u)+(\frac{1}{20}e^3-\frac{1}{40}e^5)\sin(5u),$$

$$[S_4]_{2s}=\eta\left\{-\frac{61}{2}e^2+\frac{35}{4}e^4+(61e+\frac{35}{2}e^3)\cos(u)+(-4-\frac{37}{8}e^2-\frac{37}{4}e^4)*\right.$$

$$\left.*\cos(2u)+(-\frac{7}{2}e+\frac{3}{4}e^3)\cos(3u)+(\frac{21}{16}e^2-\frac{1}{8}e^4)\cos(4u)-\frac{1}{20}e^3\cos(5u)\right\},$$

$$[S_4]_{4c}=(-\frac{99}{2}e^3+\frac{675}{32}e^5)\sin(u)+(-\frac{219}{16}e^2+\frac{459}{32}e^4)\sin(2u)+$$

$$+(\frac{83}{4}e-\frac{147}{8}e^3+\frac{45}{32}e^5)\sin(3u)+(\frac{9}{16}-\frac{69}{16}e^2+\frac{369}{128}e^4)\sin(4u),$$

$$[S_4]_{4s}=\eta\left\{-\frac{99}{4}e^4+\frac{99}{2}e^3\cos(u)+(\frac{219}{16}e^2-\frac{15}{2}e^4)\cos(2u)+\right.$$

$$\left.+(-\frac{83}{4}e+8e^3)\cos(3u)-(\frac{9}{16}+\frac{129}{32}e^2-1516e^4)\cos(4u)\right\}.$$

$$\eta=\sqrt{1-e^2}$$

The expressions up to S_3 and F_4^* are identical to those derived from the von Zeipel's method(Hori 1963). The expressions S_4 and F_5^* in this theory are diferent from those of Hori's method due to the different methods. It is shown, however, that they are mathematically equivalent with use of the relationships which are given in Hori(1970) and Yuasa(1970). The explicit expressions of x_j and y_j in terms of x_j' and y_j' are obtained from equation (3.1).

4.2. LONG PERIODIC TERMS

Removal of long periodic terms(y_2) from the Hamiltonian F^* requires another canonical transformation

$$(x_1', x_2', y_1', y_2') \longrightarrow (x_1'', x_2'', y_1'', y_2'')$$

and the new Hamiltonian will be free from y_1 and y_2 as follows

$$F^*(x_1', x_2', -, y_2') \longrightarrow F^{**}(x_1'', x_2'', -, -)$$

where

$$F^{**} = F_0^{**} + F_1^{**} + F_2^{**} + F_3^{**} + F_4^{**} + F_5^{**},$$

$$S^{*} = S_1^{*} + S_2^{*} + S_3^{*} + S_4^{*}.$$

In this subsection the orbital elements a'', e'', n'', and η''after elimination of the long periodic term are computed

$$a'' = \frac{x_1''^2}{\mu}, e'' = \sqrt{1 - (\frac{x_2''}{x_1''})^2}, n'' = \frac{\mu^2}{x_1''^3}, \eta'' = \frac{x_2''}{x_1''}.$$

As in subsection 4.1 for simplicity we omit the superscript " from the orbital elements a, e, n, η and y_2 in this section.

Following the second algorithm in Section 3 concerned with the long periodic terms, we can get the new Hamiltonian and the determining function in the following forms

$$F_0^{**} = \frac{\mu^2}{2{x_1}^2} = F_0^{*} \tag{4.2.1}$$

$$F_1^{**} = \nu x_2 = F_1^{*} \tag{4.2.2}$$

$$F_2^{**} = \frac{1}{4}\nu^2 a^2 (1 + \frac{3}{2}e^2) \tag{4.2.3}$$

$${S_1}^{*} = \frac{-15}{16}\nu^2 a^2 e^2 \sin(2y_2) \tag{4.2.4}$$

$$F_3^{**} = \frac{225}{64}\frac{\nu^3 a^2}{n} e^2 \eta \tag{4.2.5}$$

$${S_2}^{*} = \frac{-45}{64}\frac{\nu^2 a^2}{n} e^2 \eta \sin(2y_2) \tag{4.2.6}$$

$$F_4^{**} = \frac{1}{128}\frac{\nu^4 a^2}{n^2}\left[-98 + \frac{4167}{2}e^2 - \frac{12069}{8}e^4\right] \tag{4.2.7}$$

$${S_3}^{*} = \frac{1}{128}\frac{\nu^3 a^2}{n^2}\left[-963e^2 + \frac{4119}{2}e^4\right]\sin(2y_2) - \frac{1905}{32}e^4 \sin(4y_2) \tag{4.2.8}$$

$$F_5^{**} = \frac{\nu^5 a^2}{n^3}\eta\left[-\frac{97}{32} + \frac{288085}{4096}e^2 - \frac{872625}{16384}e^4\right] \tag{4.2.9}$$

$$S_4{}^* = \frac{\nu^4 a^2}{n^3}\eta\left[\left(-\frac{87275}{2048}e^2 + \frac{826975}{16384}e^4\right)\sin(2y_2) - \frac{292185}{65536}e^4\sin(4y_2)\right] \tag{4.2.10}$$

The perturbations of x_j' and y_j' can be obtained using the forms:

$$x_j' = x_j'' + \left\{x_j'', S^*\right\} + \frac{1}{2}\left\{\left\{x_j'', S^*\right\}, S^*\right\} + \ldots$$

$$y_j' = y_j'' - \left\{y_j'', S^*\right\} + \frac{1}{2}\left\{\left\{y_j'', S^*\right\}, S^*\right\} + \ldots$$

4.3. SECULAR TERMS

The Hamiltonian remaining after the second transformation determines the secular terms. The equations of motion have the form

$$\frac{dx_1''}{dt} = \frac{\partial F^{**}}{\partial y_1''},\ \frac{dy_1''}{dt} = -\frac{\partial F^{**}}{\partial x_1''}, \tag{4.3.1}$$

$$\frac{dx_2''}{dt} = \frac{\partial F^{**}}{\partial y_2''},\ \frac{dy_2''}{dt} = -\frac{\partial F^{**}}{\partial x_2''} \tag{4.3.2}$$

where

$$F^{**} = \frac{\mu^2}{2x_1''^2} + \nu x_2'' + \frac{1}{4}\nu^2 a''^2(1 + \frac{3}{2}e''^2) + \frac{225}{64}\frac{\nu^3 a''^2}{n''}e''^2\eta'' + \frac{1}{128}\frac{\nu^4 a''^2}{n''^2} *$$

$$* \left[-98 + \frac{4167}{2}e''^2 - \frac{12069}{8}e''^4\right] + \frac{\nu^5 a''^2}{n''^3}\eta''\left[-\frac{97}{32} + \frac{288085}{4096}e''^2 - \frac{872625}{16384}e''^4\right] \tag{4.3.3}$$

$$-\frac{\partial F^{**}}{\partial x_1''} = n'' - \frac{1}{4}\frac{\nu^2}{n''}(7+3e''^2) - \frac{225}{32}\frac{\nu^3}{n''^2}\eta''(1+2e''^2) - \frac{1}{128}\frac{\nu^4}{n''^3}(3187 + \frac{21267}{2}e''^2 -$$

$$- \frac{36207}{4}e''^4) - \frac{1}{1024}\frac{\nu^5}{n''^4}\eta''\left(\frac{213589}{2} + \frac{2008225}{4}e''^2 - \frac{872625}{2}e''^4\right), \tag{4.3.4}$$

$$-\frac{\partial F^{**}}{\partial x_2^{"}}+\nu = \frac{3}{4}\frac{\nu^2}{n''}\eta''+\frac{225}{32}\frac{\nu^3}{n''^2}(1-\frac{3}{2}e''^2)+\frac{1}{128}\frac{\nu^4}{n''^3}\eta''\left(4167-\frac{12069}{2}e''^2\right)+$$

$$+\frac{1}{1024}\frac{\nu^5}{n''^4}\left(\frac{294293}{2}-434220e''^2+\frac{4363125}{16}e''^4\right) \qquad (4.3.5).$$

Equations (4.3.4) and (4.3.5) define the mean motions of the mean anomaly and the longitude of perigee respectively. Since F^{**} is free from $y_1^{"}$ and $y_2^{"}$, then $x_1^{"}$ and $x_2^{"}$ are constants, and we have the mean elements

$$l^{"} = l_0 + \dot{l}^{"}t, \quad (g+h)^{"} = (g+h)_0 + \dot{(g+h)}^{"}t$$

where $\dot{l}^{"}$ and $\dot{(g+h)}^{"}$ are given by equations (4.3.4) and (4.3.5) respectively, l_0 and $(g+h)_0$ are mean elements at the epoch($t=0$). Table I gives the secular perturbations in the mean anomaly and the longitude of perigee using the mean elements in Table II.

TABLE I

Secular perturbations are given in radians per day

$O(\frac{\nu}{n})$	Mean an.	Peri.
2	-7.7364E-005	1.7623E-005
3	-2.0983E-006	2.2930E-007
4	-6.2423E-008	4.9778E-009
5	-1.2242E-009	-8.1582E-11

If two quantities A and B that are functions of orbital elements satisfiy the d'Alembert characteristics, the Poisson bracket $\{A,B\}$ keeps the d'Alembert characteristics. The original Hamiltonian F(2.4) satisfies d'Alembert characteristics and the operation appeared in Hori's algorithm is only the Poisson bracket. The transformed Hamiltonians F^* and the determining functions S, S^*, therefore, should satisfy the d'Alembert characteristics. In fact this is easily verified after F^*, S, S^* are changed in the form of the argument $iu + jy_2'$.

5. Comparison of The Analytical Solution

We compare the analytical solution obtained in this paper with the numerical integration by applying the analytical solution to a fictitious Nereid(the second satellite of Neptune) with zero inclination. The mean elements of the fictitious Nereid are given in the second column of the Table II. The orbital elements of the sun used in this integration are $a_\odot = 30.1104\text{AU}$, $e_\odot = 0.0$, $k(t=0) = 10°.0$. The small parameter in the theory, the ratio between the mean motions of the sun and the fictitious Nereid is $\nu/n = 5.98 \times 10^{-3}$. The osculating elements that are the initial conditions for the numerical integration are computed from the analytical solution(Section 4) are given in the third column of the Table II. We used for the numerical integration the extrapolation method, which has a capability of highly accurate orbital computation.

TABLE II

Mean and osculating elements

orbital elements	mean elements	osculating elements
semi-major axis(km)	5513413.256	5513226.872
eccentricity	0.751201525	0.751270690
long. of pericenter(deg)	254.809177	254.385293
mean anomaly(deg)	359.34112	0.362199662

Figures 1 and 2 show the osculating orbital elements of the fictitious Nereid over 5 years and 500 years, respectively, which are obtained from the analytical solution. Figures 1.3 and 2.3 show the periodic variation of the mean anomaly, which is obtained by subtracting the secular part of the osculating mean anomaly. Figures 3 and 4 show the differences of the osculating elements between the analytical solution and the orbit obtained by the numerical integration. The secular error in the mean anomaly is $-1°.6 \times 10^{-5}$/year. Figure 3.5 and 4.5 show the periodic residuals after taking away the secular error. This secular error is removed by the orbital adjustment which increases the semi-major axis by 160m. From this comparison the present theory has the accuracy of the level of 300 m for the semi-major axis , 3×10^{-8} for the eccentricity, and about 0.004 arc second for angle variable. The relative accuracy of them is about 4×10^{-8}, which is between $(\nu/n)^3$ and $(\nu/n)^4$.
Now we can say the present theory with zero inclination has an accuracy up to the fourth order, $(\nu/n)^4$, in periodic perturbations.

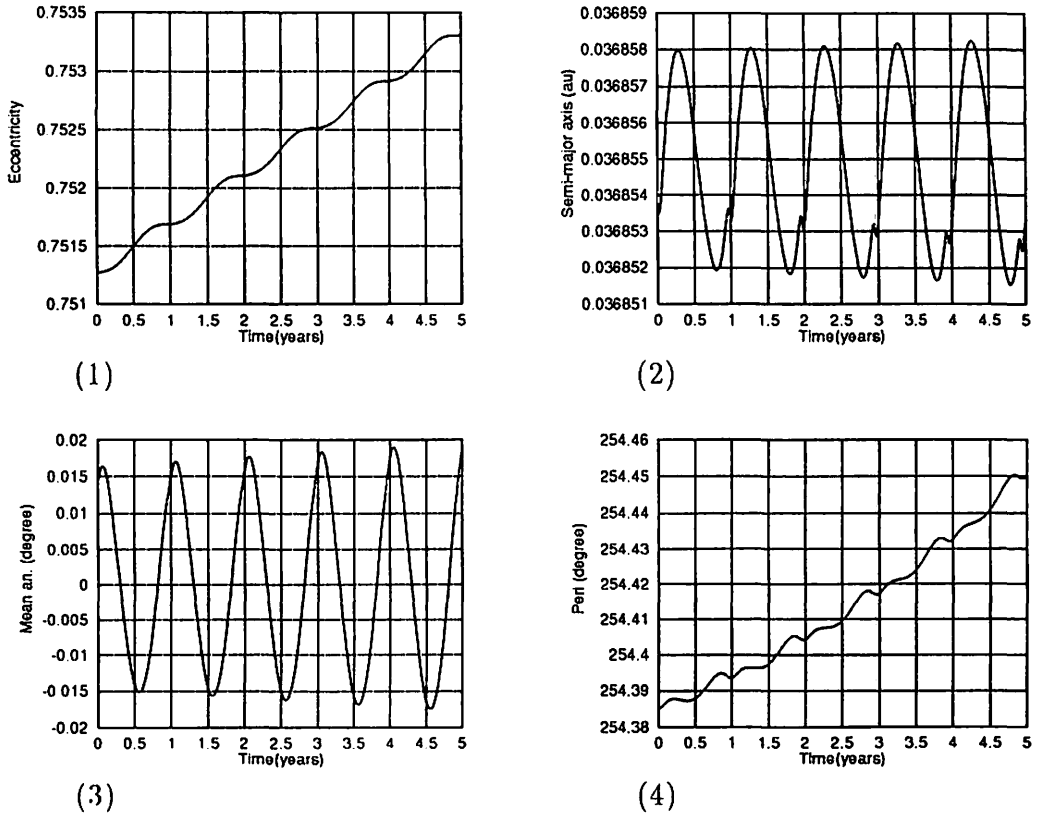


Figure 1. The osculating orbital elements of Nereid over 5 years:
(1) eccentricity,
(2) semi-major axis,
(3) periodic part of mean anomaly,
(4) longitude of pericenter.

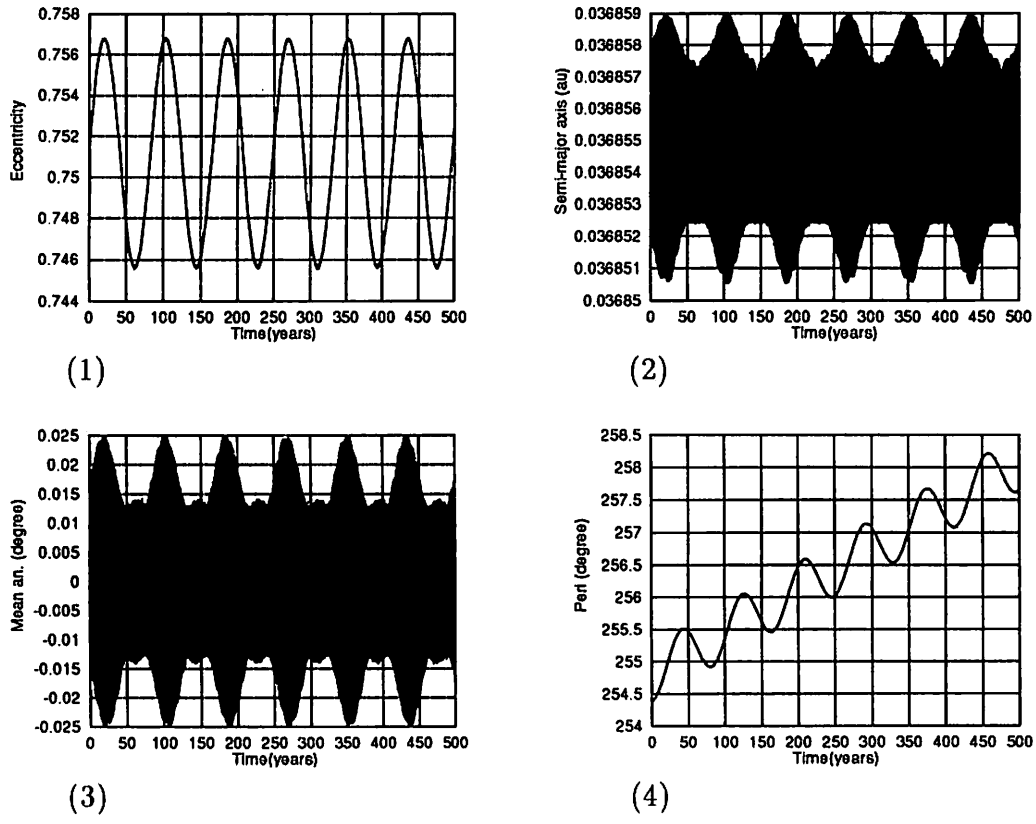
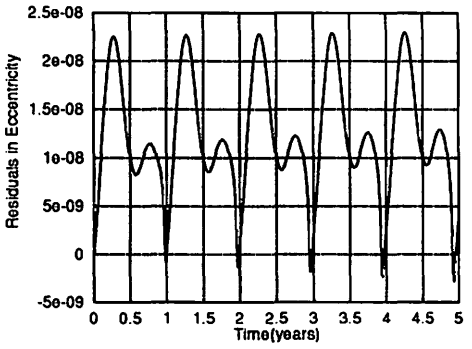
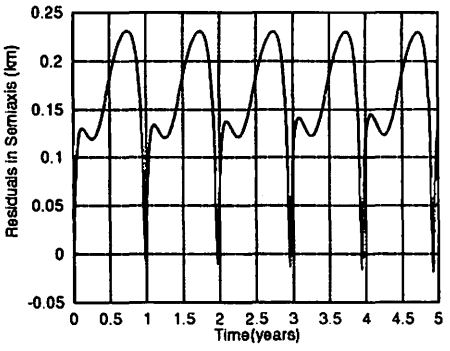


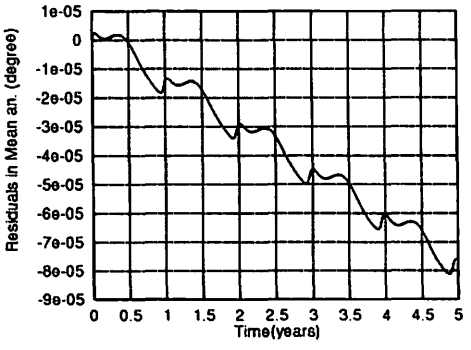
Figure 2. The osculating orbital elements of Nereid over 500 years:
(1) eccentricity,
(2) semi-major axis,
(3) periodic part of mean anomaly,
(4) longitude of pericenter.



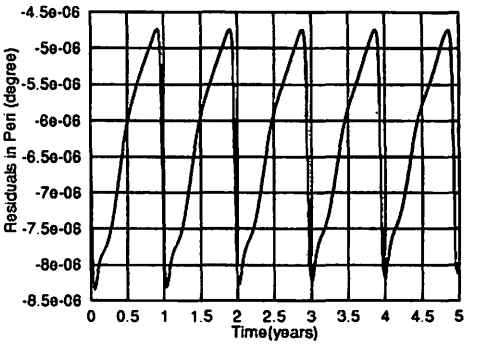
(1)



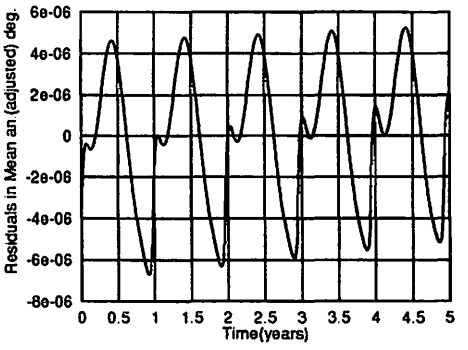
(2)



(3)

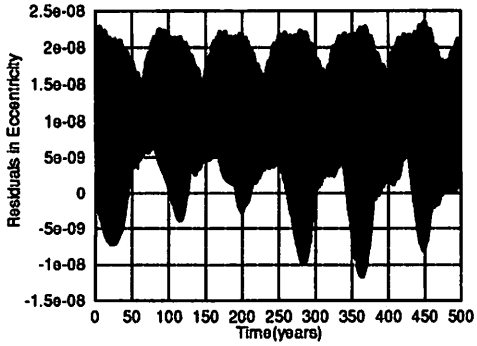


(4)

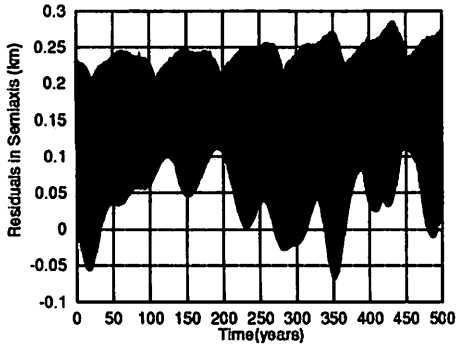


(5)

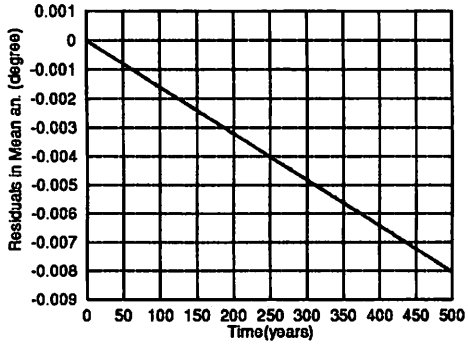
Figure 3. Difference between analytical and numerical results for the osculating orbital elements of Nereid during 5 years:
(1) eccentricity,
(2) semi-major axis,
(3) mean anomaly,
(4) longitude of pericenter,
(5) periodic part of the residuals in the mean anomaly.



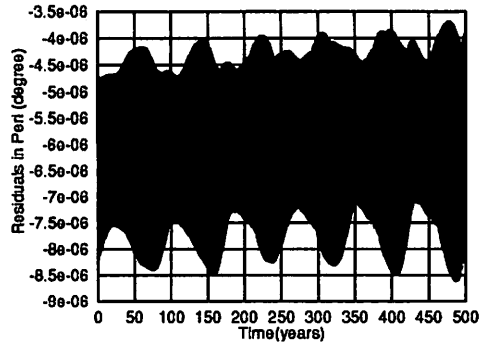
(1)



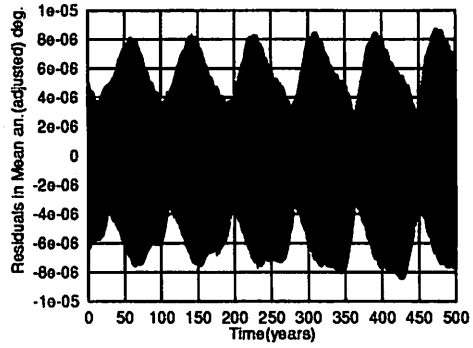
(2)



(3)



(4)



(5)

Figure 4. Difference between analytical and numerical results for the osculating orbital elements of Nereid during 500 years:
(1) eccentricity,
(2) semi-major axis,
(3) mean anomaly,
(4) longitude of pericenter,
(5) periodic part of the residuals in the mean anomaly.

In concluding the present paper, an analytical theory on the motion of a satellite with large eccentricity and zero inclination is constructed. The theory is applied to a fictitious Neptune's satellite Nereid. A comparison to the numerical integration indicates the accuracy 300 m, 3×10^{-8}, and 0.004 arc second for the semi-major axis, eccentricity and the angular variables respectively.

In order to apply the present analytical theory to actual Nereid, we have to take into account of the inclination and other perturbations such as Triton and the oblateness of Neptune, which are now in construction.

References

Brouwer, D. and Clemence, G. M.: 1961, *Methods of Celestial Mechanics*, Academic Press, New York.

Campbell, J. A. and Jefferys, W. H.: 1970,'Equivalence of The Perturbation Theories of Hori and Deprit', *Celestial Mechanics* **2**, 467-473.

Deprit, A.: 1969,'Canonical Transformations Depending on A Small Parameter', *Celestial Mechanics* **1**, 12-30.

Henrard, J. and Roels, J.: 1974,'Equivalence for Lie Transforms', *Celestial Mechanics* **10**, 497-512.

Hori, G.: 1963,'A New Approach to The Solution of The Main Problem of The Lunar Theory', *Astron. J.* **68**, 125-146.

Hori, G.: 1966,'Theory of General Perturbations With Unspecified Canonical Variables', *Publ. Astron. Soc. Japan* **18**, 287-296.

Hori, G.:1970,'Comparison of Two Perturbation Theories Based on Canonical Transformations', *Publ. Astron. Soc. Japan* **22**, 191-198.

Jacobson, R. A.: 1990,'The Orbits of The Satellites of Neptune', *Astron. Astrophys.* **231**, 241-250.

Jacobson, R. A., Riedel, J. E. and Taylor, A. H.: 1991,'The Orbits of Triton and Nereid From Spacecraft and Earthbased Observations', *Astron. Astrophys.* **247**, 565-575.

Mignard, F.: 1975,'Satellite With high Eccentricity. Application to Nereid', *Astron. Astrophys.* **43**, 359-379.

Mignard, F.: 1981,'The Mean Elements of Nereid', *Astron. J.* **86**, 1728-1729.

Rose, L. E.: 1974,'Orbit of Nereid and The Mass of Neptune', *Astron. J.* **79**, 489-490.

Segerman, A. M. and Richardson, D. L.: 1997,'An Analytical Theory For The Orbit of Nereid', *Celestial Mechanics* **66**, 321-344.

Veillet, C.: 1982,'Orbital Elements of Nereid from New Observations', *Astron. Astrophys.* **112**, 277-280.

Veillet, C. and Bois, E.: 1988,'1982-1984 Observations of Nereid. New Orbital Determination', *Astron. Astrophys.* **189**, 303-305.

van Briesbroek, G. A.: 1951,'The Orbit of Nereid, Neptune's Second Satellite', *Astron. J.* **56**, 110-111.

van Briesbroek, G. A.: 1957,'The Mass of Neptune From A New Orbit of Its Second Satellite Nereid', *Astron. J.* **62**, 272-274.

Yuasa, M.: 1970,'The Comparison of Hori's Perturbation Theory and von Zeipel's Theory', *Publ. Astron. Soc. Japan* **23**, 399-403.

カイパーベルト天体の発見条件と垂直方向分布

中村士（国立天文台）

要旨：

１９９９年１月半ば現在、発見されたカイパーベルト天体（ＫＢＯｓ）は９０個近い数に達しており、色々な統計的な議論がようやく可能なサンプル数になったと言ってよい。これらの統計的解析からＫＢＯｓのintrinsic distributionsを推定することがこの種の研究の主目的である。しかしＫＢＯｓの探査はPencil Beam Surveyと呼ばれることからも分るように、天空のごく狭い領域のみを観測しているに過ぎないので、強い観測選択効果を受けている。この効果を適切に補正してはじめて、正しいintrinsic distributionの推定ができる。本報告では、まずＫＢＯｓの発見時の条件を調べ、それに基づいて観測選択効果のシミュレーションを行なうパラメーター値を決め、観測選択効果の影響を計算した。そしてこの結果によって観測された分布に対して補正を加え、intrinsic distributionを求めた。ここでは軌道傾斜角の分布についてだけ述べる。

１．ＭＰＣ、ＭＰＥＣに掲載された発見時の日時、赤経、赤緯、見かけの等級（Ｒ）と発表された暫定的な軌道要素および絶対等級（Ｈ）から、位置推算プログラムによって黄経、黄緯、太陽離角、日心距離、地心距離、見かけの等級、などを計算した。
図１（ａ）、（ｂ）から、

・黄道が銀河と交差する黄経付近を除いて、発見はほぼ黄経の全域に分布する、
・発見は太陽離角が衝から±６０度の範囲で行なわれている、
・発見時の黄緯は、９５％が±５度の範囲に入る、
ことが分った。

２．ＫＢＯｓの適当な母集団分布を仮定し、これに基づ

く個々のＫＢＯの軌道要素をランダムに発生させて、その総数と上記の発見天域（図１ｂ）に入る数の比を観測選択効果の補正係数とした。

３．ＫＢＯｓの公転周期に比べて発見後の観測アークはごく短いので、軌道離心率（ｅ）がかなり大きなＫＢＯ以外はｅと軌道半長径の分離が良いとは考えられない。よって、ここでは衝付近の観測から最も容易に比較的正確な値が求められる軌道傾斜角（ｉ）のみに議論を限った。

４．図２（ａ）には、ｉの関数としての観測選択効果の補正係数を示す。係数の値そのものは観測視野の黄経幅に依存して不便なので、ｉ＝２．５度の係数に対する相対値で示した。ｉ＝３０度では検出確率が２．５度の時の約１／１０に低下することが分る。これは逆に、ｉ＝３０度のＫＢＯｓは観測される数の約１０倍の数が実際には存在することを意味する。図２（ｂ）の実線は、観測されたＫＢＯｓの頻度分布（ビン幅は５度）、破線は図（ａ）の係数で補正した分布でこれが真の軌道傾斜角分布に相当する。

５．補正した分布では、ｉは２０－３０度の間でへこんでいるが、これはこの範囲のＫＢＯｓがまだ数個しか発見されていないためであろう。もしそうなら、ＫＢＯｓのｉ分布は０－３０度までほとんどフラットな分布となり、Duncan-Levison(1997)がシミュレーションで計算した高軌道傾斜角のＫＢＯｓの割合より、ずっと高率で高軌道傾斜角ＫＢＯが存在すると考えられる。海王星の摂動で十分な数の高軌道傾斜ＫＢＯｓが生成されないとすれば、微惑星は始めからかなりの厚みを持って誕生した可能性も考えられる。

Intrinsic vertical distribution of Kuiper-belt objects

Tsuko Nakamura

National Astronomical Observatory, Osawa, Mikata, Tokyo 181-8588, Japan

Since the discovery in 1992 of the first member of the Kuiper belt[1] a supposedly vast reservoir of short-period comets located beyond Neptune's orbit —, nearly 90 Kuiper-belt objects have already been found[2]. With the number, we are now in a stage that a quantitative analysis in orbital statistics for the objects is feasible. As search of the Kuiper-belt objects has generally been done through narrow observation windows near the ecliptic, the apparent distributions are sometimes heavily affected by observational selections. The effects must thus be corrected to derive the intrinsic distributions, from which the origin and subsequent evolution of the Kuiper belt can rightly be discussed. Here we report the results from realistic simulations for observational selections of the bodies based on a detailed analysis of the discovery data. We find that the intrinsic number of the Kuiper-belt objects with orbital inclinations of 20–30 degrees is 5–10 times larger than the observed number. This implies that extent of the Kuiper belt vertical to the ecliptic plane is substantially denser and thicker than previously believed.

According to Minor Planet Center, International Astronomical Union[2], orbital elements for 88 Kuiper-belt objects (KBOs) are listed as of January 1999. Their discovery situations are also reported in Minor Planet Electronic Circulars (MPEC) issued from the same organization. Since the sky coverage by now available charge-coupled device (CCD) chips is limited to less than 30′ square[3,4], it is a common practice for KBO surveys to be made very near the ecliptic, aiming at captures of both low- and high-inclination KBOs there.

Figure 1 represents the distribution of discovery points on the sky, which was cal-

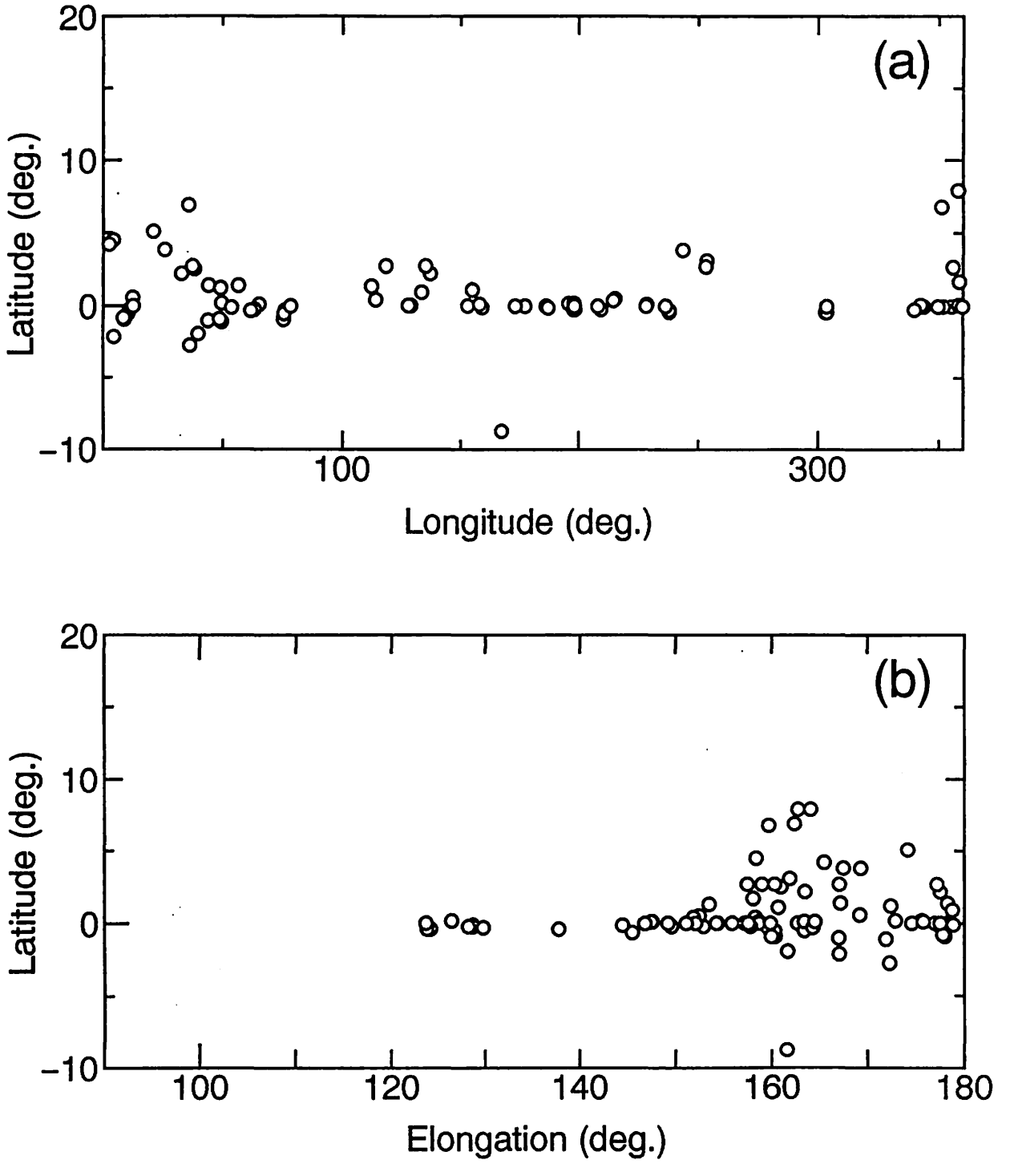


Fig.1

(a) Heliocentric longitude vs. latitude plot for discovery positions on the sky for the 88 known KBOs. The positions were calculated with a two-body ephemeris generator using MPC orbital elements in the J2000 coordinate system. The longitude is measured from the vernal equinox point and the latitude of zero represents the ecliptic plane.

(b) Elongation angle vs. latitude plot. Elongation is meant by the Sun-Earth-Object angle.

culated from MPEC data. Figure 1(a) is the heliocentric longitude vs. latitude plot for the 88 KBOs, with zero latitude denoting the ecliptic. One can see that discovery longitudes span all over the sky, except the Milky Way crossings where high star density hampers detection of faint objects. Figure 1(b) shows the elongation vs. latitude diagram for the discovered KBOs, where elongation is the Sun-Earth-Object angle. That the angle scatters over the sky as large as $\pm 60^o$ means the KBOs are not necessarily observed near opposition unlike the belt asteroids, because of remote distances. We also see from Figure 1 that discovery latitudes are confined in the region of $\pm 5^o$; only four KBOs were found outside this region. Such localization in discovery causes an observational selection heavily biased toward apparent predominance of low-inclination orbits.

In order to grasp quantitatively such selection effects, we simulated observing conditions for the KBOs rigorously; this is a byproduct from more extensive simulations for our planed survey observations of sub-km main belt asteroids with 8m-class telescopes. Considering distributions in orbital elements for the known KBOs, hypothetical KBOs were produced, using a uniform random-number generator, with the semi-major axis (a) of 42 ± 2 AU, the eccentricity of 0.1 ± 0.1, and angular elements (except inclination) of $0 - 360^o$. Inclination was varied within $\pm 1^o$ around a specified value for a run. We calculated exact "astrometric" sky positions seen from the Earth center in the J2000.0 coordinate system, by modifying a computer program which had long been used by us for ephemeris calculations of asteroids and comets. It is assumed that observations are made on the ecliptic at opposition of March 20, through a rectangular observation window. In accordance with the observing situations given in Figure 1(b), we set up a window of $120^o \times 10^o$ centered at opposition. For each hypothetical KBO, the positional calculation was made for March 20 and this process was repeated until the number of "in-window" objects amounts to 500 in typical cases, and the ratio between the total number of generated KBOs and the in-window number was regarded as a detection probability (or observational bias). Magnitude effects due to size distribution were not taken into account here.

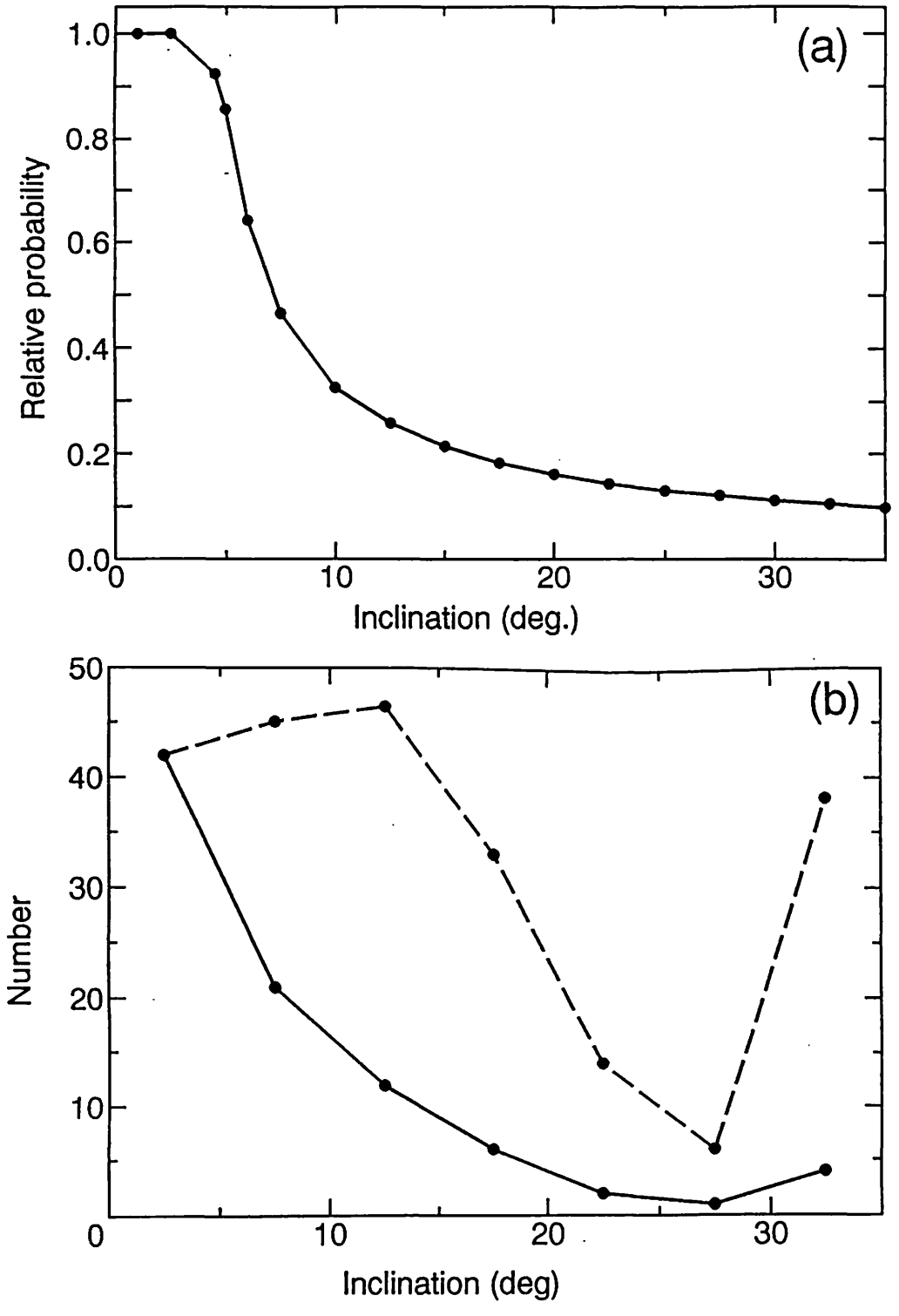


Fig. 2

(a) Relative detection probability for the KBOs as a function of inclination. The ordinate is normalized to the probability at $I = 1^o$ where the absolute probability was 0.30. Probabilities were estimated as observed through a window of $120^o \times 10^o$ on the ecliptic centered at opposition of March 20.

(b) Observed (solid curve) and corrected (dashed curve) inclination distributions

Figure 2(a) represents the relative detection probability as a function of inclination (I) for the above assumed window. The ordinate is normalized at $I = 1^o$, whose absolute probability was found to be 0.30. One can see that the probability sharply decreases at about $I = 5^o$ toward larger inclinations; this is due to the window size of $\pm 5^o$ in latitude. It is also seen that probabilities at $I = 20 - 30^o$ are 1/6–1/10 smaller than that for $I = 1^o$. This implies that if there were the same number of objects in orbits with $I = 1^o$ and $I = 20 - 30^o$, the latter objects will be detected less than the former by the above factor, at least statistically. To check robustness of the results, calculations were also made for different reasonable ranges in elongation angle, eccentricity, and for $a = 70$ and 100 AU, resulting essentially in the same values as figure 2(a).

We show in figure 2(b) the observed I-distribution for the observed 88 KBOs as a solid curve. Inclinations are divided into 5^o-interval bins and mid-values for each bin are indicated as filled circles. The observed distribution is characterized by (i) a strong concentration to low-inclination orbits, with FWHM of $6 - 7^o$ (although a few additional discoveries of low-inclination KBOs was reported as of this writing, they do not affect much the overall trend of the distribution; see MPECs), and (ii) relative abundance of the KBOs with $I \geq 30^o$. The fact (ii) is fairly certain, because inclination is one of orbital elements most easily determined only from apparent motions on the sky at opposition. Since our simulations reveal that high-inclination KBOs have much less chance of detection than low-inclination ones in the current KBO-survey scheme, this effect must be considered to obtain the true I-distribution. The dashed curve in figure 2(b) is the intrinsic distribution corrected by dividing by probabilities shown in figure 2(a). The corrected curve in figure 2(b) seems to suggest that the intrinsic I-distribution of the KBOs can be roughly flat up to 30^o, as reality of a dip seen in $25 - 30^o$ range will still be highly uncertain because of small sample number of only one there.

An approximate analytical form for inclination bias had already been discussed in papers by Jewitt and Luu[5,6]. we confirm that their formula gives basically correct

probabilities, except for the range in which inclination is approximately equal to half the latitude width of the observation window, where numerical estimates are indispensable. However, seemingly they were satisfied by applying their formula, with their CCD-chip size (e.g., 0.12^o)[5] instead of our adopted $\pm 5^o$ as the window height, for the purpose of comparing the I-distribution for the KBOs with that for short-period comets, hence without showing the corrected distribution concretely.

The dashed curve in figure 2(b) indicates that the actual Kuiper belt will extend in the direction perpendicular to the ecliptic plane much thicker than previously believed, with the depth of $2 \times 40 \tan(\sim 30^o) = \sim 46$AU. Some recent theoretical studies seem to allow the existence of such high-inclination KBOs, in addition to the primordial near-ecliptic members. Through orbital integrations of many hypothetical KBOs over 4 billion years, Duncan and Levison[7] could obtain a group of the KBOs scattered by Neptune into the orbits with inclinations of as high as $\sim 30 - 40^o$. They refer to the KBO member 1996 RQ20 as such a candidate.

Finally we note that a similar bias may affect the inclination distribution of the faint main belt asteroids more seriously than usually recognized, since they will be observed more or less near the ecliptic.

1. Jewitt, D. & Luu, J. Discovery of the candidate Kuiper belt object 1992 QB1. Nature **362**, 730-732 (1993).

2. Marsden, B.G. List of transneptunian objects (1998 Nov.); available at http://cfa-www.harvard.edu/iau/lists/TNOs.html.

3. Jewitt, D., Luu, J., & Trujillo, C. Large Kuiper belt objects: The Mauna Kea 8K CCD survey, Astr. J. **115**, 2125-2135 (1998).

4. Gladman, B. et al. Pencil-beam surveys for faint trans-Neptunian objects. Astr. J. **116**, 2042-2054 (1998).

5. Jewitt, D.C. & Luu, J.X. The solar system beyond Neptune. Astr. J. **109**, 1867-1876.

6. Jewitt, D., Luu, J. & Chen, J. The Mauna Kea–Cerro-Tololo Kuiper belt and

Centaur survey. Astr. J. **112**, 1225-1238 (1996).

7. Duncan, M.J. & Levison, H.F. A disk of scattered icy objects and the origin of Jupiter-family comets. Science **276**, 1670-1672 (1997).

すばる望遠鏡による微小小惑星のサーベイ観測計画

— 真の分布の推定法 —

Sub-km Belt Asteroid Survey Observations with SUBARU Telescope

— Estimation of the Intrinsic Distribution —

吉田二美(福岡教育大学・天文学教室)
Fumi YOSHIDA
(Dept. of Earth Science and Astronomy, Fukuoka University of Education, Japan)
中村 士(国立天文台)
Tsuko NAKAMURA (National Astronomical Observatory, Tokyo)

Abstract

Past asteroids survey could determine the size distribution down to 1km in diameter. We propose here an asteroid survey plan with 8.2m SUBARU telescope. We estimate that this survey can detect as many as 4000-10000 new asteroids whose limiting magnitude (in H) is about 21 from observations of a few nights. H=21 correspond to about D=200m for S-class and D=400m for C-class asteroids respectively. That is our survey results allow providing us a new view in size distribution of asteroids.

This report discusses observational strategy, expected results, and correction methods of observational bias for our proposed survey plan.

Here are the three main points of this study:

1) From apparent motion vectors of asteroids, the size distributions can be statistically determined as function of semi-major axis and inclination.
2) In order to estimate the intrinsic distribution from the observational data, bias correction functions were calculated by Monte Carlo simulation.
3) Our calculation takes account of non-uniform distribution in angular variables and light variation of faint asteroids. These results give considerably different bias corrections from the previous similar researches that treated only uniform distribution.

I. はじめに

小惑星の全体的なサイズ分布及び空間分布を推定するという観点から行われた最初の系統的な小惑星サーベイは、1950 年代に行われた McDonald Survey(MDS)である。このサーベイにより直径約 10km までの小惑星の累積度数分布が求められた。1960 年の Palomar-Leiden Survey(PLS)では、直径約 1km までの小惑星の累積度数分布が求められた。

小惑星のサイズ(または絶対等級)分布は次のようなべき乗則で表されると言われ、

$$N(>r)=r^{-b} \qquad (1)$$

PLSでは b=1.75 であると決定された(Soberman et al. 1971)。これは衝突破壊によって生じる破片のサイズの分布に類似している。しかしながら、理論モデルは、観測から得られる値より大きいべき指数を与えている(例えば Dohnanyi のモデルでは b=2. 5)。b の値が大きいということは、小さい小惑星の数が多いということである。一方、衝突によって砕かれた小惑星の小さい破片が大きな破片の重力に再び捕らえられ、緩く集積する Rubble pileモデルというモデルがある。このモデルでは小さな破片の数が相対的に少なくなり、b の値は小さ

くなるはずである。このような衝突の様相を探るには、現在までの観測で明らかになったものより小さい小惑星のサイズ分布を明らかにすることが重要となる。

そこで、今年2月にファーストライトを迎えたすばる望遠鏡の主焦点に広視野 CCD カメラ(Suprime-Cam)を装着すると、約 26 等の限界等級が達成でき、この観測システムで小惑星帯の小惑星を見ると、数百 m 級の小惑星を検出できる計算になる。微小な小惑星ほど継続した衝突を繰り返してできたとすれば、このサイズ領域の小惑星は大きい小惑星より小惑星帯の衝突進化の履歴をそれらのサイズ分布や空間分布に反映していると考えられ、この領域の小惑星を観測することで、必ず未発見の興味深い性質が見つかると想像される。そこで、本研究ではすばる望遠鏡で小惑星サーベイを行うことを目的に、その前段階としての準備を行った。

一つは、微小な小惑星は、その不規則な形状から大きな変光が予想されるため、追跡観測が極めて困難であることが予想される。そこで本研究では最低 2 回の観測で同定できる小惑星の見かけの運動ベクトルから、小惑星の軌道長半径(a)と軌道傾斜角(i)の関数としてサイズ分布を統計的に決定する(中村*1 1997)。

さらに、小惑星サーベイ観測は、制限された視野(観測範囲)、有限の等級までという常に制限された観測であるため、得られる観測データには偏りが生じる。ここではこの偏りの補正法を開発し、その補正量を Monte Carlo シミュレーションにより計算した。計算では、小惑星の軌道要素の角度要素についての非一様な分布と変光の効果を考慮した。

II. すばる望遠鏡による小惑星サーベイ観測計画

小惑星サーベイの場合、まず観測日時の選定が重要となる。小惑星は黄道面に集中しており、衝のとき最も明るくなる。さらに背景に恒星の少ない星野を選ぶとなると、観測時期は春分点・秋分点が衝になる秋分か春分の日が適している。

Suprime-Cam の 1 フレームの視野は、30′ × 24′ で露出時間 5 分で V バンドに対して S/N=5 の限界等級は約 26 等が達成できる。このサーベイでは1フレーム内の星野を約1時間の時間間隔をおいて 2 回撮像する。このとき達成できる掃天面積は望遠鏡の占有時間を考慮して、数夜の観測で 20 平方度程度であろう。以上のことからこのサーベイの観測視野は太陽離角 177.5° ～ 182.5° 、黄緯±2° の領域とする。

このシステムで検出できる小惑星は、小惑星帯の最も外側(3.3AU)に対して絶対等級で H～21 等までの小惑星を検出できる。H～21 は、S 型小惑星で直径およそ 200m、C 型小惑星でおよそ 400m の天体に相当する。

さらに検出できる小惑星の数をPLSとの比較から見積もると、PLSでは限界等級 V=19.2～20.2 等より明るい小惑星が1平方度当たりに約 7 個写った。これをすばる＋Suprime-Cam の限界等級であるH～21 等でまで外挿すると、このサーベイでは1平方度当たり 240～530 個の小惑星が検出できる計算になる。約 20 平方度の掃天では、検出できる小惑星の総数は約 4,800～10,600 個になる(中村*2 1997)。

図 1 に現在までに軌道がほぼ確定している小惑星約 4 万個の絶対等級の分布とすばる＋Suprime-Cam サーベイで得られるはずの小惑星の絶対等級の分布を示す。絶対等級の小さい、明るい小惑星は絶対的に数が少ないので、すばる＋Suprime-Cam サーベイで検出できる数は少なく、かわりに、今まで明らかになったものとは全く違ったサイズ領域の小惑星を検出できることがわかる。

III. 軌道要素の推定

一つの領域を時間間隔をあけて2回撮影した画像を比較することによって、小惑星の見かけの移動速度(Daily motion)とその移動方向(Position angle)を知ることができる(中村*1 1997)。小惑星帯に属する小惑星については、この二つの情報、すなわち、見かけの移動速度と移動方向から、大ざっぱではあるが、a と i の分布が図 2, 3 により知られる。また i と観測時の小惑星の位置(赤経、赤緯)の情報から小惑星の軌道が黄道面と交わる位置が計算でき、その小惑星の昇交点経度(Ω)が求まる。Ωと観測時の小惑星の位置(赤経、赤緯)より、観測時点(t_0)での小惑星の経度 $\lambda(t_0)$が知られる。

この方法では、軌道を決めるのに必要な 6 つの

軌道要素の内、4 つの要素(a、i、Ω、λ(t_0))しか決定することができないが、残りの 2 つの要素は現在観測から得られている分布が、明るい(大きい)小惑星と暗い(小さい)小惑星で似ていることから、現在の観測から外挿することができる。さらに周りの基準星との比較から小惑星の見かけの等級もわかる。これらにより、小惑星帯の全体的な空間分布とサイズ分布を統計的に求める目的においては、満足できる情報が得られると考えている。

Ⅳ. 観測選択効果の補正

Ⅳ－1 観測選択効果とは

サーベイで得られたデータは、観測視野と等級が限られているために、見かけ上真の分布から偏りが生じている。例えば、

1) 黄道面を中心にサーベイした場合、大きなiの小惑星は発見されにくい。
2) 遠くの小惑星より、近くの小惑星の方が小さいサイズのものまで見える。
3) 小惑星帯の外側で暗くて、e が小さい小惑星は、それらが限界等級内にあることはそれほどないと考えられるから発見される確率は低い。しかし e が大きければ、その小惑星が観測日にたまたま近日点近くにあれば、発見されるだろう。
4) 近くの天体は動きが速いため、画像面積当たりの等級は、露出時間と見かけ等級が等しい恒星の場合より暗くなる。
5) 動きの遅い小惑星は露出と次の露出の間に移動せず、恒星と区別ができず、発見されないかもしれない。
6) 反射率の高い小惑星は小さくても明るく見えるため、発見されやすい。

このように、ある観測で、ある小惑星は発見されやすかったり、別の小惑星は発見されにくかったりすることを観測選択効果と呼び、その結果、得られたデータに生じる偏りを観測 bias と呼ぶ。上のような観測選択効果の考察から、観測 bias は小惑星の各軌道要素、絶対等級、小惑星と地球と太陽の位置関係、観測システムの検出効率等の複雑な関数で表されると考えられる。

Ⅳ－2 観測選択効果の数理的取り扱い

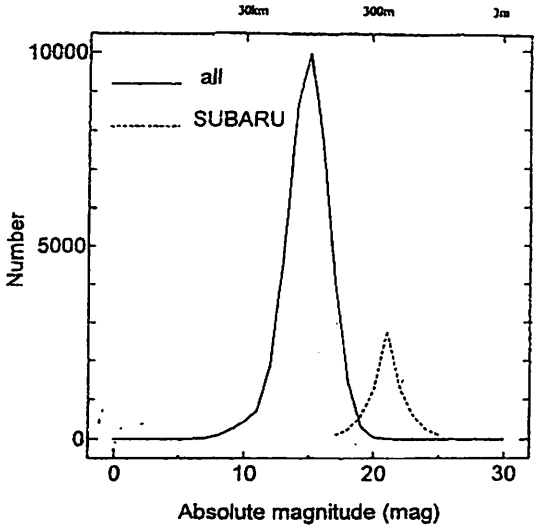


Fig.1 H-magnitude distribution for 40706 asteroids observed at more than two oppositions(solid), and the distribution expected from SUBARU survey observations(dashed).

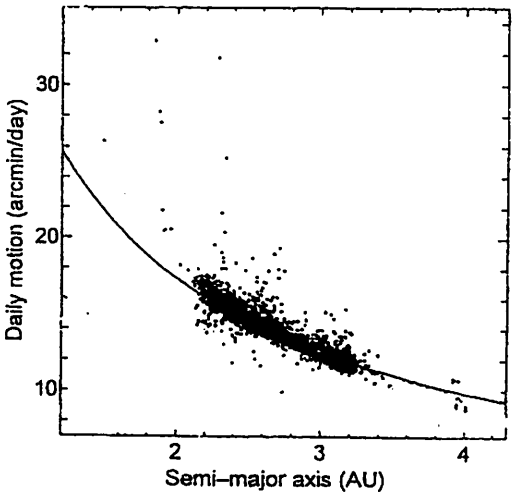


Fig.2 Semi-major axis vs. daily motion correlation.

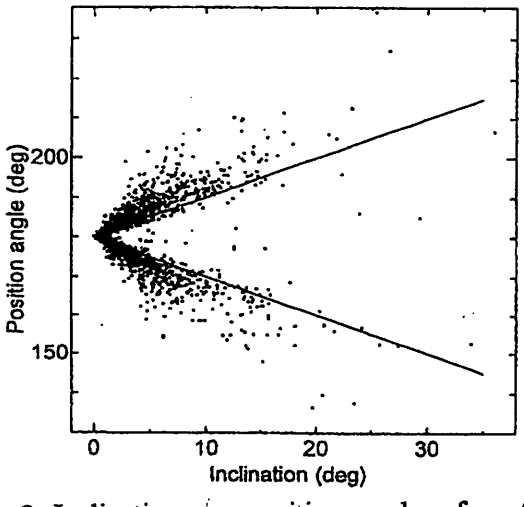


Fig.3 Inclination vs. position angle of motion correlation.

ここではすばる＋Suprime-Cam システムが検出する小惑星の発見確率が、すべての軌道要素(a, e, i, ω, Ω, M)と絶対等級Hに依るものと仮定し、数値シミュレーションにより観測 bias を見積もり、発見確率を経験式の形で求めることを目指した。

簡単には、観測 bias がある要素(E)に依るとし、観測される小惑星の数を f(E)、真の小惑星の数を F(E)とすると、 E～E+dE の間にある小惑星の数は、

観測される数: dn＝f(E)dE,
真の数:dN＝F(E)dE

で表され、ここで

$$dn = B(E)dN$$

で定義する Bias 関数、B(E)を使うと

$$B(E) = f(E) / F(E) \qquad (2)$$

と書ける。なお、

$$|B(E)| < 1$$

である。F(E)は現在までの軌道確定小惑星のサイズ分布、軌道分布より仮定し、f(E)はここでは1000個として、Ⅴ－3 に述べる数値シミュレーションから B(E)を計算する。B(E)を決めておけば、実際にすばるで観測して f(E)が得られると、(2)式の関係から真の分布 F(E)が推定できる。B(E)は f(E)と F(E)の比であるから、F(E)の分布形に直接依存しない。したがって、B(E)の関数形をシミュレーション(またはモデル)から決めることができる。

Ⅳ－3 数値シミュレーション

Bias 関数、B(E)を求めるために、次のような数値シミュレーションを行った。

(1) 各軌道要素を人工的に発生させる。
a, e, iについては、小惑星帯の内側、中央、外側で、また円軌道に近い小惑星や黄道面上の小惑星に限定して調べるために、入力した数値を基準に、ある任意の幅内でランダムに発生させた。入力値は現在知られている小惑星の軌道要素の分布をもとにしている。
ω, Ω, Mについては一様分布を仮定し、0°～360° でランダムに発生させた。

(2)(1)で決めた軌道要素をもつ小惑星の位置(赤経、赤緯)を位置推算プログラムで計算する。

(3)その小惑星がすばるの観測視野内に位置し、かつ限界等級以内にあれば、その小惑星を発見されたものとしてカウントする。

(4)観測視野内に入る小惑星の確率を発見確率として次のように定義する。

$$\text{発見確率} = \frac{\text{観測視野内に入った小惑星の数}}{\text{発生させた小惑星の総数}}$$

このようなモデル化された発見から、発見確率は観測 bias を補正する量に相当する。

Ⅳ－4 発見確率の計算

シミュレーションは次の3種類行った。

① 角度要素(ω, Ω, M)の一様分布を仮定して計算(図 4, 5, 6)。
② Ωの非一様分布(図 7)を考慮して計算(図 8, 9, 10)。
③ 変光の効果(図 11)を考慮して計算(図 12, 13, 14)。

まず①の場合、角度要素(ω, Ω, M)については一様分布を仮定し、i, a, eについてそれぞれ値を変えて上記のシミュレーションを行い、発見確率を計算した。この際小惑星の絶対等級は 10.0 等に固定し、観測視野内に入るすべての小惑星を発見できることを前提にした。こうして計算したiの関数としての発見確率を図4に示す。小惑星帯の領域別には発見確率の差はあまり見られないが、黄道面以外では発見確率が急速に低下することがわかる。

図 5 はaの関数として計算した発見確率である。実線は e=0.2±0.1、i=1±1、点線は e=0.01±

0.005、i=0.01±0.005という条件を与えて計算した。破線は小惑星の運動を平面内、円運動とした簡単なモデル計算の結果である。aの値が大きくなるほど発見確率は増加し、e の効果を入れると発見確率が低下することがわかる。

図 6 はeの関数として計算した発見確率である。観測視野内に入るすべての小惑星が見えているとした場合、発見確率は e の値によらず、一定であることがわかる。

次に②の場合であるが、現在までに軌道がほぼ決定されている小惑星のΩの分布を図 7 に示す。この図は、小惑星を絶対等級で 3 種類に分け、明るい小惑星のΩの分布は実線、暗い小惑星は破線、中間の小惑星は点線表した。Ω～0°，180° 付近で多く、90° 270° 付近が少なくなっている。これは、銀河と重なる領域で掃天が避けられていることによる観測選択効果の可能性もあるが、Ω～270° では明るい小惑星も含めて数が少なくなっていることから、真の分布を反映しているように見える。そこでシミュレーションの最初で発生させる軌道要素Ωにこの非一様な重みをかけて計算した(この際小惑星の絶対等級は①の場合と同じく 10.0 等に固定した)。Ωは黄経～0°，180° 付近に集中しているため、観測視野の中心が黄経 0°，180° の方を向いている春分の日や秋分の日に観測した方が発見確率が高くなると容易に予想されるが、その結果を図 8, 9, 10 に示す。

最後に③の場合である。微小小惑星はその不規則な形状から変光が大きいと考えられ、観測システムの限界等級付近では、この効果は特に重大になる。EMP1998 年版の変光パラメーターのデータより、比較的暗い絶対等級 10.8～19.03 等の 291 個の小惑星の変光の分布は図 11 のようになる。この分布をシミュレーションプログラムに組み込み、発見確率を再度計算した結果が図 12, 13, 14 である。この際小惑星の絶対等級は 20.0 等に固定した。変光の効果を入れた場合、入れなかった場合と比べて発見確率が大きく異なり、かなり違った観測biasを与えることがわかる。

V. まとめ

1. すばる望遠鏡に Suprime-Cam を装着して小惑星サーベイ観測を行えば、小惑星帯で直径

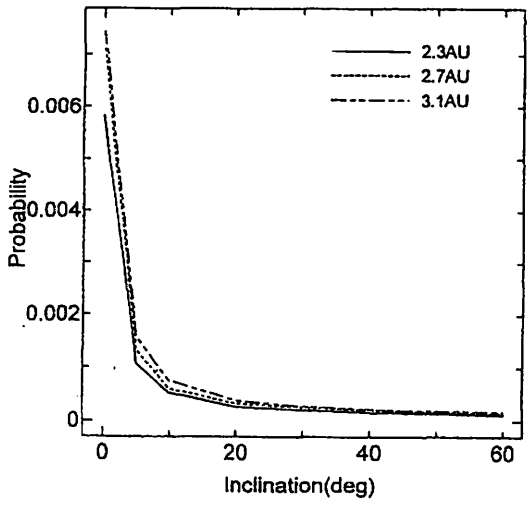


Fig4. Variation of probability as a function of inclination.

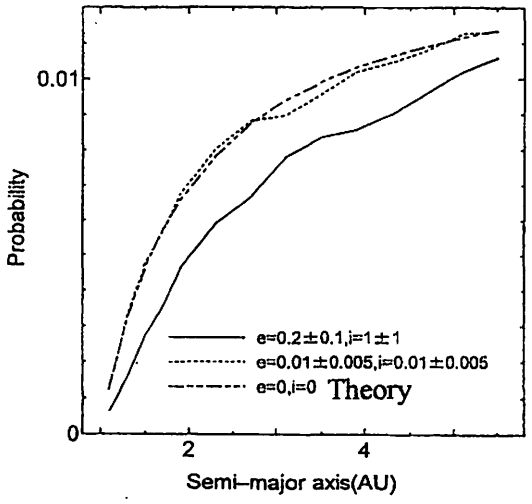


Fig5. Variation of probability as a function of semi-major axis.

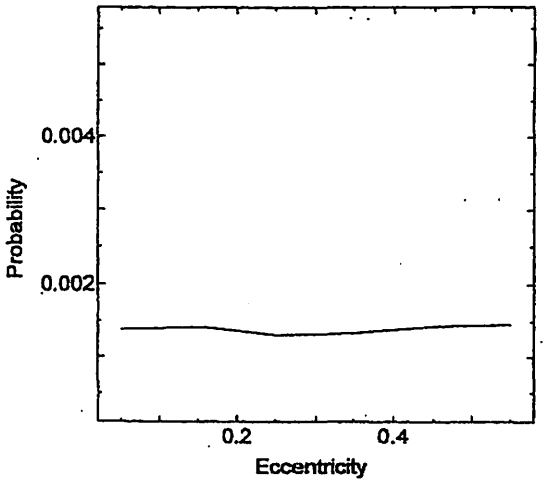


Fig.6 Variation of probability as a function of eccentricity.

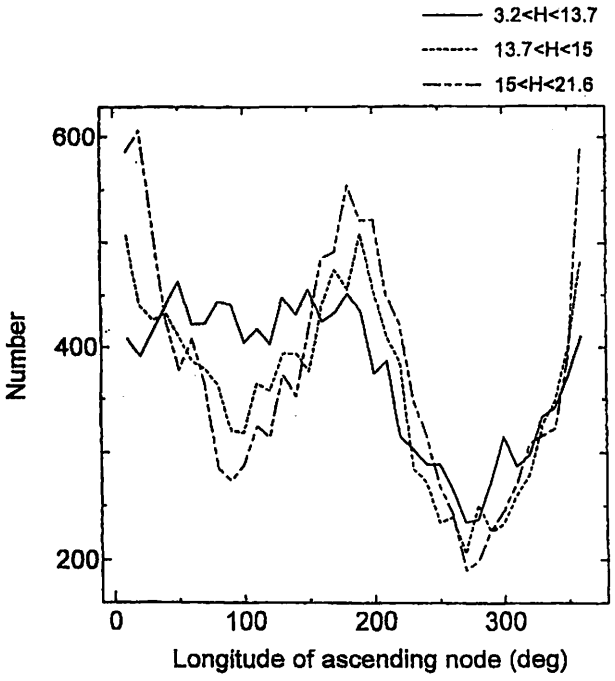


Fig.7 The non-uniform distribution in longitude of ascending node of main belt asteroids.

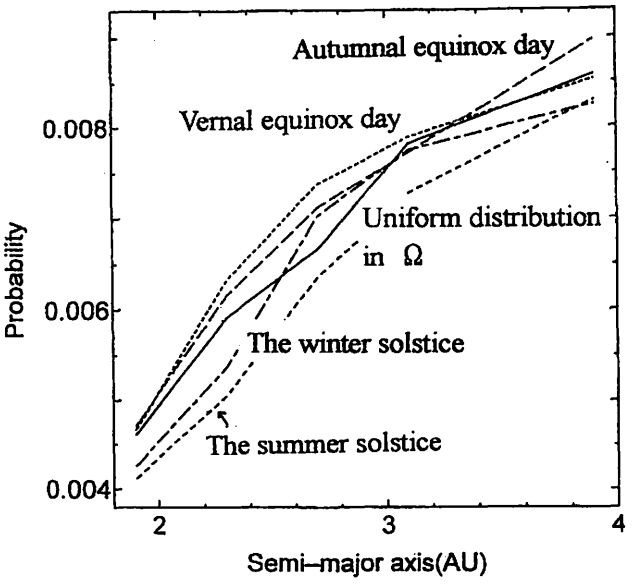


Fig.9 Variation of probability as a function of semi-major axis which is taken account of non-uniform distribution in Ω.

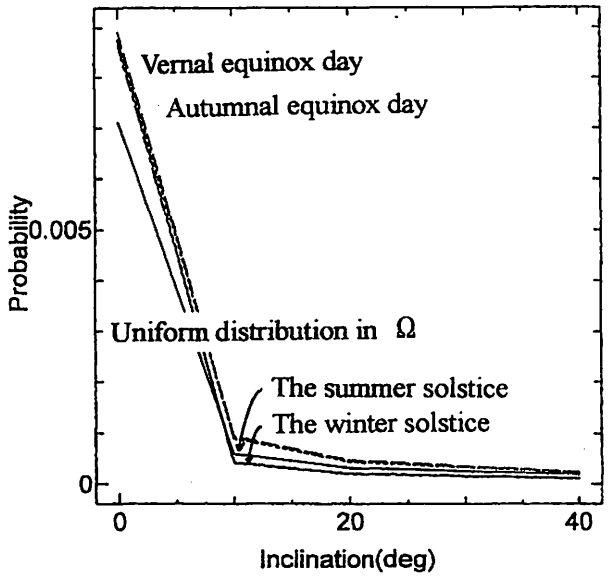


Fig.8 Variation of probability as a function of inclination which is taken account of non-uniform distribution in Ω.

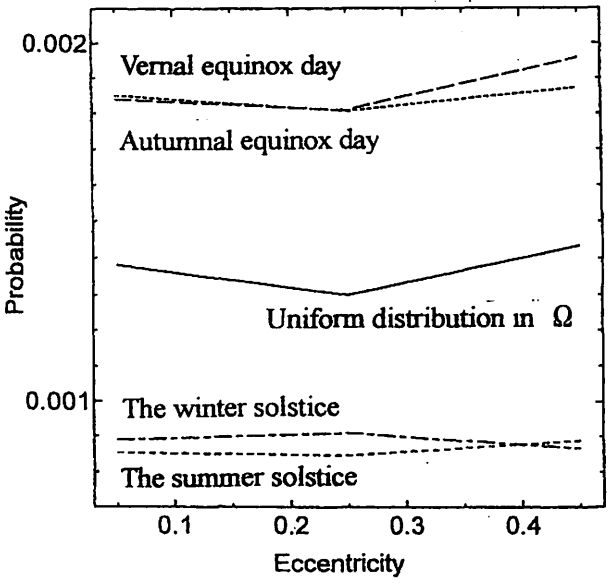


Fig.10 Variation of probability as a function of eccentricity which is taken account of non-uniform distribution in Ω.

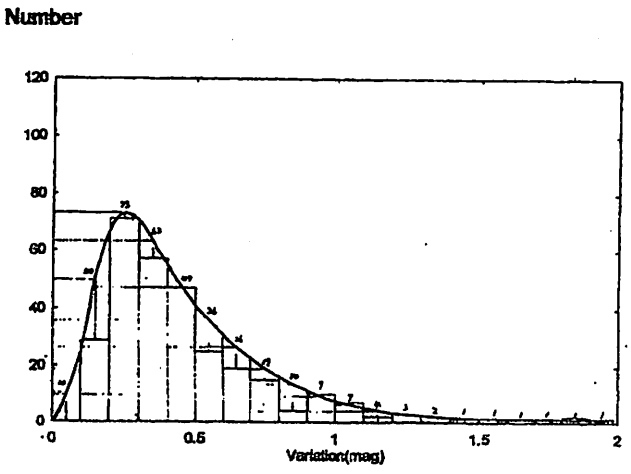


Fig.11 The distribution in light variation of known faint asteroids.

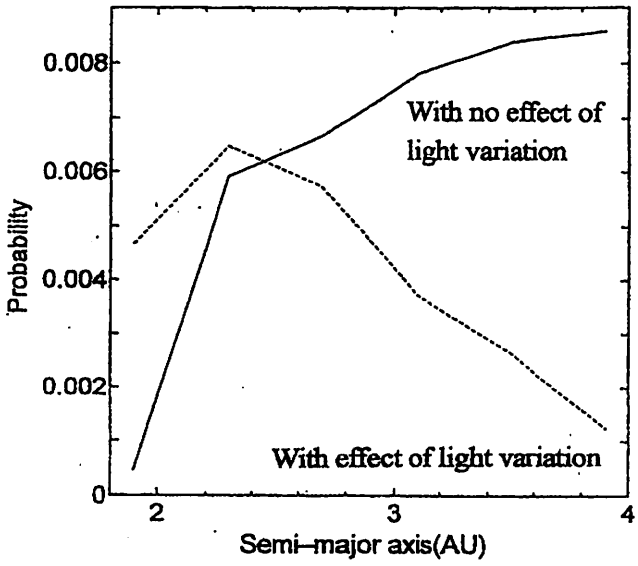


Fig.13 Variation of probability as a function of semi-major axis near limiting magnitude.

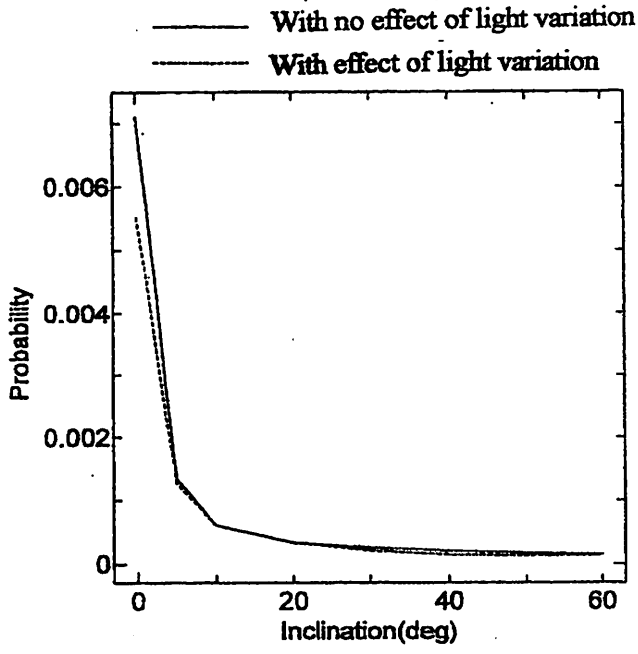


Fig.12 Variation of probability as a function of inclination near limiting magnitude.

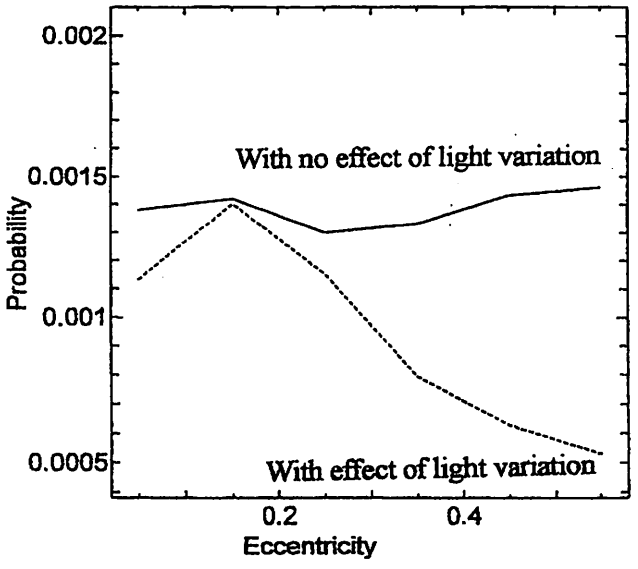


Fig.14 Variation of probability as a function of eccentricity near limiting magnitude.

数百m級の小惑星が検出できることがわかった。このことにより、現在得られている小惑星のサイズ分布及び空間分布を1桁小さいサイズ領域まで求めることができる。

2. 微小小惑星の追跡観測を避けるために、小惑星の運動ベクトルだけから、その小惑星の4つの軌道要素を推定する方法を開発した。

3. すばる＋Suprime-Camで得られる観測データの観測選択効果の補正方法を開発し、その補正量を数値シミュレーションにより計算した。
非一様な角度要素の分布（Ωについて）や観測システムの限界等級付近で重要な変光の効果を考慮すると、観測日時や限界等級により補正量が変化することが判明した。非一様分布と変光を考慮したことは、先の研究にない新しい試みである。

4. さらに、ここでは計算には入れなかったが、分光型の違いによる効果もプログラムに組み込む必要がある。というのは、大きな小惑星では小惑星帯の内側付近で相対的にS型小惑星が多く、外側付近でC型の小惑星が多いことが知られていて（Gradie & Tedesco 1982)、微小な小惑星でこの分布をどうなっているかは明らかでないが、おそらく大きな小惑星と同様、軌道長半径に沿って分光型の分布に構造があると考えられるからである。
この分布はS型とC型の小惑星がB－V＝0.8を境に明確に分かれることから（Zellner 1979)、我々の計画したサーベイをVバンドとBバンドで1回ずつ撮像し、小惑星の光度を測定することでおよそ知ることができる。

以上のようにして、補正量を様々なケースで算出し、実際の観測データに適用すれば、すばるでの小惑星サーベイ観測は、今までより1桁小さい小惑星を検出できることから、まだ発見されていない微小小惑星までを含めた小惑星帯全体のpopulationが明らかにできると予想される。
また、小惑星のサイズ分布は現在までのところ、発見された小惑星全部をひっくるめた議論しかなされていない。しかし、ここでの研究のように、小惑星の各軌道要素ごとに統計をとることによって、例えば、小惑星帯の領域ごとのサイズ分布の違いや、e の違いよるサイズ分布の相違など、小惑星の空間分布とサイズ分布を結びつけた議論が可能になろう。
ここでの結果は、2000 年以降予定しているすばるでの微小小惑星サーベイで得られるデータに適用する予定である。

参考文献

Gradie, J. C., Chapman, C. R., & Tedesco E. F., 1989, in AsteroidsII, (Binzel, R. P., Gehrels,T.& Matthews, M. S. Eds.), 316-335. Univ. of Arizona Press, Tucson.

Kuiper, G. P. et al., 1958, Astrophys.J. Suppl., 3, 289-427.

中村士*1, 1997, 第 29 回天体力学研究会集録 274-278

中村士*2, 1997, 第2回すばる望遠鏡ファーストライトシンポジウム集録 70-75

Soberman, R. K. et al.,1971, NASA SP-267, 617-631.

Van Houten, C. J. et al., 1970, Astr. Astrophys. Suppl., 2, 339-448.

Zellner, B., 1979, in Asteroids (Gehrels, T., Ed.), 783-806. Univ. of Arizona Press, Tucson.

太陽系周縁空間での小天体の軌道計算

Orbit Integration of Small Celestial Bodies in the Interstellar Space around the Solar System

三堀太平・鷲見治一（湘南工科大学工学部）
Taihei Mitsubori and Haruichi Washimi
Faculty of Engineering, Shonan Institute of Technology

Abstract

The recent development of the study of the outer heliosphere is quickly reviewed and the orbit of small celestial bodies in the interstellar space around the solar system with a scale of a few tens light year is studied by the 4th order Runge-Kutta method. It is shown by some of examples how the small bodies wander in the interstellar space under the complicated gravitational potential due to many nearby stars, before they come into the interplanetary space of our solar system.

１．まえがき

人工飛翔体の活動領域は現在のところは惑星間空間までであるが、パイオニア及びボイジャー探査機は太陽圏外圏へ近づきつつあり、近い将来は太陽圏外圏、更には銀河系（当面は太陽圏のごく近傍のみに限られるであろうが）へと拡がろうとしている。これに伴い、この分野の研究は国際的に活発に行われて来ている。本報告ではまず、この太陽圏外圏の研究の現状について簡単に紹介し、次に、今後わが国でも関心が高まるであろうと期待される太陽圏外圏を含む太陽系周縁空間に関する教育の例として、我々の大学での卒業研究（工学部情報工学科４年生）として著者の一人（三堀）が行った小天体の軌道計算及びそのグラフ表示について紹介する。

２．太陽圏外圏の研究の現状

１９５７年のスプートニク打ち上げ以来現在に至る宇宙観測は一口に云えば、「惑星探査」と表現することが出来よう。しかし、惑星系を越えてはるか銀河空間に向かう計画は早い時期から検討されていた。即ち、１９６４年にＮＡＳＡの Naugle 博士は当時目白押しの探査計画の中にあって、太陽系を越えてはるか銀河系へ到達しようとする画期的な計画、「銀河プローブ」のアイデアを打ち出した。この計画は高エネルギー研究者達を中心に強い支持を受け、大型プロジェクトとしては非常にすみやかに準備が進み、１９７２年・７３年打ち上げのパイオニア１０・１１号として実現し、さらに１９７６年のヴォイジャー１号・２号へとつながって行った。今これら４機の探査機は外惑星の探査を終え、５５ＡＵ～７０ＡＵにあり、ちょうど本研究会でも話題にな

っているカイパーベルトのあたりに到達していて、引き続き太陽圏境界を目指して約２ＡＵ／年の速度で飛行中である。

パイオニア打ち上げ以来既に２７年、ボイジャーからでも２２年（卒業研究生の年齢とほぼ同じ！）経過したわけであるが、この間これら探査機自身による深宇宙空間のプラズマ・惑星間空間磁場及び高エネルギー粒子の観測や光及び宇宙線観測、更に理論・シミュレーション解析により、星間ガスの組成・温度・相対速度や外圏の構造が序々に明らかになって来ている。太陽系を越えればそこはもう銀河系の世界であり、２１世紀の宇宙探査は「惑星探査から銀河探査へ」の飛躍の世紀とも云うべきものになることが期待される。

太陽から惑星間空間に拡がる太陽風プラズマは超音速流であり、外圏では星間ガスとの相互作用により、終端衝撃波面が生成され、そこを越えると太陽風プラズマは亜音速流に減速され、更に方向を星間ガスの下流側に転じ、全体としては吹き流しの構造を作ると考えられている。太陽風プラズマと星間ガスとの接触面はヘリオポーズと呼ばれている。終端衝撃波のスケールは上流側で１００ＡＵ、下流側で２００ＡＵ程度と予想されているが、実際のところは探査機が到達してみなければ確定しない。中性粒子や宇宙線の影響ではっきりした衝撃波構造は形成されないのではないかと考える研究者も多い。太陽風プラズマが超音速であるために、外側からの電磁流体波は伝わって来なず、このため情報の種類が限られている。即ち、外側からの信号として観測出来ているのは、主に、宇宙線、中性粒子及び中性粒子が太陽風荷電粒子と電荷交換して生まれるピックアップ粒子と呼ばれる粒子、及びKHz帯以上の高周波の電磁波に限られている。

さて、太陽圏周辺の星間ガスについては太陽との相対速度が約２０ｋｍ／ｓ、温度約７０００Ｋ、電離度３０－５０％であることが太陽系に浸入する中性ガス成分の測定や近傍の星々のスペクトル観測からわかって来ている。このように、星間ガスが不完全電離ガスであるため、ヘリオポーズから太陽圏へと浸入した中性粒子が太陽風プラズマと電荷交換をしたり、さらにそれが加速されたりと複雑な効果をもたらし、太陽圏の解析を非常に難しいものにしている。しかし、中性ガスの効果は太陽圏の構造形成に重要な役割を担っているので無視することは出来ない。

それに加えて、太陽風プラズマ自身が太陽活動１１年周期変動やもっと短い時間的な変動が常時起きているので、これに応答した変動する太陽圏を考えなければならない。

磁場の役割は銀河宇宙線の太陽系内部への流入を解析する上で重要なことは当然であるが、太陽圏外圏で太陽風の速度が急激に減速することに伴い太陽風によって運ばれて来た磁場の強度が増加し磁気圧の効果が効くことも指摘されて来ている。最近は電磁流体力学的（ＭＨＤ）解析もシミュレーションスキームの進歩と並列計算機の出現により活発になって来ている。著者の一人（鷲見）も田中高史博士（通総研）と３次元ＭＨＤ解析を行っていて、外圏における磁場構造のモデリングの研究[1] をしていて、またこの磁場構造モデルを用いて宗像一起博士（信州大）は銀河宇宙線（高エネルギー荷電粒子）の太陽圏空間での軌道計算を始める段階にまでなって来ている[2]。最新のレビューは G.P.Zank[3] に詳しく記述されている。

３．太陽系近傍における小天体の軌道計算

上述したように、太陽圏研究が盛んになって来ていて、更に銀河探査の時代がまもなく（と云っても１０年以上、もしくは数十年先のことであろうが）開かれようとしているので、一般の人々の興味も近い将来には太陽系の外へと向かうであろう。このことを期待して、我々の研究室での卒業研究として行った太陽系近傍における小天体の軌道計算及び３次元グラフ表示について述べる。なお、小天体とはその存在が恒星の運動に影響を与えないほど十分小さい天体であるという意味である。

太陽から恒星集団のつくる重力場の中を小天体がどのような軌跡を描いて太陽系へと到達するかを解析するプログラムを作成した。研究の基本的な立場は以下の通りである。

（１）太陽からの距離が 14.5 光年以内にある３０個の恒星（資料：理科年表）の作る重力場を考慮する。

（２）小天体の軌道計算は万有引力の法則のみを考慮している。ルンゲ・クッタ法を用いて小天体の運動を求める。

（３）小天体の軌道計算を数千年―１０億年にわたり行うが、その間恒星間の相対位置の変化は無視する。

（４）言語は VisualC++5.0 を使用する。

次に太陽系近傍における恒星配置図を第１図に示す。

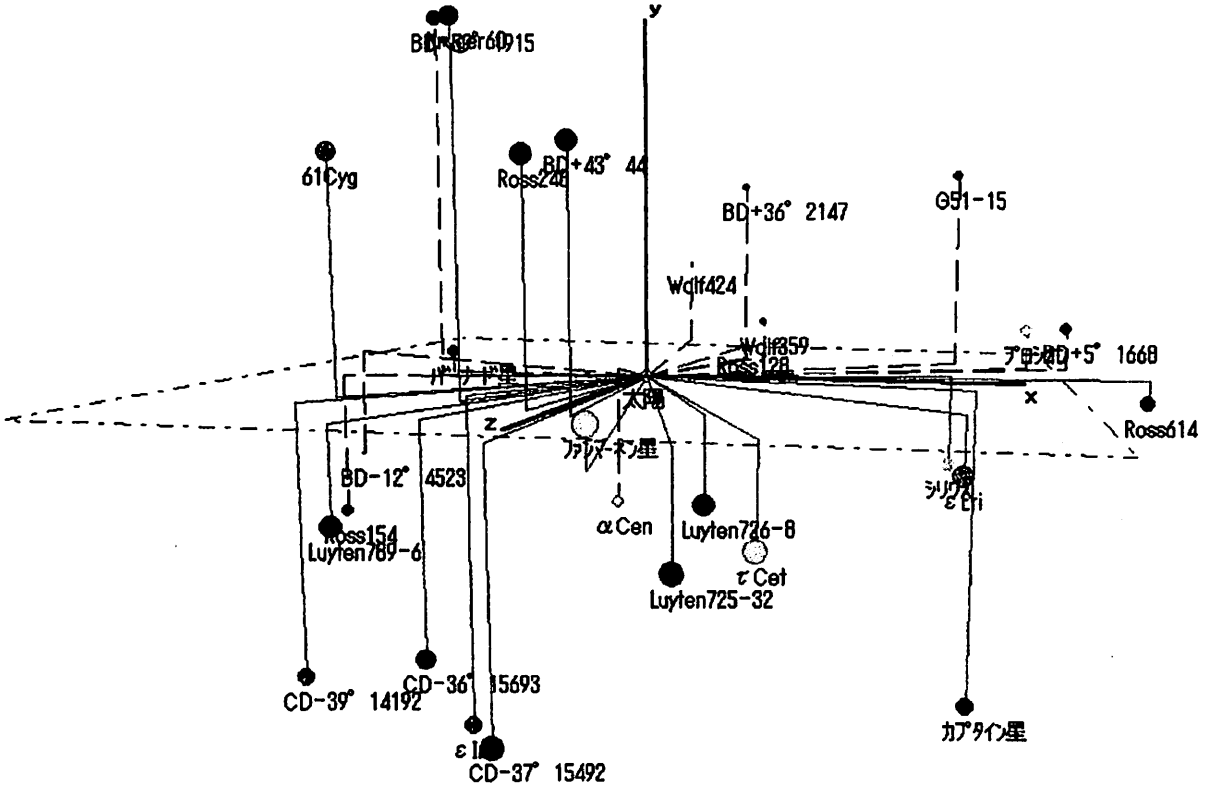


Fig. 1　Distribution of stars around the solar system

座標系としては太陽を原点に置いた３次元座標（ｘ、ｙ、ｚ）を用いる。ここで、ｘ－ｚ面は赤道座標での赤道面に、ｚ軸は赤経0°に、ｙ軸は天頂に、それぞれ合わせている。

図で丸印が恒星で、その下に書かれている文字は恒星の名前である。恒星の色は実際の色を表している（本報告は白黒であるので見えない）。大きさは手前にある恒星は大きく、奥にある恒星は小さく表示される。ｘｙｚ軸の表示も入れてある。各軸の長は、夫々、１０光年である。２点鎖線の平面は、太陽を中心として１辺の長さが２２光年の赤道面である。
各恒星から太陽までの線は、まずｙ軸に平行に出て赤道面に至り、そこからは赤道面上を太陽まで直線で結んである。恒星が太陽より手前にある場合は実線、奥にある場合は破線で奥行きを示している。

４、実行結果

まず４次精度３次元ルンゲクッタ法による重力場中での軌道計算プログラムを作成した。完成後、精度について検討したが、１恒星のまわりの円軌道でテストして、１周を12800ステップのきざみにすれば数億年にわたる計算でも十分精度が維持されることを確認した。以下に述べる計算例でもある初期位置からある初速度で計算を始め数億年にわたる軌道計算の終了後、隕石の最終位置から速度を反転させて計算を再開すればはじめの初期位置にもどること、及びステップ数を大きくしても同じ軌道を描くことを確認した。
プログラム実行上は、各ステップでの隕石の位置において、各恒星に対する等速円運動周期を求め、そのうち最も短い周期を12800で割った値を時間ステップとしている。計算速度から云えば、時間ステップを各ステップごとに決めているところを工夫すれば、精度は維持されたまま計算速度は速くなるであろうが、実用としてはパソコンの能力が向上しているので今のままでも不便は感じない。

以下に多くの恒星の作る複雑な重力場の下で小天体がどのような軌跡を描くかの例を示す。

最初の例では太陽とαCenの間を往復運動する軌道の一例を見つけたので示す。初期位置をｘ、ｙ、ｚ座標で（0.01, 0, 0）とする。単位は光年である。従って、太陽からの距離６３２．４ＡＵの位置である。初期位置は図では太陽に近すぎて表すことが出来ない。５億年後の最終位置は点で表示されている。

軌道例１　—　太陽及びαCenを周回する軌道

次の例として、初期位置、初期速度、及び経過時間は、

(1)微小天体の初期位置（0.00062, 0, 0）[光年]

(2)初速度（-5.264, -0.444, -4.167）[km/s]

(3)経過時間　5億年

。

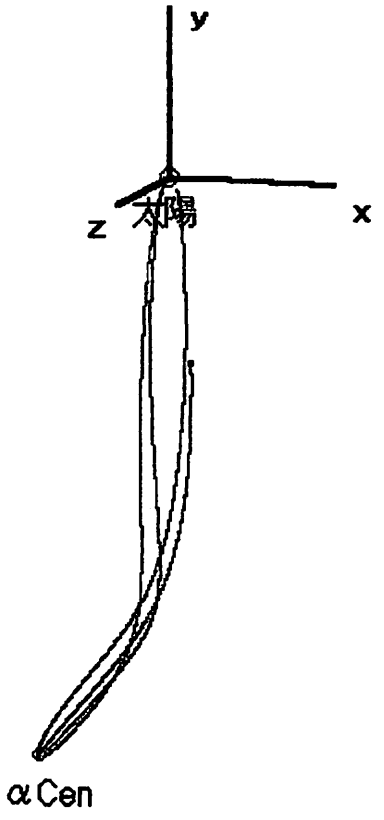


Fig. 2 Orbit example No.1 – An orbit which is located in the sun-αCen area

座標軸の長さは、それぞれ１光年である。

第２図は微小天体が太陽とαCenの間を２周近く周回していることを示している。このような軌道は非常に微妙であり、たとえばy方向の初速度 1.222km/s を 2% 減らすだけでも、微小天体は太陽の重力圏から抜け出すことができなく、また逆に、 2% 増やすと微小天体は恒星群の外へ飛んで行ってしまう。微小天体の初速度を少しでも変化させる軌道は大きく変わってしまう。この小天体の周期は約３，３００万年であるので、もし存在しているとしても周回軌道をとっていることの確認は困難であろう。

軌道例２　―　さすらいの小天体その１

小天体の初期位置、初速度及び経過時間は

(1)　小天体の初期位置（0.01, 0, 0）[光年]

(2)　初速度（-1.1389, 1.222, 0）[km/s]）

(3)　経過時間　6千万年

である。この時の軌道を第３図に示す。なお、座標軸の長さは１光年である。

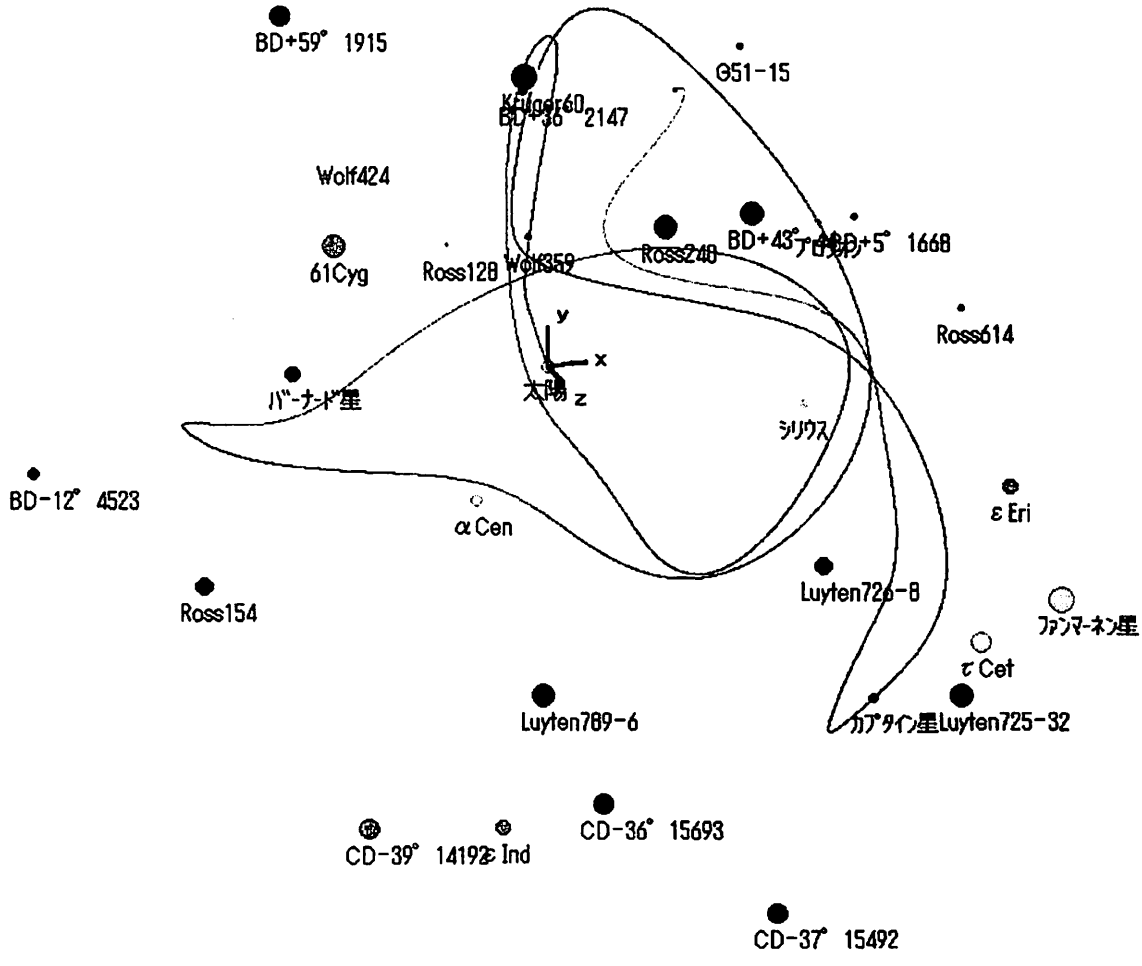


Fig. 3　Orbit example No. 2 – An orbit which traces in the interstellar space of a scale of ~ 13 light year

軌道例３　―　さすらいの小天体その２

小天体の初期位置、速度、及び経過時間

(1)　微小天体の位置（0.001，0.0，0.0）［光年］

(2)　初速度（3.891，3.597，0.0）［km/s］

(3)　経過時間　8億年

座標軸の長さは１光年である。

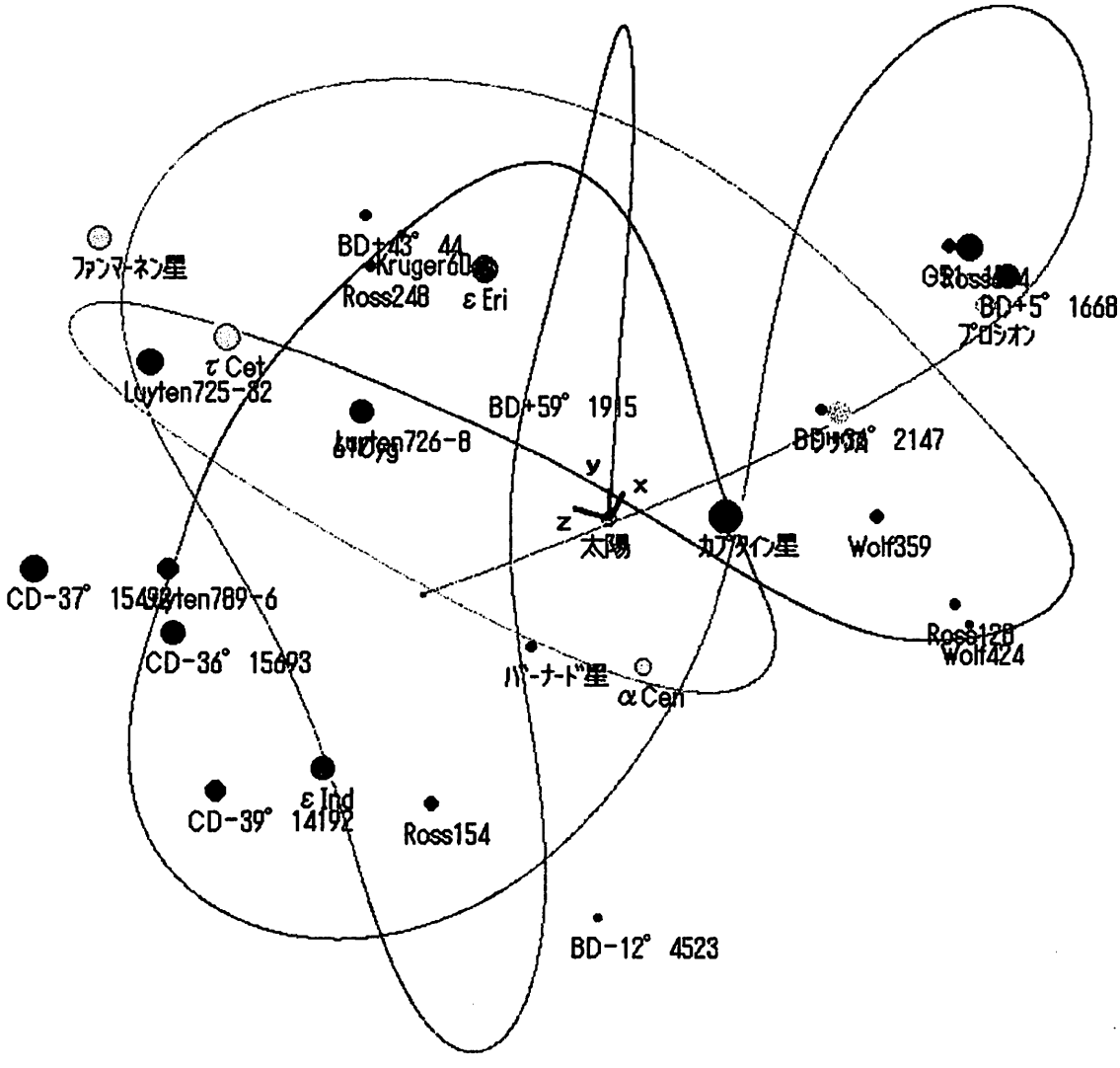


Fig. 4　Orbit example No. 3 – An orbit which traces in the interstellar space of a scale of ~18 light year

第３図と比較すると第４図の方がかなり広範囲にわたって小天体が運動している様子が分かる。

5，まとめ

太陽近傍１４．５光年以内にある３０の恒星のつくる重力場の中での小天体の軌道計算を第４節に示した３つの具体例をはじめとして、いろいろな初期位置・初期速度で調べた。これによって得られた教育的効果としては、

(1) 重力が長距離力であることがこの軌道計算により具体的に理解出来る。すなわち小天体には恒星群がつくる複雑な重力が作用し、軌道は２体問題の時とは全く異なり、第２、３，４図のように非常に複雑な曲線となる。これらの例に見られるように、数千万年もしくは１億年以上経て多くの恒星の間を渡りめぐった後太陽系へと入って来る「さすらいの小天体」と呼んでもよいような軌跡が得られることが明らかとなった。

(2) 多くの場合、太陽系へともどる周期運動は見られないが、中には第２図のように太陽及び特定の恒星の間を数千年もしくはそれ以上の時間周期で周期運動する軌道もあることが示さた。

本研究は情報工学科の卒業研究として行われたので、３Ｄグラフィックスについても力点を置いた。即ち、

(3) 恒星の位置関係を３Ｄグラフィックスで表現したことにより具体的に理解することができる。そして宇宙空間において、ある座標である速度を持った微小天体がどのような過程でどこへ到達するのかを３次元シミュレーションで表現することができた。

(4) メニューやダイアログボックス等の機能をふんだんに付け加えたことによりユーザーに恒星の位置や微小天体の軌道等について理解しやすい環境と、ソフトの使いやすさを提供した。

(5) 表示画面の拡大縮小、恒星重力圏の表示、微小天体の軌道の表示、視点の回転、連続的な視点の移動による恒星の位置及び小天体軌跡の３次元回転・動画表示が可能となっている。

なお、以上の計算では恒星の相対位置を不変としていること、及び１４．５光年以上の位置にある恒星の重力は考慮していないことを重ねてお断りしておきたい。

References

[1] H. Washimi and T. Tanaka, 3-D magnetic field and current system in the heliosphere, Space Sci. Rev., 78, 85-94, 1996.

[2] K. Munakata et al., Gaussian analysis of two hemisphere observations of galactic cosmic ray siderial anisotropies, Proc. Solar Wind 9, in print.

[3] G. P. Zank, Interaction of the solar wind with the local interstellar medium: a theoritical perspective, Space Sci. Rev., in print.

Statistical Mechanics of Self-Gravitating System : Cluster Expansion Method

Osamu Iguchi [1], Tomomi Kurokawa [2], Masahiro Morikawa [3],
Department of Physics, Ochanomizu University, 1-1 Otuka,2 Bunkyo-ku, Tokyo 112, JAPAN

Akika Nakamichi [4],
Gunma Astronomical Observatory, 6860-86 Nakayama, Takayama-mura, Agatsuma-gun, Gunma 377-0702, JAPAN

Yasuhide Sota [5], Takayuki Tatekawa [6], and Kei-ichi Maeda [7]
Department of Physics, Waseda University, 3-4-1 Okubo, Shinjuku-ku, Tokyo 169, JAPAN

Abstract

We study statistical mechanics of the self-gravitating system applying the cluster expansion method developed in solid state physics. By summing infinite series of diagrams, we derive a complex free energy whose imaginary part is related to the relaxation time of the system, and a two-point correlation function.

Introduction One of the most important subject in modern cosmology is to study how the structures found in the present Universe have been formed. Though there would be several aspects in the study of the structure formation, we believe that the self-gravity is one of the most intrinsic aspects. Then, it is important to study fundamental properties of the self-gravitating system (SGS) from a view point of statistical mechanics. There are many statistical approaches to SGS [1, 2, 3, 4]. In particular, in the structure formation of the Universe, the standard analysis based on the BBGKY hierarchy with an initial power-law spectrum of density fluctuations is important. The reduction of the BBGKY hierarchy with scaling ansatz and additional assumptions has yielded many relevant results, especially some reasonable explanation of the observed two-point correlation function, although the stability of such a similarity solution is not yet confirmed [1, 2].

There is another way to study statistical properties of SGS, in which "thermal equilibrium" is assumed[3, 4]. This would be complimentary to the above standard analysis of the structure formation. In this approach, we specify SGS as a mass distribution which consists of discrete mass points mutually interacting through the force of Newtonian gravity. The contact collisions of the ingredients will be neglected. The SGS is well known to be unstable and eventually collapses. Then it is widely believed that there is no thermal equilibrium in SGS. However relevant structures would be observed in the intermediate stage of its evolution well before the collapse. Especially we are interested in the initiation stage of the gravitational instability associated with the cooperative interaction of the self-gravitating system, which is caused by the long-range property of the force. Therefore we may have a room to construct statistical mechanics of SGS. This approach would, we believe, reveal some physics that reflects intrinsic properties of the gravity free from initial conditions. Recently, we found the approach by de Vega *et al.*[5] seems to be interesting; they consider canonical/grand canonical ensemble of SGS in order to explain a scaling relation in interstellar medium. In the thermodynamic limit, thermodynamic quantities of SGS diverge and SGS becomes unstable. De Vega *et al.* introduced another thermodynamic limit and studied the stable region of SGS in their work. Here we would like to circumvent this infrared divergence allowing a complex free energy, whose imaginary part provides dynamical information of the system.

We shall construct statistical mechanics of SGS after the model of classical electron-gas system on uniform ion background. It has the same square-inverse law of force but with opposite signature to SGS. Contrary to the ordinary belief, the correspondence, i.e. the replacement $e^2 \to -Gm^2$, where e is the charge of electron and G is the Newton's constant, is quite suggestive if we identify the temperature T as the velocity dispersion of SGS. For example, the Debye wavelength λ_D corresponds to the Jeans

[1]Email:osamu@phys.ocha.ac.jp
[2]Email:tk385@phys.ocha.ac.jp
[3]hiro@phys.ocha.ac.jp
[4]akika@astron.pref.gunma.jp
[5]sota@gravity.phys.waseda.ac.jp
[6]tatekawa@gravity.phys.waseda.ac.jp
[7]maeda@gravity.phys.waseda.ac.jp

wavelength λ_J, and the plasma frequency is related to the inverse of the free fall time, $\omega_p^2 := -4\pi Gmn = -1/\tau_{ff}^2$. We believe that these correspondences are not merely appearance. Actually, we can calculate the free energy of SGS in the same way as the classical electron gas system by applying the cluster expansion method [6, 7, 8]. Moreover, n–point correlation functions are similarly obtained.

Cluster expansion method We consider a non–relativistic gas system of N particles with the same mass m in a box of size L. We work in the canonical ensemble with the Hamiltonian of the system:

$$H = \sum_{i=1}^{N} \frac{\vec{p}_i^{\,2}}{2m} + \sum_{1\le i<j\le N} \phi_{ij}, \tag{1}$$

where ϕ_{ij} is a potential between i-th and j-th particles. The gas is in thermal equilibrium at T, where the temperature T is related to the velocity dispersion $\langle v^2\rangle$ of the system, i.e. $k_BT = m\langle v^2\rangle/3$, where k_B is the Boltzmann's constant. Fixing particle number N, volume $V = L^3$, and temperature T, the partition function is given by

$$Z(u) = \mathrm{Tr}\exp\left[-\beta H + \sum_{i=1}^{N} u(\vec{x}_i)\right] = \frac{(2\pi m/\beta)^{3N/2}V^N}{N!h^{3N}} e^{W(u)}, \tag{2}$$

$$e^{W(u)} = \int_{V^N} \frac{d^N\vec{x}}{V^N} \exp\left[\sum_{i=1}^{N} u(\vec{x}_i) + \sum_{1\le i<j\le N} \varphi_{ij}\right], \tag{3}$$

where $\beta = (k_BT)^{-1}$, h is the Planck's constant, and $\varphi_{ij} := -\beta\phi_{ij}$. We have introduced a source function $u(\vec{x}_i)$ for convenience of calculation and integrated over the momenta. By using functional derivative with respect to the source $u(\vec{x}_i)$, k–point correlation functions are given by

$$G(\vec{x}_1,\cdots,\vec{x}_k) = \lim_{u\to 0} \frac{\delta^k \ln Z(u)}{\delta u(\vec{x}_1)\cdots\delta u(\vec{x}_k)}. \tag{4}$$

We introduce the logarithm of the average of the configurational sum setting $u \to 0$:

$$W_0 := W(0) = \ln\left\langle \exp\left(\sum_{1\le i<j\le N} \varphi_{ij}\right)\right\rangle, \tag{5}$$

where $\langle\cdots\rangle = V^{-N}\int\cdots d^N\vec{x}$.

The cumulant expansion of W_0 is possible:

$$W_0 = \left\langle \exp\left(\sum_{1\le i<j\le N} \varphi_{ij}\right) - 1\right\rangle_C = \sum_{(\text{except all } \nu_i=0)} \frac{\langle \varphi_1^{\nu_1}\varphi_2^{\nu_2}\cdots\varphi_M^{\nu_M}\rangle_C}{\nu_1!\cdots\nu_M!}, \tag{6}$$

where $\langle\cdots\rangle_C$ is a cumulant and $1, 2, \cdots, M = N(N-1)/2$ are all the possible pairs of N–particles.

It is convenient to use graphical representations in order to calculate W_0. An each term in Eq.(6) corresponds to a graph which consists of vertices and lines. Each line terminates at two distinct vertices. A vertex $\vec{x}_i$ simply represents an integration position and a line connecting $\vec{x}_i$ and $\vec{x}_j$ represents an interaction φ_{ij}. All the elements in a graph are multiplied with each other and integrated over all vertices. This is the *cluster expansion method.* All the graph which contributes to $\ln Z$ is connected and does not include articulation points which divide the graph into plural pieces, because of the character of the cumulant and the translational invariance of the system.

By expanding with respect to the number density $n = N/V$, W_0 is reduced to the following form[6, 7, 8]:

$$W_0 = \sum_{k=1}^{\infty} \frac{n^{k+1}}{(k+1)!} \sum_{(\substack{\text{connected}\\ \text{irreducible}})} \int_{V^{k+1}} \prod_{1\le i<j\le k+1} \left[\frac{(\varphi_{ij})^{\nu_{ij}}}{\nu_{ij}!}\right] d\vec{x}_1 ... d\vec{x}_{k+1}, \tag{7}$$

where ν_{ij} is the number of interaction lines between the vertices i and j. The "irreducible" means that the graph cannot be separated by cutting a point in the graph.

Application to the self–gravitating system Let us consider a non–relativistic self–gravitating system where $\phi_{ij} = -Gm^2/r_{ij}$ with $r_{ij} = |\vec{x}_i - \vec{x}_j|$.

Free energy We shall include the higher orders in n for calculating W_0. Here we choose a series of graphs which is the lowest order of φ in each set of graphs containing k internal vertices. They are dominant terms at the large scale. This is equivalent to a sum of all graph which has the topology of a ring (ring approximation shown Fig.1–(a)). This approximation will not be valid for short distances where φ grows without bound.

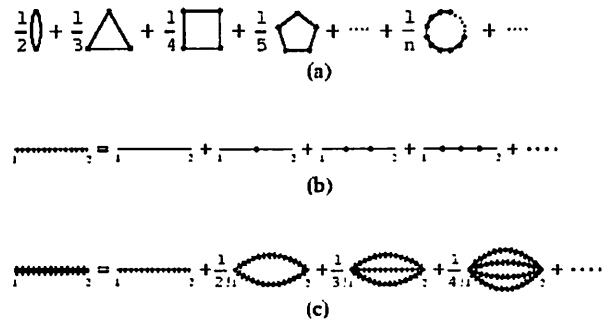


Figure 1: A series of cluster expansion: (a) ring graphs. (b) chain graphs. (c) multi–lines of the chain graphs. A symmetric factor is written in the left side of each graph.

A ring graph which contains k vertices in Eq.(7) corresponds to

$$\int_{V^k} \varphi_{12}\varphi_{23}\cdots\varphi_{k1} d\vec{x}_1 ... d\vec{x}_k = \int_{|\vec{q}|>2\pi/L} \frac{d\vec{q}}{(2\pi)^3} [\tilde{\varphi}(\vec{q})]^k, \tag{8}$$

where $\tilde{\varphi}(\vec{q}) = 4\pi\beta G m^2/\vec{q}^{\,2}$ is the Fourier transform of $\varphi(\vec{x})$. $(k-1)!/2$ terms of this kind for each k yield

$$\begin{aligned} W_0 = & \frac{V}{2}\int_{|\vec{q}|>2\pi/L} \frac{d\vec{q}}{(2\pi)^3} \sum_{k=2}^{\infty} \frac{[n\tilde{\varphi}(\vec{q})]^k}{k} \\ = & \frac{V}{4\pi^2}\int_{2\pi/L}^{\infty} dq \cdot q^2 \left[-\frac{\kappa^2}{q^2} - \ln\left(1-\frac{\kappa^2}{q^2}\right)\right] \\ = & \frac{1}{12\pi^2}\left[2\pi\kappa^2 L^2 + \kappa^3 L^3 \ln\left|\frac{\kappa L + 2\pi}{\kappa L - 2\pi}\right| + 8\pi^3 \ln\left|1 - \frac{\kappa^2 L^2}{4\pi^2}\right| + i\pi\left(\kappa^3 L^3 + 8\pi^3\right)\theta\left(\kappa L - 2\pi\right)\right], \end{aligned} \tag{9}$$

where θ is the Heaviside function and $\kappa^2 := 4\pi\beta G m^2 n = 12\pi^2/\lambda_J^2$ with $\lambda_J = \sqrt{\pi\langle v^2\rangle/(Gmn)}$ being the Jeans wavelength. For the integration in Eq.(9), we analytically continue e^2 to $-Gm^2$ from the above complex e^2 plane by the reason under Eq.(10).

We notice that the imaginary part appears for $L > 2\pi/\kappa = \lambda_J/\sqrt{3}$. This originates from the negative argument in the logarithm for small wavenumber q. The appearance of it apparently indicates the Jeans instability. The thermodynamic limit $N \to \infty, V \to \infty$ with fixed number density n, yields the free energy:

$$\begin{aligned} F &= -\beta^{-1}\ln Z(0) \\ &= -\frac{N}{\beta}\left[\frac{3}{2}\ln\left(\frac{2\pi m}{\beta h^2}\right) + 1 - \ln n - \frac{i\kappa^3}{12\pi n}\right]. \end{aligned} \tag{10}$$

The above choice of analytical continuation guarantees the positivity of the imaginary part of F: The system is truly dissipative rather than anti–dissipative.

In general, the imaginary part of a free energy is related with the decay strength of the system Γ[9, 10]: $\Gamma = (\omega\beta/\pi)\mathrm{Im}F$, where $-\omega^2$ is the negative eigenvalue at the saddle point dividing a metastable region from a stable region. Identifying ω with the inverse of the free fall time, i.e. $\omega = \tau_{ff}^{-1} = \sqrt{4\pi G m n}$, we find the decay strength Γ as

$$\Gamma = \frac{4\sqrt{3}NG^2m^2n}{\langle v^2\rangle^{3/2}}, \tag{11}$$

which is essentially the inverse of the binary relaxation time[3, 11],

$$\tau_{bc} = \frac{\langle v^2\rangle^{3/2}}{32\pi G^2m^2n\ln(N/2)}, \tag{12}$$

except replacing $\ln N$ with N. Then, we have derived this relation between the decay rate Eq.(11) and the binary relaxation time Eq.(12) based on the analogy of the Langer' work[9]. Deeper understanding of this relation, beyond our present phenomenological argument, requires the proper consideration of dynamics.

It is worth noting that the imaginary part ($\mathrm{Im}F \propto n^{3/2}$) is related to the fluctuation of the system through the fluctuation–dissipation theorem. In this context, it seems interesting to notice that the number 3 in the above $n^{3/2}$ comes from the spatial dimensionality and 2 from the inverse–square law of the gravitational force. This reminds us of the Holtsmark distribution of the gravitational force acting in the uniform self-gravitating system [11] or the stable distribution of index 3/2 [12].

Two–point correlation function Let us now turn our attention to the correlation functions. The normalized two–point correlation function which is usually used in astrophysics is

$$\xi(r) = G(\vec{x}_1, \vec{x}_2)/\left[G(\vec{x}_1)G(\vec{x}_2)\right], \tag{13}$$

where $r = |\vec{x}_1 - \vec{x}_2|$ and one–point correlation function $G(\vec{x}) = N$. A similar cluster expansion for $G(\vec{x}_1, \vec{x}_2)$ would be

$$G(\vec{x}_1, \vec{x}_2) = N(N-1)\sum_{k=0}^{\infty}\frac{n^k}{k!}\sum_{\text{connected}}\int_{V^k}\prod_{1\le i<j\le k+2}\left[\frac{(\varphi_{ij})^{\nu_{ij}}}{\nu_{ij}!}\right]d\vec{x}_3...d\vec{x}_{k+2}. \tag{14}$$

We shall include higher orders in n as previous calculation. Therefore, among each set of graphs which contain k internal vertices (excluding both $\vec{x}_1$ and $\vec{x}_2$), we choose the lowest order skeleton graph in φ_{ij}. This is equivalent to a sum of all the graph which has the topology of a chain shown in Fig.1–(b).

$$\begin{aligned}
\overline{\varphi}_{12} &= \varphi_{12} + \int_V \varphi_{13}\varphi_{32}d\vec{x}_3 \quad + \int_{V^2}\varphi_{13}\varphi_{34}\varphi_{42}d\vec{x}_3d\vec{x}_4 + \cdots \\
&= \int\frac{d\vec{q}}{(2\pi)^3}\frac{\tilde{\varphi}(\vec{q})}{1-n\tilde{\varphi}(\vec{q})}\exp\left[i\vec{q}(\vec{x}_1-\vec{x}_2)\right] \\
&= \frac{2\beta Gm^2}{\pi r}\int_0^{\infty}d\lambda\frac{\lambda\sin\lambda}{\lambda^2-\kappa^2r^2} \\
&= \beta Gm^2\frac{\cos(\kappa r)}{r}.
\end{aligned} \tag{15}$$

Since we have analytically continued from the above complex e^2 plane in Eq.(9), we should choose the pole at $-\kappa r$ for the integration. Summing over multi–lines of chain graphs shown in Fig.1–(c), we obtain the following function:

$$\begin{aligned}
\widehat{\varphi}_{12} &= \overline{\varphi}_{12} + \frac{1}{2!}(\overline{\varphi}_{12})^2 + \frac{1}{3!}(\overline{\varphi}_{12})^3 + \frac{1}{4!}(\overline{\varphi}_{12})^4 + \cdots \\
&= \exp(\overline{\varphi}_{12}) - 1.
\end{aligned} \tag{16}$$

This form results from the summation over all the possible dominant graphs at the large scale. Moreover, we should consider mass renormalization for the external vertex. In principle, this is given by sum of all

the graph which consists of a cluster around the external vertex $\vec{x}_1$. Since infinite summation of this class of graphs at present is technically impossible, we phenomenologically introduce an effective mass. Since the gravitational attractive force balances with the stirring force arising from the velocity dispersion at the length scale κ^{-1}, the mass inside of this scale is thought to behave collectively. Thus we estimate the effective mass $m^* = (4\pi/3)(\kappa^{-1})^3 nm$, which should be replaced with the mass m at $\vec{x}_1$ or at $\vec{x}_2$ but not both. Strictly speaking, m^* is a parameter in the present phenomenological argument. By replacing m at $\vec{x}_1$ in Eq.(15) with m^* at $\vec{x}_1$, the two-point correlation function becomes $G(\vec{x}_1, \vec{x}_2) = N^2 \, \widehat{\varphi}_{12}|_{m_{\text{at}\,\vec{x}_1} \to m^*_{\text{at}\,\vec{x}_1}}$. Finally we obtain the normalized two-point correlation function:

$$\xi(r) = \exp\left[\frac{\cos(\kappa r)}{3\kappa r}\right] - 1. \tag{17}$$

The scaling property is manifest: Small scale SGS with large κ has the same correlation function as that of large SGS with small κ.

This function $\xi(r)$ has interesting features. As a function of $s = \kappa r$, it has a unique inflection point at $s = 0.440$ when plotted in the Log–Log graph. The slope at $s = 0.44$ is -1.667 and the magnitude there is 0.986. Therefore, the inflection point $r_c \sim 0.44\kappa^{-1} \sim 0.04\lambda_J$ is regarded as the correlation length of SGS.

This two-point correlation function is shown in Fig.2 with typical observational results[13, 14] on top of it. The observation does not exclude our correlation function except small scale region where the interaction φ exceeds 1 and our approximation is no longer valid. However this comparison with observation is not intended as the demonstration of our approach but simply as an illustration. Actually, a cluster of galaxies is an extended object and the treatment of it as a point mass, in the strict sense, is a crude approximation.

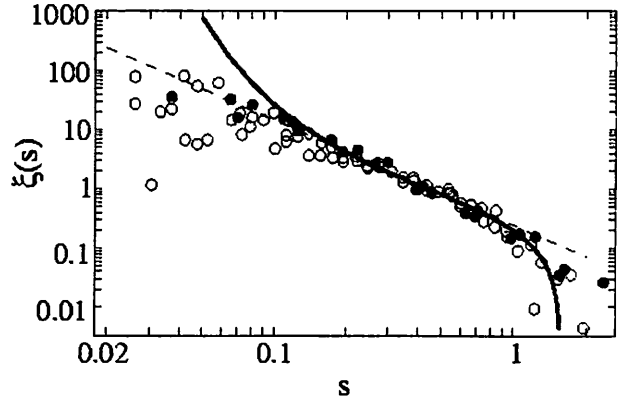


Figure 2: A Log–Log graph of two-pint correlation function: $\xi(s) = \exp[\cos(s)/(3s)] - 1$ (solid line), $(s/0.45)^{-1.77}$ (broken line), observational data of galaxies [13] (open circle), and observational data of clusters of galaxies [14] (filled circle). These correlation lengths of the observational data are rescaled. The error bars ignored here are found in [14].

From the observational data[13, 14], the correlation lengths of galaxies and of clusters of galaxies directly read off as $r_c^{galaxy} \sim 6.2h^{-1}$Mpc and $r_c^{cluster} \sim 15h^{-1}$Mpc respectively. Since $r_c \sim 0.44\kappa^{-1}$ for the two-point correlation function we obtained, we rescaled the observational data: $r \to s = \kappa r$ with $\kappa = 1/14$ hMpc^{-1} for galaxies and $\kappa = 1/34$ hMpc^{-1} for clusters of galaxies, where $h = H_0/\left(100\text{kms}^{-1}\text{Mpc}^{-1}\right)$ with H_0 being the Hubble constant at the present time. These values of κ correspond to $\lambda_J^{galaxy} \sim 152h^{-1}$Mpc and $\lambda_J^{cluster} \sim 370h^{-1}$Mpc from the relation given immediately after

Eq.(9). On the other hand, typical Jeans lengths calculating from the standard observations for galaxies and for clusters of galaxies are respectively

$$\lambda_J^{galaxy} = 123\mathrm{Mpc}\left[\left(\frac{\langle v^2\rangle}{(500\mathrm{kms}^{-1})^2}\right)\left(\frac{2\times 10^{44}\mathrm{g}}{m}\right)\times\left(\frac{3.7\times 10^{-2}h^3\mathrm{Mpc}^{-3}}{n}\right)\right]^{-1/2}, \quad (18)$$

$$\lambda_J^{cluster} = 450\mathrm{Mpc}\left[\left(\frac{\langle v^2\rangle}{(1000\mathrm{kms}^{-1})^2}\right)\left(\frac{2\times 10^{47}\mathrm{g}}{m}\right)\left(\frac{1.1\times 10^{-5}h^3\mathrm{Mpc}^{-3}}{n}\right)\right]^{-1/2}. \quad (19)$$

Here, we have used typical values for the normalization though there are some amount of ambiguity especially in the mass estimation which comes from unknown dark matter contributions. The above values of the Jeans lengths are not so far from the above values through two–point correlation function we obtained.

summary and outlook Studying the canonical ensemble of the self–gravitating system (SGS), we obtained the complex free energy by summing an infinite series of graph in the cluster expansion method for SGS, Eq.(10). The imaginary part of the free energy yields the decay strength of SGS, Eq.(11). Similar summation of an infinite series of graph yields the universal two–point correlation function Eq.(17) which scales essentially with the Jeans wavelength. The correlation length is linearly proportional to the mean separation of ingredients.

We would like to report further development soon, including (a) the higher–point correlation functions, (b) much profound calculation on the complex free energy in relation with the dynamics, (c) systematic argument on the mass renormalization m^*, (d) the observational tests of our arguments, and (e) the effects of the cosmic expansion and of the Dark matter.

References

[1] P. J. E. Peebles, *The Large–scale Structure of the Universe* (Princeton Univ. Press, Princeton, 1980).

[2] T. Padmanabhan, *Structure formation in the universe* (Cambridge Univ. Press, Cambridge, 1993).

[3] W. C. Saslaw, *Gravitational physics of stellar and galactic system* (Cambridge Univ. Press, Cambridge, 1985).

[4] T. Padmanabhan, *Phys. Rep.* **188**, 286 (1990).

[5] H. J. de Vega, N. Sánchez, and F. Combes, Nature **383**, 56 (1996); Phys. Rev. D **54**, 6008 (1996); ApJ. **500**, 8 (1998); preprint, astro–ph/9807048.

[6] M. Wortis, in *Phase Transitions and Critical Phenomena* ed. C. Domb and J. L. Lebowitz, (Academic Press London, 1983), Vol.3, p.113.

[7] J. E. Mayer and M. G. Mayer, *Statistical mechanics*, 2nd ed. (John Wiley and Sons, London, 1977).

[8] R. Abe, *Statistical mechanics*, 2nd ed. (Tokyo Univ. Press (*in Japanese*), 1991).

[9] J. S. Langer, Ann. Phys. **54**, 258 (1969).

[10] I. Affleck, Phys. Rev. Lett. **46**, 388 (1981).

[11] S. Chandrasekhar, Rev. Mod. Phys. **15**, 1 (1943).

[12] W. Feller, *An Introduction to Probability Theory and Its Applications* 2nd ed. (Wiley and Sons, Inc. New York, 1966), Vol. II.

[13] M. Davis, *13th Texas Symposium on Relativistic Astrophysics* (World Scientific, 1986), p.289.

[14] A. Dalton *at el.*, Mon. Not. R. Astron. Soc. **271**, L47 (1994).

フラクタル的初期密度ゆらぎと構造形成
Primordial Fractal Density Perturbation and Structure Formation of the Universe

立川 崇之*, 前田 恵一†
Takayuki Tatekawa and Kei-ichi Maeda
Department of Physics, Waseda University, 3-4-1 Okubo, Shinjuku-ku, Tokyo 169-8555, JAPAN
(April 21, 1999)

現在の銀河分布の観測から銀河や銀河団の二点相関関数が距離のべきになるといった、スケール則が見い出されている。一方で宇宙背景輻射のゆらぎがフラクタル的であるという解析も存在する。これは初期にフラクタル的な密度ゆらぎがあり、現在のフラクタル的な大規模構造に成長した可能性を示唆する。そこで一様・等方宇宙におけるフラクタル的な密度ゆらぎから成長した構造のフラクタル性、特にフラクタル次元の時間変化について解析する。今回は簡単なモデルとして、一次元シート系における構造形成を取り扱う。

According to recent observations of the 2-point correlation function of galaxy distribution, the structure of the present Universe shows a scaling law. Further more, E. M. de Gouveia Dal Pino et al.(1995) reported that temperature fluctuation of Cosmic Microwave Background Radiation(CMBR) may have fractality. This analysis implies primordial density fluctuations would be fractal. Then such small fractal perturbation may develop the present large scale structure with a scaling law. We analyze fractal structure developed from primordial fractal density fluctuation in homogeneous and isotropic Universe. In particular we analyze time variation of fractal dimension. In this paper, we study only 1-D sheet model.

I. INTRODUCTION

現在の銀河分布の観測から、その二点相関関数は、銀河～銀河団スケールにおいて距離のべきに比例する事が知られている [8] [23]。この二点相関のべき則は銀河がフラクタル分布をしている可能性を示唆するが、COBE による宇宙背景輻射の観測からは、宇宙全体のスケールで宇宙はほぼ一様等方であった事が分かっている。では、現在の銀河分布がもし本当にフラクタル的であるとすると、そのような分布は一様な宇宙からどのようにして形成されたのであろうか。

この問いに関して、密度ゆらぎが非常に小さい線形領域では、摂動を用いた解析が可能であるが、非線形領域については十分な解析法が確立されていない。Davis と Peebles(1977) は、統計的考察から一様等方宇宙での非線形領域における二点相関関数の自己相似解を与えている [8]。初期条件として特徴的なスケールをもたない密度ゆらぎを与え、この密度ゆらぎに対してまず線形領域において解析的に二点相関関数を求める。非線形領域においても、スケールによらない解が存在するという仮定のもとに解を調べ、初期の密度ゆらぎのスペクトルに依存する形で、べきの値を求めている。ただ、この解が力学的に安定な解かどうかは分かっていない。

ゆらぎの時間発展としては、宇宙初期の微少な密度ゆらぎが重力不安定により成長して銀河などの構造を形成したと考える理論、つまり現在、標準的な銀河形成論である、重力不安定説を採用する。重力不安定説にもとづく解析としては、N 体シミュレーションがよく用いられ、さまざまなモデルに対して、非線形領域における二点相関関数がスケール則となる解も得られている [7] [20] が、本研究で考えているフラクタル次元を解析する場合には十分な解像度が得られないと思われる。そこで本研究では高い解像度を得るために半解析的なアプローチを用いる。

高い解像度を得る解析的な方法としては Zel'dovich 近似を用いる。Zel'dovich 近似ではゆらぎの振舞いを、ダスト粒子の Lagrange 座標における微小変位としてとらえ、この微小変位の線形近似を用いるため、密度ゆらぎそのものは線形近似にはなっておらず、非線形の領域でもよい近似となることが簡単な場合の厳密解との比較によって知られている [24]。特に、空間一次元モデルにおいてこのアプローチを用いた場合、解は厳密解となる。

一方で初期条件として考える時期の情報は、宇宙背景輻射の温度ゆらぎから得られる。宇宙背景輻射は、Pensiaz と Wilson によりいたる方向からもたらされるマイクロ波として発見され、長い間一様等方なものと見られていた。だが、観測衛星 COBE によりこの背景輻射の温度に 10^{-5} 程度のわずかなゆらぎが存在することが発見された。

*E-mail: tatekawa@gravity.phys.waseda.ac.jp
†E-mail: maeda@gravity.phys.waseda.ac.jp

このCOBEによる宇宙背景輻射の温度ゆらぎについて、輻射温度の等温線の周囲長と、その内部の面積との関係にフラクタル的なべき法則が見い出されたという解析も報告されている [5]。初期の密度ゆらぎとして断熱ゆらぎを考えると、温度ゆらぎと密度ゆらぎは比例関係にある。宇宙背景輻射にフラクタル的な関係が見い出されたならば、断熱ゆらぎを考慮した場合には再結合時に密度ゆらぎがフラクタル的であったことを示唆する。

本研究では銀河分布のスケール則が生じた一つの可能性として、初期の密度ゆらぎにフラクタル性がある場合の構造形成問題を考察する。一様・等方で平坦な膨張宇宙において、初期の密度ゆらぎによる重力不安定から構造が形成されたと考え、解析的な方法による時間発展を追う。

ここでは簡単なモデルとして、空間一次元モデルを扱う。フラクタル的な初期ゆらぎとして、線分を一定の比率で分割、除去するCantor集合を用いる。数学的にはこの操作を無限回くり返さなければならないが、実際の宇宙ではフラクタル的だとは考えらるスケールには下限が存在するので、途中でこの操作を打ち切って用いる。

時間とともに重力不安定により、銀河、銀河団などの非線形構造が形成されるが、密度の高い領域にこのような構造が出来ているとみなし、一定の密度以上になった領域に対し、フラクタル次元と二点相関関数の計算を行った。また構造が形成された時期の非線形構造のフラクタル次元と、初期条件として与えた密度ゆらぎのフラクタル次元との比較も行った。また、十分に時間が経過すると非線形構造のフラクタル次元は時間発展に対して一定値に落ち着いているように見え、その値は初期密度ゆらぎのフラクタル次元に近い値となった。また二点相関関数の計算に関しては規則的なパターンを与えるCantor集合をモデルとして用いたために、線形領域ではパターンの周期的な特徴が表れてしまい、べき則が存在しないが、十分に非線形な領域では構造はいくつかの集団にまとまっており、この集団のスケール以下では二点相関関数はべき則に従った。このべきは、非線形構造のフラクタル次元を反映する値となった。

II. 銀河・銀河団の分布

現在の銀河分布の観測では、Gellerらの観測 [11] やLas Campanas Redshift Survey [16] による銀河団の天球上での分布を見ると、特に際立った分布の片寄りが見られない。宇宙は超銀河団を超えるような非常に大きいスケールでは、一様等方であるとみなせると考えられる。そこで、時空の振る舞いを見る際には第ゼロ次近似として、一様等方なモデルを考えることができる。このモデルはRobertson-Walker計量で与えられる。

$$ds^2 = -dt^2 + a(t)^2 \left(\frac{dr^2}{1 - Kr^2} + r^2 d\theta^2 + r^2 \sin^2 \theta d\phi^2 \right) \tag{1}$$

ここで現われた$a(t)$およびKは空間座標(r, θ, ϕ)によらないパラメーターで、この2つのパラメーターにより宇宙のモデルが特定される。$a(t)$は宇宙のスケールファクターと呼ばれ、宇宙の大きさを表わす。Kは空間の曲率の符号を表わすパラメーターである。本研究では簡単のため平坦な宇宙を考える。この場合は$K = 0$とする。時空の振る舞いを見る時には、宇宙に存在する物質でどのような物質が優勢であるかを考えなければならない。ここでは星や銀河を構成するような、運動エネルギーの大きさが静止質量エネルギーに比べて無視できる物質を考える事にする。このような物質が優勢である場合、スケールファクターは

$$a \propto t^{2/3} \tag{2}$$

と記述できる。つまり宇宙は永遠に膨張するが、その膨張速度は減速していくということである。このような時空における構造の形成を考える。

第ゼロ近似で宇宙は一様・等方とみなすことができると述べた。観測してみると銀河は宇宙に一様に分布しているように見えるが、統計的手法を用いて分布の特徴を解析してみると一様ではない。

銀河の分布を統計的に解析する手法として二点相関関数が用いられる。これはある微小領域dV_1とdV_2が距離rだけ離れている時に、双方に銀河が見い出される確率を

$$dP = n(1 + \xi(r)) dV_1 dV_2 \tag{3}$$

と記述する。nは銀河の平均個数密度である。この平均分布からのずれ$\xi(r)$を二点相関関数という。銀河が全くランダムに分布していると$\xi = 0$となる。

銀河分布については、かつては銀河の分布は天球上の二次元的な分布しか得られなかったため、個々の奥行きの情報が得られなかった。よって天球面上に射影された二次元分布から実際の三次元分布を推測するしかなかった。天球面での二点の角度相関から三次元での分布を与える手法が考え出され、東辻と木原 [23] は二点相関関数が

$$\xi(r) \propto r^{-\gamma}, \gamma \simeq 1.8 \tag{4}$$

という距離のべきに比例することを明らかにした。

最近の観測においても、二点相関関数が距離のべきに従う事が示されている [11] [13] [14]。DavisとPeebles(1977)は、統計的考察から一様等方宇宙での非線形領域における二点相関関数の自己相似解を求め、このべき則を導いている [8]。

彼らの方法は、まず一様等方宇宙で Newton 的な重力相互作用をおよぼす質量 m の質点からなる物質を考えこれらの物質に関して、相空間での質点の分布の時間変化を記述する BBGKY 方程式を導出し、それに対するスケーリング解を求める。その結果を使うと、$\delta \equiv \rho - \langle\rho\rangle / \langle\rho\rangle$ という密度揺らぎを定義した際に、$\delta > 1$ という物質が集中した領域で、二点相関関数

$$\xi(r) \propto r^{-\gamma} \tag{5}$$

のべき γ が、初期密度ゆらぎのパワースペクトル $P(k)$

$$P(k) \propto k^n \tag{6}$$

のべき n によって次のような関係で与えられるというものである。

$$\gamma = \frac{9+3n}{5+n} \tag{7}$$

ただし、この解を求める際にいくつかの仮定を用いている。仮定の中で注目すべきものは、構造の固有運動と宇宙膨張による速度がつり合って、全体として速度が 0 になるというものである。この仮定が妥当なものかが分からず、この解が安定に存在するかがわかっていない。そこで力学的に時間発展を行い、二点相関関数がべき則に従うかを調べる必要がある。

III. 宇宙背景輻射の温度ゆらぎ

1965 年に Penzias と Wilson によって、宇宙にはマイクロ波の背景輻射が発見され、その電波は銀河系、太陽系などからの輻射や宇宙背景輻射に対する地球の固有運動の影響を取り除くと非常に等方的であり、2.7K のプランク分布を示していることがわかった。この背景輻射は COBE によって精密に調査され温度が 2.725 ± 0.020K であり、また固有運動の影響を取り除いた CMB の天球上での温度ゆらぎは、天球上の 7 度角の分解能において $\delta T/T \simeq 10^{-5}$ というものであったということが明らかにされた。

この背景輻射の温度ゆらぎに対し、E. M. de Gouveia Dal Pino ら [5] は次のような解析を行った。背景輻射の温度ゆらぎは銀河面などの一部を除いて、ほぼ全域に渡って観測されている。この観測結果により、天球上に等温線を引くことができる。この等温線の長さと等温線で囲まれる領域の面積との関係について解析した。つまり全天を非常に細かいメッシュで区切り、ある等温線に着目して、等温線を含むメッシュの数と等温線で囲まれる領域の関係を調べていった。等温線の周囲長と内部の面積の関係を調べた。

$$a^{1/2} = Fp^{1/D} \tag{8}$$

a : 等温線で囲まれる領域の面積

p : 等温線の長さ

F : 比例定数

E. M. de Gouveia Dal Pino らの解析の結果、$D = 1.43 \pm 0.07$ という値が得られ、温度ゆらぎにフラクタル性が存在する事を主張している。

宇宙初期の密度ゆらぎとして断熱ゆらぎを仮定すると

$$\frac{\delta T}{T} = \frac{1}{3}\frac{\delta \rho}{\rho} \tag{9}$$

なる関係が成り立つ。この関係は、電磁波と相互作用する物質であるバリオンに対してのみ成り立つ。ここから、背景輻射がフラクタル的であったとすると宇宙初期のバリオンの密度ゆらぎもフラクタル的な分布であったことが示唆される。

IV. フラクタル

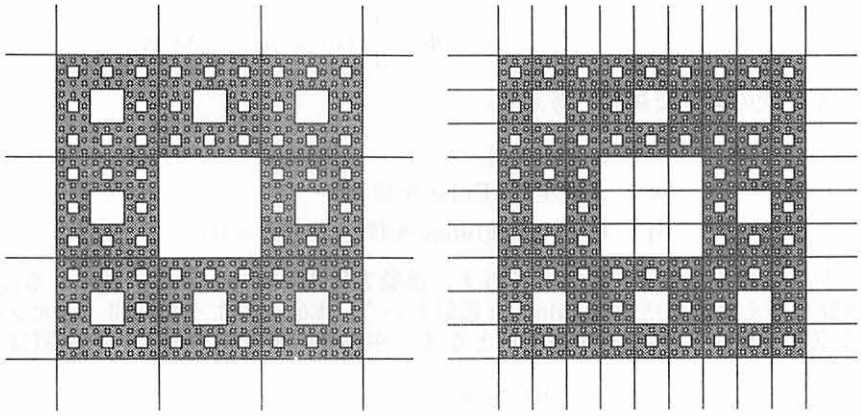
フラクタルとは、特徴的な長さを持たない図形や構造、現象などを総称したもので、自己相似性が重要な特徴として挙げられる。着目している構造について、ある一部分がそれよりも大きい部分、または全体と相似関係にあるということである。自然現象においてフラクタル性を確かめるにはフラクタル次元を計算する必要がある。

フラクタル次元は数学的に厳密な定義があるが、自然現象を解析する際には非常に扱いにくい。そこで次元の定義を実用的なものに改良する必要がある。改良された次元の定義はいくつかあるが、ここでは本研究で用いたものだけを紹介する [10], [22]。

a. Box-counting の方法：　数学的には対象物をある半径の球で覆いつくし、この球の半径を無限小にした極限での球の個数と半径の関係からフラクタル次元を定義する。それと数学的に同等な定義として、d 次元 Euclid 空間に存在する図形を一辺の長さが r の d 次元立方体で分割し、考えている図形の一部を含むような立方体の数 $N(r)$ を数え上げていく方法がある。r を十分小さく取ると

$$N(r) \propto r^{-D} \tag{10}$$

なる関係が満たされる場合、この図形のフラクタル次元は D 次元であるという。この方法を Box–counting の方法という。この方法は曲線や立体などに適用されるだけではなく、例えば川のように分岐をたくさん含むような図形の解析にも適用できる。本研究でもこの Box–counting の方法を用いてフラクタル次元の計算を行う。図 1 ではフラクタル図形である Sierpinski gasket を分割したものを例として示す。



(a)　　(b)

FIG. 1. (a) $N\left(\frac{1}{3}\right) = 8$ (b) $N\left(\frac{1}{9}\right) = 64$

b. 相関次元：　二点相関関数からもフラクタル次元が定義できる。相関関数が指数型（$\exp(-r/r_0)$）やガウス型（$\exp(-r^2/2r_0^2)$）である場合は特徴的な距離 r_0 が存在するためにフラクタルにはならない。相関関数が距離のベキになっている場合は特徴的な距離が存在しないために相関の落ち方がいつも同じ割合になる。たとえば

$$\xi(r) \propto r^{-\gamma} \tag{11}$$

となっていたとすれば、距離が二倍になれば相関が $1/2^\gamma$ 倍になるという関係がどのような距離でも満たされる。ベキの指数 γ とフラクタル次元 D との間には次のような関係がある [10] [22]。

$$\gamma = d - D \tag{12}$$
$$d \ : \text{空間の次元}$$

V. 重力不安定による構造形成

宇宙の構造の作る重力ポテンシャルが非相対論的で、物質が完全流体であると見なした時は、Newton 力学をもとにした次の三つの方程式が構造の進化に対する基礎方程式になる。

$$\left(\frac{\partial \rho}{\partial t}\right)_r + \nabla_r \cdot (\rho \mathbf{u}) = 0 \tag{13}$$

$$\left(\frac{\partial \mathbf{u}}{\partial t}\right)_r + (\mathbf{u} \cdot \nabla)_r \mathbf{u} = -\frac{1}{\rho}\nabla_r P + \frac{\mathbf{f}_{ex}}{\rho} \tag{14}$$

$$\mathbf{f}_{ex} = -\nabla_r \Phi, \nabla^2 \Phi = 4\pi G\rho \tag{15}$$

時間微分の添字は、物理的な座標系での時間微分である。Hubble の法則より、宇宙は膨張しているので、宇宙膨張に乗った新しい座標系（背景宇宙の共動座標系）

$$\mathbf{x} = \frac{\mathbf{r}}{a(t)}, a(t) : \text{scale factor} \tag{16}$$

で方程式を書き直すと、次のような基礎方程式が得られる。

$$\frac{\partial \delta}{\partial t} + \frac{1}{a}\nabla \cdot \{\mathbf{v}(1+\delta)\} = 0 \tag{17}$$

$$\frac{\partial \mathbf{v}}{\partial t} + \frac{1}{a}(\mathbf{v}\cdot\nabla)\mathbf{v} + \frac{\dot{a}}{a}\mathbf{v} = -\frac{1}{a}\nabla\phi - \frac{1}{a\rho}C_s^2\rho_b\nabla\delta \tag{18}$$

$$\nabla^2\phi = 4\pi G\rho_b a^2\delta \tag{19}$$

$$\delta \equiv \frac{\rho - \rho_b}{\rho_b}$$

$$\phi = \Phi - \frac{2}{3}\pi G a^2\mathbf{x}^2\rho_b + \frac{1}{6}\Lambda a^2\mathbf{x}^2$$

これらの方程式系において、次のような摂動を考える。

$$\mathbf{x} = \mathbf{q} + \mathbf{S}(t, \mathbf{q}) \tag{20}$$

$\mathbf{x}$: 共動座標 (Eular 座標)

$\mathbf{q}$: 粒子の Lagrange 座標, $(\nabla \times \mathbf{S} = 0)$

つまり、粒子の一様分布からの微小変位を摂動として与え、運動方程式を解く。この摂動による近似法を Lagrange 近似といい、特に一次まで摂動を考える近似法を Zel'dovich 近似という。本研究のような空間一次元シート系では、Lagrange 近似の一次の近似である Zel'dovich 近似の解が厳密解となる。平坦な背景宇宙の場合、厳密解は

$$x = q + a(t)S_1(q) + a(t)^{-3/2}S_2(q) \tag{21}$$

で考えられる。S_1、S_2 は初期条件から決まる。

ところで Lagrange 近似では shell crossing という問題が生じる。これは密度ゆらぎが成長していくと密度無限大の点が現われ、以降の時間発展が追えなくなるという問題である。その後の現象を扱うためには、今考えている一次元のシート系では、一般に N 体シミュレーションによりそのシート同士の相互作用を考慮する必要がある。本研究では簡単のため衝突に際してシート同士はすり抜けるものとして取り扱う。このようなモデルにおいては、シートが交差した時に密度無限大の点を生じるが、その後も解析的に取り扱う事ができ、十分な時間発展を行う事ができる [12]。

VI. 構造のフラクタル次元の時間変化

本研究では簡単な場合として一次元シートモデルを取り扱った。つまり、シート状の物質がシートと垂直方向にのみ運動するモデルである。宇宙背景輻射の温度揺らぎに対する解析からは、バリオンの密度揺らぎだけがフラクタル的であるということになるが、本研究では構造を形成する全ての物質についてフラクタル的な密度ゆらぎが存在したと仮定する。初期にフラクタル的な密度ゆらぎを与えた場合に、時間発展でどのような構造が形成されていくかについて、非線形構造のフラクタル次元と二点相関関数を求めた。

密度ゆらぎとしては、フラクタル的なものを与える必要がある。そこで本研究では初期条件として Cantor 集合を応用したものを用いた。ただし、数学的に定義された Cantor 集合では無限小までフラクタル性が存在するが、現実の宇宙ではそのような事はあり得ないと考え下限を与えた。モデル全体のスケールを L とした時に、下限となるスケールが $10^{-3}L$ 程度になるようなモデルを考えた。つまり Cantor 集合を形成する操作を行う際に、線分を取り除く回数を有限でとどめる。以後、このスケール L は宇宙の膨張に乗った共動座標を表わす事にする。そうすれば、時間発展の結果によりスケールが変化しても座標変数の値を大きく変える必要がなくなる。本研究では 1/10 ずつ線分を取り除く操作を 7 回繰り返した集合を用いた。この Cantor 集合のフラクタル次元は

$$D = \frac{\log 2}{\log 20/9} \simeq 0.868 \tag{22}$$

である。

線分を取り除いた領域を密度ゆらぎが負 (δ_-)、線分が残った領域を密度ゆらぎが正 (δ_+) としたモデルを与えた。δ_+, δ_- は一定値としている。この線分を取り除く操作を 7 回繰り返した集合は、今考えている Cantor 集合の規則に従って分布していると考える。つまり、長さ L の集合が間隔 $2L/9$ だけ離れて存在しており、これらのセットがまた $(2/9)^2L$ だけ離れて存在しているということにする。この際に L 以上のスケールのパターン同士の間の密度揺らぎは 0 とする。時間発展の初期にはそれぞれのパターンが独立に構造を形成していくが、十分に時間が経過すると複数のパターンが合体する事がある。

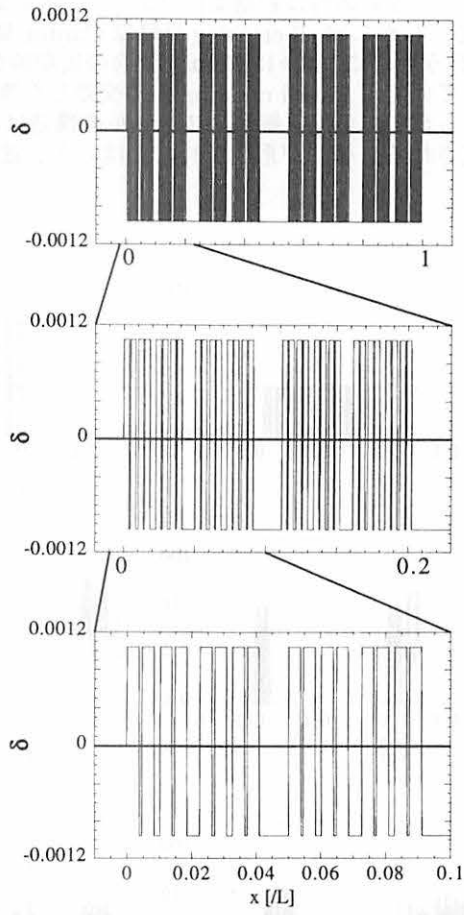


FIG. 2. Primordial density fluctuaion

初期を $a=1$ とした。これを再結合時とすると、現在は $a=10^3$ に対応する。本研究では全ての物質がフラクタル的な密度ゆらぎをなしていたと考え、また $a=10^3$ で観測されるような構造が形成されるために必要な密度ゆらぎの大きさを考慮し、

$$|\delta_+| \simeq |\delta_-| \simeq 10^{-3} \tag{23}$$

とした。そこで初期の密度ゆらぎは

$$\delta_+ = 1.0434062 \times 10^{-3} \tag{24}$$

$$\delta_- = -9.565938 \times 10^{-4} \tag{25}$$

を与えた。この初期条件を与えた場合、shell crossing が現れる時刻は

$$a^* = 958.3994 \tag{26}$$

となる。$a=10^3$ の頃に shell crossing が発生し、銀河、銀河団などの非線形な構造が形成される。本研究ではさらに時間発展をすすめ、$a=2\times10^4$ まで時間を進めた。

まず、構造のフラクタル次元を計算する。フラクタル次元の計算方法としては、密度ゆらぎが $\delta \geq 1$ である領域を Box–counting の方法で拾い集める方法を用いた。十分に時間が経てば銀河、銀河団などの構造が形成されるが、このような構造が形成された領域は構造形成の理論でいう非線形領域であるため、$\delta \geq 1$ の領域に着目した。初期 $(a=1)$ の密度ゆらぎは $|\delta_0| \simeq 10^{-3}$ であるので、$\delta \geq 1$ である領域が発生する時期は $a \simeq 5\times10^2$ である。よって、$a \geq 5\times10^2$ でのみフラクタル次元を計算した。

Box–counting の方法を用いる際、shell crossing までは構造の最小のスケール、すなわちモデルとして与えた Cantor 集合の最小の線分の長さを下限として計算した。この線分の長さは自己重力による収縮により時間変化する。shell crossing 後は構造が複雑に入り組むようになるので、最小のスケールとしては $2^{-16}L$ を下限として計算した。上限は Box の数が少なくなるために生じる誤差を避けるため、$10^{-1}L$ 程度とした。すなわちモデル全体を 10 個程度に区切るスケールである。これは図 3 で示す非線形構造で、1 つの集団のスケールにほぼ相当する。

図3を見ると、shell crossing前である$a = 500$の時は初期条件であるCantor集合の徴候を残しており、shell crossing直後である$a = 1000$でもその徴候はある。しかしshell crossing以降はCantor集合より作られたモデルの最小スケールの構造が互いに重なりあって相互作用を行うために、徐々にCantor集合の周期的なパターンは崩れていく。shell crossingまでは物質は加速しながら構造を形成していくが、shell crossing後は交差した物質同士の相互作用のために、減速していくためにだんだんと非線形構造がつくられてていく。細かいスケールの構造は次々と合体していき、$a = 5000$程度でいくつかの巨大な集団による構造が形成される。構造が形成されたとはいえ、速度は0にはなっていないので、その形を変えながら発展していく。

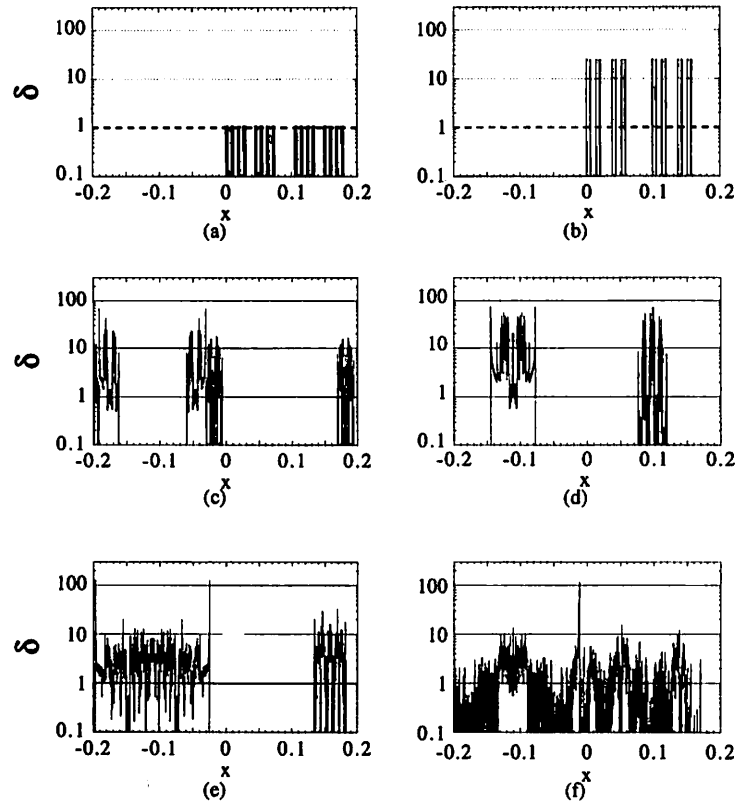


FIG. 3. Time evolution of density fluctuation. (In this figure, part of model is shown.) (a)$a = 500$, (b)$a = 1000$, (c)$a = 6000$ (d)$a = 10000$, (e)$a = 15000$, (f)$a = 20000$

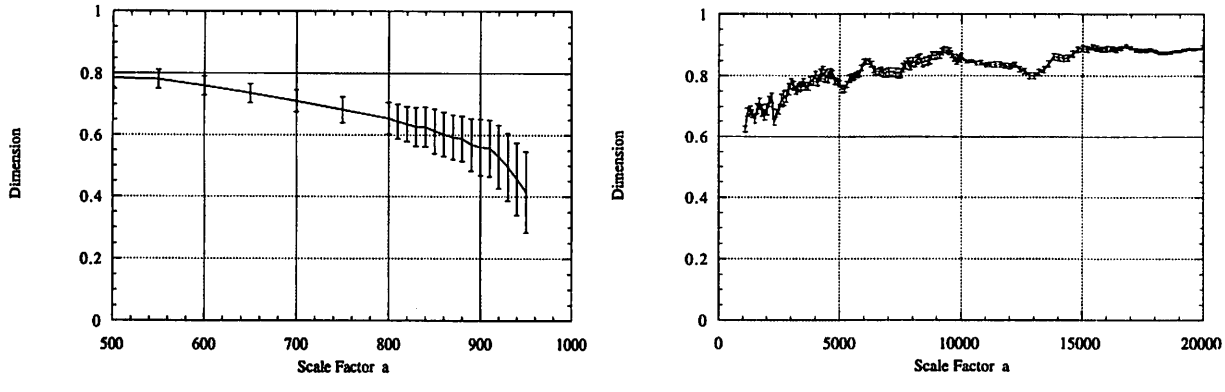


FIG. 4. Fractal dimension of nonlinear structure

この結果を見ると、まずshell crossing前は時間発展によりフラクタル次元が低くなっていく様に見えるが、誤差が非常に大きいため構造をフラクタル的であるとみなすには問題があると思われる。shell crossing後は振動をしながら徐々

にフラクタル次元が高くなっていき、$a = 15000$ 以降はほぼ一定になっているようである。次元の誤差も非常に小さくなっているので、フラクタル的な非線形構造が出来ているとみなせる。この時のフラクタル次元は初期密度ゆらぎのフラクタル次元 $D = 0.868$ に近い値をとる。十分に時間が経過した後の非線形構造のフラクタル次元は、初期密度ゆらぎのフラクタル次元を反映するかどうかは、いくつかの初期条件で計算している。

また、非線形構造の二点相関関数も計算する。二点相関関数 ξ の定義は (3) の通り、距離 r だけ離れた二点に銀河が存在する確率から求める。本研究では $\delta \geq 1$ である場所に構造が出来ていると考え、このような領域の分布から二点相関関数を計算する。このため密度ゆらぎ $\delta(x)$ の二点相関関数

$$\xi_\delta(r) \equiv \int \delta(x)\delta(x+r)dx \tag{27}$$

とは異なる結果を与える。

各時刻での構造の分布に対して、定義から二点相関関数を求める。shell crossing 直後である $a = 1000$ では、Cantor 集合の規則性のために r の増加に対して一旦は負になった二点相関が再び正になるという徴候が周期的に見られる。その後は Cantor 集合の規則性が徐々に崩れていくために、二点相関関数の周期性は見られなくなっていく。

それで十分に時間が経過した後の二点相関関数を計算すると、図 3 で示したように、この集団のスケール以下では関数が正の値を与える事が分かった。IV で述べたように、フラクタル分布をしている構造の二点相関関数は次のような関係が成り立つから

$$\xi(r) \propto r^{-(d-D)} \tag{28}$$

d : フラクタル構造が存在する空間の位相次元

D : フラクタル構造の次元

このスケールにおいてべき則が存在するかどうか調べた。例えば本研究のように空間一次元モデル ($d = 1$) において、構造のフラクタル次元が初期密度ゆらぎのフラクタル次元 ($D = 0.868$) をそのまま反映したとすると、二点相関関数は

$$\xi(r) \propto r^{-0.132} \tag{29}$$

となる。本研究のモデルの二点相関関数のグラフは図 5 のようになる。

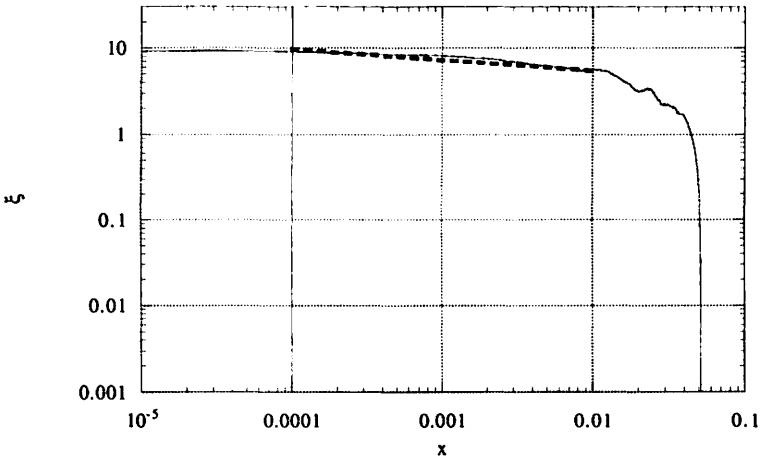


FIG. 5. 2–point correlation function of nonlinear structure ($a = 10000$)

この二点相関関数がべき則に従っているかどうかを見るため、図 4 における各時刻のフラクタル次元から (28) の関係によりべきを求め、このべきにあうかどうか調べた。二点相関関数はそれぞれの集団のスケールで負の値に落ちる。べき則を見る際はこのスケールではなく、それより小さいスケールで調べる。集団の大きさは $0.05L$ 程度であるが、このスケールではべき則を満たしているとは到底言えない。この中に $0.01L$ 程度の構造がいくつか存在する。この $0.01L$ に相当する構造のスケール以下で二点相関関数のべきを解析した。$a = 10000$ の場合に、

$$\xi(x) \propto x^\gamma,\ \gamma = -0.130 \pm 0.005 \tag{30}$$

という値が得られた。他の時刻についても調べ、結果を図 6 で示した。

図を見ると、二点相関関数のべきは非線形構造のフラクタル次元を反映していると考えられる。ただし、この二点相関関数は $\xi \simeq 1$ 付近では、はっきりしたべき則を持たない。これは、初期密度ゆらぎが Cantor 集合をもとに生成され、スペクトルが波数のべきに従うようなスケールフリーな分布をしていないためと考えられる。

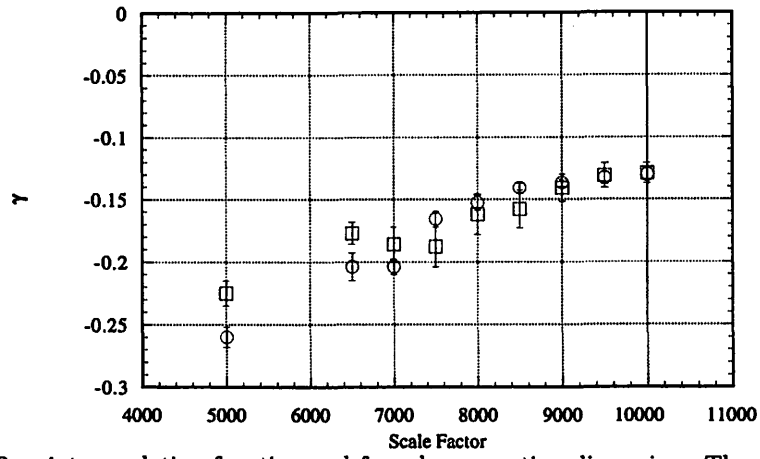


FIG. 6. Power index of 2-point correlation function and from box-counting dimension. The circle show power of 2-point correlation, the box show power from fractal dimension.

また、Davis と Peebles のスケーリング解を与える際の仮定として、初期密度ゆらぎの相関に強い制限が与えられている。三点以上の相関はないという仮定である。ところが本研究で考えた密度ゆらぎは同じパターンが繰り返し現れるように非常に規則性が高い。このことも、二点相関関数が $\xi \simeq 1$ 付近では、はっきりしたべき則を持たない事を説明する理由と考えられる。

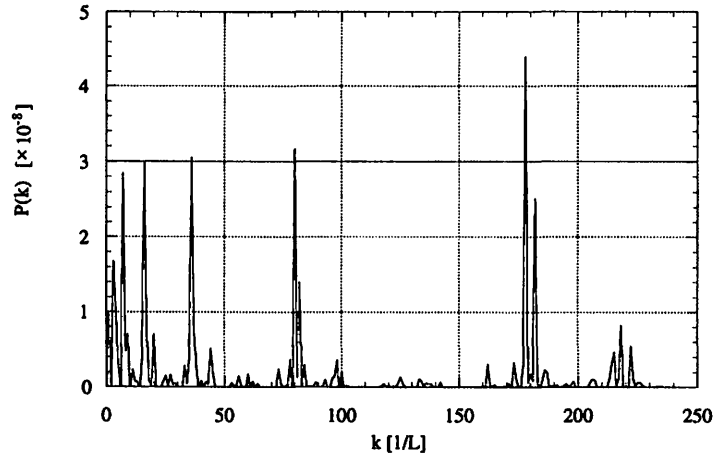


FIG. 7. Spectrum of Initial density fluctuation

VII. CONCLUSION

本研究では線分を 1/10 ずつ取り除く Cantor 集合を用いて初期密度ゆらぎを与えた。この初期密度ゆらぎのフラクタル次元は 1 に近い。十分に時間が経過した後の構造のフラクタル次元は、初期密度ゆらぎのフラクタル次元に近い値を与えたが、それが異なる初期条件を設定して初期密度ゆらぎのフラクタル次元を変えた場合も同じ事がいえるのか、それとも一次元シート系での重力不安定による構造形成の性質で同じようなフラクタル次元を与える非線形構造を構成するのかという点を解析すべきである。

また、二点相関関数は非線形構造のフラクタル次元を反映するものとなった。ただし、十分に時間が経過しないとこの傾向は見られない。なぜなら、モデルとして与えた初期密度ゆらぎが規則的すぎるために、最初のうちはこの規則性が二点相関関数に色濃く現れてしまうためである。これをなくすには、初期密度ゆらぎとして与えるフラクタルのモデルにランダム性を与える必要がある。例えば random Cantor 集合やランダムウォークを応用したモデルが考えられる [10]。

現実的な問題に照らし合わせるには空間三次元モデルを考える必要がある。空間三次元モデルを取り扱う場合は、現在用いている方法には問題がある。空間一次元モデルの場合は Zel'dovich 近似が厳密解を与えるために詳細な解析を行う事が可能であったが、空間二次元以上では Zel'dovich 近似は線形摂動解でしかない。より次数の高い Lagrange 近似を用いても、非線形構造を十分に記述するとはいいがたい。空間三次元モデルを取り扱う場合には高い解像度を得るこ

とができ、なおかつ高い精度で近似が成り立つ方法を用いなければならない。最近になって Couchman & Peebles [4] による繰り込みを用いた N 体シミュレーションが提唱されているが、この手法には Davis と Peebles のスケーリング解の関係式を用いているため、妥当な方法かという点で疑問が生じる。だが、フラクタルという規則的な階層構造を利用して、繰り込み群的な手法を用いて空間と時間のスケーリングを行い、少ない計算量で高い解像度を得る方法を見い出すという考え方は有効と思われる。

---

[1] C. L. Bennett, A. J. Banday, K. M. Gorski, G. Hinshaw, P. Jackson, P. Keegstra, A. Kogut, G. F. Smoot, D. T. Wilkinson, and E. L. Wright, ApJ Letters **464**(1996) L1

[2] F. R. Bouchet, R. Juszkiewicz, S. Colombi, and R. Pellat, ApJ **394** (1992) L5

[3] F. R. Bouchet, S. Colombi, E. Hivon, and R. Juszkiewicz, Astron. Astrophys., **296** (1995) 575B

[4] H. M. P. Couchman and P. J. E. Peebles, ApJ **497** (1998) 499

[5] E. M. de Gouveia Dal Pino, A. Hetem, J. E. Horvath, C. A. W. de Souza, T. Villela, and J. C. N. de Araujo, ApJ **442** (1995) L45

[6] G. B. Dalton, R. A. C. Croft, G. Efstathiou, W. J. Sutherland, S. J. Maddox, and M. Davis, MNRAS **271** (1994) L47

[7] M. Davis, G. Efstathiou, C. S. Frenk, and S. D. M. White, ApJ **292** (1985) 371

[8] M. Davis and P. J. E. Peebles, ApJ Suppl. **34** (1977) 425

[9] G. de Vaucouleurs, Science **167** (1970) 1203

[10] K. Falconer, *Fractal Geometry*, (Wiley 1990)

[11] M. J. Geller and J. P. Huchra, Science **246** (1989) 897

[12] N. Gouda and T. Nakamura, Prog. Theor. Phys. **81** (1989) 633

[13] L. Guzzo, J. G. Bartlett, A. Cappi, et al. (The ESP Team), *Proceedings of XIV IAP Colloquium "Wide Field Surveys in Cosmology"* , eds. S. Colombi and Y. Mellier, Paris, May 26-30 1998
L. Guzzo et al., astro-ph/9901378

[14] Y. P. Jing, H. J. Mo, and G. Börner, ApJ **494** (1998) 1

[15] L. Kofman, D. Pogosyan, S. F. Shandarin, and A. L. Melott, ApJ **393** (1992) 437

[16] S. A. Shectman et al., ApJ **470** (1996) 17
The Las Campanas Redshift Survey WebPage :
http://manaslu.astro.utoronto.ca:80/~lin/lcrs.html

[17] P. J. E. Peebles, *The Large Scale Structure of the Universe*, (Princeton University Press 1981)

[18] P. J. E. Peebles, *Principles of Physical Cosmology*, (Princeton University Press 1993)

[19] V. Sahni and P. Coles, Phys. Rep. **262** (1995) 1

[20] 須藤　靖, パリティ物理学コース クローズアップ ダークマターと銀河宇宙, (丸善 1993)

[21] J. Loveday and J. Pier, 14th IAP meeting "Wide field surveys in cosmology", ed. Y. Mellier et al.
Sloan Digital Sky Survey WebPage : http://www.sdss.org/

[22] 高安　秀樹, フラクタル, (朝倉書店 1986)
高安　秀樹, フラクタル科学, (朝倉書店 1987)

[23] H. Totsuji and T. Kihara, Pub. Astron. Soc. Japan, **21** (1969) 221

[24] A. Yoshisato, T. Matsubara, and M. Morikawa, ApJ **498** (1998) 49

[25] Ya. B. Zeldovich, Astron. Astrophys. **5** (1970) 84

Reformulation of Scattering Problem in an Expanding Universe

Eliani ARDI[1,2], Toshio TSUCHIYA[1] and Shogo INAGAKI[1]

Department of Astronomy, Kyoto University, Japan
Department of Astronomy, Institute of Technology Bandung, Indonesia

Abstract

We reformulate the scattering problem so that the time change of the cross section due to cosmic expansion can be taken into account in evaluating scattering cross section and relaxation time of galaxies. We calculate numerically the orbits of encounters to various sets of orbital parameters while specifying the initial states of the encounters: initial separation, impact parameter, and initial relative velocity between two galaxies. Combining the results of these studies, we obtain the two-body relaxation time as a function of the relative velocity v_0 at the initial epoch and the number density n of galaxies within proto-clusters of galaxies. The relaxation time is roughly the free-fall time for a small relative velocity, $v_0 \leq 300$ km s$^{-1}(m/10^{11}M_\odot)^{1/3}[11/(1+z)]^{1/2}$, while it is proportional to ${v_0}^3$ for a high relative velocity. The effects of the cosmic expansion make the relaxation time longer than that in non-expanding systems, especially for proto–clusters of galaxies with the number density $n < 3 \times 10^3$ Mpc^{-3} $(10^{11}M_\odot/m)[(1+z)/11]^3$, where m is the mass of a galaxy and z is its redshift.

1. Introduction

There have been many studies devoted to determining the relaxation time in non-expanding systems (Chandrasekhar 1942, Spitzer 1969, Spitzer and Hart 1971). However, they cannot be applied to study the relaxation time of large-scale expanding systems because the effects of cosmic expansion are not taken into account. The aim of this paper is to examine the fundamental physics of the effects of cosmic expansion on a gravitationally clustered system which is initially started from a random distribution of galaxies.

A previous study on the distribution of stochastic forces during the formation of a cluster of galaxies (Ardi, Inagaki 1997) showed that the force from the nearest-neighbor galaxy is the dominant one acting on each galaxy. Therefore, in this paper we discuss the two-body relaxation time in the expanding universe.

We start by studying only large-angle scattering for the following reasons. First, the equation of motion cannot be solved analytically, so that statistical examination of the cumulative effect of small-angle scattering becomes more difficult in the expanding case. Second, in the non-expanding case the relaxation time determined by large-angle scattering is longer than that determined by the small-angle one only by some logarithmic factors. Hence, the large-angle scattering does not give a very bad estimate of the relaxation time. We also know that in the non-expanding case, the relative importance of large-angle scattering on the relaxation time increases with a decrease in N. In a cluster of galaxies $N \leq 10^3$, which is smaller than $N \approx 10^5$ in globular clusters. Thus, neglecting small-angle scattering may not be so bad compared with

in the non-expanding case. Third, because expansion decreases the relative velocities between particles, a test particle tends to be scattered by a large-angle in the expanding case. We really found that some encounters form binaries, although in the non-expanding case they are not formed with the same collision parameters. This suggests that in the expanding case, the contribution of large-angle scattering on the relaxation time is higher than in the non-expanding case.

2. Reformulation of the Scattering Problem

In the case of static stellar systems, the relaxation time is represented by the mean free time of a typical particle (i.e. star) which is moving through the medium, and is given by

$$t_{\rm relax} = \frac{1}{nQv}, \tag{1}$$

where n, Q, and v denote the number density of the field particles, the cross section of scattering by a field particle, and the relative velocity between a test particle and the field particle, respectively. In the expanding universe, however, this formula should be modified because the velocity of a test particle decreases with time, even for free motion. Consequently, the scattering cross section changes with time. We thus need to reformulate the scattering problem so that the effects of expansion of the medium can be taken into account in evaluating the scattering cross section and the relaxation time.

In an ordinary scattering problem, the volume in which a test particle is scattered by a field particle is Qvt, and increases linearly with time, except for cases of extremely low velocity (free-fall) and of short time t. In the expanding universe, however, the corresponding volume, which is referred to hereafter as the *scattered volume* $V_{\rm scat}$, no longer has the same parameter dependence as that in the non-expanding case, because the velocity of the test particle approaching the field particle changes with time due to expansion of the medium (see below for more details). The mean free time is given by the time when the volume $V_{\rm scat}$ becomes large enough to contain one field particle. Determination of the scattered volume is as follows.

We consider two identical point masses which move in the expanding universe under mutual gravitational interaction. Let $\boldsymbol{x}_1$, $\boldsymbol{x}_2$, and m denote the positions and masses of the two particles. The equation of motion describing the relative motion of these two particles is given by (e.g. Peebles 1980)

$$\ddot{\boldsymbol{x}} + 2\frac{\dot{a}}{a}\dot{\boldsymbol{x}} = -\frac{2Gm}{a^3}\frac{\boldsymbol{x}}{|\boldsymbol{x}|^3}, \tag{2}$$

where $\boldsymbol{x} = \boldsymbol{x}_2 - \boldsymbol{x}_1$, G is the gravitational constant and a is the scale factor of the expansion of universe. We assume the matter-dominant Einstein–de Sitter universe as the model of the expanding universe.

Particle 2 is initially (at $t = t_0$) located at (x_0, y_0) with velocity $\dot{x} = -v_0$, $\dot{y} = 0$. If it moves freely without feeling the gravitational force of particle 1, the solution of equation (2) is

$$\begin{cases} x_{\rm free}(t) &= x_0 - 3v_0t_0\left\{1 - \left(\frac{t_0}{t}\right)^{1/3}\right\}, \\ y_{\rm free}(t) &= y_0, \\ v_{x,\rm free}(t) &= \dot{x}(t) = v_0\left(\frac{t_0}{t}\right)^{4/3}, \\ v_{y,\rm free}(t) &= \dot{y}(t) = 0. \end{cases} \tag{3}$$

With the gravitational force of the particle 1, however, the actual orbit of particle 2 deviates from that of the free motion described by equations (3). Once the deviation becomes large enough (we will give a numerical criterion below), particle 2 is called *scattered* and that time is referred to as the scattering time t_{sc}, which is a function of x_0, y_0, and v_0.

The scattered volume V_{scat} at an arbitrary time t is defined such that any particle (with a given v_0) in V_{scat} is scattered in a time t_{sc} (which depends on x_0 and y_0) less than t, and any particle outside the volume cannot be scattered in a time less than t :

$$V_{scat}(v_0, t) = \int_{t_{sc}<t} 2\,\pi\, y_0\, dy_0\, dx_0, \tag{4}$$

This volume increases with time, and the relaxation time t_{relax} can be evaluated as the time when the scattered volume contains one field particle, i.e.,

$$nV_{scat}(v_0, t_{relax}) = 1. \tag{5}$$

3. Analytic Evaluation of the Scattered Volume

3.1.High-Speed Approximation

Consider the orbit of particle 2 relative to particle 1. Particle 2 moves freely on a straight line [equations (3)] until it reaches a point ($x \approx 0$) very close to particle 1, and is then scattered instantly. Time t_{sc} and velocity v_{sc} of particle 2 at this moment are given by

$$\left(\frac{t_0}{t_{sc}}\right)^{1/3} = 1 - \frac{x_0}{3v_0t_0}, \qquad \text{and} \qquad v_{sc} = v_0\left(\frac{t_0}{t_{sc}}\right)^{4/3}. \tag{6}$$

The gravitational scattering after this moment is supposed to be the same as that of the Rutherford scattering. This gives an initial separation of x_0 and the largest impact parameter y_{max} of particles which are scattered larger than 90° in angle at $t = t_{sc}$ as

$$\begin{aligned} x_0(t_{sc}) &= 3v_0t_0(1 - (t_0/t_{sc})^{1/3}), \\ y_{max}(t_{sc}) &= \frac{2Gm}{a_0^3v_0^2}\left(\frac{t_{sc}}{t_0}\right)^{2/3}. \end{aligned} \tag{7}$$

Following formula (4), the scattered volume is

$$V_{scat}(v_0, t) = \int_{t_0}^{t} \pi y_{max}(t_{sc})^2 \frac{dx_0}{dt_{sc}} dt_{sc} = \frac{4\pi G^2m^2}{a_0^6v_0^3}(t - t_0). \tag{8}$$

3.2.Free-Fall Approximation

In the expanding universe, even when the relative velocity in the comoving coordinates is zero, there is the Hubble recession velocity in the proper coordinates. Therefore, the free-fall approximation is valid only in the vicinity of the nearest field particle, where the kinetic energy of the recession is much smaller than the gravitational energy between two particles,

$$\frac{1}{2}(\dot{a}_0x_0)^2 = \frac{1}{2}\left(\frac{2}{3}\frac{a_0}{t_0}x_0\right)^2 \ll \frac{2Gm}{a_0x_0}. \tag{9}$$

The free-fall time $t_{\rm ff}$ is given approximately by

$$t_{\rm ff} \approx \frac{\pi}{4}\sqrt{\frac{a_0^3 x_0^3}{Gm}}. \tag{10}$$

Therefore, when the free-fall approximation is adequate, the scattered volume $V_{\rm scat}$ becomes

$$V_{\rm scat}(v_0, t) = \frac{4\pi x_0^3}{3} \approx \frac{64Gm}{3\pi a_0^3}(t - t_0)^2. \tag{11}$$

4. Numerical Determination of Relaxation Time

We discuss our numerical integration of the equation of motion (2) using the fourth-order Runge-Kutta method with adaptive step-size control (Press et al. 1992). The initial parameters for calculating orbits of a test particle are three, i.e. x_0, y_0, and v_0 at t_0. First we give a value of v_0, then thousands of pairs of (x_0, y_0) are taken from mesh points in x–y plane and the integration is performed for each case.

In our calculation, all quantities were normalized by setting $G = 2m = t_0 = 1$ as units of simulations. Assuming that the unit of time is 2.74×10^8 yr (the redshift $z = 10$) and a typical galaxy with mass $m = 10^{11} M_\odot$, the units of length and velocity are 40 kpc and 147 km s^{-1}, respectively.

In order to specify numerically the condition of "strong scattering", we introduce a velocity deviation $\Delta v(t)$ and a position deviation $\Delta r(t)$:

$$\begin{aligned} \frac{\Delta v^2}{v_{\rm free}^2}(t) &= \frac{[v_x(t) - v_{x,\rm free}(t)]^2 + v_y(t)^2}{v_{x,\rm free}^2}, \\ \frac{\Delta r^2}{r_{\rm free}^2}(t) &= \frac{[x(t) - x_{\rm free}(t)]^2 + [y(t) - y_{\rm free}(t)]^2}{x_{\rm free}^2 + y_{\rm free}^2}, \end{aligned} \tag{12}$$

where $v_{x,\rm free}$, $x_{\rm free}$, and $y_{\rm free}$ are the solutions of the free motion given by equations (3).

The deviations increase with time, and when both of them become as large as unity, the particle is regarded as *scattered.* Hence, the time of scattering $t_{\rm sc}$ is determined by

$$\min\left[\left(\frac{\Delta v}{v}\right)(t_{\rm sc}),\ 2\left(\frac{\Delta r}{r}\right)(t_{\rm sc})\right] = 1. \tag{13}$$

In the case of a high-speed encounter, the criterion of the velocity deviation is crucial: the condition $\Delta v/v = 1$ discriminates large-angle scattering ($\chi \approx 45°$). When the velocity is small, on the other hand, the relative change of the velocity quickly becomes large, because of the small denominator [see equation (12)]. In that case the criterion of the position deviation is employed. Factor 2 of the position deviation in equation (13) is introduced to define the free-fall time as the time when the particle falls to half of the initial radius.

Figure 1 shows three examples of orbits with the same initial velocity, $v_0 = 4$, but with different initial positions (x_0, y_0). Since orbit 1 has a small impact parameter, it is scattered by an angle larger than 45°. On the other hand, orbit 2 is not close enough to experience large-angle scattering. Orbit 3 has a larger impact parameter and a larger initial separation, thus the particle takes longer time to reach $x = 0$ than do the others. Owing to cosmic expansion, the velocity decreases considerably with time. If there is no gravitational force of particle 1,

the incident particle 2 stops near $x = 0$, keeping $y \sim y_0$. The gravity of particle 1, however, attracts particle 2, and finally particle 2 is caught in a bound orbit. For this orbit, the criterion of the position deviation is satisfied when the particle reaches near $x = 0$.

The growth of velocity and position deviations is shown, respectively, in figures 2 and 3. Orbit 1 reaches both $\Delta v/v = 1$ and $2\Delta r/r = 1$ at a very early epoch; i.e., $t_{sc} \approx 10$ in our simulation units. At this moment the particle reaches $x \approx 0$ and is strongly scattered. On the other hand, orbit 2 never satisfies both criteria. The particle is scattered only in a small-angle, and then moves away in a straight line. Orbit 3 satisfies the condition $\Delta v/v = 1$ at $t \approx 160$, which is long before the time when the condition $2\Delta r/r = 1$ is satisfied at $t \approx 180$. The reason why $\Delta v/v = 1$ is realized much earlier than $2\Delta r/r = 1$ is because as long as the gravitational force of particle 1 is negligible, particle 2 moves in free-motion. During that time, the velocity suffers the effect of cosmic expansion. The velocity of the free-motion thus tends to zero, so that the velocity deviation increases rapidly, although, the actual orbit of particle 2 does not change much from that of free-motion.

Figure 4 shows the loci of constant scattering time on the initial separation versus the impact parameter plane when the relative velocity v_0 is fixed. From the top to the bottom panels $v_0 =$ 1, 4, and 8. They represent, in turn, typical cases of small, medium and high relative velocities. In the case $v_0 = 1$ [panel (a)], the initial relative velocity is small, and thus the test particles are 'scattered' through free-fall motion. Test particles which are initially located in negative separations need a longer time to be scattered than those located at positive separation. For example, a particle with the initial position $(-1.5, 0)$ needs the same 'scattered' time ($t_{sc} =$ 512) as that of a particle whose initial position is (5,0). This is because the test particle moves away from its nearest particle (particle 1) first, until it loses its relative velocity, and is then pulled back by the gravitational force of particle 1 through free-fall motion. From panel (b), we notice that the initial velocity $v_0 = 4.0$ is large enough for test particles with small initial separations and small impact parameters to be strongly scattered within a short time. We also notice that test particles which are initially located at negative separation do not interact with the nearest field particles because the initial velocity, $v_0 = 4.0$, is large enough to escape from the gravitational force of their nearest particles. Panel (c) shows that initial relative velocity v_0 $= 8.0$ is so large that particles are scattered before they suffer the effect of cosmic expansion.

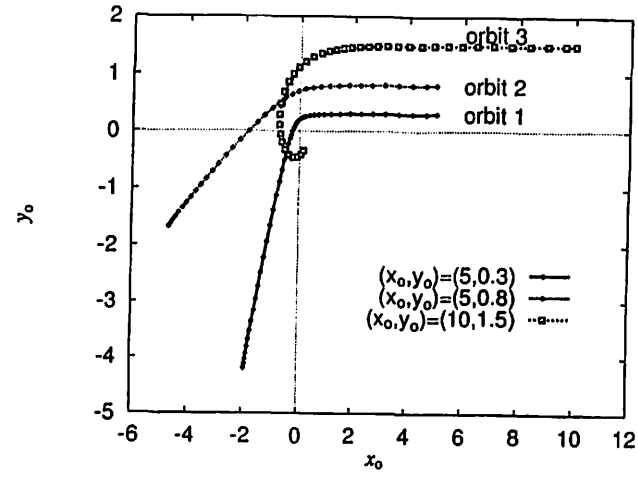


Fig1. Three orbits represent a *scattered* Rutherford orbit (orbit 1), *unscattered* Rutherford orbit (orbit 2) and a *scattered* bound orbit (orbit 3).

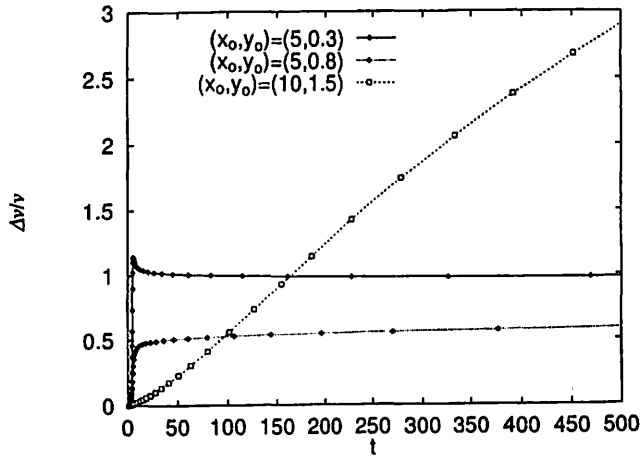


Fig 2. Growth of velocity deviation $\Delta v/v$ for the three kinds of orbits as a function of time.

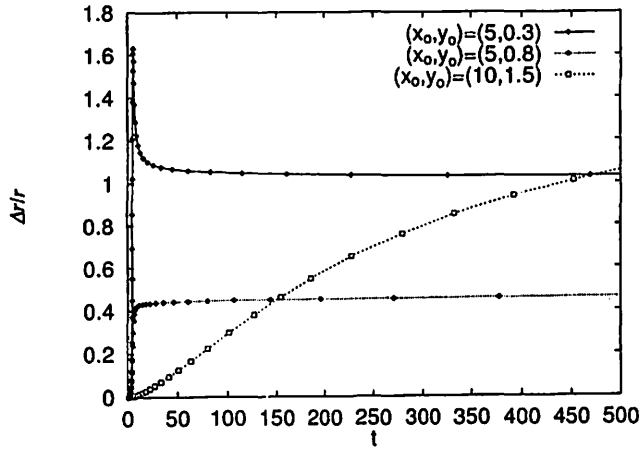


Fig 3. Growth of position deviation $\Delta r/r$ for the three kinds of orbits as a function of time.

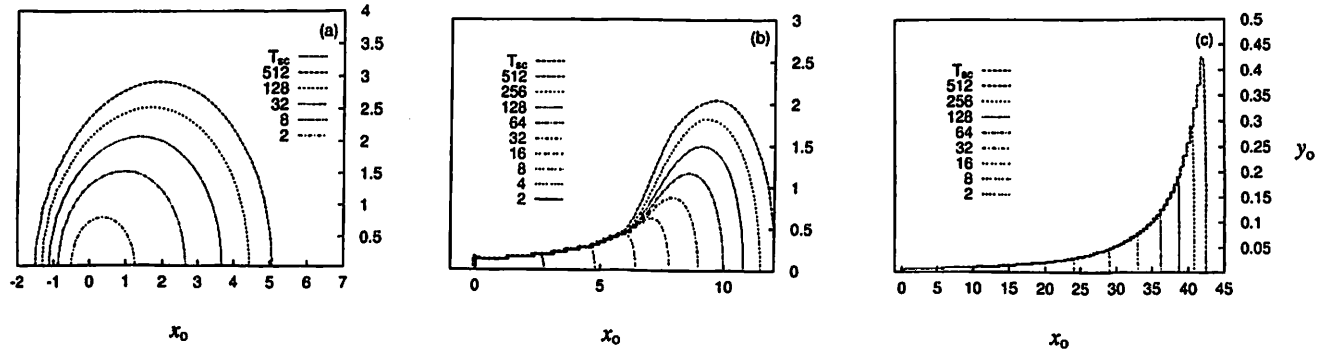


Fig 4. Loci of constant scattering time in initial separation-impact parameter (x_o-y_o) plane for (a) low (v_o =1.0) (b) intermediate (v_o=4.0) and (c) high (v_o=8.0) relative velocities.

The time dependence of the scattering volume $V_{\rm scat}$ is presented in figure 5 for some given initial velocities. The numerical results of simulations agree well with those derived analytically in for limiting cases: for an extremely high relative velocity, $V_{\rm scat} \propto t$, and $V_{\rm scat} \propto t^2$ for free-fall motion. When v_0 is small, $V_{\rm scat}$ grows as $V_{\rm scat} \propto t^2$. This is because for small t the particles contributing to $V_{\rm scat}$ are only those with small x_0 and they make free-falls. The scattering volume increases with time, but the rate of increase decreases for large t, which is especially prominent in cases of small v_0. This is an effect of expansion of the universe and can be understood as follows. As t increases, the contribution of particles with large separation on the scattering volume $V_{\rm scat}$ increases. Owing to the effects of expansion, however, the approaches of such test particles toward the nearest field particle are delayed. It takes a very long time before the test particle approaches the nearest field particle in the comoving coordinates. This leads to a decrease in the increasing rate of $V_{\rm scat}$. This trend begins to appear at an earlier epoch when v_0 is small. For a comparison, the $V_{\rm scat}$–t relations in non-expanding cases are shown by the parallel dotted lines.

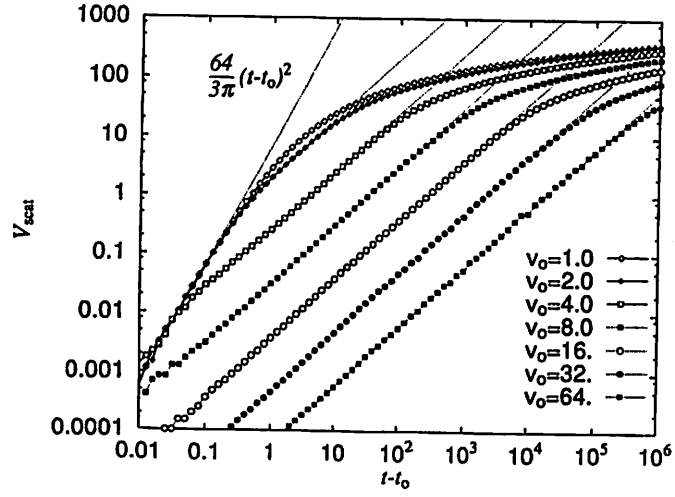


Fig 5. Scattered volume as a function of time for various relative velocities. The dotted line on the most upper-left side is the analytical results obtained under the approximation of *free-fall* motion. Other dotted lines show the scattered volume calculated in the non-expanding

systems.

By using the results shown in figure 5, we derive the relation between the relaxation time and the relative velocities for various number densities. The results are shown in figure 6. The lowest line is produced for the highest number density $n = 6.4$, followed upward by the number densities, decreasing by factor 0.25 in turn. For low relative velocities, $v_0 \leq 2$, the relaxation time is roughly equal to the free-fall time. The free-fall time, which is on the order of the dynamical time of the system, can be expressed as $t_{\rm ff} \approx 0.4n^{-1/2}$ when $n > 0.1$. As the velocity becomes higher, the relaxation time grows as $t_{\rm relax} \approx 0.064{v_0}^3 n^{-1}$. Here, the relaxation time behaves the same as in a non-expanding system.

We should pay attention to the fact that the increase in the relaxation time along with a decrease in the number density is especially prominent when the number density is low, say $n \leq 0.1$. This can be explained as follows. Following the equation of motion (2), the condition when the kinetic energy of the expansion is much greater than the gravitational energy between two particles is

$$x_0^3 > \frac{9Gmt_0^2}{a_0^3}. \tag{14}$$

This is the opposite condition of the free-fall approximation. The above mean separation gives the number density $n < 0.2$. Therefore, effect of expansion becomes dominant when $n < 0.2$ in our simulation unit or $n < 3 \times 10^3$ Mpc^{-3} $(10^{11}\ M_\odot/m)$ $[(1+z)/11]^3$. Here, m is the mass of a galaxy and z is its redshift.

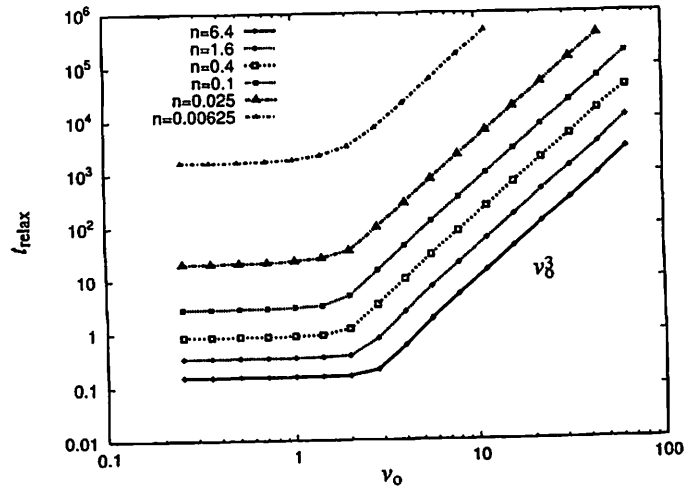


Fig 6. Relation between relaxation time and relative velocity for various number densities. From upper to lower curves, the number density in adjacent curves increases by factor 4.

5. Discussion and Conclusions

The effects of expansion are clearly shown especially when the relative velocities are small: the increase of $V_{\rm scat}$ with t is suppressed. As t increases, particles with a large separation contribute to $V_{\rm scat}$, but approaches of such particles to a field particle are decelerated by the cosmic expansion, leading to a slow increase in $V_{\rm scat}$.

We find that for low relative velocities $v_0 \leq 300$ km s$^{-1}(m/10^{11}M\odot)^{1/3}[11/(1+z)]^{1/2}$ and for a high number density, $n > 3 \times 10^3$ Mpc^{-3} $(10^{11}\ M_\odot/m)$ $[(1+z)/11]^3$, the relaxation time can be regarded as the free-fall time and is independent of v_0. The free-fall time, which is shorter than the age of the universe, can be expressed as

$$t_{\rm ff} = 0.11\ {\rm Gyr}\left(\frac{3\times 10^3 {\rm Mpc}^{-3}}{n}\right)^{1/2}\left(\frac{10^{11}M_\odot}{m}\right)^{1/2}\left(\frac{11}{1+z}\right)^{3/2}. \qquad (15)$$

As the relative velocity becomes higher, the relaxation time increases as $t_{\rm relax} \approx 0.064 {v_0}^3 n^{-1}$, which is the same as in the non-expanding case. We compare this, for example, with the relaxation time determined by Spitzer and Hart (1971). When $N = 100$, it leads to $t_{\rm relax} = 0.070 {v_0}^3 n^{-1}$.

In usual Rutherford scattering, the relaxation time is in proportion to n^{-1}, where n is the number density of the field particles. In this paper, however, we find that the relaxation time in the expanding universe increases along with a decrease in n with a higher power than n^{-1} when $n < 3 \times 10^3$ Mpc^{-3} $(10^{11}\ M_\odot/m)$ $[(1+z)/11]^3$. This means that the relaxation time becomes longer for a system of particles with a large mean separation.

Let us apply our results to proto-clusters of galaxies at a redshift of $z \approx 10$, as an example. A present typical rich cluster has a central number density of 10^3 Mpc^{-3}. Thus, the proto-cluster of galaxies may have a number density 8–times smaller at the begining and $n \approx 125$ Mpc^{-3}.

Assuming that the model of our universe is the Einstein–de Sitter one, and taking the age of the universe $t_{\rm age}$ to be 10^{10}yr (Peebles et.al. 1989), we have the unit of time $t_0 = 2.74 \times 10^8$ yr as the age of the universe at $z = 10$. Substituting the values of the gravitational constant and the mass of galaxies, we find that the units of length and velocity are 40 kpc and 147 km s^{-1}, respectively. The number density of the galaxies is about 0.008 in our simulation units.

Since it is difficult to predict the typical relative velocity between galaxies in a pre-collapsing cluster both observationally and theoretically, we employ, as a typical value of the large-scale velocity field, 600 km s^{-1} (see Bahcall et al. 1994), which is about 4.0 in our simulation units, as the highest extreme case. In this case, the effect of expansion is noticeable if the density of the sytems is much smaller than 0.2 (see figure 7), and for a number density of 0.008 the relaxation time in the expanding universe is one magnitude longer than that in the case of the expansion being neglected. Therefore, the cosmic expansion delays the relaxation very much.

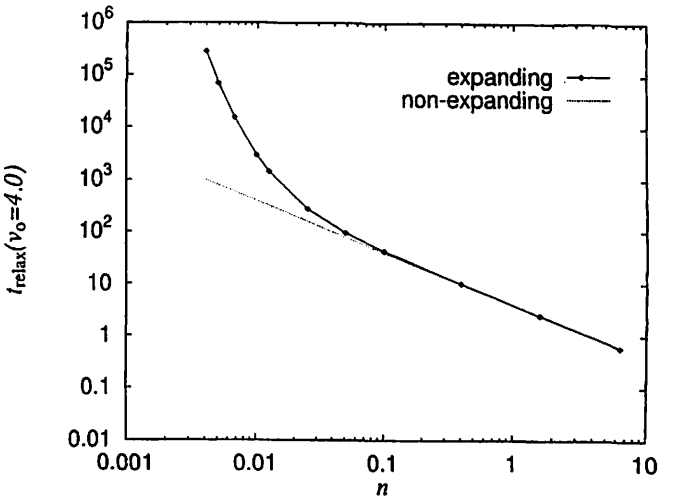


Fig 7. Relation between relaxation time and number density for an intermediate relative velocity $v_o = 4$. For comparison, the results in the non-expanding case is shown by dashed curve.

A realistic local velocity dispersion is probably smaller than the above velocity. If the velocity is less than 150 km s^{-1}, which means that it is less than 1.0 in the simulation units, our results in figure 6 show that the relaxation time for the system can be regarded as the free-fall time. The free-fall time in a non-expanding universe can be evaluated as 0.54 Gyr. Hence, the system could relax during the collapse of the cluster. On the other hand, the free-fall time in the expanding universe is found about 12 Gyr from figure 6. Therefore, the cosmic expansion is crucial for the relaxation of rich clusters in the early epoch. If we consider poorer clusters, the effect of cosmic expansion becomes bigger.

References

Ardi E., Inagaki S. 1997, PASJ 49, 187

Bahcall N.A., Gramann M., Cen R. 1994, ApJ 436, 23

Chandrasekhar S. 1942, Principles of Stellar Dynamics (The University of Chicago Press) pp 48–79

Peebles P.J.E. 1980, The Large Scale Structure of the Universe (Pricenton University Press, Pricenton) pp 41–44

Peebles P.J.E., Mellot A.L., Holmes M.R., Jiang L.R. 1989, ApJ 345, 108

Press W. H., Teukolsky S. A., Vetterling W. T., Flannery B. P. 1992, Numerical Recipes, 2nd ed. (Cambridge University Press) ch. 16

Spitzer L.Jr. 1962, Physics of Fully Ionized Gases (John Wiley, New York) pp 76–81

Spitzer L.Jr. 1969, ApJ, 158, L139

Spitzer L.Jr., Hart M.H. 1971, ApJ 164, 399

The Pseudoparticle Multipole Method

Junichiro Makino
Department of Astronomy, School of Science
The University of Tokyo
7-3-1 Hongo, Bunkyo-ku, Tokyo 113, Japan
and Atsushi Kawai
College of Arts and Science
The University of Tokyo
3-8-1 Komaba, Meguro-ku, Tokyo 153, Japan

要旨

Barnes-Hut (BH) treecode and Fast multipole method (FMM) are algorithms which can calculate the gravitational interaction between N particles in $O(N \log N)$ or $O(N)$ time. Both methods were proposed in 1980s and the treecode has become commonly used in cosomlogical simulations and simulations of galactic dynamics, where required accracy was low. For simulations which require high accuracy, however, the treecode has not been used because the implementation of high-order treecode is complex and calculation cost is high.

In this paper, we describe a new approach to implement high-order treecode and FMM, which we call Pseudoparticle Multipole Method (P^2M^2). In this method, we represent the multipole expansion back again by particles. As a result, the translation formulae for expansion coefficients are greatly simplified. Moreover, with this new method we can use GRAPE to accelerate the evaluation of the multipole expansion. Compated to the program on a workstation, the P^2M^2 treecode with GRAPE can achieve up to 100 times speedup.

1 はじめに

ツリー法 [1, 2] は N 個の質点間の重力相互作用を $O(N \log N)$ の計算コストで計算する方法であり、宇宙論的 N 体（あるいはSPH）計算や、銀河の形成、相互作用のシミュレーションでは広く使われている。計算法の原理については例えば [8] やそのなかであげた文献を参考にしていただくとして、基本的な発想は、遠くの粒子を適当にまとめて、それらからの重力を多重極展開におきかえて評価するというものである。遠くにいくほどまとめ方を大きくしていけるので、粒子一つへの重力を計算する計算量は $O(\log N)$ となり、全粒子への重力を計算するのに $N \log N$ のオーダーの計算量で済む。直接計算では計算量は N^2 に比例するので、粒子数が十分大きいところで ツリー法 はまともにやるのに比べてはほとんど N に比例して速いということになる。

もちろん、$N \log N$ の前につく係数は決して小さくはないが、それほど高い計算精度を要求しなれば、N が数百程度で直接計算よりも速くなる。[1]

さて、ツリー法は提案から 10 年以上過ぎ、銀河形成や銀河・銀河相互作用、あるいは宇宙論などの N 体計算や SPH を使った自己重力流体計算に広く使われるようになっている。これらの計算では、無衝突系を扱うという性格上、計算精度をあげるためには粒子数を増やして particle noise (ある

[1] 昨年までの「天体力学研究会」では、N というのは大きくても 10 程度であると想像され、このような N が大きいところの話というのはそぐわなかった感もあるが、「天体・N 体力学研究会」ではこのような話題も場違いではないものと信じる。

いは2体緩和) を押えることが、粒子間相互作用の精度を上げることよりも重要であり、そのためにツリー法のほうが直接計算よりもよい結果が得られるためである。

これに対して、散開星団、球状星団の熱的進化や、あるいは惑星形成のような、2体緩和によって系が進化する過程を扱う場合には、ツリー法はほとんど使われていうない。これには以下のような事情があった。まず、長時間シミュレーションが必要になるために、扱える粒子数が減る。また、計算精度も高くする必要がある。そうすると、ツリー法の計算量と直接計算の計算量の差が小さくなるか、場合によっては逆転してしまう。このために、また、GRAPE[11, 9] によって直接計算がかなり高速化されたということもあり、これらの系には基本的には直接計算が用いられてきた。

必ずしも一般的な事情ではないが、GRAPE を作ってきた我々の観点からはまた以下のような問題があった。GRAPE を使ってツリー法を加速することは可能ではある [7] が、GRAPE で計算できるのは粒子（質点）の重力場だけである。従って、遠くの粒子をまとめる時も、それらを重心で置き換えることしかできず、近似の次数でいえばダイポールよりも上げることは出来なかった。このために、計算精度をあげようとすると非常に急速に計算時間が増えてしまい、ツリー法を使うメリットがなくなってしまっていた。

我々は、高精度のツリー法や FMM （Fast Multipole Method, 高速多重極展開法）のような高速計算法で精度をあげるときに必要になる高次の多重極展開を、そのままの形ではなく粒子に戻って表現する方法（疑似粒子多重極法、Pseudoparticle Multipole Method）を開発した。これにより、多重極展開の評価も GRAPE を使って行なえるようになり、高精度の計算が飛躍的に加速された。さらに、実は粒子の形で表現することで、従来よりも実装が容易になり、任意次数のツリー法が簡単に実現出来るようになった。

以下、まずツリー法についてごく簡単に説明したあと、疑似粒子多重極法の基本的な思想について述べ、それから数学的な定式化を与える。次に、計算精度と計算量、計算時間についての数値実験の結果をまとめる。

2 ツリー法の原理

計算したいものは要するに以下のような常微分方程式系の右辺である。

$$\frac{d^2\boldsymbol{x}_i}{dt^2} = G\sum_{j=1,j\neq i}^{N} m_j \frac{\boldsymbol{x}_j - \boldsymbol{x}_i}{|\boldsymbol{x}_j - \boldsymbol{x}_i|^3}, \tag{1}$$

ここで N は粒子の数、$\boldsymbol{x}_i, m_i$ はそれぞれ粒子 i の位置、質量であり、G は重力定数である。

右辺を評価するもっとも単純な方法は、右辺をそのまま力任せに計算することである。粒子数が数百程度であれば、この方法（単に直接計算という）よりも速くするのは難しい。しかし、直接計算では $O(N^2)$ で計算量が増えていくので、粒子数がある程度大きくなると直接計算は現実的ではなくなる。計算量を減らすには、重力を及ぼす遠くの粒子を適当にまとめることが考えられる。まとめておいて多重極展開を作っておけば、精度は必要に応じてあげられるわけである。これがツリー法の基本的な考えである。

粒子がある空間を分割することで多重極展開や球面調和関数展開を使えるようにする。もっとも簡単に、空間を規則的な格子で分割した場合を考えてみると、粒子は、自分と、自分に隣接するセルに入っている粒子からの力は直接計算し、それ以外のセルからの力は多重極展開で計算することになる。粒子数を N、格子のセルの数を n とすれば、計算量は $N^2/n + Nn$ の程度となり、最適な n をとれば $O(N^{3/2})$ となる。

このやり方で無駄なところは、自分から見て非常に遠くにある粒子も近くにあるものも全部同じ大きさのセルにわけていることである。計算精度のことを考えるなら、遠くにいくほどセルが大きくなってよい。これを実現するために、正方形（3次元なら立方体）を再帰的に分割していくことでツ

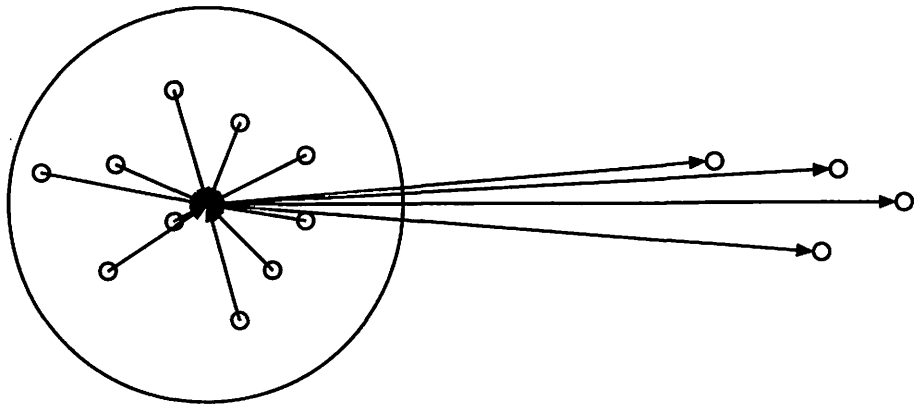


図 1: ツリー法の基本的アイディア。

リー構造を作る。ある粒子への力を計算するには、トップレベルからツリーを降りていって、大きさに比べて十分遠い、つまり誤差が小さくなったとみなせるところで止めればよい。十分遠いかどうかの判定には、通常は 見込み角と呼ばれる量

$$\theta = l/r \tag{2}$$

を使う。ここで r はセルの中心から力を評価する粒子までの距離、l はセルの一辺の長さである。

ここで、計算量と計算精度の間の関係が原理的にはどのようなものになるかを簡単に述べる。細かく考えるといろいろややこしいので、とりあえず一つのセルからの力の誤差について考えよう。展開次数を p とすれば、あるセルからの力の相対誤差は

$$e_{cell} \sim C_{p+1}(r/l)^{-(p+1)} = C_{p+1}\theta^{-(p+1)} \tag{3}$$

の程度になるはずである。ここで C_{p+1} は l^{p+1} で規格化した $p+1$ 次の多重極展開の係数の値から決まる定数である。規格化したので C_{p+1} には上限があり、上の式からツリー法の誤差限界を与えることができる。

さて、計算量の方はどうなるかといえば、極めて大雑把には

$$C \propto \theta^{-3}p^2 \tag{4}$$

で与えられるはずである。ここで、θ^{-3} に比例するのは、θ に比例してセルが小さくなるのと、3 次元なのでセルの大きさの 3 乗に反比例してセルの数が増えるからである。また、多重極展開の係数の数が p^2 なので、1 つのセルからの力の計算量は p^2 に比例する。

ここで C_p が p によらないとか p が大きいとか仮定すれば、計算精度を決めた時に計算量がもっとも小さくなる θ の値は、計算精度に無関係に $e^{-2/3} \sim 0.5$ となる。つまり、精度をあげるためには次数 p をあげていくのが有効であり、見込み角を変えることで精度を調整しようとするのは計算量から見ると損になるのである。

しかし、現在広く使われているツリー法の実装は、$p = 2$ の固定のものがほとんどである。また、GRAPE 上でのツリー法の実装は、GRAPE が計算するのが質点からの重力であるために原理的に $p = 1$ しか扱えない。このために、精度をあげるのは困難であったわけである。

3 P^2M^2 法

高次の項を扱うのに GRAPE を使う方法はないかと考えると、「多重極展開を粒子で表現する」ということができればよい。もともと多重極展開は実際の粒子分布の作る場を近似しているわけだ

から、適当に粒子（以下、こちらを本当の粒子と区別するために疑似粒子と呼ぶ）をおくことで多重極展開を表現しなおすということは原理的には可能なはずである。この時に、元の粒子よりも少ない数での疑似粒子で、ある次数までは元の粒子と同じ多重極展開を持たせることが出来れば、あとはそれらの疑似粒子からの重力を普通の粒子と同じように計算することでツリー法が実現できることになる。

重力の場合、いくつかの粒子の重心は $p=1$ まで展開が合うという意味でそのような粒子のおき方の例になっている。

次に4重極項を考えてみる。重心回りの4重極項モーメントテンソルはトレースレスなので、対角化すれば独立な要素は2個である。言い換えれば、疑似粒子は第一軸と第二軸の張る平面上におけばよく、3個の等質量の粒子を2等辺3角形におくことでで4重極項は表現出来る。

もっと高次の場合でも、このように粒子配置を構成するのも原理的には可能なはずだが、しかし、次数をあげていくと解くべき方程式がどんどんややこしくなるし、だんだんどうやって式を立てればいいかも良くわからなくなる（わかる人は是非やってみて欲しい）。従って、以下ではもうちょっと安直なアプローチを考える。

結局やりたいことはなんであったかというと、多重極展開の係数が実粒子の分布と等しくなるような疑似粒子の分布を見つけだすということである。ここで、疑似粒子をおく位置を球面上に限ってしまうと、疑似粒子の方の分布の多重極展開は、もちろん疑似粒子の質量分布の球面調和関数展開そのものであるから、別にこういう制限をつけても疑似粒子のおき方を決めることは依然として可能である。もちろん、位置の自由度のうちの1つを殺しているので、その分たくさん粒子がいることになるが、とりあえず大した問題ではないと信じよう。この時には、結局球面調和関数展開の逆変換で球面上の質量分布が求まり、それを適当に離散的な質点で近似することで疑似粒子を決めるということになる。

さて、依然として問題であるのは、どうやって球面上に粒子をおくかである。これは、上で求まった球面上の質量分布の、なんらかの意味での最良近似であるようなものを有限個の粒子で表現するという問題になる。ここで注意して欲しいのは、粒子一つの多重極展開は位置に対しては非線形な関数であるが、質量に対しては線形な関数であることである。従って、位置を固定して質量だけを変えるようにして、p^2 個の粒子を適当において質量だけを変えてやることで、必ず p^2 個の展開係数の値を合わせることが出来るはずである。

が、落ち着いて考えてみるとこれはまだいろいろ面倒そうである。計算の順番が、まず展開係数を求めてから、その展開係数を満たすような質量分布を決めるということになる。もちろん後半の計算では解くべき線形方程式の係数はいつも同じなので、LU分解しておくか逆行列を求めておいて、それから計算すれば一応はいいことになる。ここでは、点の数が増えることはあまり気にしないで定式化やプログラムが簡単になる方向で考える。

形式的には、するべきことは実粒子分布を多重極展開し、その係数を逆球面調和関数展開して仮想粒子の質量に戻すということである。仮想粒子の数が無限に多い連続極限では、形式的には逆変換の式は極めて簡単で

$$\rho = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \alpha_l^m Y^{*-m}_l \tag{5}$$

と書ける。ここで α_l^m は球面調和関数展開の係数であり、関数 $Y_l^m(\theta,\phi)$ は球面調和関数である。簡単になるのは球面上で球面調和関数が完全正規直交系を作っているからである。有限個の仮想粒子ですます時も、もしもその上で球面調和関数を有限次数で切ったものが直交系をなしていれば、上の式と本質的に同じ簡単な変換式がつかえることになる。

p 次までの球面調和関数が有限個の仮想粒子の質量が作る空間の上で直交系をなすためには、その空間のなかで内積演算が定義されてそれが直交性を満たしていればよい。そのためには、結局 $2p$ 次

の多項式の積分が厳密に求められればよい。そのように仮想粒子を置く方法はいろいろありえて、例えば θ 方向は p 次ルジャンドル多項式の零点に置き、 ϕ 方向は $2p+1$ 個を等間隔に置くというものでも実現できる。が、この場合、仮想粒子の重みづけが θ の値によって変わる。

もうすこしましな置き方はないかと思うわけだが、これは実は非常に昔から計算幾何学のほうで研究があり、 spherical designs という名前で知られている。これはまさに我々が必要としている、

- ある次数までの多項式の球面上の積分を厳密に与える。
- 数値積分の重みがすべての点で等しい

という性質を満たすものである。つまり、 t 次の spherical design というのは数学的には以下のように定義される。

単位球面 S 上の座標点 $r_i(1 \leq i \leq K)$ の集合で、 t 次以下の任意の多項式 $g(\boldsymbol{r})$ に対して

$$\int_S g(\boldsymbol{r})ds = \frac{4\pi}{K}\sum_{i=1}^{K} g(\boldsymbol{r}) \tag{6}$$

が成り立つようなもの。

これには古典的な研究 [10] と、最近になって詳細に評価しなおしたもの [6] があって、若干結果が違うが 20 次程度のものまで求められている。また、座標値そのものは WWW 上[2]に公開されているので、ファイルに落すのは非常に簡単である。次数 p が小さい時には、正多面体の頂点といったものになっているが、それ以上では基本的には切頭立方体の組合せといった方法で構成されている。点の数は結局ルジャンドル多項式を使う場合とあまり変わらないが、数値積分の重みがすべての点で等しいためにプログラムが簡単で数値的にも性質がよい。

さて、 spherical design を使うことに決めると、結局実粒子から仮想粒子への変換は以下のように計算できることになる。

まず、球面調和関数展開の係数 α_l^m は、

$$\alpha_l^m = \sum_{i=1}^{N} m_i r_i^l Y_l^{-m}(\theta_i, \phi_i) \tag{7}$$

で与えられる。ここで m_i は粒子 i の質量であり、 (r_i, θ_i, ϕ_i) はその位置の極座標表示である。関数 $Y_l^m(\theta, \phi)$ は球面調和関数であり、

$$Y_l^m = (-1)^m \sqrt{\frac{2l+1}{4\pi}\frac{(l-|m|)!}{(l+|m|)!}} P_l^m(\cos\theta) e^{im\phi} \tag{8}$$

で与えられる。ここで P_l^m はルジャンドル陪関数である。これと同じ展開係数をもつ半径 a の球面上での質量分布 $\rho(\theta, \phi)$ は

$$\alpha_l^m = \int_S \rho(\theta, \phi) Y_l^{-m}(\theta, \phi) dS, \tag{9}$$

を満たす必要がある。球面調和関数の直交性から、この質量分布 ρ は

$$\rho = \sum_{l=0}^{\infty}\sum_{m=-l}^{l} \alpha_l^m {Y^*}_l^{-m} \tag{10}$$

で与えられることになる。ここで積分を先ほどの K 個の仮想粒子上での数値積分で置き換えれば、仮想粒子 j の質量が

$$m_j = \frac{1}{4\pi K}\sum_{l=0}^{p}\sum_{m=-l}^{l} \alpha_l^m {Y^*}_l^{-m} \tag{11}$$

[2] http://www.research.att.com/~njas/sphdesigns/index.html

であたえられることになる。ここで、展開が p 次で打ち切られていることに注意してほしい。

さて、これで必要な変換は与えられたが、いちいち係数を求めてからそれをもう一回変換するのは美しくない。実粒子の質量から直接仮想粒子の質量を求めることを考えてみると、これは i, l, m の3重の総和になって一見計算が大変なような気がする。しかし、実は三角関数と同様に球面調和関数にも加法定理

$$P_l(\cos\gamma) = \frac{4\pi}{2l+1} \sum_{m=-l}^{l} Y_l^m(\theta,\phi) Y_l^{-m}(\theta',\phi') \tag{12}$$

というものがあり [ここで γ は方向 (θ,ϕ) と (θ',ϕ') の間の角度]、これを使えば m についての和は消えて、結局

$$m_j = \frac{2l+1}{K} \sum_{i=1}^{N} m_i \sum_{l=0}^{p} (r_i/r)^l P_l(\cos\gamma) \tag{13}$$

ということになる。

ここで注意してほしいことは、式(13)は実粒子から仮想粒子への変換だけでなく、ツリーの上のレベルに上がっていく仮想粒子から仮想粒子への変換にも同様に使えることである。つまり、式(13)で必要な変換はすべて表現されていることである。もうひとつ重要なことは、l の総和の式は結局 r_i/r と $\cos\gamma$ の l 次多項式なので、まともに計算すれば $O(l^2)$、漸化式を使えば $O(l)$, 区分近似多項式を準備しておけば $O(1)$ で計算できるということである。

なお、ツリー法の場合にはこれで話は終りであり、あとはセルから各粒子への力を計算するときに疑似粒子からの力の和として計算するというだけである。従って、従来のツリー法にわずかな手直しをするだけで任意次数のものが作れる。

FMM[4, 5]についても同様なことが考えられるが、ここでは省略する。その理由は、現在までに発表されたほとんどの論文で、計算精度が同程度ならばツリー法のほうが FMM よりも有意に速いという結果が示されているためである [3]。

ツリーは計算量が $O(N\log N)$ で FMM は $O(N)$ なのだから、FMM のほうが速そうなものであるが、実際にはそうはならない。ここではこの理由に深入りはしないが、簡単にいうとツリーでは多重極を評価するところだけで誤差が入るのに対し、FMM では多重極変換を局所変換に変換して、さらにそれ移動してというふうに段階が増え、それぞれの段階で誤差が入ってくるためである。

4 テスト計算

上の定式化に基づいた $\mathrm{P^2M^2}$ を FORTRAN で書かれたツリー法に組み込んで、計算時間、計算精度の評価を行なってみた。まず、汎用スカラー計算機での結果を示し、次に GRAPE-4 での結果を示す。

4.1 汎用計算機での結果

図2に見込み角 θ と計算誤差の関係を示す。ここでは計算誤差は加速度の相対誤差の2乗平均として定義される。これは時間積分の精度には直接つながらないという意味であまりよい量ではないが、他の多くの論文でこれがもちいられているのでここではそれに従う。粒子分布は球のなかの一様分布で、粒子数は 262,144 である。

図2からわかるように、次数が高いほど計算精度は上がる。計算精度の見込み角への依存性は大雑把にいって $e \sim \theta^{p+1.5}$ の程度になる。計算量と見込み角の関係を示したものが図3で、いうまでもなく θ が小さいほど、また次数が高いほど計算量は増える。なお、ここで計算量として書いているのは計算時間ではなく、ある粒子が力を受けている粒子（実粒子と疑似粒子をあわせて）の個数の平均値である。

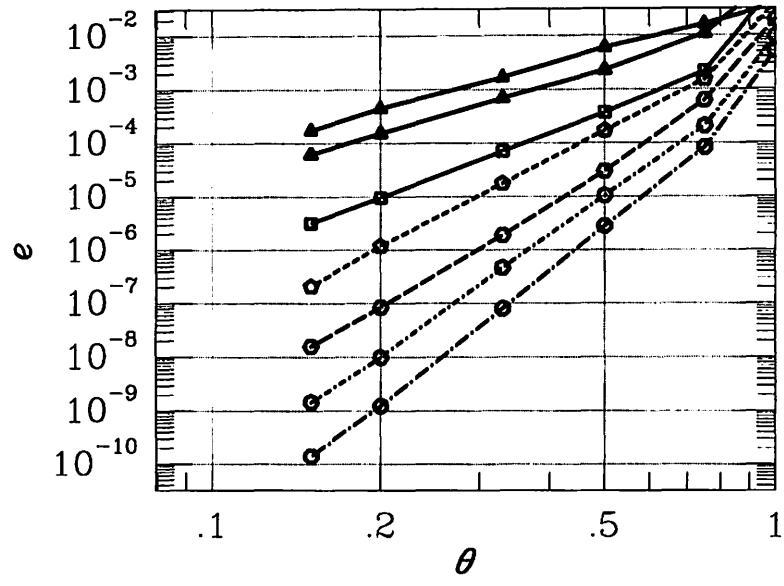


図 2: P^2M^2 法での見込み角と計算精度の関係。上から従来のツリー法、$p = 1, \cdots 6$。

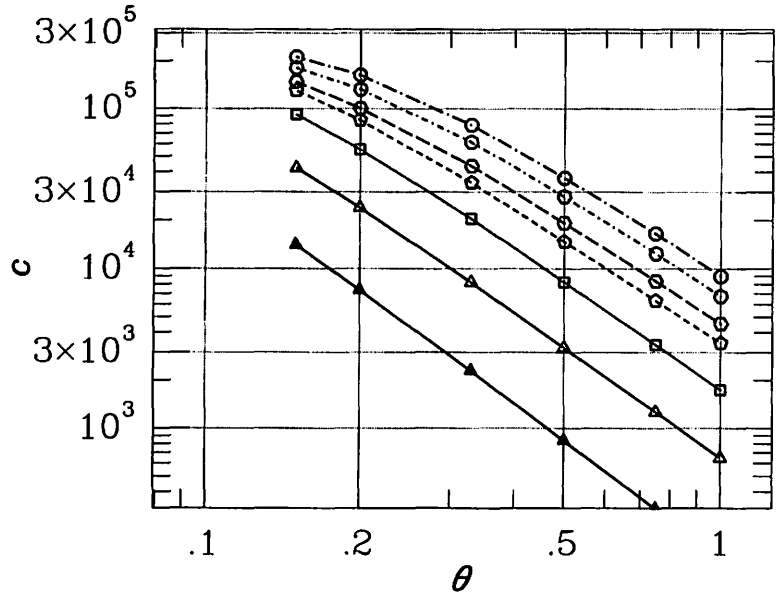


図 3: P^2M^2 法での見込み角と計算量の関係。線、シンボルの意味は図 2 と同じ。

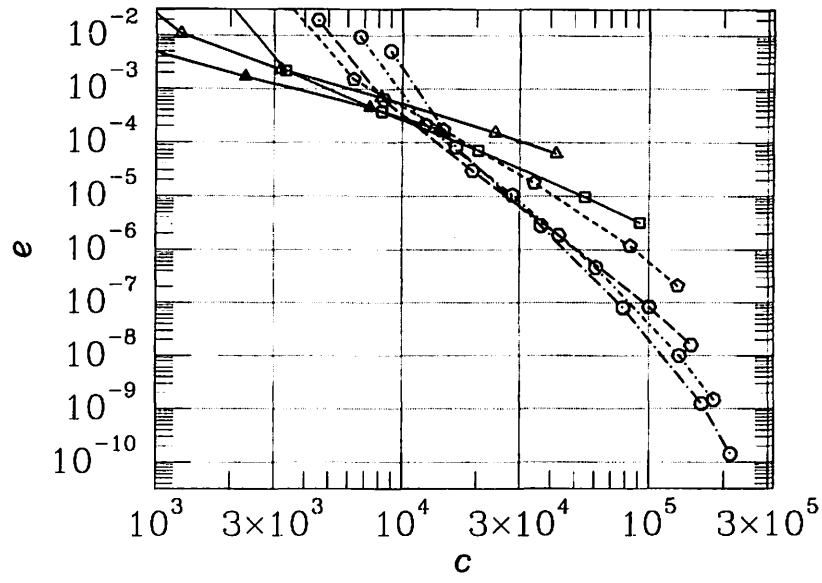


図 4: P^2M^2 法での計算量と計算精度の関係。線、シンボルの意味は図 2 と同じ。

さて、先に述べたように、必要な計算精度を決めれば最適な（計算量を最小にする）次数と見込み角の組合せが存在しているはずである。これは、計算量と計算精度の関係を見ればわかるので、それを図にしたものが図 4である。要求する計算精度を上げていくと、最適次数は上がっていくことがわかる。逆に、見込み角はほぼ一定のままであることに注意して欲しい。

なお、$p = 1$ の場合はダイポールまで表現されていることになり、これは従来のツリー法では重心だけを使っている場合に相当する。計算量は疑似粒子を 4 個必要とするために 4 倍となっていて、この場合や、その次の $p = 2$ の場合（疑似粒子 12 個が必要）については、汎用計算機で計算するかぎり従来のツリー法に比べていいことはない。

4.2 GRAPE-4 での実装

GRAPE を使う場合の計算コードも、従来のものを少し手直しするだけで構成できる。GRAPE を使う場合の基本的な考えは、ホストでツリーをたどっていくのをあるグループに対して行ない、それで出来た相互作用リストとグループの中の粒子のリストを GRAPE に送って力を計算してもらおうというものであった [7]。このために、一粒子あたりの計算量はグループの中の粒子の数の数倍程度増えている。GRAPE が相対的に高速なので、この部分はあまり問題にならないわけである。疑似粒子を使った場合でも、この GRAPE を使うために計算量が増えている分は変わらないので、結局次数を上げてもそれほど計算量が増えないということが期待できる。

計算精度、計算時間等の測定結果を図 5に示す。粒子分布は前と同じであり、汎用計算機は Alpha 21164 300 MHz (DEC Alphaserver 8400), GRAPE は GRAPE-4 ボード 1 枚（32 chip, 10 Gflops）である。上で期待したとおり、次数を上げても、また見込み角を小さくしても汎用計算機の場合のようには計算量が増えない。結果として、精度を上げるほど GRAPE を使うメリットは大きくなり、10^{-6} 程度の精度を要求するなら計算速度の差は 100 倍近くになる。GRAPE を使う場合には、もともと計算量が増えているということもあって疑似粒子法の計算量が多いという欠点が目立たなくなっていることがわかる。

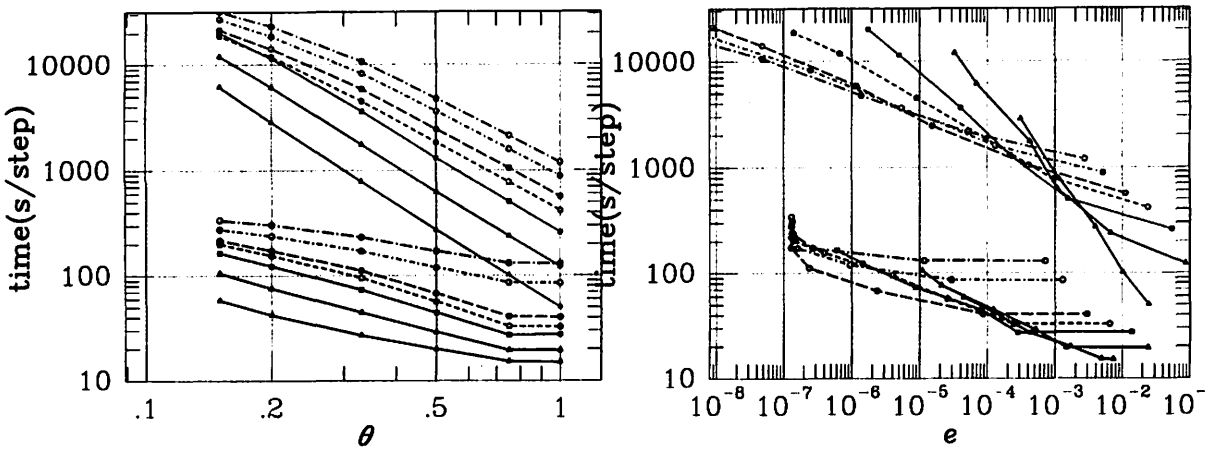


図 5: 汎用計算機と GRAPE-4 での（左）見込み角と計算時間、（右）計算時間と計算精度の関係。上は汎用計算機、下は GRAPE-4. シンボルの意味は図 2 と同じ。

5 まとめと議論

ここでは、高精度のツリー法を実現するために必要なポテンシャルの多重極展開を、疑似粒子をつかって表現する疑似粒子多重極法 (P^2M^2 法) について、その原理、定式化、数値実験の結果を概説した。この方法は、普通の計算機ではプログラムが簡単になるという利点はあるが計算量が増えるためかならずしも実用的ではないかもしれない。しかし、粒子間の重力を高速に計算する専用計算機 GRAPE を使って多重極展開が計算できるため、劇的な高速化が可能であるというのが大きな特長である。

参考文献

[1] A. W. Appel: "An efficient program for many-body simulation," *SIAM Journal on Scientific and Statistical Computing* 6, 85–103 (1985).

[2] J. Barnes and P. Hut: "A hiearchical $O(N \log N)$ force calculation algorithm," *Nature* 324, 446–449 (1986).

[3] D. Blackston and T. Suel: " Highly portable and efficient implementations of parallel adaptive n-body methods," *Proceedings of SC97.* ACM 1997, (CD–ROM).

[4] L. Greengard and V. A. Rokhlin: "A fast algorithm for particle simulations," *Journal of Computational Physics* 73, 325–348 (1987).

[5] L. Greengard and V. A. Rokhlin: "Rapid evaluation of potential fields in three dimensions," *Vortex Methods* (edited by Anderson C. and Greengard, C.), number 1360 in Lecture Notes in Mathematics, Springer-Verlag, Berlin, 121–141 (1988).

[6] R. H. Hardin and N. J. Sloane: "McLaren's improved snub cube and other new spherical designs in three dimensions," *Discrete and Computational Geometry,* 15, 429 (1996).

[7] J. Makino: "Treecode with a special-purpose processor," *Publ. Astron. Soc. J.* 43, 621–638 (1991).

[8] 牧野淳一郎: "高速多重極展開法とツリー法 — 多体シミュレーションのための高速算法," 応用数理 8, 277-286 (1998).

[9] J. Makino and M. Taiji, *Special Purpose Computers for Scientific Simulations - The GRAPE systems*, John Wiley and Sons, Chichester (1998).

[10] A. D. McLaren: "Optimal numerical integration on a sphere," *Math. Comput.*, 17, 361–383 (1963).

[11] D. Sugimoto, Y. Chikada, J. Makino, T. Ito, T. Ebisuzaki, and M. Umemura: "A special-purpose computer for gravitational many-body problems," *Nature* 345, 33-45 (1990).

非線形 Schrödinger 方程式に対する symplectic 数値解法
Symplectic Integrators for Nonlinear Schrödinger Equation

吉田 春夫 (国立天文台), 佐々 成正 (日本原子力研究所)
H. Yoshida (NAOJ) and N. Sasa (JAERI CCSE)

1 はじめに

非線形 Schrödinger 方程式 (NLS) とは 2 変数の関数 $u = u(x,t)$ に対する

$$iu_t + u_{xx} + 2|u|^2 u = 0 \qquad (1)$$

の形の偏微分方程式である．この非線形 Schrödinger 方程式は無限次元の Hamilton 系と解釈され，さらに可積分系であることが知られている．また多くの物理現象を記述することも知られている．例えば [1] を参照せよ．ここで可積分とは

- 初期値問題が逆散乱法によって解ける
- N-soliton 解が存在する
- 無限個の保存量（積分）が存在する

ことを意味している．可積分性の厳密な定義は無限次元の Hamilton 系においてはまだ無いようである．ここで Hamilton 系としての定式化はハミルトニアン

$$\begin{aligned} H &= -i\int(|u_x|^2 - |u|^4)dx \\ &= -i\int(u_x u_x^* - (uu^*)^2)dx \end{aligned}$$

から Hamilton 方程式

$$u_t = \frac{\delta H}{\delta u^*}, \quad u_t^* = -\frac{\delta H}{\delta u} \qquad (2)$$

として (1) が導出されることを意味する．ここで記号 $\frac{\delta}{\delta u}$ は汎関数微分で，一般の汎関数 $I = \int f(u, u_x, u_{xx}, \ldots)dx$ に対して

$$\frac{\delta}{\delta u} I = \frac{\partial f}{\partial u} - \frac{d}{dx}\left(\frac{\partial f}{\partial u_x}\right) + \frac{d^2}{dx^2}\left(\frac{\partial f}{\partial u_{xx}}\right) \cdots$$

で定義される．非線形 Schrödinger 方程式 (1) は可積分系であり，$t = 0$ での初期値 $u(x,0)$ が与えられた時，任意時刻での解 $u = u(x,t)$ が「原理的」には解析的に求まる．しかしソリトン解のような特殊解以外ではあまり実用にはならないという現実がある．一方で種々の広い応用があることから，非線形 Schrödinger 方程式に対して多くの数値解法が開発され適用されてきた．例えば [6] を参照せよ．

本稿では非線形 Schrödinger 方程式 (1) に対する高速な数値解法として以前から知られてきた split-step Fourier 法が Hamilton 系に対する symplectic（シンプレクティック）解法に他ならないことを示し，さらに symplectic 解法についての既存の理論からより高次（高精度）の split-step Fourier 法が簡単に構成できることを示す．これらはいずれも数値解法としては陽的解法である．そして最後に実際の数値例を 1 ソリトン，2 ソリトン解の場合に示す．

2 Split-step Fourier 法

非線形 Schrödinger 方程式 (1) を導くハミルトニアンを

$$H = H_1 + H_2,$$
$$H_1 = -i\int |u_x|^2 dx, \quad H_2 = i\int |u|^4 dx$$

と分解する．各々のハミルトニアンから導かれる Hamilton 方程式 (2) はそれぞれ

$$\text{(i)} \quad iu_t + u_{xx} = 0 \qquad (3)$$
$$\text{(ii)} \quad iu_t + 2|u|^2 u = 0 \qquad (4)$$

である．第1の方程式は線形の偏微分方程式であり Fourier 変換を施すことによって簡単に解くことができる．第2の方程式は常微分方程式でこれも簡単に解ける．(3) の解は連続変数の場合，$u(x,t)$ の Fourier 変換$\hat{u}(k,t)$ を

$$u(x,t) = \int \hat{u}(k,t)\exp(ikx)dk$$

で定義すれば$\hat{u}(k,t)$ の満たす常微分方程式

$$\hat{u}_t = -ik^2\hat{u}$$

から

$$\hat{u}(k,t) = \exp(-ik^2t)\hat{u}(k,0)$$

が得られ，これから

$$u(x,t) = \int \exp(-ik^2t + ikx)\hat{u}(k,0)dk$$

がその解となる．また常微分方程式 (4) の解が

$$u(x,t) = u(x,0)\exp(2i|u(x,0)|^2t)$$

となることをチェックすることは易しい．

いま物理空間 x および運動量空間 kの離散化を行い，その上での従属変数を

$$u(x,t) = (u_1(t), u_2(t), \ldots, u_N(t)) \qquad (5)$$
$$\hat{u}(k,t) = (\hat{u}_1(t), \hat{u}_2(t), \ldots, \hat{u}_N(t))$$

とする．そして離散 Fourier 変換 Fを

$$u_j = F(\hat{u}) = \frac{1}{\sqrt{N}}\sum_m \hat{u}_m \exp\left(\frac{2ijm}{N}\right)$$
$$\hat{u}_m = F^{-1}(u) = \frac{1}{\sqrt{N}}\sum_j u_j \exp\left(-\frac{2ijm}{N}\right)$$

で定義する．特に離散点の数を $N = 2^n$とすれば高速 Fourier 変換 (FFT) のアルゴリズムが活用でき，速度が向上する．これら離散化された変数での時間発展は (3) に対しては

$$u_j(t+\Delta t) = F[\exp(-ik_n^2\Delta t)F^{-1}(u(t))]$$

および (4) に対しては

$$u_j(t+\Delta t) = u_j(t)\exp(2i|u_j(t)|^2\Delta t)$$

である．これら2つの写像の単なる合成として split-step Fourier 法は定義される．Taha and Ablowitz[6] には split-step Fourier 法を他の代表的な数値解法と比較したときの同一要求精度に対する計算時間の比較がなされている．それによれば 1984 年の時点で split-step Fourier 法は最高速の解法であった．実はこの高速性は split-step Fourier 法が Hamilton 系に対する専用数値解法である symplectic 数値解法であることがその基礎にある．実際 split-step Fourier 法の symplectic 数値解法としての解釈は Herbst, Varadi and Ablowitz[3] によってもなされている．

3 Symplectic 数値解法

Hamilton 力学系

$$\frac{dq}{dt} = \frac{\partial H}{\partial p},\quad \frac{dp}{dt} = -\frac{\partial H}{\partial q}$$

の任意の時刻間の解の時間発展 $(q,p) \to (q',p')$ はエネルギー（ハミルトニアン）を保存する，つまり

$$H(q,p) = H(q',p')$$

であるとともに，symplectic 2次形式 $dp\wedge dq$を保存すること（symplectic 性），つまり

$$dp \wedge dq = dp' \wedge dq'$$

となることが知られている．この symplectic 2次形式は相空間の要素を各 (p_i,q_i) 平面に射影して得られる向きづけられた面積要素の和を表す．写像の symplectic 性は自由度1の系（2次元写像）においては面積保存と等価である．一般のハミルトン系に対して $H(p,q) = \text{const.}$と $dp\wedge dq = \text{const.}$を同時に満足する数値解法は存在しえない．symplectic 数値解法とは各ステップ $(q_m,p_m) \to (q_{m+1},p_{m+1})$ が厳密に symplectic 写像となるものとして定義される．そしてエネルギーの保存に関しても優れた特性を有する．symplectic 数値解法についての一般的な参考文献としては [2], [5], [9], [10] を参照せよ．

一般の Hamilton 系に対して，常に陰的な symplectic 解法が構成できる．例えば1次の陰的解法はΔt をτと書いて

$$q' = q + \tau\left(\frac{\partial H}{\partial p}\right)_{q=q'}, p' = p - \tau\left(\frac{\partial H}{\partial q}\right)_{q=q'}$$

で実現される．高次の陰的解法については例えば[9]を参照せよ．一方，特殊なクラスのHamilton系に対しては以上の陰的解法に加えて，陽的なsymplectic解法が可能である．ここで「特殊」とはハミルトニアンが

$$H = H_1 + H_2 \tag{6}$$

の和に書けH_1, H_2が共に単独では可積分（解が簡単に書ける）となるものである．この形のハミルトニアンの代表的な例は

$$H = T(p) + V(q) \tag{7}$$

であり，$H = T(p)$および$H = V(q)$をハミルトニアンとする系の解は明らかに相空間の直線運動を表す．そして系(7)に対する1次の陽的symplectic解法は，2つの直線運動の単なる合成として

$$q' = q + \tau \left(\frac{\partial T}{\partial p}\right)_{p=p}, p' = p - \tau \left(\frac{\partial V}{\partial q}\right)_{q=q'}$$

あるいは

$$p' = p - \tau \left(\frac{\partial V}{\partial q}\right)_{q=q}, q' = q + \tau \left(\frac{\partial T}{\partial p}\right)_{p=p'}$$

で実現される．

(7)以外の形の(6)の形のHamilton系においては，H_1およびH_2は単独では可積分であるが，各々の時間発展を記述する作用・角変数は異なることが一般的である．例えば摂動を受けたケプラー問題（2体問題）

$$H_1 = \frac{1}{2}\boldsymbol{p}^2 - \frac{1}{r}, \quad H_2 = R(x, y, z)$$

に於いては，H_1系に対する作用・角変数はDelauney変数と呼ばれるもので，H_2系に対するものは通常の直交座標である．そのためこのような系に対するsymplectic解法はMixed Variable Symplectic(MVS)解法の名で呼ばれることが多い[4]．

以上は数値解法として1次の解法であるが，2次以上のSymplectic解法も容易に作ることができる．(6)に対する1次の陽的解法はまずH_1をハミルトニアンとしてステップτだけ時間発展させ，次にH_2をハミルトニアンとしてステップτだけ時間発展させたものの合成（あるいはその逆）として1ステップが定義されている．これは互いに非可換な作用素A, Bの指数関数に関する関係式

$$\exp[\tau(A+B)] = \exp[\tau A]\exp[\tau B] + O(\tau^2) \tag{8}$$

がその基礎になっている．ここでA, Bは特にH_1, H_2とのPoisson括弧をつくる微分作用素

$$A := \{\cdot, H_1\}, \quad B := \{\cdot, H_2\}$$

を表す．(8)の左辺の$\exp[\tau(A+B)]$は真の解の時間発展，右辺の$\exp[\tau A]\exp[\tau B]$はH_1とH_2によるフローの合成を表し，その差が$O(\tau^2)$であることから，

$$S_1(\tau) := \exp[\tau A]\exp[\tau B]$$

が1次の精度（解法）であることが保証される．またこの定義自身，任意のハミルトニアンの解，およびその合成はsymplecticであることから，上に述べた1次の解法がSymplectic解法となることの根拠も与えている．

この指数関数表記を使えば2次の陽解法は

$$\begin{aligned}&\exp[\tau(A+B)]\\ &= \exp[\frac{\tau}{2}A]\exp[\tau B]\exp[\frac{\tau}{2}A] + O(\tau^3)\\ &= S_2(\tau) + O(\tau^3)\end{aligned}$$

でできる．この意味はH_1で$(\tau/2)$，H_2で(τ)，そしてH_1で$(\tau/2)$時間発展させて1ステップを構成するという意味である．この2次の陽解法はSymplectic解法となるのみならず，時間反転対称性を持つ解法

$$S_2(\tau)S_2(-\tau) = S_2(-\tau)S_2(\tau) = identity$$

となることも注意しておく．

4次以上の時間反転対称なSymplectic解法は上の2次の解法の対称的な合成で得られることが知られている[8]．例えば4次の解法は3つの2次の解法の合成

$$S_4(\tau) := S_2(x\tau)S_2((1-2x)\tau)S_2(x\tau)$$

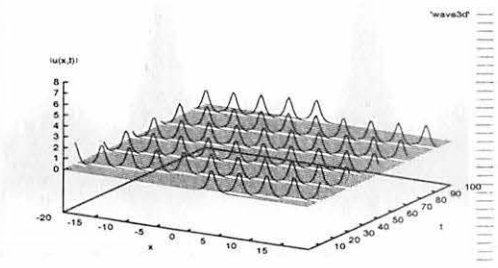


図 1: 1 ソリトン解

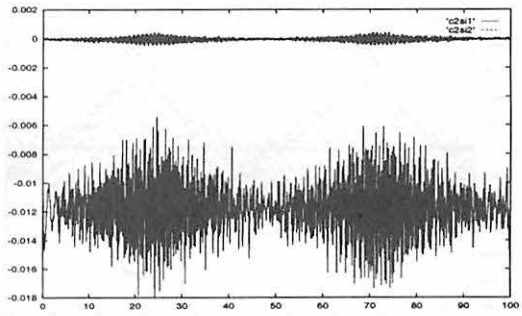


図 2: SI1 と SI2 による I_2の誤差

で得られる．ここで $x = 1/(2 - 2^{1/3})$ である．同様に 6 次の解法は

$$\begin{aligned} S_6(\tau) &:= S_2(y_3\tau)S_2(y_2\tau)S_2(y_1\tau)S_2(y_0\tau) \\ &\quad \times S_2(y_1\tau)S_2(y_2\tau)S_2(y_3\tau) \end{aligned}$$

で良い．ここで数値定数 y_0, y_1, y_2, y_3は

$$\begin{aligned} y_1 &= -1.1776\ 79984\ 17887 \\ y_2 &= 0.23557\ 32133\ 59357 \\ y_3 &= 0.78451\ 36104\ 77560 \\ y_0 &= 1 - 2(y_1 + y_2 + y_3) \end{aligned}$$

で与えられるもので，ある連立非線形代数方程式の数値解として求められている．8 次の解法は [8] で与えられているものは誤差定数が大きくあまり勧められない．

4 Split-step Fourier 法の symplectic 解法としての解釈と高次解法

空間方向に離散化した変数 (5) に対して正準共役な変数の組 (u_j, v_j) を

$$(u_j, v_j) := (u_j, u_j^*)$$

で定義する．さらに離散フーリエ変換

$$u_j = \frac{1}{\sqrt{N}} \sum_n \hat{u}_n \exp(ik_n x_j)$$

$$k_n = \frac{2\pi}{L} n, \quad x_j = \frac{L}{N} j$$

は Poisson 括弧が

$$\begin{aligned} \{\hat{u}_n, \hat{u}_l^*\} &:= \sum_j \left(\frac{\partial \hat{u}_n}{\partial u_j} \frac{\partial \hat{u}_l^*}{\partial u_j^*} - \frac{\partial \hat{u}_n}{\partial u_j^*} \frac{\partial \hat{u}_l^*}{\partial u_j} \right) \\ &= \frac{1}{N} \sum_j \exp(-ik_n x_j) \exp(ik_l x_j) \\ &= \delta_{n,l} \end{aligned}$$

となることから正準変換であることが示される．つまり $(\hat{u}_n, \hat{u}_n^*)$ も正準共役な変数の組である．よって離散化された変数によってハミルトニアンは

$$\begin{aligned} H &= H_1 + H_2 \\ H_1 &= -i \sum_n k_n^2 |\hat{u}_n|^2 \Leftarrow -i \int |u_x|^2 dx \\ H_2 &= i \sum_j |u_j|^4 \Leftarrow i \int |u|^4 dx \end{aligned}$$

と書ける．H_1系のハミルトニアン

$$H_1 = -i \sum_n k_n^2 |\hat{u}_n|^2$$

から導かれる Hamilton 方程式は

$$\frac{d}{dt}\hat{u}_n = -ik_n^2 \hat{u}_n, \quad \frac{d}{dt}\hat{u}_n^* = ik_n^2 \hat{u}_n^*$$

であり，その解は直ちに

$$\hat{u}_n(t + \Delta t) = \exp(-ik_n^2 \Delta t)\hat{u}_n(t)$$

となる．一方 H_2系のハミルトニアン

$$H_2 = i \sum_j |u_j|^4$$

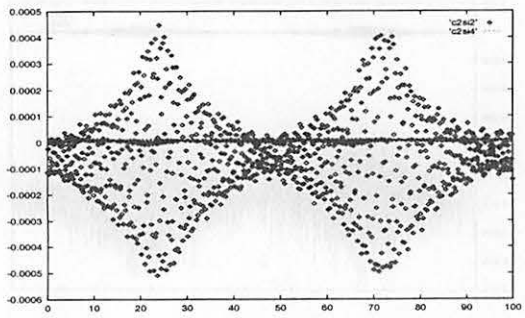


図 3: SI2 と SI4 による I_2の誤差

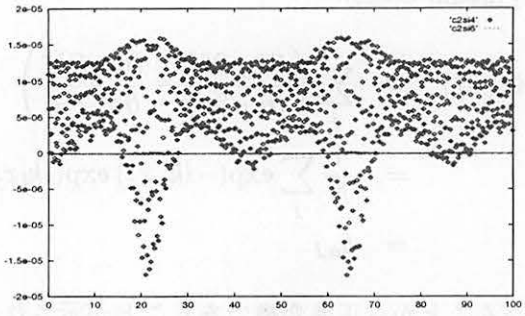


図 4: SI4 と SI6 による I_2の誤差

から導かれる Hamilton 方程式は

$$\frac{d}{dt}u_j = 2i|u_j|^2 u_j, \quad \frac{d}{dt}u_j^* = -2i|u_j|^2 u_j^*$$

であり，その解も容易に

$$u_j(t+\Delta t) = u_j(t)\exp(2i|u_j(t)|^2\Delta t)$$

と求められる．前節の議論から，これらの単なる合成によって 1 次，2 次，4 次，6 次，... の symplectic 解法が構成できるわけである．ただし H_1と H_2の時間発展を記述する変数は異なるから Mixed Variable Symplectic 解法である．特に 1 次のものは既知の split-step Fourier 法である．

5 数値例

得られた symplectic 解法を以下の条件の下で実際に非線形 Schrödinger 方程式の数値解法として使った結果を以下に示す．

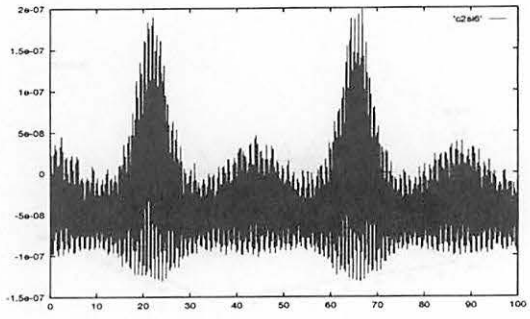


図 5: SI6 による I_2の誤差

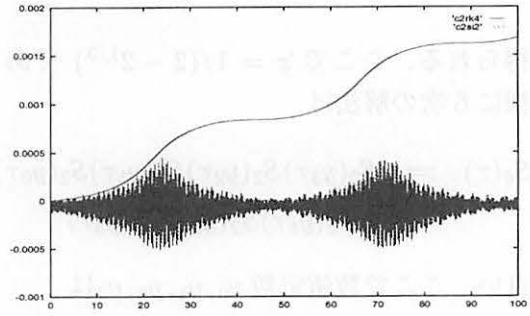


図 6: SI2 と RK4 による I_2の誤差

- 空間：$-20 \le x \le 20, (L = 40)$ を $N = 128$ 分割（周期境界条件）
- 時間：$0 \le t \le 100$ を時間刻み $\Delta t = \tau = 0.01$ で 10000 ステップ
- 初期条件：[6] に与えられた 1 ソリトンおよび 2 ソリトン
- 数値解法：
  - SIn = n 次の symplectic $(n = 1, 2, 4, 6)$
  - RK4 = 4 次の Runge-Kutta
- 精度を比較するための保存量（エネルギー）:

$$I_2 = \int (|u_x|^2 - |u|^4)dx$$
$$\Rightarrow \sum_n k_n^2 |\hat{u}_n|^2 - \sum_j |u_j|^4$$

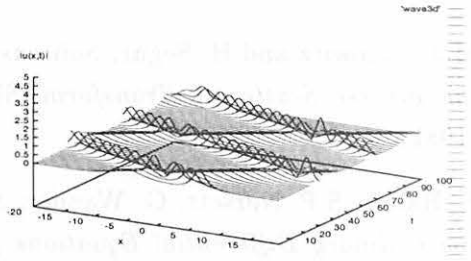


図 7: 2 ソリトン解

なお非線形 Schrödinger 方程式が持つ別の保存量（確率）

$$I_1 = \int |u|^2 dx \Rightarrow \sum_j |u_j|^2$$

は split-step 法によって厳密に保存されることが知られており [7]，同じ理由でここに述べた symplectic 解法によっても厳密に保存される．よって数値的に得られる I_1 の誤差はすべて丸め誤差に起因する．

5.1 1 ソリトン解

図 1 は与えた初期条件から $0 \leq t \leq 100$ の範囲で数値的に得られた 1 ソリトン解である．この範囲では数値解法の違いに依らず同じ図が得られる．ソリトンが円周上を約 5 周する様子がわかる．図 2 は SI1 と SI2 によるエネルギー I_2 の相対誤差

$$\Delta I_2(t) = (I_2(t) - I_2(0))/I_2(0)$$

を時間 t の関数としてプロットしたものである．以下，順に SI2 と SI4，SI4 と SI6，そして SI6 による誤差を図 3 から図 5 に示した．また図 6 は SI2 と RK4 による誤差を比較したものである．symplectic 解法でない RK4 によるものでは誤差は一方的に増大するが，symplectic 解法によるものでは誤差は一定の範囲に留まる．そして解法の次数を上げることによってその範囲がどんどん小さくなっていく様子が見て取れる．

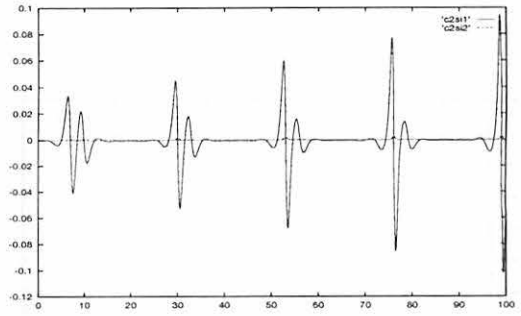


図 8: SI1 と SI2 による I_2 の誤差

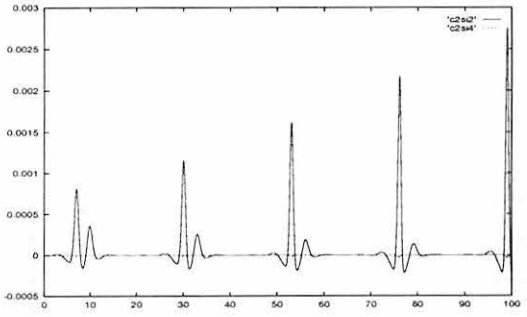


図 9: SI2 と SI4 による I_2 の誤差

5.2 2 ソリトン解

図 7 は与えた初期条件から $0 \leq t \leq 100$ の範囲で数値的に得られた 2 ソリトン解である．この範囲では数値解法の違いに依らず同じ図が得られる．ソリトンが円周上で 5 回ほど衝突している様子がわかる．この 2 ソリトン解に対して図 2 から図 5 に対応するものを図 8 から図 11 に示した．ソリトンの衝突というイベントに対応して誤差は一時的には大きくなるが，すぐに元のレベルに復帰することがわかる．これも symplectic 解法による特徴的な現象である．

図 8 から図 11 では誤差の最大幅が増大しているように見えるので，念のため図 12 では 10 倍の時間 $0 \leq t \leq 1000$ における SI2 による I_2 の誤差を示した．図 12 が示すようにやはり symplectic 解法によってエネルギーの誤差は単調に増大しないことがわかる．

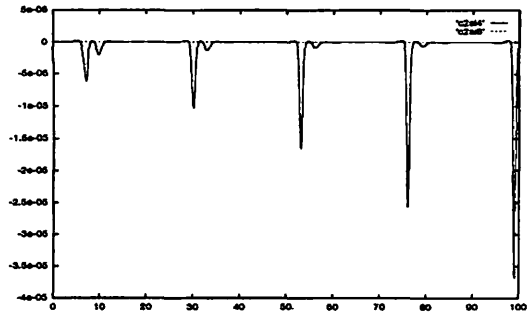


図 10: SI4 と SI6 による I_2の誤差

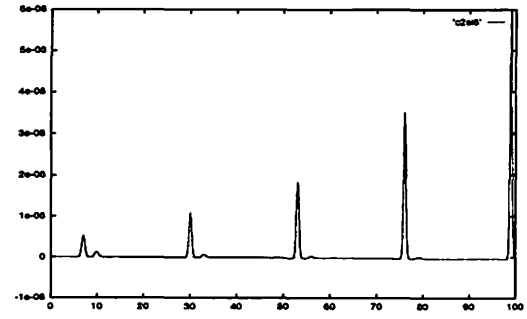


図 11: SI6 による I_2の誤差

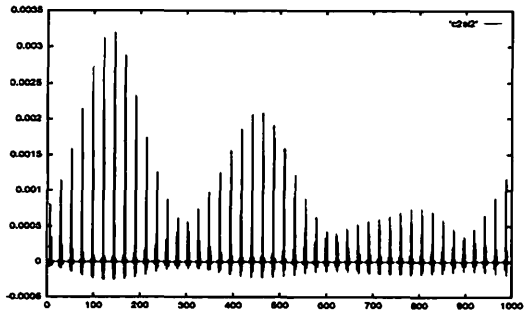


図 12: SI2 による I_2の誤差，$t = 1000$ まで

参考文献

[1] M.J. Ablowitz and H. Segur, *Solitons and the Inverse Scattering Transform*, SIAM (1981).

[2] E. Hairer, S.P. Nørsett, G. Wanner, *Solving Ordinary Differential Equations (2nd revised edition)*, Springer (1993).

[3] B.M. Herbst, F. Varadi, M.J. Ablowitz, Symplectic methods for the nonlinear Schrödinger equation, *Math. Comput. Simulation* **37** (1994) 353–369.

[4] P. Saha and S. Tremaine, Symplectic integrators for solar system dynamics, *Astron. Journal* **104** (1992) 1633–1640.

[5] J.M. Sanz-Serna, and M.P. Calvo, *Numerical Hamiltonian Problems*, Chapman and Hall (1994).

[6] T.B. Taha, and M.J. Ablowitz, Analytical and numerical aspects of certain nonlinear evolution equations. II. Numerical nonlinear Schrödinger equation, *J. Computational Phys.* **55** (1984) 203–230.

[7] J.A.C. Weideman, and B.M. Herbst, Split-step methods for the solution of the nonlinear Schrödinger equation, *SIAM J. Numer. Anal.* **23** (1986) 485–507.

[8] H. Yoshida, Construction of higher order symplectic integrators, *Phys. Lett.***A150** (1990) 262–268.

[9] H. Yoshida, Recent progress in the theory and application of symplectic integrators, *Celest. Mech.* **56** (1993) 27–43.

[10] 吉田春夫, シンプレクティック数値解法, 数理科学 **384** (1995) 37–46.

A METHOD FOR SYMPLECTIC INTEGRATION OF SATELLITE ORBITS

SEPPO MIKKOLA
National Astronomical Observatory,
Mitaka, Tokyo 181, Japan
Permanent address: Tuorla Observatory,
University of Turku, 21500 Piikkiö, Finland

April 6, 1999

Abstract

The use time transformations for constructing efficient symplectic methods for computing the long term behavior of perturbed two-body systems are discussed. Main applications are for artificial satellite orbits. The suggested methods are efficient for all eccentricities.

Key words: symplectic integration – satellite orbits

1 Introduction

Since the publication of the Wisdom-Holman mapping method (Wisdom and Holman 1991, Kinoshita, Yoshida and Nakai 1991), symplectic methods have gained increasing popularity. Recently Mikkola (1997) suggested the use of time transformation to improve the treatment of highly eccentric orbits. This has been further studied by Rauch and Holman (1999) and found to be an essential improvement in some respects. Recently interest has arisen in the application of symplectic methods for artificial satellite orbits (Mikkola, Palmer and Hashida 1999).

The problem of computing satellite orbits differ in some respects from Solar System calculations. The oblateness potential is usually the most important perturbation and this is strong near the Earth. Also air drag may be present. Drag forces are non-canonical and the equations of motion are not symplectic any more, but this is inconsequential if

the drag is weak (Mikkola 1998). Thus the most important new feature in the problem is the presence of strong perturbation near pericentre.

2 Time transformations

Recently I considered (Mikkola 1997) the use of time transformation in constructing symplectic integration methods for the few body problem. For the two-body problem the suggested transformation takes the form $dt = rds$, where s is the new independent variable, r the distance and t the physical time. This is approximately equivalent to using the eccentric anomaly as the independent variable. However, if we consider eccentric orbits of artificial satellites, then almost all the perturbation occurs near pericentre. Since in symplectic integration the steplenght in the independent variable must be kept constant, the use of the true anomaly as an independent variable concentrates the perturbation evaluations near pericentre and thus improves precision.

It is possible to parameterize the time transformation so as to include all the alternatives, mean, eccentric and true anomaly, as well as any linear combination of them in one expression:

$$ds = dt/\mu(r) \quad (1)$$

$$\mu(r) = 1/(B_0 + B_1/r + B_2/r^2), \quad (2)$$

where B_0, B_1, B_2 are adjustable parameters (constants). Let us consider the Hamiltonian $H = K + R$, where $K = \mathbf{p}^2/2 - m/r$ is the Keplerian part and $R = R(\mathbf{r}, t)$ is the perturbing potential. Following Mikkola (1997) we apply the Poincare transformation to obtain the time transformed Hamiltonian Γ in the form

$$\Gamma = \mu(r)(K + p_0 + R) = \mu(r)(K + p_0) + \mu(r)R(\mathbf{r}, t), \quad (3)$$

where p_0 is the momentum conjugate to the physical time, which now is a coordinate while s is the independent variable. For p_0 the initial value $p_0 = -(K + R)_{/t=0}$ must be used to make the Hamiltonian $\Gamma(t) = 0$ along the orbit and consequently ensure the correctness of the Hamiltonian equations of motion:

$\mathbf{r}' = \partial\Gamma/\partial\mathbf{p}$, $\mathbf{p}' = -\partial\Gamma/\partial\mathbf{r}$, $t' = \partial\Gamma/\partial p_0$ and $p_0' = -\partial\Gamma/\partial t$,
(prime denoting derivative with respect to the new independent variable s).

If the generalized leapfrog method is used an integration step may be written as

$$\Gamma_0(h_s/2)\Gamma_1(h_s)\Gamma_0(h_s/2), \quad (4)$$

where $\Gamma_0 = \mu(r)(\mathbf{p}^2/2 - m/r + p_0)$ represents the motion of the system according to the 'unperturbed' part of the Hamiltonian (3) and $\Gamma_1 = \mu(r)R(\mathbf{r}, t)$ represents the kick of the

momenta due to the perturbation. The argument tells the length of the motion in terms of the independent variable step h_s.

The Γ_1 step here is simple, but the motion under the Hamiltonian Γ_0 is not completely trivial. To find the solution it is convenient to transform back to the physical time system. The result is

$$H_0 = \frac{1}{\mu(r)}[\mu(r)(K + p_0) - \gamma_0] = K + p_0 - \frac{\gamma_0}{\mu(r)}, \tag{5}$$

where γ_0 is the constant (over one step) numerical value of Γ_0 computed at the beginning of the step. When we now substitute the suggested expression for the time transformation function $\mu(r)$ we get

$$\begin{aligned} H_0 &= \mathbf{p}^2/2 - m/r + p_0 - \gamma_0(B_0 + B_2/r + B_2/r^2) \\ &= \mathbf{p}^2/2 - (m + \gamma_0 B_1)/r - \gamma_0 B_2/r^2 + p_0 - \gamma_0 B_0. \end{aligned} \tag{6}$$

We note that if $B_2 = 0$, this has the form of a two-body Hamiltonian with a modified mass. However, when $B_2 \neq 0$ a small extra term with $1/r^2$ dependence appears. In polar coordinates (r, θ) the above Hamiltonian can be written

$$H_0 = \frac{1}{2}(p_r^2 + (p_\theta^2 - 2\gamma_0 B_2)/r^2) - \tilde{m}/r, \tag{7}$$

where the unnecessary constant term has been dropped and $\tilde{m} = m + \gamma_0 B_1$ is the modified mass. The generating function $S = \psi p_\psi$ transforms the Hamiltonian again into the two-body form

$$H_0 = \frac{1}{2}(p_r^2 + p_\psi^2/r^2) - \tilde{m}/r. \tag{8}$$

After defining a special function $\mathrm{at}_1(z)$ by the expression

$$\begin{aligned} \mathrm{at}_1(z) &= \frac{\mathrm{atan}(\sqrt{z})}{\sqrt{z}} \\ &= \frac{1}{2}\log\left(\frac{1+\sqrt{-z}}{1-\sqrt{-z}}\right)/\sqrt{-z} \\ &= 1 - z/3 + z^2/5 - z^3/7 + z^4/9 - z^5/11 + -.., \end{aligned} \tag{9}$$

[where the series expansion is used when the argument z is small (the normal situation), and for larger absolute values of z the closed form expressions are chosen] we can list the equations needed for computing the solution.

3 Collection of formulae

Let $\mathbf{r}_0$, $\mathbf{v}_0$ and t_0 be the values of position vector, velocity vector and time in the beginning of a Γ_0 step, then the equations below can be used to obtain new values of these quantities at the end of the step, which has the length h_s in the independent variable. Note that the variable p_0 does not chance during this step. The special function at_1 is defined in (9), and the c_ν's are the Stumpff (1962) functions.

$$\mu = 1/(B_0 + B_1/r_0 + B_2/r_0^2) \tag{10}$$
$$\gamma_0 = (\mathbf{v}_0^2/2 - m/r_0 + p_0)\mu \tag{11}$$
$$\tilde{m} = m + \gamma_0 B_1 \tag{12}$$
$$\epsilon = \gamma_0 B_2 \tag{13}$$
$$\mathbf{p}_\theta = \mathbf{r}_0 \times \mathbf{v}_0 \tag{14}$$
$$p_\theta^2 = |\mathbf{p}_\theta|^2 \tag{15}$$
$$\eta = \mathbf{r}_0 \cdot \mathbf{v}_0 \tag{16}$$
$$\dot{r}_0 = \eta/r_0 \tag{17}$$
$$p_\psi^2 = p_\theta^2 - 2\epsilon \tag{18}$$
$$w^2 = \dot{r}_0^2 + p_\psi^2/r_0^2 \tag{19}$$
$$\beta = 2\tilde{m}/r_0 - w^2 \tag{20}$$
$$\zeta = \tilde{m} - r_0\beta \tag{21}$$

Starting value for h

$$h = h_s/(B_0 r_0 + B_1 + B_2/r_0) \tag{22}$$

Beginning of iteration loop (for h)

$$G_\nu = h^\nu c_\nu(\beta h^2) \tag{23}$$
$$r = r_0 + \eta G_1 + \zeta G_2 \tag{24}$$
$$f = 1 - \tilde{m} G_2/r_0 \tag{25}$$
$$g = r_0 G_1 + \eta G_2 \tag{26}$$
$$\xi = g/(r_0^2 f + \eta g) \tag{27}$$
$$y_\psi = \mathrm{atan}(p_\psi \xi)/p_\psi = \xi\, \mathrm{at}_1(p_\psi^2 \xi^2) \tag{28}$$
$$\Delta t = r_0 h + \eta G_2 + \zeta G_3 \tag{29}$$
$$d = B_0 r + B_1 + B_2/r \tag{30}$$
$$h = h - (B_0 \Delta t + B_1 h + B_2 y_\psi - h_s)/d \tag{31}$$

End of iteration loop (exit at convergence)

$$\dot{r} = (\eta G_0 + \zeta G_1)/r \tag{32}$$

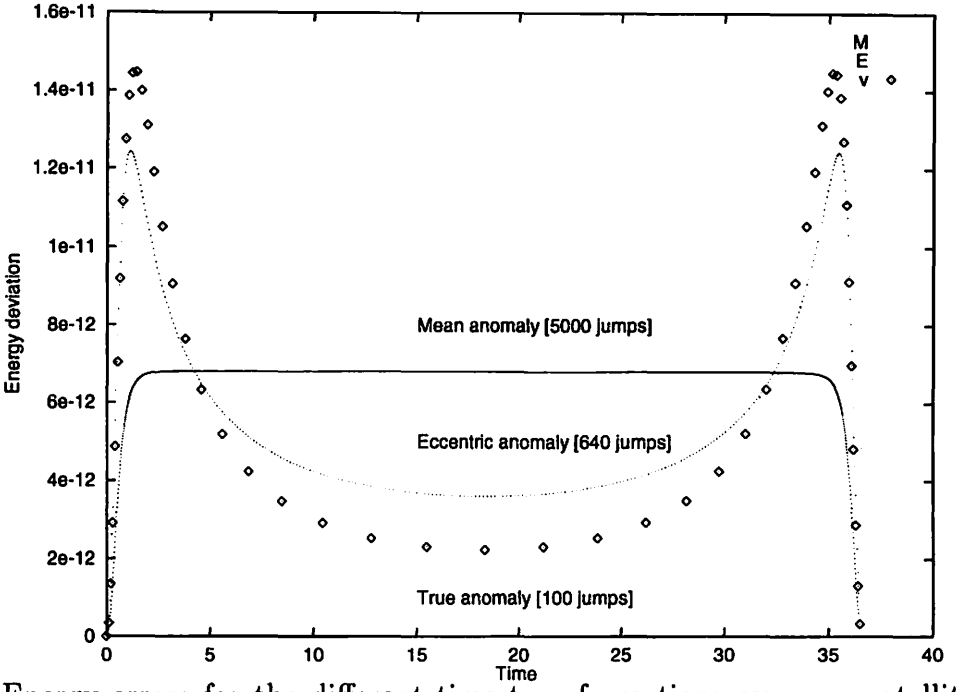


Figure 1: Energy errors for the different time transformations over one satellite period. The perturbing potential used was $.5 \times 10^{-3}/r^3$. For this potential the true anomaly is clearly the most efficient transformation. Units are those of the example orbit given in the main text.

$$\theta^2 = p_\theta^2 y_\psi^2 \tag{33}$$

$$\mathbf{r} = \frac{r}{r_0}(c_0(\theta^2)\mathbf{r}_0 + y_\psi c_1(\theta^2)\mathbf{p}_\theta \times \mathbf{r}_0) \tag{34}$$

$$\mathbf{v} = \frac{\dot{r}}{r}\mathbf{r} + \frac{1}{r^2}\mathbf{p}_\theta \times \mathbf{r} \tag{35}$$

$$t = t_0 + \Delta t \tag{36}$$

Finally, what remains to be explained here is the advancement of the system under the perturbation part $\mu(r)R(\mathbf{r},t)$ of the Hamiltonian. Since this depends only on coordinates (time is also a coordinate now), the motion is linear in momentum space and coordinates remain constant. We thus have the momentum jumps

$$\delta\mathbf{v} = -h_h \frac{\partial(\mu(r)R(\mathbf{r},t))}{\partial\mathbf{r}}, \quad \delta p_0 = -h_s \frac{\partial(\mu(r)R(\mathbf{r},t))}{\partial t}, \tag{37}$$

which are to be added to the most recent values of $\mathbf{v}$ and p_0.

4 Experiments in a toy model

Here we demonstrate in a simple way the effect of various time transformations to the computing accuracy. In Figure 1 we plot the energy errors in an equatorial satellite orbit

with eccentricity $e = 0.69$, and the perigee distance $q = 1$, computed over (approximately) one period. An improved leapfrog was used in the computation (Mikkola 1999). Only the so called J_2 term was taken into account. Thus in this case the perturbing potential was simply $J_2/(2r^3)$ ($J_2 = 0.001$). The stepsizes in the different computations were adjusted so as to get roughly similar error levels, this resulted into the values: 5000 jumps for using no time transformation (labeled Mean anomaly), 640 jumps for $dt = rds$ (labeled Eccentric anomaly) and only 100 jumps for $dt = r^2 ds$ (labeled True anomaly).

5 Conclusion

The use of time transformations, other than the traditional eccentric anomaly, are promising alternatives in the symplectic integration of eccentric orbits of satellites. In the case of J_2 perturbations only, the true anomaly offers the highest efficiency, while in very elongated orbits with tidal perturbation the eccentric anomaly or a mixed anomaly (several of the B-coefficients nonzero) is likely to be a better alternative. Finally it must be stressed that in real satellite orbit integrations the perturbing force is complicated and computationally costly. Thus the possibility of decreasing the necessary number of force evaluations (velocity jumps) is important.

References

Kinoshita, H. , Yoshida, H. and Nakai. H. :(1991) 'Symplectic Integrators and their Application in Dynamical Astronomy', *Celest. Mech. Dyn.Ast.* **50**, 59–71.

Mikkola, S. (1997) 'Practical symplectic methods with time transformation for the few-body problem', *Celestial Mechanics and Dynamical Astronomy*, **67**: 145–165.

Mikkola, S. (1998) 'Non-canonical perturbations in symplectic integration', *Celestial Mechanics and Dynamical Astronomy*, **68**: 249–255

Mikkola, S., Palmer, P. L. and Hashida. Y. : 1999, 'A Symplectic Orbital Estimator for Direct Tracking on Satellites', submitted.

Mikkola, S. (1999) 'Efficient symplectic integration of satellite orbits', submitted.

Rauch, K. P. and Holman, M.:(1999) 'Dynamical Chaos in the Wisdom-Holman Integrator: Origins and Solution', Astron J. **117**, 1087-1102.

Stumpff, K.:(1962) *'Himmelsmechanik'*, Band I, VEB Deutscher Verlag der Wissenschaften, Berlin.

Wisdom, J. and Holman, M. :(1991) 'Symplectic Maps for the N-Body Problem.' *Astron. J.* **102**, 1520.

Super Implicit Multistep Methods

Toshio Fukushima
National Astronomical Observatory of Japan
(Mail) 2-21-1, Ohsawa, Mitaka, Tokyo 181-8588, Japan
(E-mail) Toshio.Fukushima@nao.ac.jp

ABSTRACT

We developed a new type of implicit formulas of linear multistep methods to integrate ordinary differential equations (ODE) numerically. The formulas, which we call to be *super-implicit*, are of more implicitness than the so-called implicit formulas in the sense that they require the knowledge of functions not only at the past and present timesteps but also at the future ones. By using the operator formalism, we derived the expressions of generating functions of Cowell- and Adams-type super-implicit formulas of arbitrary degree of implicitness together with those of some auxiliary formulas such as to be used in the starting procedures and to evaluate derivatives numerically. By combining some of these formulas, we created several sets of implicit formulas to integrate ODEs for an interval of finite length. We adapted the Picard method (Fukushima 1997a) to solve these nonlinear equations. The resulting method works as a one-step method integrating a large (say, a few to some tens orbital periods) time interval as a whole. As an example, we developed a series of schemes to integrate a special second order ODE by using the symmetric form of Cowell-type formulas. The resulting schemes are of high order, up to 12th, and are of the same properties that the integration error grows linearly with respect to time as the symmetric multistep methods have (Quinlan & Tremaine 1990; Fukushima 1998, 1999).

1. Introduction

The symmetric multistep methods (Lambert & Watson 1976; Quinlan & Tremaine 1990; Fukushima 1998) are known to be of a good property that orbital integration errors grow more slowly with respect to time than the case of ordinary integrators such as the Runge-Kutta formulas, extrapolation methods, and asymmetric multistep methods like Adams- and Störmer-Cowell formuals. Unfortunately, high[1] order symmetric formulas

[1]8th and higher for explicit formulas, and 10th and higher for implicit formulas

seem to suffer numerical instabilities caused by the resonances among the frequencies embbeded in the integration formulas and/or the frequencies proper to the problem to be integrated (Fukushima 1998; Fukushima 1999; Quinlan 1999). The frequencies embbeded in the symmetric formulas come from the spurious roots of the characteristic polynomial for position, $\sigma(z)$.

One remedy is to explore the symmetric formulas whose $\sigma(z)$ containing the principal (double) root, +1, only. However, the order of such formulas are limited; 2 for the explicit,

$$x_{n+1} = 2x_n - x_{n-1} + h^2 f_n \tag{1}$$

and 4 for the implicit formulas;

$$x_{n+1} = 2x_n - x_{n-1} + h^2 \left(\frac{1}{12} f_{n+1} + \frac{5}{6} f_n + \frac{1}{12} f_{n-1} \right) \tag{2}$$

In order to overcome this *order barrier*, we consider here a new extension of multistep methods; *super-implicit* formulas. The super-implicit formulas are the formulas requiring the knowledge (the f-value in this case) of not only the past and present timesteps but also the future ones. On the other hand, the explicit formulas require the knowledge of past only while the implicit formulas do those of past and present. An example of super-implicit formula is

$$x_{n+1} = 2x_n - x_{n-1} + h^2 \left(\frac{-1}{240} f_{n+2} + \frac{1}{10} f_{n+1} + \frac{97}{120} f_n + \frac{1}{10} f_{n-1} + \frac{-1}{240} f_{n-2} \right) \tag{3}$$

which is of 6th order. Remark that, in this formula, the f-value at a future timestep, f_{n+2}, is necessary to compute the position at the present timestep, x_{n+1}.

In this short article, we report that the super-implicit formulas are of the same good property as the explicit and implicit symmetric formulas. First, in Section 2, we will describe the general formulas by using the operator formalism of difference method. Next, in Section 3, we will present how the Picard method solves the super-implicit formulas. Then, in Section 4, we will show some numerical tests; numerical integrations of Keplerian orbits. Finally, in Section 5, we will conclude the new method and state its future issues.

2. Formulas

The super-implicit formulas are a kind of generalization of explicit and implicit formulas. Therefore, we derive them here for special second order ODEs by extending the operator

formalism to develop general explicit and implicit formulas. Of course, the concept of super-implicitness is not proper to the integration of special second order ODEs. Their applications to the formulas to evaluate velocities and the integration formulas for general first order ODEs will be described in Appendices.

2.1. Running Formulas of Cowell Methods

Let's begin by the formulas for a special second-order ordinary differential equation

$$\frac{\mathrm{d}^2x}{\mathrm{d}t^2} = f(t, x). \tag{4}$$

A *truely* general linear multistep method to solve this is expressed as

$$x_{n+1} = -\sum_{j=1}^{k} \alpha_j x_{n+1-j} + h^2 \sum_{j=0}^{\ell} \beta_j f_{n+1+m-j}, \tag{5}$$

where h is the stepsize and we assume that

$$\alpha_k \beta_0 \beta_\ell \neq 0. \tag{6}$$

We call the formula to be explicit when $m < 0$, implicit when $m = 0$, and super-implicit when $m > 0$. Also we name m the degree of implicitness. Once α-coefficients are given, the corresponding β-coefficients are automatically determined if the parameters m and ℓ are fixed and if one requires that the maximum order is achieved by the method. In this sense, the β-coefficients are additionally indexed by the pair (m, ℓ) as $\beta_j^{(m,\ell)}$. The order of the method is equal to the number of free coefficients plus one[2]. Therefore the order becomes the number of β-coefficients plus one, $\ell + 2$.

For the moment, we concentrate ourselves with the Störmer-Cowell type formulas. Namely $k = 2$ and α-coefficients are

$$\alpha_1 = -2, \qquad \alpha_2 = 1. \tag{7}$$

Thus Eq.(5) becomes

---

[2]This additional order comes from the property of the α-coefficients such that Eq.(5) satisfies a linear solution, $x_n = x_0 + nhv_0$.

$$x_{n+1} - 2x_n + x_{n-1} = h^2 \sum_{j=0}^{\ell} \beta_j^{(m,\ell)} f_{n+1+m-j}, \tag{8}$$

Now, the β-coefficients are derived by the operator formalism. Let us introduce the shift operator E, the backward difference operator ∇, and the derivative operator D;

$$Ex_n \equiv x_{n+1}, \qquad \nabla f_n \equiv f_n - f_{n-1}, \qquad D^2 x_n \equiv f_n. \tag{9}$$

The relations among these operators are symbolically written as

$$E = e^{hD}, \qquad \nabla = 1 - E^{-1}, \qquad \text{or} \qquad hD = -\log(1-\nabla). \tag{10}$$

Then we learn the following translations

$$x_{n+1} - 2x_n + x_{n-1} \to \nabla^2 x_{n+1}, \qquad f_{n+1+m-j} \to E^{m-j} f_{n+1}, \qquad f_{n+1} \to D^2 x_{n+1}, \tag{11}$$

which transform Eq.(8) in the language of operators as

$$\nabla^2 = (hD)^2 \sum_{j=0}^{\ell} \beta_j^{(m,\ell)} E^{m-j}. \tag{12}$$

We interprete this as an ℓ-th order truncation of the expansion series with respect to ∇ as

$$\sum_{j=0}^{\ell} \beta_j^{(m,\ell)} E^{m-j} \approx (1-\nabla)^{-m} \sum_{j=0}^{\infty} \kappa_j^{(m)} \nabla^j. \tag{13}$$

If we introduce the generating function of the expansion coefficients $\kappa_j^{(m)}$ as

$$K_m(z) = \sum_{j=0}^{\infty} \kappa_j^{(m)} z^j, \tag{14}$$

then Eq.(12) is rewritten as

$$K_m(\nabla) \approx \left(\frac{\nabla}{hD}\right)^2 (1-\nabla)^m, \tag{15}$$

Thus the generating function is written explicitly as

$$K_m(z) = \left[\frac{z}{\log(1-z)}\right]^2 (1-z)^m. \tag{16}$$

The case $m = -1$ corresponds to the explicit formula, namely Störmer formulas. While the case $m = 0$ does to the implicit formula, i.e. Cowell formulas (Hairer *et al.* 1993). To obtain the κ-coefficients explicitly, we recommend the usage of *Mathematica* or other software of formula processing[3]. After the κ-coefficients are obtained, the β-coefficients are computed as

$$\beta_k^{(m,\ell)} = (-1)^k \sum_{j=k}^{\ell} \left(\begin{array}{c} j \\ k \end{array} \right) \kappa_j^{(m)} \tag{22}$$

2.2. Boundary Formulas of Cowell Methods

Consider how to start Cowell Methods. One option is, of course, to prepare some starting values by using one of one-step methods such as Runge-Kutta methods or the extrapolation methods. Another way is to provide the formulas deriving necessary starting values. In the case of Cowell methods, we should consider how to find x_1 from the given initial condition, x_0 and v_0. Once x_1 is known, it is easy to construct the formulas to obtain x_2 and the following values; just using the running formula of the form, Eq.(5).

---

[3]Of course, we may obtain them by the recurence relations (Hairer *et al.* 1993). The generating functions for non-zero m are easily derived from $K_0(z)$ as

$$K_m(z) = K_0(z)(1-z)^m. \tag{17}$$

This is rewritten in the recurrence form as

$$K_m(z) = K_{m-1}(z)(1-z), \tag{18}$$

which follows the following recurrence formula for coefficients

$$\kappa_0^{(m)} = \kappa_0^{(m-1)} = 1, \qquad \kappa_j^{(m)} = \kappa_j^{(m-1)} - \kappa_{j-1}^{(m-1)}, \qquad (j > 0). \tag{19}$$

Remark that there exist symmetric relations as

$$\kappa_{2j}^{(m)} = \kappa_{2j}^{(2j-3-m)}, \qquad \kappa_{2j+1}^{(m)} = -\kappa_{2j+1}^{(2j-2-m)}, \qquad (m \geq j-1). \tag{20}$$

As a result

$$\kappa_{2m+3}^{(m)} = 0. \tag{21}$$

To do this task by means of super-implicit approach, let us seek for formulas of the form

$$x_1 = x_0 + hv_0 + h^2 \sum_{j=0}^{\ell} \gamma_j^{(\ell)} f_j \tag{23}$$

By using the operator D, we rewrite this as

$$e^{hD} - 1 - hD = (hD)^2 \sum_{j=0}^{\ell} \gamma_j^{(\ell)} e^{jhD}. \tag{24}$$

We interprete this as a truncation of the expansion series with respect to the forward difference operator

$$\Delta \equiv e^{hD} - 1 \tag{25}$$

as

$$\Delta - hD = (hD)^2 \sum_{j=0}^{\infty} b_j \Delta^j \tag{26}$$

By using the rewriting

$$hD = \log(1 + \Delta) \tag{27}$$

we know that the corresponding generating function is given as

$$\begin{aligned} G(z) \equiv \sum_{j=0}^{\infty} g_j z^j &= \frac{z - \log(1+z)}{[\log(1+z)]^2} \\ &= \frac{1}{2} + \frac{1}{6}z - \frac{1}{24}z^2 + \frac{1}{45}z^3 - \frac{7}{480}z^4 + \frac{107}{10080}z^5 - \frac{199}{24192}z^6 \\ &+ \frac{6031}{907200}z^7 - \frac{5741}{1036800}z^8 + \frac{1129981}{239500800}z^9 - \frac{435569}{106444800}z^{10} + \cdots \end{aligned} \tag{28}$$

After the g-coefficients are obtained, the corresponding boundary γ-coefficients are computed as

$$\gamma_k^{(\ell)} = \sum_{j=k}^{\ell} (-1)^{(j-k)} \left(\begin{array}{c} j \\ k \end{array} \right) g_j \tag{29}$$

3. Method

3.1. Set of Formulas

Let us construct a self-consistent set of (implicit and) super-implicit formulas for a special second order ODE. To do this, we consider a long time interval $(t_0, t_0 + H)$ and name it as a block. Then we divide the block into N equal-length subintervals $(t_0, t_1, \cdots, t_N)$ as

$$t_j = t_0 + jh, \tag{30}$$

where

$$h = \frac{H}{N} \tag{31}$$

Then such set must be of the form

$$x_n = X_n \left(x_0, v_0; x_1, x_2, \cdots, x_N; f_0, f_1, \cdots, f_N\right), \qquad (n = 1, \cdots, N) \tag{32}$$

where we assume that the initial conditions are given as

$$x\left(t_0\right) = x_0, \qquad v\left(t_0\right) = v_0. \tag{33}$$

By (1) fixing the (even) order as $p = \ell + 2 = 2m + 4$, (2) requiring no usage of one-step methods as the starting procedure, (3) aiming to use the symmetric form in the running mode, and (4) trying to make the whole set symmetric, we prepare such set as follows;

1. First Step $(n = 0)$

   We rewrite the boundary formula, Eq.(23), as

$$x_1 = x_0 + hv_0 + h^2 \sum_{j=0}^{\ell} b_j^{(0)} f_j \tag{34}$$

2. Starting Steps $(n = 1, \cdots, m)$

   We rewrite the asymmetric Cowell-type running formulas, Eq.(8), as

$$x_{n+1} = 2x_n - x_{n-1} + h^2 \sum_{j=0}^{\ell} b_j^{(n)} f_j, \qquad (n = 1, \cdots, m). \tag{35}$$

Table 1: Coefficients of Super-Implicit Cowell Method

m	n	b_0	b_1	b_2	b_3	b_4	b_5	b_6
0	0	$\frac{7}{24}$	$\frac{1}{4}$	$\frac{-1}{24}$				
	1	$\frac{1}{12}$	$\frac{5}{6}$	$\frac{1}{12}$				
1	0	$\frac{367}{1440}$	$\frac{3}{8}$	$\frac{-47}{240}$	$\frac{29}{360}$	$\frac{-7}{480}$		
	1	$\frac{19}{240}$	$\frac{17}{20}$	$\frac{7}{120}$	$\frac{1}{60}$	$\frac{-1}{240}$		
	2	$\frac{-1}{240}$	$\frac{1}{10}$	$\frac{97}{120}$	$\frac{1}{10}$	$\frac{-1}{240}$		
2	0	$\frac{28549}{120960}$	$\frac{275}{576}$	$\frac{-5717}{13440}$	$\frac{10621}{30240}$	$\frac{-7703}{40320}$	$\frac{403}{6720}$	$\frac{-199}{24192}$
	1	$\frac{863}{12096}$	$\frac{8999}{10080}$	$\frac{-769}{20160}$	$\frac{1987}{15120}$	$\frac{-1609}{20160}$	$\frac{263}{10080}$	$\frac{-221}{60480}$
	2	$\frac{-221}{60480}$	$\frac{977}{10080}$	$\frac{16451}{20160}$	$\frac{1357}{15120}$	$\frac{71}{20160}$	$\frac{-31}{10080}$	$\frac{31}{60480}$
	3	$\frac{31}{60480}$	$\frac{-73}{10080}$	$\frac{2171}{20160}$	$\frac{12067}{15120}$	$\frac{2171}{20160}$	$\frac{-73}{10080}$	$\frac{31}{60480}$
3	0	$\frac{324901}{1451520}$	$\frac{8183}{14400}$	$\frac{-653203}{907200}$	$\frac{50689}{56700}$	$\frac{-196277}{241920}$	$\frac{92473}{181440}$	$\frac{-95167}{453600}$
	1	$\frac{33953}{518400}$	$\frac{424759}{453600}$	$\frac{-81629}{453600}$	$\frac{11143}{28350}$	$\frac{-27533}{72576}$	$\frac{110563}{453600}$	$\frac{-23017}{226800}$
	2	$\frac{-9829}{3628800}$	$\frac{81533}{907200}$	$\frac{761057}{907200}$	$\frac{43151}{907200}$	$\frac{3757}{72576}$	$\frac{-34549}{907200}$	$\frac{14717}{907200}$
	3	$\frac{1571}{362880}$	$\frac{-107}{16200}$	$\frac{11959}{113400}$	$\frac{364033}{453600}$	$\frac{7411}{72576}$	$\frac{-631}{226800}$	$\frac{-779}{453600}$
	4	$\frac{-289}{3628800}$	$\frac{149}{129600}$	$\frac{-8593}{907200}$	$\frac{101741}{907200}$	$\frac{57517}{72576}$	$\frac{101741}{907200}$	$\frac{-8593}{907200}$
4	0	$\frac{41198923}{191600640}$	$\frac{4671}{7168}$	$\frac{-342968359}{319334400}$	$\frac{14221079}{7983360}$	$\frac{-13395017}{5913600}$	$\frac{170469863}{79833600}$	$\frac{-5199599}{3548160}$
	1	$\frac{3250433}{53222400}$	$\frac{3124027}{3193344}$	$\frac{-57128921}{159667200}$	$\frac{16745741}{19958400}$	$\frac{-88645069}{79833600}$	$\frac{42375577}{39916800}$	$\frac{-2342533}{3193344}$
	2	$\frac{-330157}{159667200}$	$\frac{6691513}{79833600}$	$\frac{27608543}{31933440}$	$\frac{-331627}{19958400}$	$\frac{12507059}{79833600}$	$\frac{-6189401}{39916800}$	$\frac{8484887}{79833600}$
	3	$\frac{45911}{159667200}$	$\frac{-417589}{79833600}$	$\frac{15908131}{159667200}$	$\frac{93191}{114048}$	$\frac{6248807}{79833600}$	$\frac{950809}{39916800}$	$\frac{-1773361}{79833600}$
	4	$\frac{-3499}{53222400}$	$\frac{11527}{11404800}$	$\frac{-1412513}{159667200}$	$\frac{2205017}{19958400}$	$\frac{12700339}{15966720}$	$\frac{4336807}{39916800}$	$\frac{-523189}{79833600}$
	5	$\frac{317}{22809600}$	$\frac{-17453}{79833600}$	$\frac{40489}{22809600}$	$\frac{-222331}{19958400}$	$\frac{9186203}{79833600}$	$\frac{31494553}{39916800}$	$\frac{9186203}{79833600}$

Table 2: Coefficients of Super-Implicit Cowell Method (Contd.)

m	n	b_7	b_8	b_9	b_{10}
3	0	$\frac{7703}{151200}$	$\frac{-5741}{1036800}$		
	1	$\frac{5627}{226800}$	$\frac{-9829}{3628800}$		
	2	$\frac{-3607}{907200}$	$\frac{1571}{3628800}$		
	3	$\frac{289}{453600}$	$\frac{-289}{3628800}$		
	4	$\frac{149}{129600}$	$\frac{-289}{3628800}$		
4	0	$\frac{9471361}{13305600}$	$\frac{-14825963}{63866880}$	$\frac{21860567}{479001600}$	$\frac{-435569}{106444800}$
	1	$\frac{7139837}{19958400}$	$\frac{-18674153}{159667200}$	$\frac{1838819}{79833600}$	$\frac{-330157}{159667200}$
	2	$\frac{-204371}{3991680}$	$\frac{2642791}{159667200}$	$\frac{-257759}{79833600}$	$\frac{45911}{159667200}$
	3	$\frac{227393}{19958400}$	$\frac{-23981}{6386688}$	$\frac{58843}{79833600}$	$\frac{-3499}{53222400}$
	4	$\frac{-1477}{2851200}$	$\frac{87139}{159667200}$	$\frac{-317}{2280960}$	$\frac{317}{22809600}$
	5	$\frac{-222331}{19958400}$	$\frac{40489}{22809600}$	$\frac{-17453}{79833600}$	$\frac{317}{22809600}$

Remark that the index of f-values required runs only in the range $(0, \ell)$.

3. Middle Steps $(n = m+1, \cdots, N-m-1)$

   We rewrite the symmetric Cowell-type running formula, Eq.(A2), as

$$x_{n+1} = 2x_n - x_{n-1} + h^2 \left[b_{m+1}^{(m+1)} f_n + \sum_{j=1}^{m+1} b_{m+1-j}^{(m+1)} \left(f_{n+j} + f_{n-j} \right) \right],$$

$$(n = m+1, \cdots, N-m-1). \tag{36}$$

4. Ending Steps $(n = N-m, \cdots, N-1)$

   We rewrite the asymmetric Cowell-type running formulas, Eq.(8), in the reverse form as

$$x_{n+1} = 2x_n - x_{n-1} + h^2 \sum_{j=0}^{\ell} b_j^{(N-n)} f_{N-j}, \qquad (n = N-m, \cdots, N-1). \tag{37}$$

5. End Velocity

   In order to derive the formula to obtain the end velocity, v_N, we solved the boundary formula, Eq.(23), with respect to the velocity as

$$v_N = \frac{x_N - x_{N-1}}{h} - h \sum_{j=0}^{\ell} b_j^{(0)} f_{N-j} \tag{38}$$

Here the new coefficients, $b_j^{(n)}$, are expressed as

$$b_j^{(0)} = \gamma_j^{(2m+2)}, \qquad b_j^{(n)} = \beta_{2m+2-j}^{(2m+1-n, 2m+2)}, \qquad (n = 1, \cdots, m+1). \tag{39}$$

The b-coefficients are shown for $m = 0, 1, 2, 3, 4$ in Tables 1 and 2. As an example, we illustrate the set of 6th order ($m = 1$) super implicit formulas;

$$x_1 = x_0 + hv_0 + h^2 \left(\frac{367}{1440} f_0 + \frac{3}{8} f_1 - \frac{47}{240} f_2 + \frac{29}{360} f_3 - \frac{7}{480} f_4 \right),$$

$$x_2 = 2x_1 - x_0 + h^2 \left(\frac{19}{240} f_0 + \frac{17}{20} f_1 + \frac{7}{120} f_2 + \frac{1}{60} f_3 - \frac{1}{240} f_4 \right),$$

$$x_{n+1} = 2x_n - x_{n-1} + h^2 \left[\frac{97}{120} f_n + \frac{1}{10} \left(f_{n-1} + f_{n+1} \right) - \frac{1}{240} \left(f_{n-2} + f_{n+2} \right) \right],$$

$(n = 2, \cdots, N-2)$

$$x_N = 2x_{N-1} - x_{N-2} + h^2 \left(-\frac{1}{240} f_{N-4} + \frac{1}{60} f_{N-3} + \frac{7}{120} f_{N-2} + \frac{17}{20} f_{N-1} + \frac{19}{240} f_N \right),$$

$$v_N = \frac{x_N - x_{N-1}}{h} + h \left(\frac{7}{480} f_{N-4} - \frac{29}{360} f_{N-3} + \frac{47}{240} f_{N-2} - \frac{3}{8} f_{N-1} - \frac{367}{1440} f_N \right). \tag{40}$$

Similar symmetric sets of formulas of higher orders up to 12th are easily computed from the tables shown in this article.

3.2. Technique to Reduce Round-Off Error

The formulas described in the previous subsection suffer round-off errors since they contain the differences, ∇^2. A common technique to reduce the round-off errors is to introduce intermediate variables corresponding to ∇x_n and to let them convey the information of lower bits, which would be lost if using the raw values x_n only (Hairer *et al.* 1993).

Let us follow the technique and introduce the intermediate variables

$$u_n \equiv \frac{\nabla x_n}{h} = \frac{x_n - x_{n-1}}{h} \tag{41}$$

Then the formulas are rewritten in term of the u-variables as

$$u_1 = v_0 + h \sum_{j=0}^{\ell} b_j^{(0)} f_j, \tag{42}$$

$$u_{n+1} = u_n + h \sum_{j=0}^{\ell} b_j^{(n)} f_j, \qquad (n = 1, \cdots, m), \tag{43}$$

$$u_{n+1} = u_n + h \sum_{j=0}^{\ell} b_j^{(m+1)} f_{n-m-1+j}, \qquad (n = m+1, \cdots, N-m-1), \tag{44}$$

$$u_{n+1} = u_n + h \sum_{j=0}^{\ell} b_j^{(N-n)} f_{N-j}, \qquad (n = N-m, \cdots, N-1), \tag{45}$$

while the x-variables are derived from a simple summation of the u-variables as

$$x_n = x_{n-1} + hu_n, \qquad (n = 1, \cdots, N). \tag{46}$$

and the end-velocity formula is expressed as

$$v_N = u_N - h \sum_{j=0}^{\ell} b_j^{(0)} f_{N-j}. \tag{47}$$

Remark that this last expression contains no small divisor $1/h$, which maginifies the round-off errors appear in the difference $x_N - x_{N-1}$ significantly.

3.3. Picard Method

Consider how to solve the self-consistent set of (implicit and) super-implicit formulas obtained in the previous subsection. Remark that, except the last one, the end-velocity formula, they can be finally written in a vector mapping form as

$$\mathbf{x} = \mathbf{X}\left(\mathbf{f}\left(\mathbf{t}, \mathbf{x}\right)\right), \tag{48}$$

where

$$\mathbf{x} = \left(x_1, x_2, \cdots, x_N\right), \qquad \mathbf{t} = \left(t_1, t_2, \cdots, t_N\right), \tag{49}$$

and $\mathbf{f}$ is the vector notation of f-values as

$$\mathbf{f}\left(\mathbf{t}, \mathbf{x}\right) = \left(f_1\left(t_1, x_1\right), f_2\left(t_2, x_2\right), \cdots, f_N\left(t_N, x_N\right)\right). \tag{50}$$

The mapping is regarded as a large set of nonlinear equation with respect to $\mathbf{x}$. We solve this set by a simple fixed-point iteration;

$$\mathbf{x}^{(i+1)} = \mathbf{X}\left(\mathbf{f}\left(\mathbf{t}, \mathbf{x}^{(i)}\right)\right), \tag{51}$$

from a certain initial guess $\mathbf{x}^{(0)}$. This is the Picard method (Fukushima 1997a). As for the initial guess of the Picard iteration, we adopted the integrated result of the explicit symmetric Cowell method,

$$x_{n+1} = 2x_n - x_{n-1} + h^2 f_n. \tag{52}$$

with the starter based on Taylor's expansion,

$$x_1 = x_0 + hv_0 + h^2\frac{f_0}{2}, \tag{53}$$

which coresponds to the similar end-velocity formula

$$v_N = \frac{x_N - x_{n-1}}{h} + h\frac{f_N}{2}. \tag{54}$$

This method is of 2nd order.

Once the solution is obtained, the end-velocity is calculated from the end-velocity formula. This whole process, the combination of Picard method and the end-velocity formula, is considered as a transformation from the initial state, (x_0, v_0), to the end state, (x_N, v_N). Then, by the symmetries of the formulas adopted, this transformation is symmetric. Namely the reverse process, $(x_N, v_N) \to (x_0, v_0)$ is identical with the original one, $(x_0, v_0) \to (x_N, v_N)$.

4. Numerical Tests

In order to test the Picard method of super-implicit Cowell formulas of the order 4, 6, 8, 10, and 12, we did a basic test; to integrate Keplerian orbits.

First comes the question of the convergence of Picard iterations. Experiments showed us that the number of iterations required are (1) hardly dependent on the order of methods, p, (2) weakly dependent on the stepsize h, and (3) strongly dependent on the blocksize H. Figure 1 illustrates the H-dependence of the number of iterations required for the 12th order method to integrate a Keplerian orbit with the eccentricity $e = 0.2$ where h is fixed as $2\pi/(256)$. Remark the the case with the blocksize of 16 orbital periods did not converge at all. For the blocksize of one orbital period, the number of iterations required is around 20. This means that $20 \times 256 \sim 5000$ function evaluations are required in integrating one orbital period. One may be surprised at this large computational cost. This is true if one sticks to use the scalar computers. In this sense, the super-implicit methods are less practical in the scalar computers.

However, we remark that each Picard iteration can be done *parallely*, or more precisely speaking, *in vector form* (Fukushima 1997b). This means that if one uses a parallel and/or vector computer, the wall-clock time required will be greatly reduced. The expected wall-clock time would be proportional to the number of iterations. Thus, if ignoring

Figure 1. Convergence of Picard Iteration

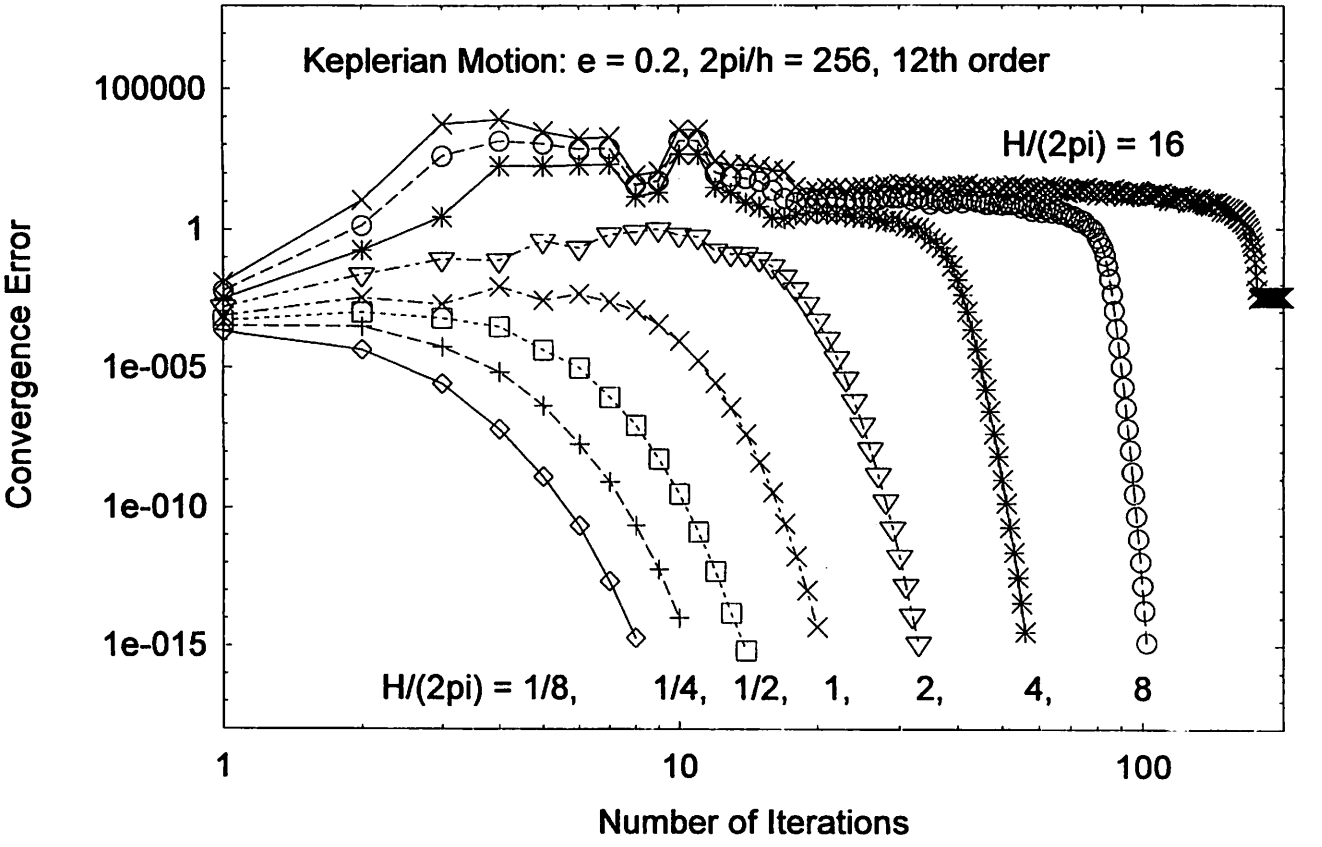
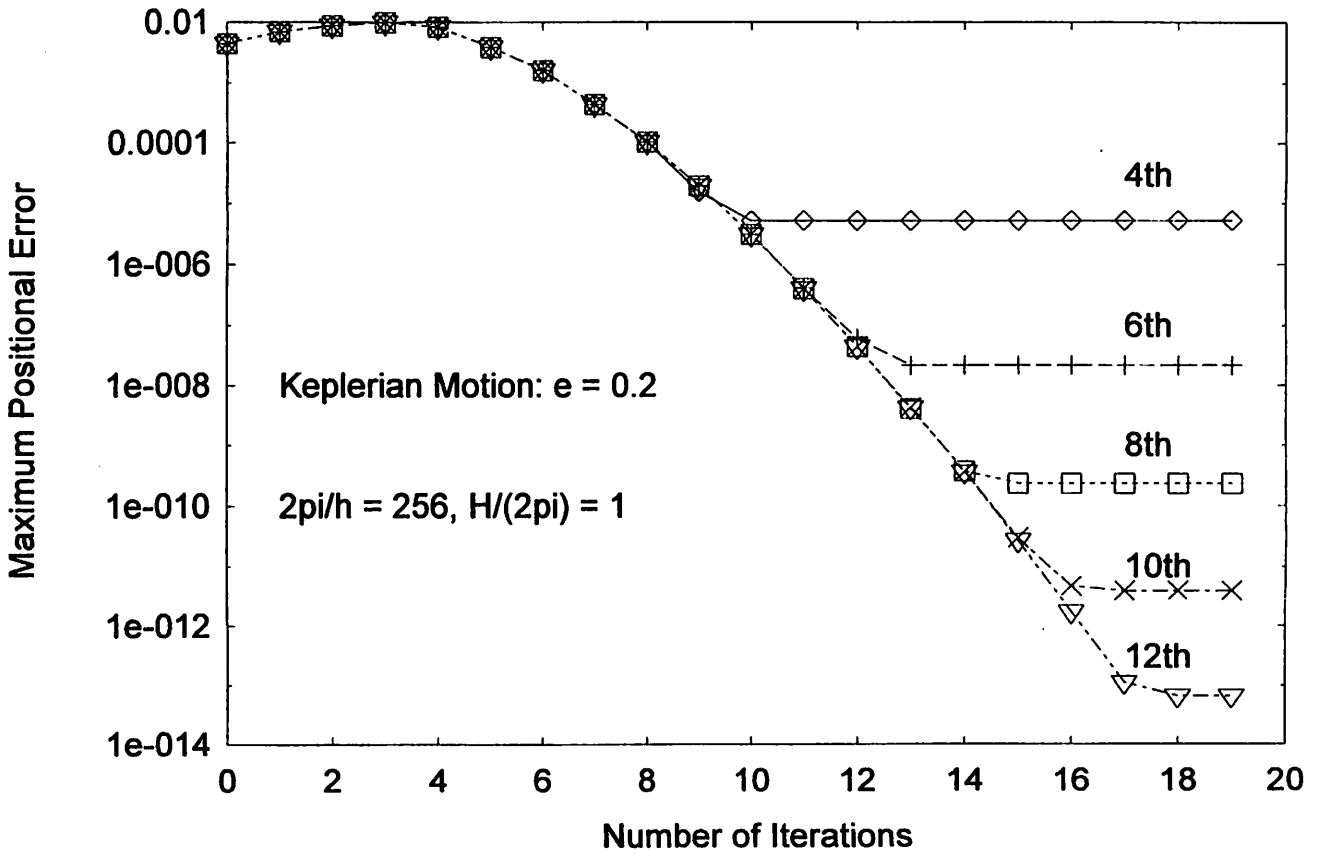


Figure 2. Error Variation during Picard Iteration



overhead times, we anticipate that a parallel/vector computing will cost around 20 function evaluation times per orbit in this case.

Figure 2 shows the maximum of the positional error in a single block as a function of number of Picard iterations. Here we define the positional error as

$$\Delta r \equiv \left| \mathbf{r}_{\text{Numerical Integration}} - \mathbf{r}_{\text{Analytical Solution}} \right| \tag{55}$$

and took the parameters as It is quite interesting that, for the first several iterations, the error rather increases. This is independent with the order of methods. Then it turns to decrease and asymptotes to some fixed values depending on the order of methods. At the same time the convergence error which we defined as difference between subsequent two solutions in the Picard iteration,

$$\delta r^{(i)} \equiv \left| \mathbf{r}^{(i)} - \mathbf{r}^{(i-1)} \right|, \tag{56}$$

does continuously decreases independently with the order of the methods.

Next, let us investigate the integration error in a single block, $(t_0, t_0 + H)$. Figure 3 shows the positional error as a function of time in the first block. Here we integrated a Keplerian orbit of $e = 0.2$ started from the pericenter passage by means of the 6th order super-implicit multistep method described previously for the parameters $h = \frac{2\pi}{256} \sim 2.454 \times 10^{-2}$ and $H = 4\pi$. It is interesting that the positional error seems to vary like a periodic function plus a linear trend. Namely the error after averaging over one orbital period seems to grow linearly as $< \Delta r > \approx c_0 t$ where the magnitude of the linear trend per one orbital period is estimated as $c_0 \sim 1.32 \times 10^{-8}$.

Figure 4 illustrates how this integration error grows for a longer time span; 20 orbital periods. This figure shows that the amplitude of the period term also grows linearly. As a result, one can aproximate the integration error as

$$\Delta r \approx c_0 t - c_1 t \sin M \tag{57}$$

where M is the mean anomaly and the linear trend of the amplitude of the periodic term per one orbital period is estimated as $c_1 \sim 0.28 \times 10^{-8}$.

Figure 5 shows the order dependence of integration error for a longer integration period; 128 orbital periods. There shown are the results for the 4th, 6th, 8th, 10th, and 12th order super-implicit lienar multistep methods. This time the parameters are taken as $h = \frac{2\pi}{100} \sim 6.28 \times 10^{-2}$ and $H = 4\pi$. Remark that the error of the 12th order method are of the same order as that of the 10th order method.

Figure 3. Positional Error in a Single Block

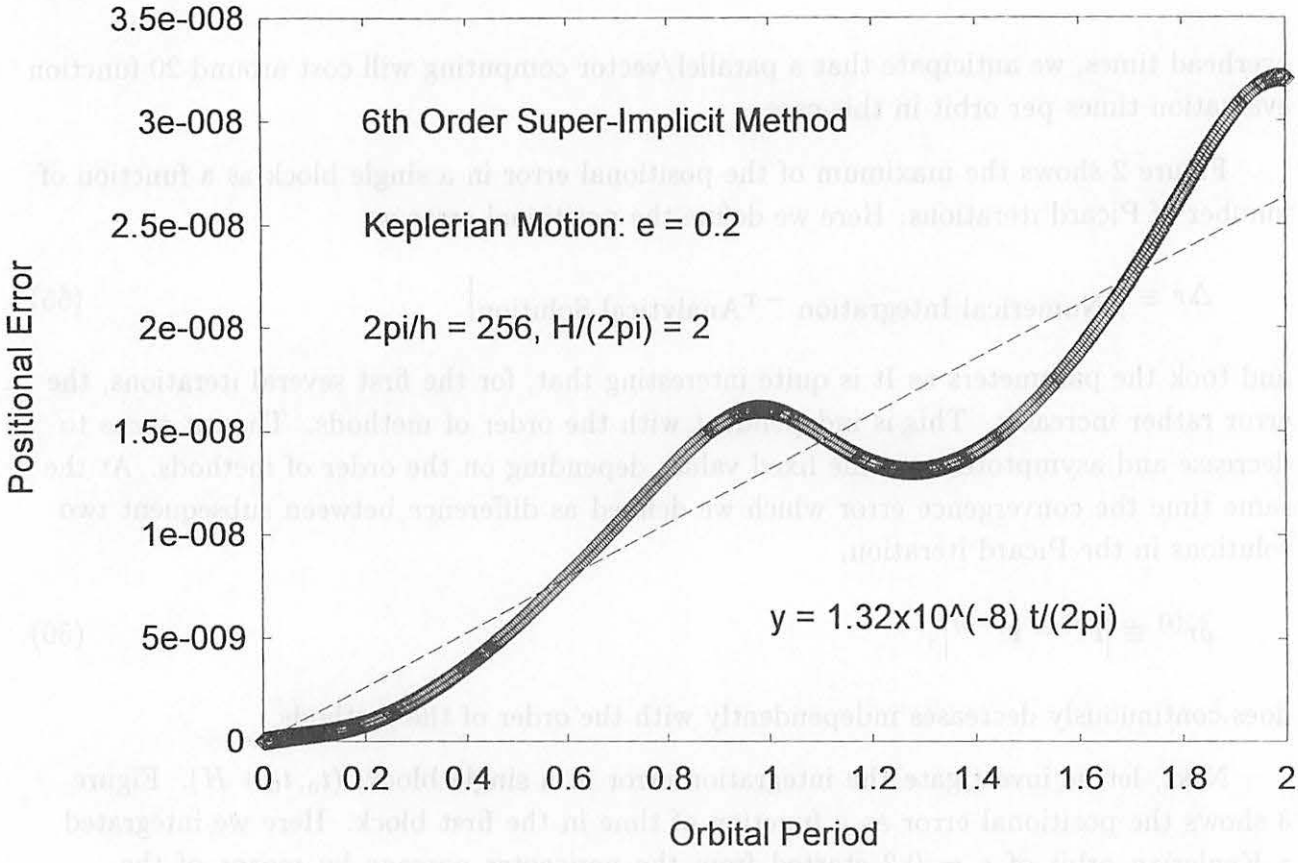
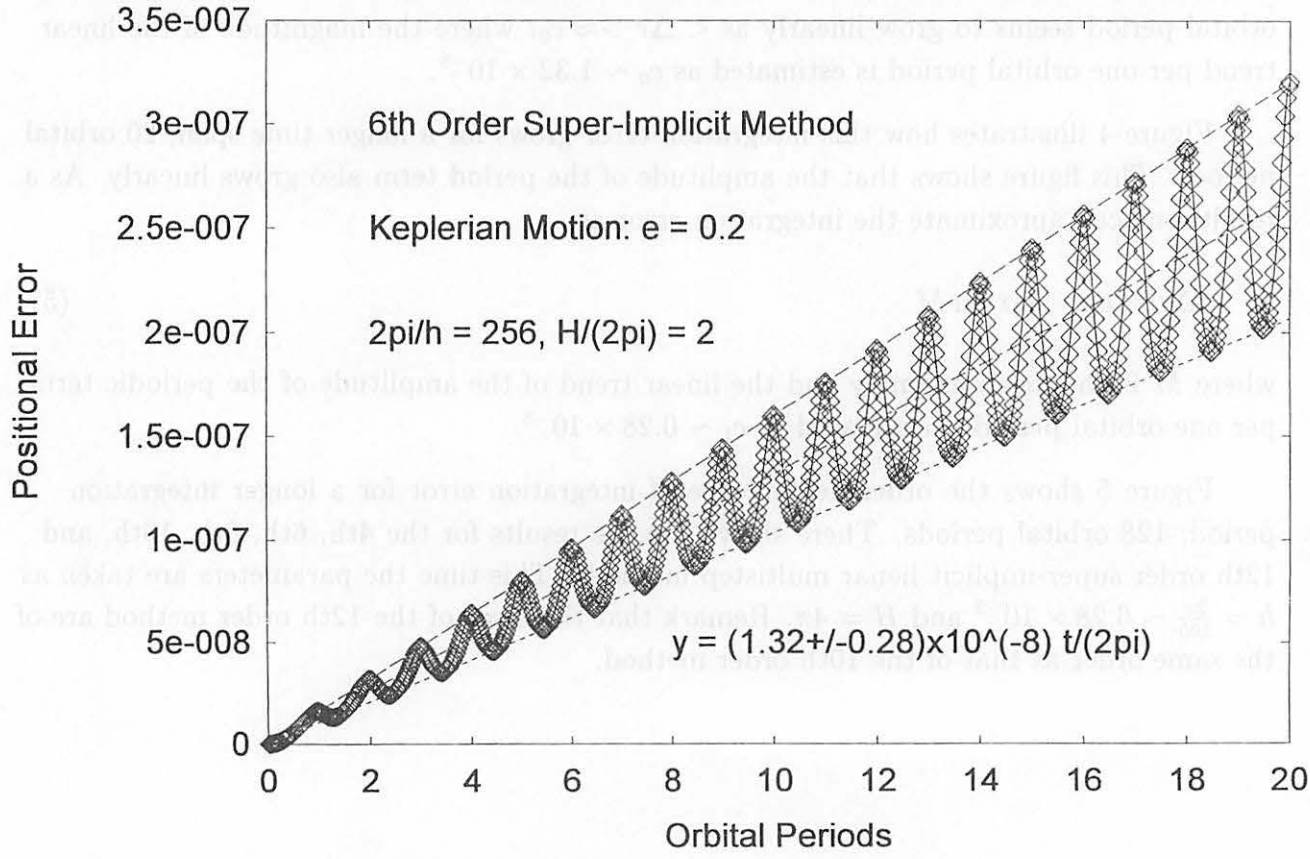


Figure 4. Short-Term Behavior of Positional Error



Finally Figure 6 illustrates the h-dependence of integration errors for the 12th order method. This shows that the super-implicit methods suffer no stepsize resonance/instability.

5. Conclusion

We created a new concept in the theory of linear multistep methods to integrate ODEs numerically; the *super-implicit* formulas. The formulas have more implicitness than the existing implicit formuals in the sense they require the information of not only the present timesteps but also those of future ones. By means of operator formalisms, we introduced the general expressions of Cowell-type super-implicit formulas for special second order ODEs, as well as the starting (or boundary) formulas and the formulas to evaluate velocities at intermediary steps. Also we derived the general formulas of Adams-type super-implicit formulas for general first order ODEs. By combining some of these formulas symmetrically, we developed the sets of self-consistent formulas of the order of 6, 8, 10, and 12 for special second order ODEs. We stated clearly how these nonlinear equations are solved by the Picard method. These methods are considered as one-step integrators of large stepsize such as a few to some tens orbital periods. As a result the method is expected to be accelerated by the factor of some to several tens by using parallel/vector computers. The methods are time-symmetric in the sense that they transform the final state vector obtained back to the initial one exactly. As a result, the truncation errors produced by these integrators for orbital dynamics are expected to grow only linearly with respect time. Also, by its nature, these integrators are expected to suffer no stepsize-dependent resonance/instabilities. We confirmed this through test integrations of Kepler motions for long periods.

REFERENCES

Hairer, E., Nϕrsett, S.P., and Wanner, G., 1993, *Solving Ordinary Differential Equations I (2nd ed.)*, Springer-Verlag, Berlin.

Henrici, P., 1962, *Discrete Variable Methods in Ordinary Differential Equations*, Wiley, New York.

Fukushima, T.: 1997a, AJ, 113, 1909-1914.

Fukushima, T.: 1997b, AJ, 113, 2325-2328.

Fukushima, T.: 1998, *in Proc. 30th Symp. on Cele. Mech. (Fukushima et al. eds)*, 229-247.

Fukushima, T.: 1999, *in Proc. IAU Coll. No.173*, to be printed.

Figure 5. Long-Term Behavior of Integration Error

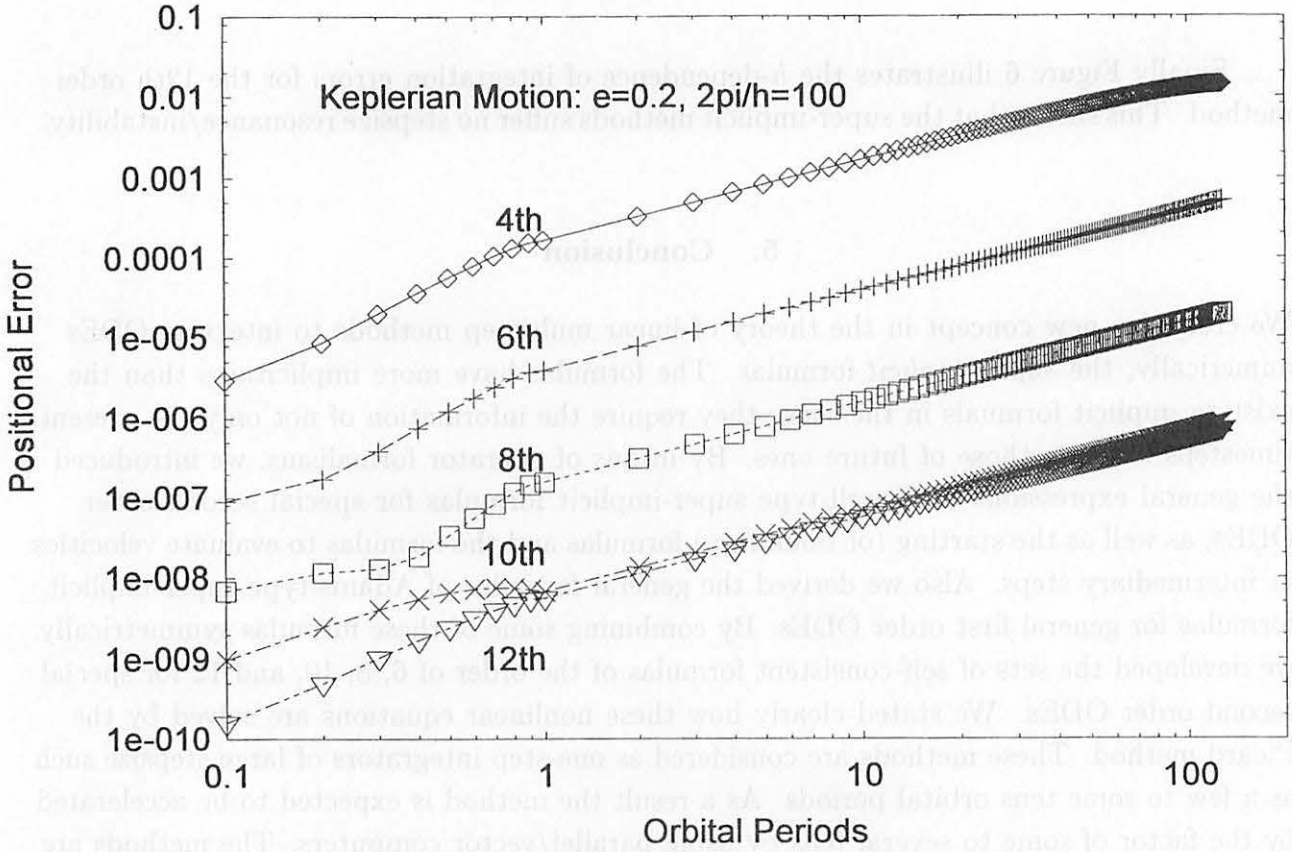
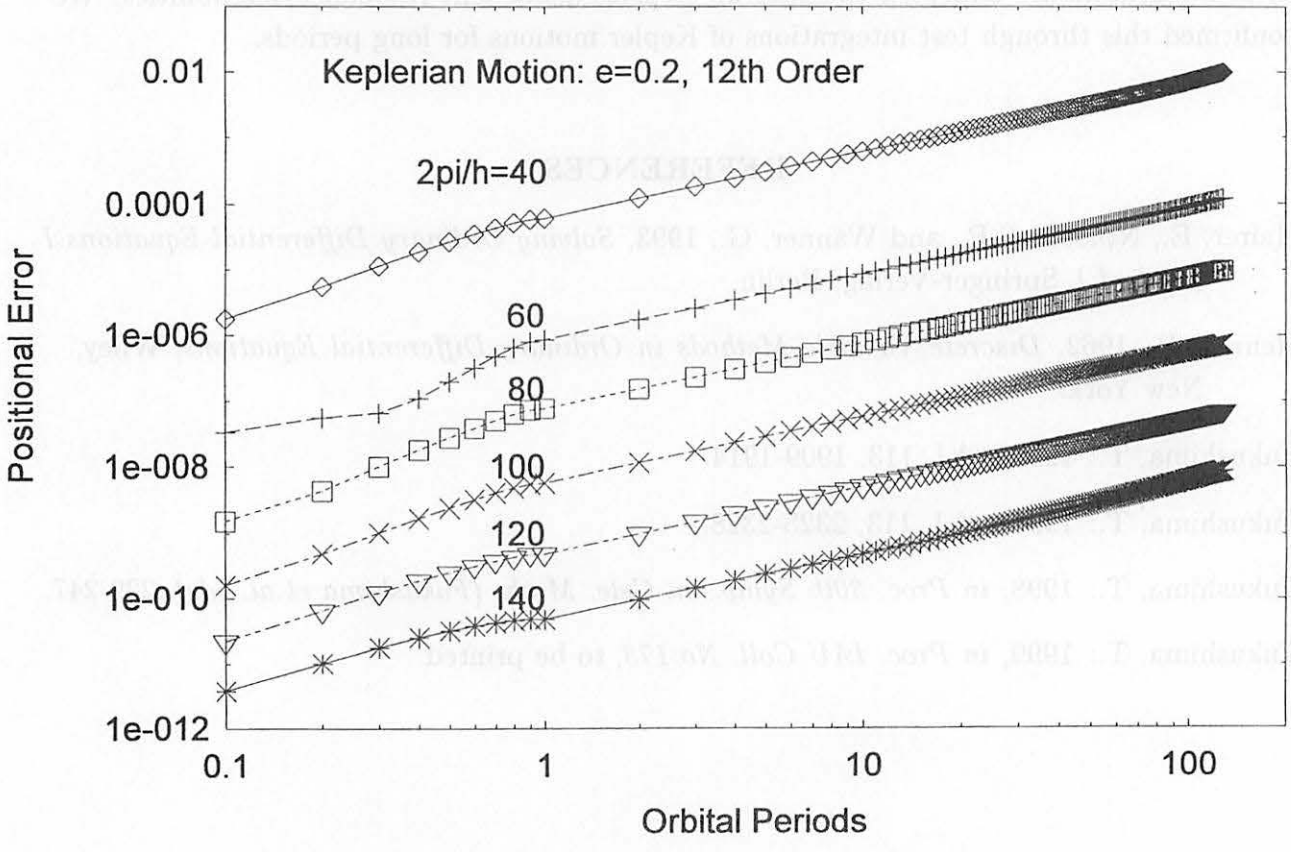


Figure 6. Step-Size Dependence of Long-Term Behavior



Lambert, J.D., and Watson, I.A., 1976, J. Inst. Maths Applics, 18, 189-202.

Quinlan, G.D., 1999, submitted to AJ.

Quinlan, G.D., and Tremaine, S., 1990, AJ, 100, 1694-1700.

APPENDICES

A. Symmetric Running Formulas of Cowell Methods

Consider a situation where β-coefficients are symmetric. In this case, $\ell = 2m + 2$ and

$$\beta_k^{(m,2m+2)} = \beta_{2m+2-k}^{(m,2m+2)}, \qquad (k = 0, \cdots, m) \tag{A1}$$

and therefore, Eq.(5) becomes

$$x_{n+1} = 2x_n - x_{n-1} + h^2 \left[\beta_{m+1}^{(m,2m+2)} f_n + \sum_{j=1}^{m+1} \beta_{m+1-j}^{(m,2m+2)} \left(f_{n+j} + f_{n-j} \right) \right] \tag{A2}$$

Of course, the symmetric β-coefficients can be calculated by the formulation given in the main text. However, by using the symmetry, we can derive them in a different and smarter way.

Let's begin by introducing the central difference operator

$$Y \equiv E - 2 + E^{-1} = 2(\cosh hD - 1). \tag{A3}$$

By using Y, we rewrite Eq.(A2) in the form of operators as

$$Y = (hD)^2 \left[\beta_{m+1}^{(m,2m+2)} + \sum_{j=1}^{m+1} \beta_{m+1-j}^{(m,2m+2)} \left(E^j + E^{-j} \right) \right] \tag{A4}$$

We interprete this as a truncation of the expansion series with respect to Y as

---

This preprint was prepared with the AAS LaTeX macros v4.0.

$$\beta_{m+1}^{(m,2m+2)} + \sum_{j=1}^{m} \beta_{m+1-j}^{(m,2m+2)} \left(E^j + E^{-j}\right) \approx \sum_{j=0}^{\infty} \eta_j Y^j \tag{A5}$$

Then the corresponding generating function is given as

$$H(z) \equiv \sum_{j=0}^{\infty} \eta_j z^j = \frac{z}{\left[\cosh^{-1}\left(1+\frac{z}{2}\right)\right]^2}$$

$$= 1 + \frac{1}{12}z - \frac{1}{240}z^2 + \frac{31}{60480}z^3 - \frac{289}{3628800}z^4 + \cdots \tag{A6}$$

After the η-coefficients are obtained, the symmetric β-coefficients are computed as

$$\beta_k^{(m,2m+2)} = \sum_{i=0}^{k} (-1)^{(k-i)} \begin{pmatrix} 2(m+1-i) \\ k-i \end{pmatrix} \eta_{m+1-i}, \qquad (k = 0, \cdots, m+1) \tag{A7}$$

B. Velocity Formulas

Although the information on velocities are not needed in the numerical integration based on Cowell Methods, it is quite useful to obtain the velocities at the intermediate steps where the position and accelerations have been already obtained. One possibility to realize this is to use Adams-type integration formula

$$v_{n+1} = v_n + hF_n\left(f_0, f_1, \cdots\right) \tag{B1}$$

Unfortunately, this type of integrtion formulas suffer the accumulation of truncation errors. Rather, a difference-type formula like

$$v_n = \frac{1}{h} X_n\left(x_0, x_1, \cdots\right) + hF_n\left(f_0, f_1, \cdots\right) \tag{B2}$$

is preferrable although the first part X_n must contain differences, ∇x_j, which causes round-off errors. To avoid this as much as possible, we must minimize the number of differences; to reduce it to one. Namely

$$X_n = \nabla x_n \tag{B3}$$

Then, the resulting form is generally expressed as

$$v_n = \frac{x_n - x_{n-1}}{h} + h \sum_{j=0}^{k} \lambda_j^{(m,k)} f_{n+m-j} \tag{B4}$$

By using the difference and other operators, we rewrite this as

$$D = \frac{\nabla}{h} + hD^2 \sum_{j=0}^{k} \lambda_j^{(m,k)} E^{m-j}. \tag{B5}$$

We interprete this as a truncation of the expansion series with respect to the backward difference operator, ∇, as

$$D = \frac{\nabla}{h} + hD^2 E^m \sum_{j=0}^{\infty} \ell_j^{(m)} \nabla^j. \tag{B6}$$

By using the relations among the operators again, we obtain the generating function corresponding to the ℓ-coefficients as

$$L_m(z) \equiv \sum_{j=0}^{\infty} \ell_j^{(m)} z^j = \left\{ \frac{z + \log(1-z)}{[\log(1-z)]^2} \right\} (1-z)^m \tag{B7}$$

After the ℓ-coefficients are obtained, the corresponding λ-coefficients are computed as

$$\lambda_k^{(m,i)} = (-1)^k \sum_{j=k}^{i} \begin{pmatrix} j \\ k \end{pmatrix} \ell_j^{(m)} \tag{B8}$$

C. Symmetric Velocity Formulas

Consider symmetric formulas for evaluating velocities. Among such formulas, one option to minimize the number of differences is to take the central difference once as

$$X_n = \frac{1}{2}(x_{n+1} - x_{n-1}) \tag{C1}$$

Further, we assume that F_n takes the symmetric form, which minimizes the truncation error if the order is specified. Then, the resulting form is

$$v_n = \frac{x_{n+1} - x_{n-1}}{2} + \frac{h}{2} \sum_{j=1}^{m} \delta_j^{(m)} (f_{n+j} - f_{n-j}) \tag{C2}$$

By using the difference operator, D, we rewrite this as

$$D = \frac{e^{hD} - e^{-hD}}{2h} + hD^2 \sum_{j=1}^{m} \delta_j^{(m)} \left(e^{jhD} - e^{-jhD} \right). \tag{C3}$$

We interprete this as a truncation of the expansion series with respect to the central difference operator, Y, as

$$D = \frac{\sinh hD}{2h} \left[1 - (hD)^2 \sum_{j=0}^{\infty} d_j Y^j \right]. \tag{C4}$$

By using the relation

$$hD = \cosh^{-1}\left(1 + \frac{Y}{2}\right), \qquad \sinh hD = \sqrt{Y + \frac{Y^2}{4}}, \tag{C5}$$

we know that the generating function corresponding to the d-coefficients is given as

$$\begin{aligned} D(z) \equiv \sum_{j=0}^{\infty} d_j z^j &= \frac{\sqrt{z + \frac{z^2}{4}} - \cosh^{-1}\left(1 + \frac{z}{2}\right)}{\left[\cosh^{-1}\left(1 + \frac{z}{2}\right)\right]^2 \sqrt{z + \frac{z^2}{4}}} \\ &= \frac{1}{6} - \frac{7}{360} z + \frac{37}{10080} z^2 - \frac{199}{259200} z^3 + \cdots \end{aligned} \tag{C6}$$

After the d-coefficients are obtained, the corresponding δ-coefficients are computed as

$$\delta_k^{(m)} = d_{k-1} - \binom{2k}{k} d_k + \sum_{j=k+1}^{m} (-1)^j d_j \left[\binom{2j}{j+1-k} - \binom{2j}{j-1-k} \right] \tag{C7}$$

D. Running Formulas of Adams Methods

Consider the formulas for a general first-order ordinary differential equation

$$\frac{\mathrm{d}y}{\mathrm{d}t} = f(t, y). \tag{D1}$$

A *truely* general linear multistep method to solve this is expressed as

$$y_{n+1} = -\sum_{j=1}^{k} \alpha_j y_{n+1-j} + h \sum_{j=0}^{\ell} \beta_j f_{n+1+m-j}, \tag{D2}$$

where we assume again that $\alpha_k \beta_0 \beta_\ell \neq 0$. The order of the method is equal to the number of free coefficients. Therefore the order becomes the number of β-coefficients, $\ell + 1$.

For the moment, we concentrate ourselves with the Adams type formulas. Namely $k = 1$ and α-coefficients are $\alpha_1 = -1$. Thus Eq.(D2) becomes

$$y_{n+1} - y_n = h \sum_{j=0}^{\ell} \beta_j^{(m,\ell)} f_{n+1+m-j}, \tag{D3}$$

By using the operators, we learn the following translations

$$y_{n+1} - y_n \to \nabla y_{n+1}, \qquad f_{n+1+m-j} \to E^{m-j} f_{n+1}, \qquad f_{n+1} \to D y_{n+1}, \tag{D4}$$

which transform Eq.(D3) in the language of operators as

$$\nabla = hD \sum_{j=0}^{\ell} \beta_j^{(m,\ell)} E^{m-j}. \tag{D5}$$

We interprete this as an ℓ-th order truncation of the expansion series with respect to ∇ as

$$\sum_{j=0}^{\ell} \beta_j^{(m,\ell)} E^{m-j} \approx (1 - \nabla)^{-m} \sum_{j=0}^{\infty} \sigma_j^{(m)} \nabla^j. \tag{D6}$$

If we introduce the generating function of the expansion coefficients $\sigma_j^{(m)}$ as

$$S_m(z) = \sum_{j=0}^{\infty} \sigma_j^{(m)} z^j, \tag{D7}$$

then it is written explicitly as

$$S_m(z) = \frac{-z}{\log(1-z)} (1-z)^m. \tag{D8}$$

The case $m = -1$ corresponds to the explicit formula, namely the Adams-Bashforth formulas. While the case $m = 0$ does to the implicit formula, i.e. the Adams-Moulton formulas (Hairer *et al.* 1993). To obtain the σ-coefficients explicitly, we again recommend the usage of *Mathematica* or other formula processors. After the σ-coefficients are obtained, the β-coefficients are computed as

$$\beta_k^{(m,\ell)} = (-1)^k \sum_{j=k}^{\ell} \left(\begin{array}{c} j \\ k \end{array} \right) \sigma_j^{(m)} \tag{D9}$$

E. Symmetric Running Formulas of Adams Methods

Consider a situation where β-coefficients are symmetric. In this case, $\ell = 2m+1$ and

$$\beta_k^{(m,2m+1)} = \beta_{2m+1-k}^{(m,2m+1)}, \qquad (k = 0, \cdots, m) \tag{E1}$$

and therefore, Eq.(D2) becomes

$$y_{n+1} = y_n + h \sum_{j=0}^{m} \beta_{m-j}^{(m,2m+1)} \left(f_{n+1+j} + f_{n-j} \right) \tag{E2}$$

Of course, the symmetric β-coefficients are calculated by the formulation given in the previous section. However, by using the symmetry, we can derive them in a different and smarter way.

Let's consider a following expansion

$$y_{n+1} = y_n + \frac{h}{2} \left(f_{n+1} + f_n \right) + h \sum_{j=0}^{\infty} \tau_j \nabla^{2j+1} \left(f_{n+2+j} - f_{n+j} \right). \tag{E3}$$

Then its truncation is easily shown to be of the same form as Eq.(E2). We rewrite this expansion form in the form of operators as

$$\sinh \frac{hD}{2} = \frac{hD}{2} \cosh \frac{hD}{2} + (hD) \sinh hD \sum_{j=0}^{\infty} \tau_j \left(2 \sinh \frac{hD}{2} \right)^{2j+1} \tag{E4}$$

Then the corresponding generating function is given as

$$T(z) \equiv \sum_{j=0}^{\infty} \tau_j z^j = \frac{\sqrt{\frac{z}{4+z}} - \sinh^{-1} \frac{\sqrt{z}}{2}}{2z \sinh^{-1} \frac{\sqrt{z}}{2}}$$

$$= \frac{-1}{24} + \frac{11}{1440} z - \frac{191}{120960} z^2 + \frac{2497}{7257600} z^3 - \frac{14797}{191600640} z^4 + \cdots \tag{E5}$$

High accuracy method for solving N-body problems

Nobuo ISHIBASHI
nobuo@info.kanagawa-u.ac.jp
Department of Information
and Computer Sciences
Kanagawa University
Tsuchiya, Hiratsuka,
Kanagawa, 259-1293, JAPAN

March 3, 1999

Abstract

We have solved trajectories of point mass by using high order B-spline approximation functions. This method is equivalent to arbitral order implicit differential method. The order of approximation function is $3 \sim 45$-th order in practical use. We have also estimated an upper limit of numerical (and truncation) errors. This method is useful in the case of stiff equations.

1 Introduction

In this paper, we propose a high accuracy method to solve the trajectories of particles with B-spline functions. This higher order B-spline approximation function (the term "B"means "basis") is made with linear summation of piecewise polynomial basis functions, and it is different from the well known cubic interpolation function. We use, as basis functions, k-th order piecewise polynomials ($k = 3 \sim 32$). By using *de Boor-Cox*'s recurrence formula, we can generate these basis functions with high accuracy and efficiency[1].

N-th particles are given by interacting with gravity The equation of motion is,

$$\ddot{\mathbf{r}}_i = -G\sum_{i\neq j} m_j \frac{\mathbf{r}_{ij}}{r_{ij}^3}, \qquad (i \neq j), \tag{1}$$

where, $\mathbf{r}_i$ is a trajectory of i-th particles, m_i is a math of particles, and

$$\begin{aligned} \mathbf{r}_{ij} &= \mathbf{r}_i - \mathbf{r}_j \\ r_{ij} &= \left\{ \sum_{l=x,y,z} (r_{i,l} - r_{j,l}) \right\}^{\frac{1}{2}}. \end{aligned}$$

$\ddot{\mathbf{r}}_i$ means second derivative of i-th trajectory of particles. The 3-rd derivative of equation of motions are

$$\frac{\mathrm{d}^3\mathbf{r}_i}{\mathrm{d}t^3} = -G\sum_{i\neq j} m_j \left\{ \frac{\dot{\mathbf{r}}_{ij}}{r_{ij}^3} - 3\frac{\mathbf{r}_{ij}\dot{r}_{ij}}{r_{ij}^4} \right\} \qquad (i \neq j), \tag{2}$$

where,

$$\dot{r}_{ij} = \frac{\mathbf{r}_{ij}\cdot\dot{\mathbf{r}}_{ij}}{r_{ij}}, \tag{3}$$

and 4-th derivative of equations are

$$\frac{\mathrm{d}^4\mathbf{r}_i}{\mathrm{d}t^4} = -G\sum_{i\neq j} m_j \left\{ \frac{\ddot{\mathbf{r}}_{ij}}{r_{ij}^3} - 6\frac{\dot{\mathbf{r}}_{ij}\dot{r}_{ij}}{r_{ij}^4} - 3\frac{\mathbf{r}_{ij}\ddot{r}_{ij}}{r_{ij}^4} + 12\frac{\mathbf{r}_{ij}\dot{r}_{ij}^2}{r_{ij}^5} \right\} \qquad (i =\neq j), \tag{4}$$

where,

$$\ddot{r}_{ij} = \frac{\mathbf{r}_{ij}\cdot\ddot{\mathbf{r}}_{ij}}{r_{ij}}, \tag{5}$$

and as so on. $\frac{\mathrm{d}^k\mathbf{r}_i}{\mathrm{d}t^k}$ are also existing. Especially in the case that r_{ij} decreases, the value of $\frac{\mathrm{d}^k\mathbf{r}_i}{\mathrm{d}t^k}$ increase. If we require high accuracy numerical solutions, it is

KEYWORDS: B-Spline, ordinary differential equations, collocation, high accuracy

necessary to use higher order approximation functions. We can obtain approximation function that agree with $\frac{\mathrm{d}^k \mathbf{r}_i}{\mathrm{d}t^k}$ by using $k-1$-th order spline approximation functions.

2 Spline approximation functions

We define the approximation function for each i-th trajectory $\mathbf{r}_i(t)$ of particles by n spline basis functions $B_{j\{il\},k}(l = x, y, z)$ with $k-1$-th order, and n expansion coefficients $C_{j\{il\}}$,

$$\hat{r}_{\{il\}}(t) \equiv \sum_{j=1}^{n} B_{j\{il\},k} C_{j\{il\}} \qquad i = 1, \cdots, N \quad l = x, y, z. \tag{6}$$

(We write $B_{n,k}$ as B_n after this.) $C_{j\{il\}}$ is obtained uniquely by giving n conditions for trajectory of i-th particles for each x, y, z. In matrix form, at some discrete time t_i,

$$\begin{pmatrix} B_{1\{il\},k}(t_1) & B_{2\{il\},k}(t_1) & \cdots & B_{n,\{il\},k}(t_1) \\ \vdots & & & \\ B_{1\{il\},k}(t_n) & B_{2\{il\},k}(t_n) & \cdots & B_{n,\{il\},k}(t_n) \end{pmatrix} \begin{pmatrix} C_{1,\{il\},k} \\ \vdots \\ C_{n,\{il\},k} \end{pmatrix} = \begin{pmatrix} \hat{r}_{1\{il\},k}(t_1) \\ \vdots \\ \hat{r}_{n\{il\},k}(t_n) \end{pmatrix} \tag{7}$$

The derivative of trajectories by time is obtained by

$$\frac{\mathrm{d}^m \hat{r}_{il}(t)}{\mathrm{d}t^m} = \sum_{i=1}^{n} \frac{\mathrm{d}^m B_{j\{il\}}}{\mathrm{d}t^m} C_{j\{il\}} \qquad i = 1, \cdots, N \quad l = x, y, z \quad m \leq k. \tag{8}$$

From equation of motions (1), i-th approximation function $\hat{r}(t_n)$ at t_n should satisfy the following equations

$$F_{\{il\}}(t_n) = \ddot{\hat{\mathbf{r}}}_i(t_n) + G \sum_{i \neq j} m_j \frac{\hat{\mathbf{r}}(t_n)_{ij}}{\hat{r}_{ij}(t_n)^3} = \mathbf{0} \qquad i \neq j, \tag{9}$$

at all t_n.

In order to satisfy initial position $r_{\{il\}}(t = 0)$ and initial velocity $\dot{r}_{\{il\}}(t = 0)$, the following relation are required.

$$\sum_{j=1}^{n} B_{j\{il\}}(t = 0) C_{j\{il\}} = r_{j\{il\}}(t = 0) \qquad i = 1, \cdots, N, l = x, y, z, \tag{10}$$

and

$$\sum_{j=1}^{n} \dot{B}_{j\{il\}}(t = 0) C_{j\{il\}} = \dot{r}_{j\{il\}}(t = 0) \qquad i = 1, \cdots, N, l = x, y, z. \tag{11}$$

Therefore

$$F_{1\{il\}} = \sum_{j=1}^{n} B_{j\{il\}}(t=0)C_{j\{il\}} - r_{\{il\}}(t=0) = 0 \qquad i = 1,\cdots,N, l = x,y,z \quad (12)$$

$$F_{2\{il\}} = \sum_{j=1}^{n} \dot{B}_{j\{il\}}(t=0)C_{j\{il\}} - \dot{r}_{\{il\}}(t=0) = 0 \qquad i = 1,\cdots,N, l = x,y,z \quad (13)$$

should also be satisfied.

The derivative of **F** at t_n by $C_{j\{il\}}$ in equation (9) are,

$$J_{n\{il\}} = \frac{\partial F_{\{il\}}}{\partial C_{n\{il\}}} = \frac{\mathrm{d}F_{\{il\}}}{\mathrm{d}\hat{r}_{\{il\}}} \underbrace{\frac{\partial \hat{r}}{\partial C_{n\{il\}}}}_{=B_{n\{il\}}(t_n)} + \frac{\mathrm{d}F_{\{il\}}}{\mathrm{d}\ddot{\hat{r}}_{\{il\}}} \underbrace{\frac{\partial \ddot{\hat{r}}_{\{il\}}}{\partial C_{n\{il\}}}}_{=\ddot{B}_{n\{il\}}(t_n)} \qquad i = 3,\cdots,N, l = x,y,z. \quad (14)$$

We obtain expansion coefficient by multi dimensional Newton method with equation (14). In practice, Jacobi matrix for N-body system is

$$J_{n\{il\}} = \ddot{B}_{n\{il\}}(t_n) + G\sum_{i\neq j} m_j \left\{ \frac{1}{\hat{r}_{ij}(t_n)^3} - 3\frac{(\hat{r}_{il}(t_n) - \hat{r}_{jl}(t_n))^2}{\hat{r}_{ij}(t_n)^5} \right\} B_{j\{il\}} \quad i = 1,\cdots,N, l = x,y,z. \quad (15)$$

From equation(12),

$$J_{1\{il\}} = B_j(t=0), \quad i = 1,\cdots,N, \quad l = x,y,z \quad (16)$$

should be required. And from equation (13),

$$J_{2\{il\}} = \dot{B}_j(t=0), \quad i = 1,\cdots,N, \quad k = x,y,z \quad (17)$$

should be required to obtain initial conditions. We define $C^M_{j\{il\}}$ as expansion coefficients after M times iteration, then $M+1$-th coefficients are obtained by

$$\begin{pmatrix} C^{M+1}_{1\{il\}} \\ \vdots \\ C^{M+1}_{n\{il\}} \end{pmatrix} = \begin{pmatrix} C^{M}_{1\{il\}} \\ \vdots \\ C^{M}_{n\{il\}} \end{pmatrix} - J^{-1}_{j\{ij\}} \begin{pmatrix} F_{\{il\}}(t_1) \\ \vdots \\ F_{\{il\}}(t_n) \end{pmatrix}. \quad (18)$$

Then we can obtain expansion coefficients by iteration of

$$J_{j\{il\}} \begin{pmatrix} h_{1\{il\}} \\ \vdots \\ h_{n\{il\}} \end{pmatrix} = - \begin{pmatrix} F_{\{il\}}(t_1) \\ \vdots \\ F_{\{il\}}(t_n) \end{pmatrix}, \quad (19)$$

where,

$$C^{M+1}_{j\{il\}} - C^{M}_{j\{il\}} = h_{j\{il\}}$$

If h are converged, expansion coefficients are obtained.

Convergence of expansion coefficients depends on initial value of trajectories. In this case, we propose to use 2-nd order functions,

$$\hat{r}_{\{il\}}(t_n) = r_{\{il\}} + \dot{r}_{\{il\}} t_n + \frac{1}{2}\ddot{r}_{\{il\}} t_n^2. \tag{20}$$

We obtain expansion coefficients $C_{n\{il\}}$ for initial values by

$$\begin{pmatrix} B_{1\{il\}}(t_1) & B_{2\{il\}}(t_1) & \cdots & \\ B_{1\{il\}}(t_2) & & & \\ \vdots & & & \\ & \cdots & & B_{n\{il\}}(t_n) \end{pmatrix} \begin{pmatrix} C_{1\{il\}} \\ C_{2\{il\}} \\ \vdots \\ C_{n\{il\}} \end{pmatrix} = \begin{pmatrix} \hat{r}_{\{il\}}(t_1) \\ \hat{r}_{\{il\}}(t_2) \\ \vdots \\ \hat{r}_{\{il\}}(t_n) \end{pmatrix}. \tag{21}$$

If we use $k-1$-th order approximation function, the simultaneous convergence of $C_1, \cdots, C_n$ are not required. We should obtain first k-th expansion coefficients. We can obtain expansion coefficients one after another by using Gauss-seidel like iteration methods.

$B_{j,k}(t)$ are $(k-1)$-th piecewise polynomials, given by the iteration formula, called *de Boor-Cox*'s recurrence formula [1]

$$B_{j,k}(t) = \frac{t - q_j}{q_{j+k-1} - q_j} B_{j,k-1}(t) + \frac{q_{j+k} - t}{q_{j+k} - q_{j+1}} B_{j+1,k-1}(t) \quad (k > 1), \tag{22}$$

where $q_j (qJ \leq q_{j+1}$ are knot points which define the spline basis functions. The initial basis functions, corresponding to $k = 1$, are given by

$$B_{j,1}(t) = \begin{cases} 1 & (q_j \leq t < q_{j+1}), \\ 0 & (x < q_j, t \geq q_{j+1}). \end{cases} \tag{23}$$

Collocation points t_j are ordered as $t_n < t_{j+1}$ $(j = 1, \cdots, n-1)$ There is a freedom in taking the sequence of t_j and q_j subject to the Schoenberg-Whitney condition,

$$q_j < t_j < q_{j+k} \quad (j = 1, \cdots, n). \tag{24}$$

It is known that one can approximate the solution well when this condition is satisfied [1]. When the knot point q_j are monotonically increasing, a good approximation is obtained by choosing collocation points t_i according to

$$t_i = \frac{1}{k+1} \sum_{j=0}^{k} q_{i+j}. \tag{25}$$

This satisfies the Schoenberg-Whitney condition (24).

In this paper, we adopt equidistant point at the beginning, for example,

$$
\begin{aligned}
q_1 = q_2 = \cdots = q_k &= t_{min} - \frac{t_n - t_1}{(n-k+1)k}, \\
q_i &= \frac{t_n - t_1}{n-k+1}(i-k) + t_{min} \quad (i = k+1, n), \\
q_{n+1} = q_{n+2} = \cdots = q_{n+k} &= t_{max} + \frac{t_n - t_1}{(n-k+1)k},
\end{aligned} \tag{26}
$$

for fixed boundary problem, where t_{min} and t_{max} are the boundary points. We locate the first k knot points on this side of these boundaries, and k knot points on the other side of last boundary points. In this case, we can get k-th order extrapolation functions from t_{max} to $t_{max} + \frac{t_n - t_1}{(n-k+1)k}$. In this method, we can easily obtain the k-th order extrapolation function between the time t_{max} and $t_{max} + \frac{t_n - t_1}{(n-k+1)k}$, and this is very useful as initial condition after t_{max}. Some basis functions are shown in figure 1. Usually B is a band matrix with band width $k+1$, i.e. Fig.1

$$
\begin{aligned}
B_{1,k}(t) = B_{2,k}(t) = \cdots = B_{i-1,k}(t) = 0, \\
B_{i+k+2,k}(t) = B_{i+k+3,k}(t) = \cdots = B_{n,k}(t) = 0,
\end{aligned} \tag{27}
$$

where $q_i \le t \le q_{i+1}$.

Thus the approximation function can be calculated by the sum of products of $k+1$-th coefficients and basis functions:

$$
\hat{r}(t) = \sum_{i=left(t)}^{left(t)+k} B_{i,k}(t) \cdot C_i, \tag{28}
$$

where $left(t)$ stands for a function whose value is j if $q_j \le t < q_{j+1}$ is satisfied.

The partial derivative of the expansion of the approximant $\hat{r}(t)$ is given by

$$
\frac{\mathrm{d}\hat{r}(t)}{\mathrm{d}t} = \sum_{i=1}^{n} \frac{\mathrm{d}B_{i,k}(t)}{\mathrm{d}t} C_i. \tag{29}
$$

The derivatives of the basis functions of the k-th order iteration are given by the basis of the $k-1$-th iteration,

$$
\frac{\mathrm{d}B_{i,k}}{\mathrm{d}t} = (k-1)\left\{ \frac{B_{i,k-1}(t)}{q_{i+k-1} - q_i} - \frac{B_{i+1,k-1}(t)}{q_{i+k} - q_{i+1}} \right\}. \tag{30}
$$

Fig.2

As the same way, k-th derivative of basis function $\frac{\mathrm{d}^k B_{i,k}}{\mathrm{d}t^k}$

3 Truncation errors

The truncation errors is estimated by

$$
\|\hat{r} - r\| \le \left(\frac{\delta t}{2}\right)^{\max(k+1,l)} \left\{1 + \|\mathcal{L}^k\|\right\} \|r^{(k+1)}\|, \tag{31}
$$

where $\hat{r}$ is an approximation function and δt is a distance between neighboring knot points, l is a continuity of function $f \in C^l$, and $||\mathcal{L}^m||$ denotes the Lebesgue constant given in Ref. [19],[18]. $\mathcal{L} = 1 \sim 10$ in this cases. If there is no collision, $l = \infty$ in this case.

Although the spline approximation is higher order approximation, Runge's phenomena in Lagrange's approximation does not appear [6]. Because the spline approximation uses the basis with compact spatial support, the effects of locally violent points do not propagate immediately as far as the approximation function is determined by local rules. Since there is no need to make a uniform mesh, it is possible to concentrate the mesh points around a locally violent region.

4 Round off errors

From the above discussion, we see that the spline approximation has very small truncation errors and that these errors occur locally. It is not easy to estimate the round-off errors when the equation is written by a summation as in equation (6), but since the expansion coefficients are obtained by solving the matrix equation (7), we can estimate the round-off errors by analyzing these matrices.

We first calculate the condition number of matrix B, and then show how errors in the coefficients of matrix $\mathbf{C}$, which are caused by errors in either B or $\mathbf{f}$, are related to this number[20]. We choose to define a norm of a matrix A by

$$||\mathbf{A}||_2 = \max_i |\lambda_i|, \tag{32}$$

where λ_i^2 are the eigenvalues of the matrix $A^T A$.

$$||B||_2 \cdot ||B^{-1}||_2 = \frac{\max|\lambda_i|}{\min|\lambda_i|} \equiv p. \tag{33}$$

This is called the Von Neumann p-condition number. We take it as the definition of the condition number for matrix B.

We write numerical error of vector $\mathbf{h}$, ΔF, matrices ΔJ, round off error of evaluation functions $\delta\mathbf{F}$, and evaluate the relative error of expansion coefficients in equation (19)

$$(J + \Delta J)(\mathbf{h} + \Delta\mathbf{h}) = \mathbf{F} + \delta\mathbf{F}. \tag{34}$$

For any norm, we have by definition

$$||J|| \cdot ||\mathbf{h}|| \geq ||\mathbf{F}|| \tag{35}$$

From equation (34)

$$||\Delta\mathbf{h}|| \leq ||J^{-1}||||\Delta J||||\Delta\mathbf{h}|| + ||J^{-1}||||\Delta J||||\mathbf{h}|| + ||J^{-1}||||\delta\mathbf{F}||. \tag{36}$$

Hence, we have the inequality

$$\begin{aligned}\frac{\|\Delta \mathbf{h}\|}{\|\mathbf{h}\|} &\leq \frac{\|J\|\|J^{-1}\|}{1-\|J\|\|J^{-1}\|}\left(\frac{\|\Delta J\|}{\|J\|}+\frac{\|\delta \mathbf{F}\|}{\|J\|\|\mathbf{h}\|}\right) \\ &\leq \frac{\|J\|\|J^{-1}\|}{1-\|J\|\|J^{-1}\|}\left(\frac{\|\Delta J\|}{\|J\|}+\frac{\|\delta \mathbf{F}\|}{\|\mathbf{F}\|}\right) \\ &= \frac{p_J}{1-\varepsilon_J}\left(\frac{\|\Delta J\|}{\|J\|}+\frac{\|\delta \mathbf{F}\|}{\|\mathbf{F}\|}\right),\end{aligned} \tag{37}$$

where

$$\varepsilon_J = \|J^{-1}\|\|\Delta J\| = p_J\frac{\|\Delta J\|}{\|J\|}, \tag{38}$$

so that the growth of the relative error in $\mathbf{h}$ is bounded above by condition number of J times sum of the relative error in equation (32).

The relative error of vector $\mathbf{r}$, $\Delta\mathbf{r}$ is given by

$$(B+\Delta B)(\mathbf{C}+\Delta\mathbf{C}+\delta\mathbf{C}) = \mathbf{r}+\Delta\mathbf{r}. \tag{39}$$

Since $\|\Delta C\| = \|\Delta h\|$, and from equation (39), we obtain

$$\begin{aligned}\frac{\|\Delta \mathbf{r}\|}{\|\mathbf{r}\|} &\leq \left(\|B\|\|B^{-1}\|+\|\Delta h\|\|B^{-1}\|\right)\left(\frac{\|\Delta\mathbf{C}\|}{\|\mathbf{C}\|}+\frac{\|\delta\mathbf{C}\|}{\|\mathbf{C}\|}\right) \\ &= (p_B+\varepsilon_B)\left(\frac{\|\Delta\mathbf{C}\|}{\|\mathbf{C}\|}+\frac{\|\delta\mathbf{C}\|}{\|\mathbf{C}\|}\right)\end{aligned}. \tag{40}$$

From equation (37) we obtain

$$\frac{\|\Delta\mathbf{r}\|}{\|\mathbf{r}\|} \leq \frac{p_B p_J}{1-\varepsilon_J}\left(\frac{\|\Delta J\|}{\|J\|}+\frac{\|\delta\mathbf{F}\|}{\|\mathbf{F}\|}\right)+p_B\frac{\|\delta\mathbf{C}\|}{\|\mathbf{C}\|}+\varepsilon_B\left\{\frac{p_J}{1-\varepsilon_J}\left(\frac{\|\Delta J\|}{\|J\|}+\frac{\|\delta\mathbf{F}\|}{\|\mathbf{F}\|}\right)+\frac{\|\delta\mathbf{C}\|}{\|\mathbf{C}\|}\right\}. \tag{41}$$

where $p_B$is B a condition number of matrix B, $\varepsilon_B = p_B\frac{\|\Delta B\|}{\|B\|}$. $p_B p_J$ is a dominant value of numerical errors.

In de Boor-Cox's recurrence formula(22), the following relation are satisfied

$$\frac{t-q_j}{q_{j+k-1}-q_j} \leq 1, \quad \text{and} \quad \frac{q_{j+k}-t}{q_{j+k}-qj_{j+1}} \leq 1.$$

If $B_{j,k-n}(t)$ have some numerical error, it will be canceled while calculating basis functions, and the practical upper limit of round off error is $\sqrt{\frac{k(k+1)}{2}} \times \epsilon$, where ϵ is a computer epsilon. Thus the upper limit of $\|\Delta B\|$ is computer epsilon $\times k$ in practical use. In the case that both q_i and t are binary descriptive values, there are no errors in $B_{j,k}$, and $\|\Delta B\| = 0$. $\|\Delta\mathbf{h}\|$ is determined according to the judgment standard of newton method, and is a size level of the computer epsilon $\times$ size of matrix.

We show the dependence of the condition number on the number of knot points and on the order of the basis in fig. 3 and fig. 4. Since the condition number

increases with the order of the basis, the error of the expansion coefficients relative to the error in the approximation function will also increase with k. The condition number increases as $10^{(-0.133+0.231\times k)}$ for $N = 64$.

If we reuse double precision variables with IEEE format, the figure part have 52 bits and exponent part have 12 bits. Therefor the computer epsilon is $2^{-52} = 2.220E - 016$. Thus the maximum value for round-off errors are approximately $2.2 \times 10^{-16.133+0.231\times k}$.

Order of basis functions is important especially in high accuracy calculation.

For example, the round off error with $N = 64$, $k = 10$ is about 3.337×10^{-14}. From equation (31), the truncation error is about $const \times (1/2^7)^{k+1}$, where $const$ is $1 \sim 10$. The truncation error for $k = 10$ becomes $const * 6.62 \times 10^{-24}$. In this case, there is no needs to use high order approximation function with double precision.

The round off error with $N = 64$, $k = 3$, is about 8.0×10^{-16}. The truncation error is about $const \times 3.7 \times 10^{-9}$. In this case, higher order approximation function is useful with double precision.

Fig.3

Fig.4

5 Detection of errors

If we evaluate the equation (31) and equation (41), the upper bound of the error on collocation point can be calculated.

It is clear that $\frac{d^k r_{\{il\}}}{dt^k}$ grows as r_{ij} decreases. In other word, we only have to use δt properly according to the size of $\frac{d^2 r_{\{il\}}}{dt^2}$.

Division points $T_{L1}, T_{L2}, \cdots$, are determined so that the interval (or one unit) contains 64 basis functions. One can cope with the stiff equations by using more fine mesh point calculation where the second derivative are large as shown in table (5).

L1	T_1								T_2							
L2	T_1								T_2							
L3	T_1				T_2				T_3				T_4			
L4	T_1		T_2		T_3		T_4		T_5		T_6		T_7		T_8	
L5	T_1	T_2	T_3	T_4	T_5	T_6	T_7	T_8	T_9	T_{10}	T_{11}	T_{12}	T_{13}	T_{14}	T_{15}	T_{16}
⋮																

For instance, $\hat{r}_{\{jl\}}$ is by $\hat{r}_{\}il\}}$ and is not stiff and does not exist in stiff ($\frac{d^2}{dt^2}$ is small). The approximate value made from T_{L1} is used for $\hat{r}_{\{il\}}$. The thing to decide the approximation function made from T_{L3} can be done for $\hat{r}_{\{jl\}}$. (Naturally, it is similar concerning $\hat{r}_{\{ix\}}, \hat{r}_{\{iy\}}$ of $\hat{r}$.)

In the case of the sun (almost geostationary), the earth (cycle on year), the moon (cycle a month), and Pluto (cycle 248 years), it is necessary to calculate all in a usual method.

When this method is required, it is enough as shown in the following tables.

	cycle	Level	calculation points
moon	1month	L10	(1024) x 64
earth	1year	L6	(64) x 64
Pluto	248years	L1	64

Moreover, a vertical axis to the orbit side can almost be calculated enough with the calculation point of L_1.

References

[1] Carl de Boor, *A Practical Guide to spline* (Springer-Verlag,New York,1978).

[2] Paul de Faget de Casteljau, *Shape Mathematics and Cad* (Kogan Page, University of Metz, France, 1986).

[3] P. J. Barry and R. N. Goldman, Computer Aided design **3** 114(1988).

[4] M. Daniel and J. C. Daubisse, Computer Aided design **6** 121(1989).

[5] W. J. Goedheer and J. H. H. M. Potters, J. Comp. phys **61**,269(1985)

[6] William H. Press, Brian P. Flannery, Saul A. Teukolsky, and William T. Vetterling, *Numerical Recipes* (Cambridge University Press,Cambridge,1990).

[7] Thomas R. Lucas and George W. Reddien,JR, SIAM J. Numer. Anal **9** 341(1972).

[8] N. Ishibashi and K.Kitahara, J. Phys. Soc. Japan **61** 2795(1992).

[9] P.M.Prenter, *Spline and Variational Methods* (John Wiley & sons, New York,1989)

[10] A. M. Device, J. Comp. Phys **27** 123 (1978).

[11] J. Elschner, Numer. Math. **43** 265(1984).

[12] K. S. Thomas, Numer. Math. **36** 193(1981).

[13] G. Schmidt, Zeitschrift für Analysis und ihre Anwend ungen Bd. 3(4) 24(1984).

[14] *Spline Analyses*, M. H. Schultz (Prentice-Hall,New York,1973).

[15] Ichirou Kawakami, Masamitu Aizawa, Kiyomitu Suzuki, *Numerical Solution to Hyperbolic Systems of Conservation Laws in term of the Spline Functions* conference of numerical methods in Nagoya.

[16] P. C. Jain and B. L. Lohar, Comp. & Math. with Apples., **5** 179(1979).

[17] Gordon D. Smith, *Numerical Solution of Partial Differential Equations* (Clarendon Press,Oxford,1965).

[18] M. Reimer, Numerical Math. 44,417(1984).

[19] F. Richards, Journal of Approx. Theory **14**,83(1975).

[20] Ichirou Kawakami, *The basic theory of numerical analysis*, (Nihon University,Tokyo,1994).

[21] Joan R. Westlake, *A Handbook of Numerical Matrix Inversion and Solution of Linear Equations*, (Control Data Corporation, New York,1968).

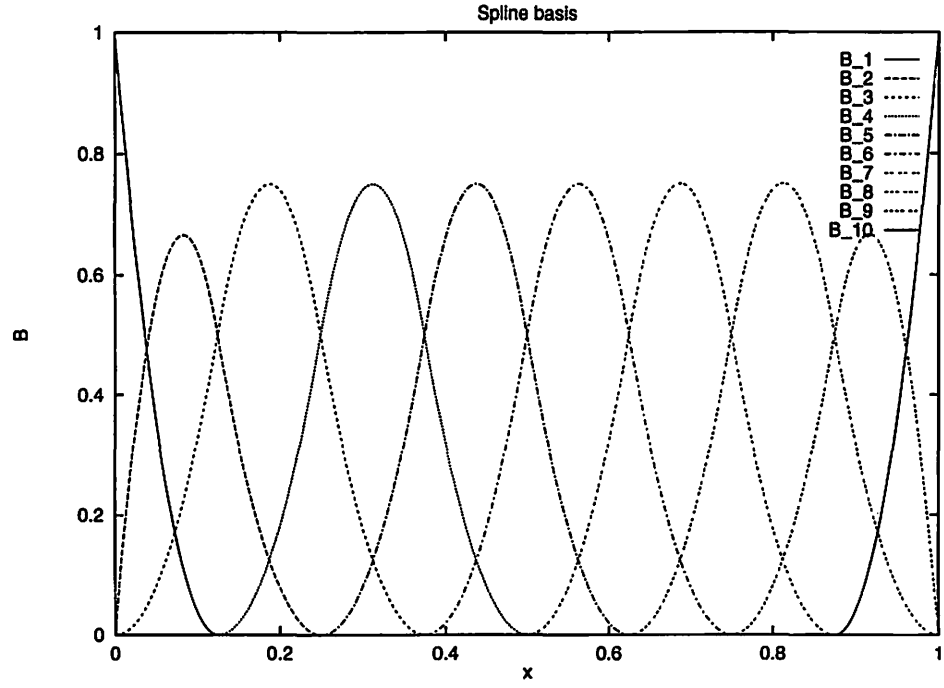


Figure 1: The spline basis functions with order 3. The number of the knot points is 10. Figure 1) shows the basis functions for fixed boundary problems. Spatial region is $[0,1]$.

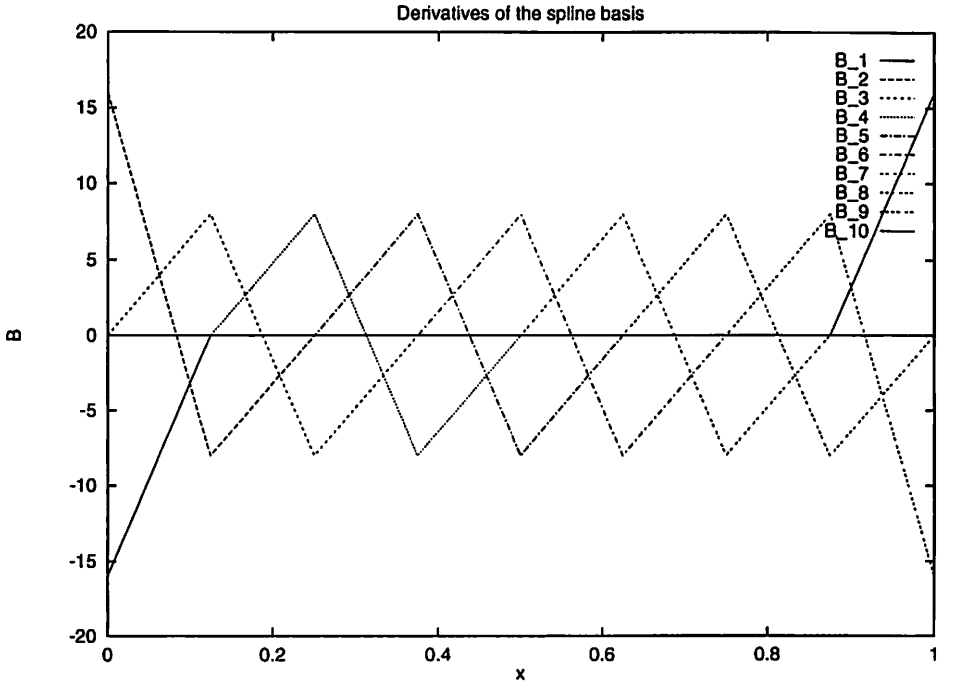


Figure 2: The derivatives of the 3rd order basis functions. Fig. 2 a) shows the derivatives for fixed boundary problems. The region of both systems is the same as in fig. 1.

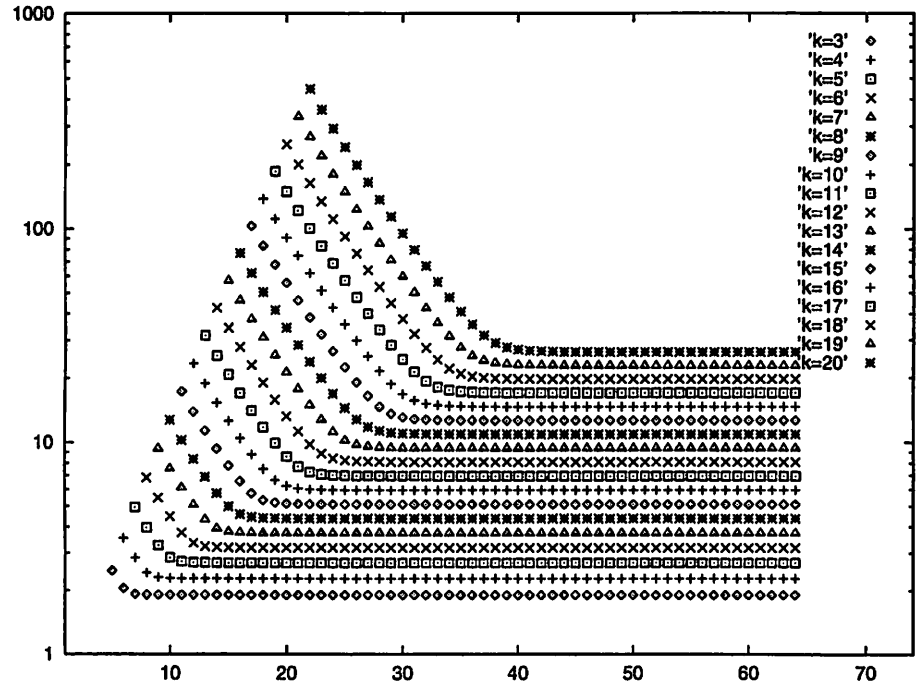


Figure 3: Dependence of the condition number p on the number of knots N. The condition number tends to some limit as N increases. The limit value increases when the order of the basis increases.

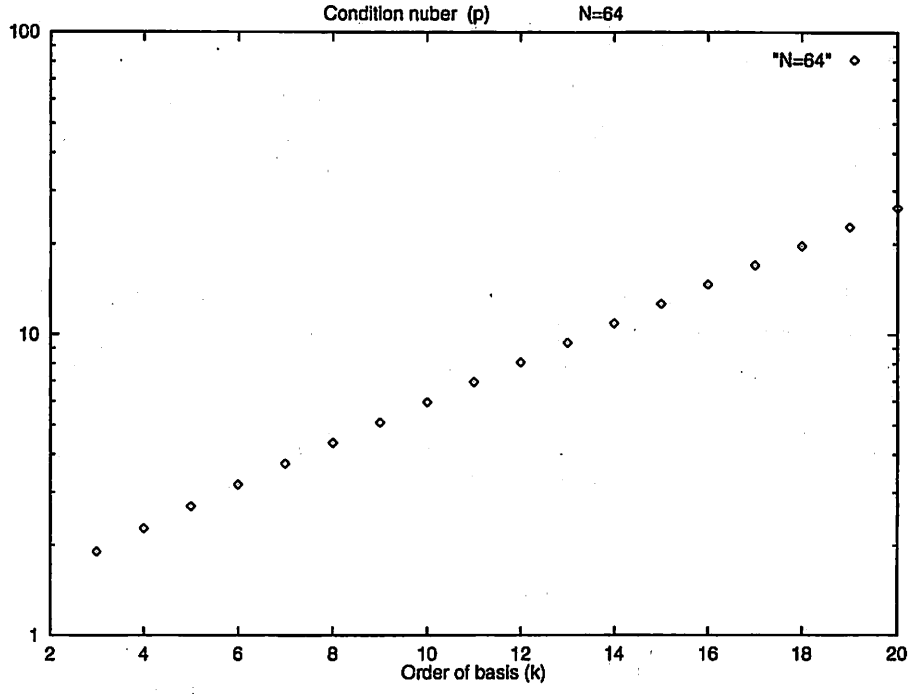


Figure 4: Dependence of the condition number on the order of the basis functions. The condition number increases when the order of the basis functions increases.

Relative Orbit Estimations for Close Geosynchronous Satellites

S. Kawase and F. Sawada

Communications Research Laboratory
Kashima, Ibaraki 314-0012 Japan
kawase@crl.go.jp / sawada@crl.go.jp

A new technique for monitoring the relative orbital motions of closely operating satellites is described. A radio-interferometer receives signals from two satellites so as to observe their differential line-of-sight directions, and a Kalman filter estimates the relative orbital state of the satellites. Estimation performance indicates a potential accuracy of 100m for relative satellite position, and this makes satellite near-miss warning and collision avoidance feasible.

Introduction

Due to the worldwide growth of satellite telecommunication activities, new launches of artificial satellites into geosynchronous orbits continue as illustrated in Fig. 1. The geosynchronous orbit is a unique existence, as it is a circular orbit with a particular height due above the earth's equator. As more and more satellites come into this orbit, spatial separations between satellites reduce. This tendency is seen in Fig. 2, where the occurrence of small separations between adjacent satellites as 0.1 degree or less has doubled in a few years.

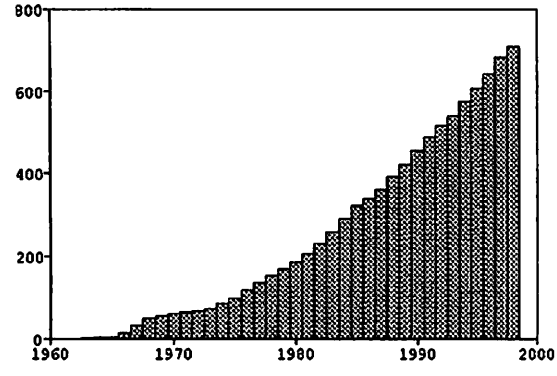


Fig.1 Geosynchronous launch, yearly cumulative

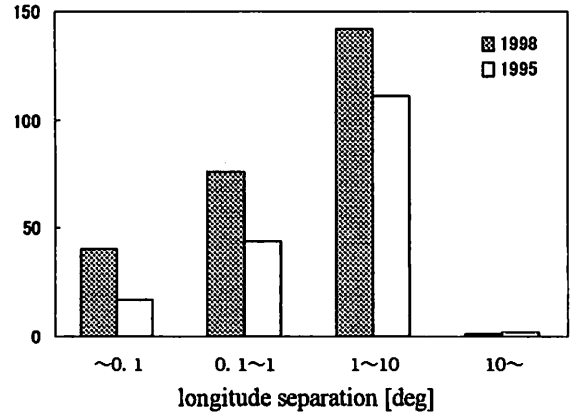


Fig.2 Small separations between stallites increase

Satellites operating closely to each other raise a potential risk of collision[1]. Though the actual collision probability is not high, any event of collision will generate a number of fragments which will then dissipate along the orbit to damage the environment of our precious orbit. So, any collision must be avoided, and this necessitates developing a technique for warning every possible near-miss between the satellites. Near-miss warning has to be accurate enough, since a poor accuracy must be compensated for by an enlarged error-margin of warning zone around each satellite and this would result in frequent triggering of unessential avoidance actions.

The present paper describes a newly developed technique for accurate near-miss warning. A radio-interferometer tracks two satellites so as to observe differential line-of-sight direction of the satellites, from which relative motions of the satellites are estimated. The relative motion estimation is accurate, because the differential observation eliminates any errors arising commonly for each satellite. The following sections describe the interferometer concept along with the relative orbit estimation filtering and its performance.

Differential Interferometer

The radio-interferometer has been set up at Kashima Space Communications Center, Communications Research Laboratory, as illustrated in Fig. 3. Three antennas of 1.2m diameter are directed to a group of TV satellites operating in 11GHz bands at 110 degree east of the geosynchronous orbit. The antenna baselines are arranged so that AB measures the azimuth and BC the elevation. Antenna sites are connected to a measuring site by cables for transmitting IF signals and distributing down-conversion reference signals, and in this respect our interferometer looks simply traditional.

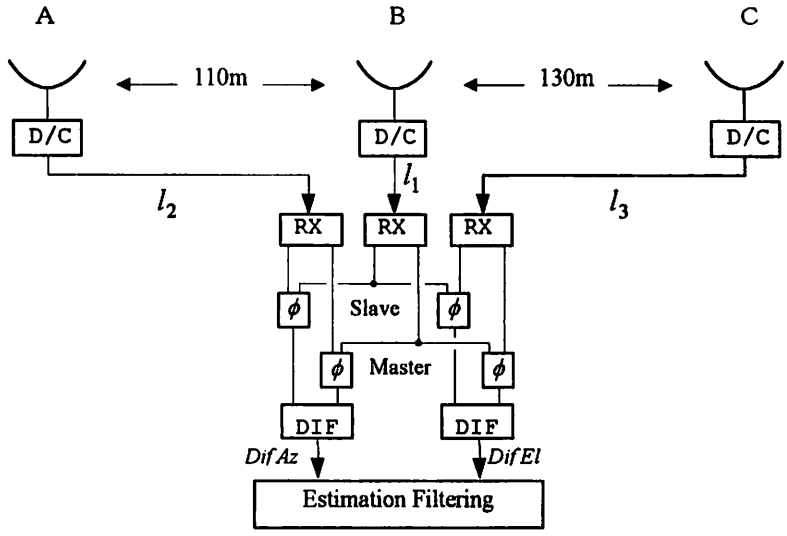


Fig. 3 Outline of radio-interferometer

What is particular with our interferometer is its capability of differential observation. At the measuring site, two beacon signals are detected separately as coming from two satellites. This choice of two is any out of the grouped satellites, and the chosen two are called "Master" and "Slave" satellites for convenience (which one is Master is arbitrary). Then we have three beacon signals of the Master satellite as coming through sites A, B, and C. We measure the phases of the A-site signal and of the C-site signal against the B-site signal. The same phase-measurement proceeds also for Slave satellite signal. Then evaluate the difference between the Master and Slave, for baseline AB and for BC. This provides, schematically speaking, differential azimuth and elevation observations of the two satellites, and the differential procedure will eliminate any observation errors that arise commonly for each satellite. Particularly important here is that the phase-measurement for the Master and that for the Slave take place strictly at the same moment by means of digital signal processing, and this proves essential for accurate differential observation as seen in

the next section.

Observation Data

Observation of satellite line-of-sight direction is subject to errors arising from atmospherical fluctuations. As a matter of fact, non-differential phase measurements with our azimuth baseline appear, for the Master and Slave, with fluctuating errors as shown in Fig. 4. By contrast, the differential data "diff", Slave minus Master, become smooth. Time scale is magnified in Fig. 5 to clarify the differential procedure; the observed data fluctuate from minute to minute almost identically for Master and for Slave. Owing to our interferometer design of measuring the phases simultaneously for two satellites, the common fluctuations are eliminated from the differential data. The same effect of differential observation operates also for the elevation baseline.

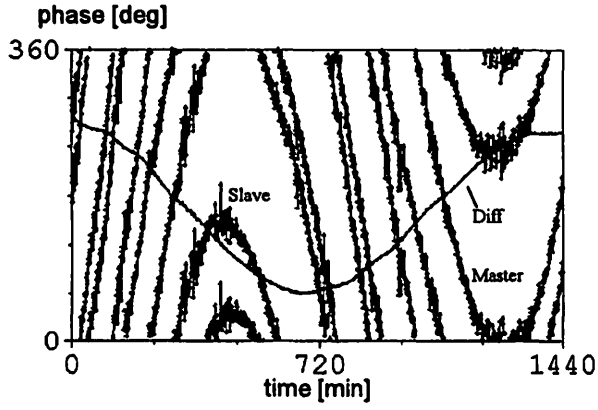


Fig.4 Non-differential /differential data

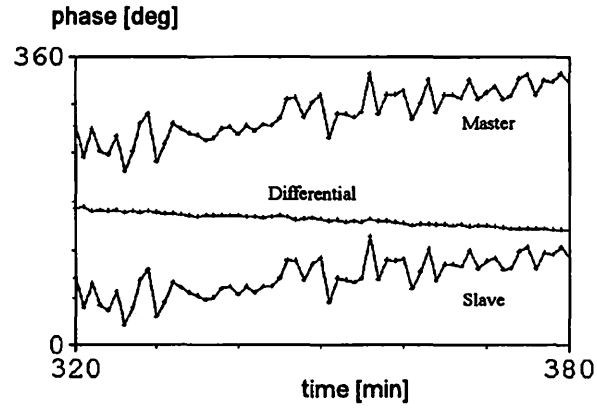


Fig.5 Differential effect

Observation Modeling

The path-length from a satellite to one antenna of an interferometer will differ from the path-length to another antenna, by a p, as illustrated in Fig. 6. Phase measurement for one satellite is then modeled as

$$\phi = (2\pi / \lambda)p + (2\pi / \lambda')(l_1 - l_2)$$

where λ and λ' are the wavelengths in a vacuum and in cables, and l_1 and l_2 are the lengths of IF transmission cables – so the second term is a "cable-term." The cable-term includes equivalently the delay in the antenna and receiving equipment. We have two satellites at different positions with different beacon wavelengths. Differential phase measurement, Slave minus Master, is then modeled, for the azimuth baseline, in the form of

$$\Delta_a\phi = \frac{2\pi}{\lambda_S} p_S - \frac{2\pi}{\lambda_M} p_M + \left(\frac{2\pi}{\lambda'_S} - \frac{2\pi}{\lambda'_M}\right)(l_1 - l_2) \qquad ..(1)$$

where subscript M and S are for Master and for Slave, and the cable lengths refer to Fig. 3. Also for the elevation baseline, we have

$$\Delta_e\phi = \frac{2\pi}{\lambda_S} q_S - \frac{2\pi}{\lambda_M} q_M + \left(\frac{2\pi}{\lambda'_S} - \frac{2\pi}{\lambda'_M}\right)(l_1 - l_3) \qquad ..(2)$$

with q's denoting the path-length differences for this baseline. We should set to

$l_1 = l_2 = l_3$, so as to make the cable terms in (1) and (2) vanish. Since we prepare identical antennas and equipment for the three sites, our problem is to make the lengths of IF cables equal. If the phase measurement should be accurate to $\delta\phi$, the three cable lengths must then be equal to a precision of δl, with

$$\delta\phi = \left(\frac{2\pi}{\lambda'_S} - \frac{2\pi}{\lambda'_M} \right) \delta l$$

In our experiment, $\delta\phi$ is one degree, the IF is in a 100MHz band, and two beacons differ by not more than 10MHz; hence δl is 8cm. Actually, cables were pre-cut to that precision of equal length before they were installed.

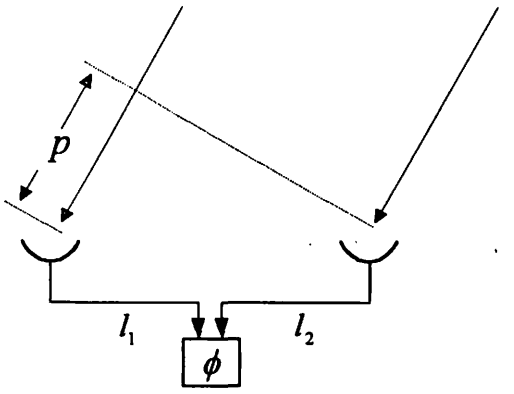


Fig.6 Phase measurement modeling

Relative Orbit Estimations

Two approaches are possible for estimating the relative satellite orbital motions. One is to assume that both satellite orbits are unknown. The other, which is simpler, assumes that the Master orbit is known and the Slave orbit is unknown. Our experiment chose the latter.

The Slave satellite's orbital motion is defined by inertial position and velocity and by the satellite's area/mass-ratio parameter; these are combined to make Slave satellite's state vector. An observed set of differential azimuth and elevation data makes an observation vector. A Kalman filter is then organized in an established procedure so as to update the state vector each time we obtain a new observation vector. The satellite dynamics in the filtering considers the solar/lunar gravity, the earth's potential, and the solar radiation pressure. The Slave's state is thus updated, while the Master's state is predicted from its assumed orbital elements. Thus we have a set of two orbital states: updated Slave's and predicted Master's, and this set provides the estimate of the relative orbital motion. The resultant estimate of relative motion is accurate as corresponding to the precision of the differential interferometer observations.

Estimation Performance

The relative orbit estimation performs as shown in Fig. 7. Parameter estimation for the solar pressure parameter (effective area/mass ratio) indicates the filter's convergence after one day of tracking. The O-C magnitude after the convergence corresponds to an accuracy of 80m of relative satellite position.

Relative position accuracy will go down when predicted, and so this is critical to our

near-miss warning. One can evaluate the degree of the accuracy reduction by switching the state-update on and off in the filtering, and this evaluation indicated that the relative position accuracy of around 100m is available after one-day's prediction. This accuracy of near-miss warning enables a collision avoidance strategy at a reasonable cost of satellite propellant consumption and reasonable extra operational workload[2].

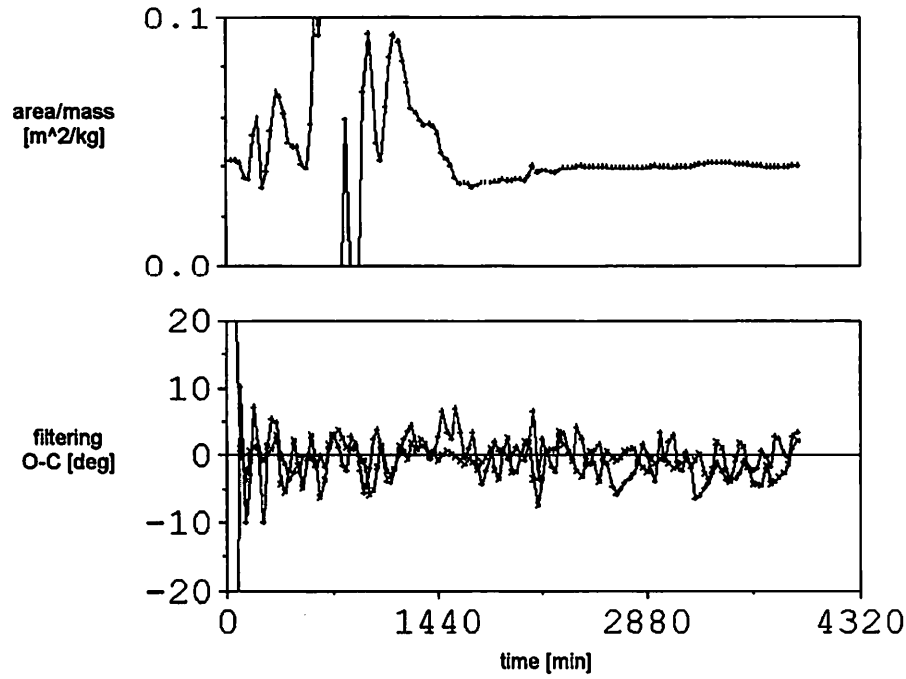


Fig.7 Relative orbit estimation performance

Summary

Relative orbit estimations using differential interferometer is a feasible approach to accurate near-miss warning and hence to collision avoidance for closely operating geosynchronous satellites. Choice of two satellites is arbitrary as long as the interferometer receives the satellite signals, so the present technique can be applied to any group of close *N* satellites in geosynchronous orbits.

References

[1] Cobotov, V. A. and Johnson, C. G.: Effects of Satellite bunching on the probability of collision in geosynchronous orbit, Jaournal of spacecraft and rockets, Vol.31, 1994, pp.895-899.

[2] Sawada, F. and Kawase, S.: Near-Miss Analysis for Uncoordinatedly Co-located Geosynchronous Satellites, 12th International Symposium on Space Flight Dynamics, Darmstadt, Germany, June 1997.

静止軌道近傍の光学観測

(Optical survey in a neighorhood of the geostationary orbit)

鷲尾智幸[1], 梅原広明[2], 木村和宏[2], 川瀬成一郎[2]

(Tomoyuki WASHIO[1], Hiroaki UMEHARA[2]*, Kazuhiro KIMURA[2], Seiichirou KAWASE[2])

[1] 豊橋技術科学大学 機械システム工学系

(Department of Mechanical Engineering, Toyohashi University of Technology)

Tempaku-cho, Toyohashi, Aichi 441-8580, Japan

[2] 郵政省通信総合研究所 鹿島宇宙通信センター

(Kashima Space Research Center, Communications Research Laboratory)

Hirai, Kashima, Ibaraki 314-0012, Japan (http://www.crl.go.jp/ka/)

Abstract. The space within a latitude ±0.2° of the geostationary orbit is surveyed between 68° and 190° east longitude by using a 35-cm reflecting telescope with a software detecting geostationary objects at Kashima Space Research Center. 67 satellites recognized by North American Air Defence Command (NORAD) are confirmed, and 2 satellites unrecognized by NORAD are found. There are also several objects crossing the viewing field. A low drift-rate object considered to be a space debris is followed for three days. By the orbital determination and its error estimation, it is found that at least two days are needed to monitor the motion of a near-geostationary debris.

1 イントロダクション

静止軌道上では, 衛星が増加し続け混雑の一途をたどっている. したがって, 衛星どうしの衝突や電波干渉を回避するため, 現在, 各衛星には静止軌道近傍において存在が許された「割り当て範囲」が決められている. 各衛星の運用が安全に行なわれるためには, 割り当て範囲を遵守して位置制御されているかを監視することや, 現在の衛星の分布を広域的に把握すること, 衛星近傍にスペースデブリが存在するか否かなどの観測を行なう必要がある.

本観測では, 光学望遠鏡を用いて, 静止軌道付近の物体の位置観測を行なった. なお, 他に, 衛星からの電波を受信し位置を測定する方法がある. 光学観測の長所は, 電波を発しないスペースデブリの探索に有効であることと, 大気揺らぎによる観測誤差が含まれないこと (背景星から位置を測定する場合), 観測方法が直接的であること (電波受信は間接観測といえる) である.

そこで, 静止軌道上およびその周辺にある天体のサーベイを行い, 静止軌道の混み具合の調査を行なった. また, スペースデブリの捜索と軌道決定を行なった.

2 観測装置の仕様

通信総合研究所宇宙制御技術研究室では, 1998 年 3 月に, 静止衛星などの人工天体を主な観測対象とした光学観測装置を開発した. 装置は, 光学系・架台部・撮像部・画像処理部から構成される. 観測の流れは以下の通りである. 所望の方向に望遠鏡を向け, 静止軌道近傍の物体を点像に, 恒星を線像に撮像するために, 望遠鏡を固定する (架台部). 恒星軌跡に時刻の情報を盛り込むため, 決められた時刻にメカニカルシャッターを開閉し, 恒星の日周運動による軌跡に刻印を入れながら撮像を行なう (撮

*Give any contact to Hiroaki UMEHARA (ume@crl.go.jp).

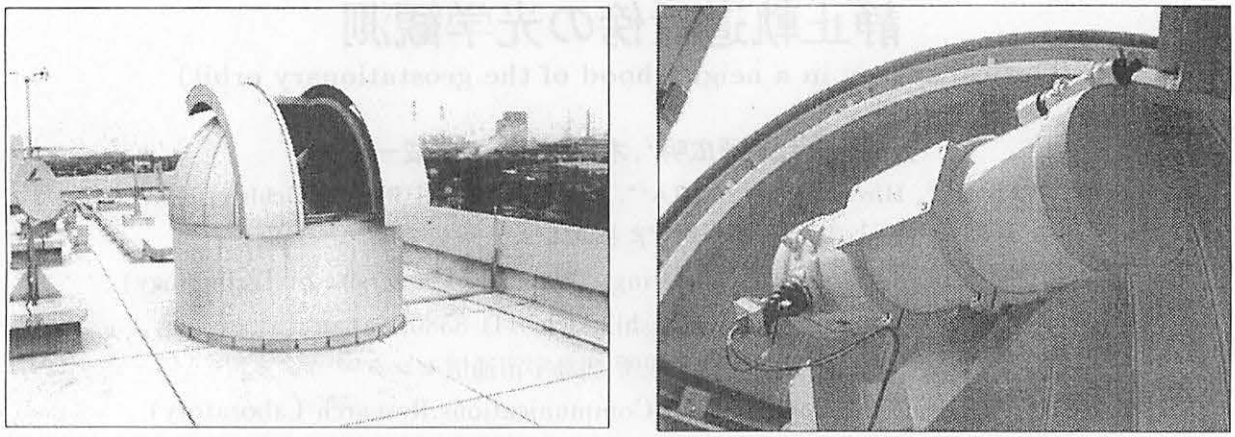


Figure 1: The telescope system at Kashima Research Space Center, Communication Research Laboratory.

像部). CCD 画像をコンピュータに取り込み, シャッター開閉パターンの情報などから恒星と静止天体の識別を行ない, 撮像時の概略方向と時刻およびガイドスターカタログ [1] により恒星を同定し, 静止天体の位置 (赤経・赤緯) を算出する (画像処理部). 一連の作業は, ほぼ自動化されている. 仕様の詳細は以下の通りである.

- 光学系:
  - 口径:350mm, 焦点距離:1248mm. 静止軌道上, 直径 1m 程度の物体を捕らえられる.
  - 視野:0.422deg × 0.634deg. 割り当て範囲 (0.2deg × 0.2deg) を一画面に捉えられる.
- 架台部: 中高度以上の移動衛星にも対応できるような駆動装置を有する.
- 撮像部:
  - CCD カメラ:160 万 (1536 × 1024) 画素. 撮像後ただちにデータ処理ができる規模とした.
  - シャッター精度 : 30ms 以下 (UTC 比).
- 画像処理部: ガイドスターカタログ [1] との比較により, 1/1000 ° の精度で物体の方向を同定することができる.

3 静止軌道サーベイ

1999 年 1 月 12 日から 37 日間, 東経 68 ° から 190 ° までの静止軌道をスキャンした. この区間は, 日本およびアジア各国の衛星が集中しており, 比較的軌道が密な部分である. 異なる画角で同一の物体を少なくとも二回捉えられるよう望遠鏡を 0.25 ° 刻みで動かし, そのつど露出 20 秒で撮影し衛星の有無を調べた. 存在した場合には位置を測定した. 図 2 に, サーベイした衛星位置を記す. 縦軸・横軸はそれぞれ, 衛星直下点の緯度・経度を表す. なお, 同じ図にスペースデブリの追跡結果を重ねて示しているが, これについては次節で詳述する.

視野内に捉えたデブリ以外の動く衛星は, 視野外に去るまで追跡し, 赤道に最も近かった点の緯度・経度をその衛星の静止軌道上における緯度・経度とした. 全ての衛星は必ず昇降点赤経 90 ° 付近 (オリオン星雲付近) を通過する. したがって, 日周運動をするオリオン星雲に追随するよう望遠鏡を指向すればいつかは写る. 軌道傾斜角の小さい静止衛星以外は, すべてこの方法で捕捉した. これは, デブリを捜索するときに効果的であると思われる.

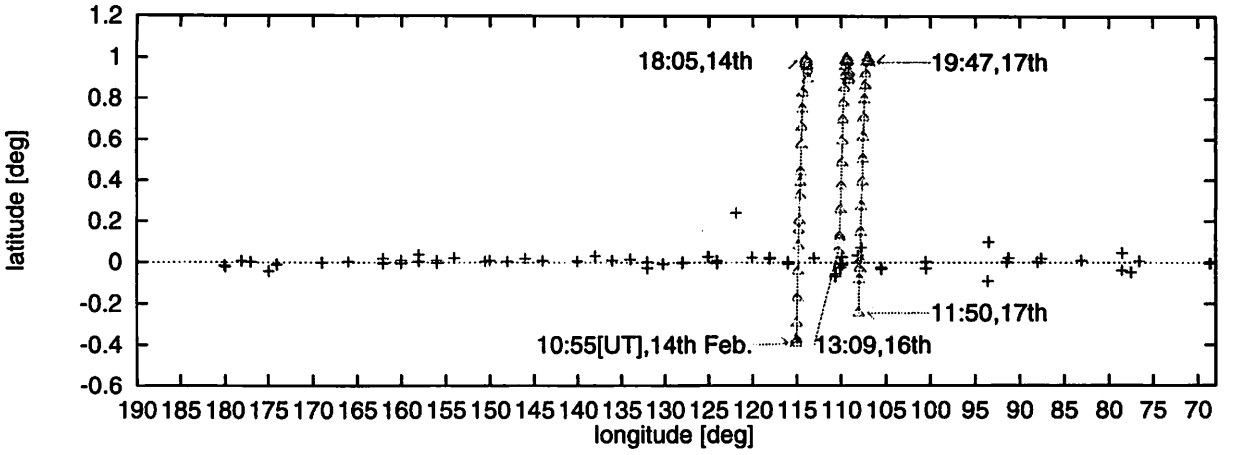


Figure 2: Distribution of objects around the geostationary orbit from 12 Jan. to 17 Feb. 1999.

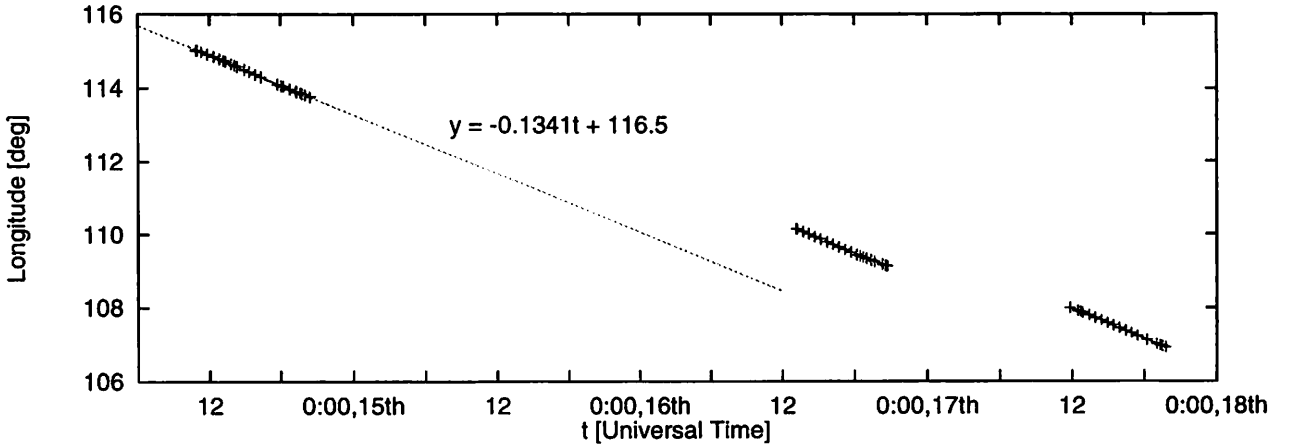


Figure 3: Time-dependent longitude of a debris on 14, 16, and 17 February 1999. The line segment is the result of fitting the plots sampled on the 14th.

観測区間内で69天体を補足した．ところどころ衛星どうし近接している部分や，逆に隙間が開いている部分があるが，全体的に見れば規則正しく並んでいる，といえる．

どの方向にどういった衛星が存在するかについては，北米航空宇宙防衛司令部が静止軌道リストをインターネット上に公開している [2]．そのリストと本観測データを照合して衛星の判別をした結果，東経175°および180°付近にはGEOリスト（1999.2.19現在）には載っていない衛星が1機ずつあった．何日か継続して観測したところ，これらの衛星は同一視野内に収まる程度の動きしかしていないため，軌道傾斜角、drift rate共に0.1以下と思われる．driftが大きいとデブリと見なされGEOリストから外されるが，少なくともこの2機はデブリとは言い難い．打ち上げて間も無いためにNORADが未だ把握していない静止衛星であると思われる．

4 スペースデブリの追跡

スペースデブリとは，人工衛星及びロケット運用上の廃棄物，任務を終えた人工衛星およびロケットそのもの，並びにそれらの破片・塗料片等，宇宙空間に存在する不要物の総称である．今回、偶然捉えた1つのデブリを延べ3日間に渡って追尾し，軌道決定および誤差評価を行なった．

2月14日10時54分40秒(UT)，望遠鏡を東経115° 00'，南緯0° 24'に指向していたとき，ある物体がCCD視野内に検出された．付近には軌道傾斜角の大きいCHINASAT 5という衛星があり，

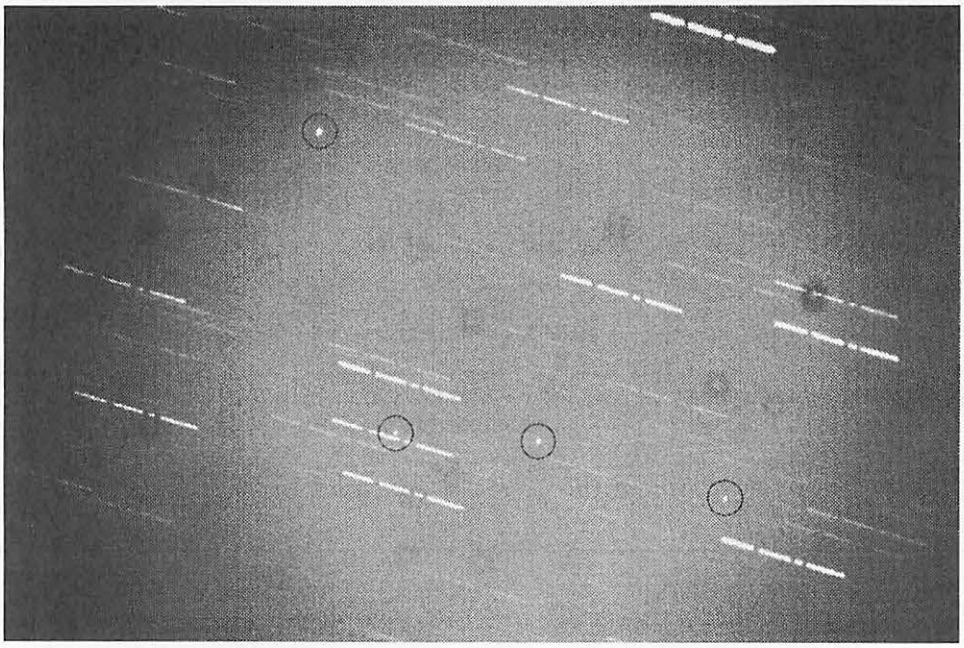


Figure 4: Telescope image around 110° E, 0° N at 13:13:30'(UT) with an exposure of 20 seconds on 16 February 1999. The three white dots in the lower part of the frame are B-SAT 1B, B-SAT 1A, and BS-3N (starting from the left). The upper dot is considered to be a debris since this object was ejected outside the frame 30 minutes later.

当初はそれであるかと思われた．しかし，軌道傾斜角に比して余りにも移動速度が遅いために，別の可能性があると判断し，付近を捜索したところ，ごく近傍の東経 115 ° 30' , 南緯 0 ° 24' に新たな衛星が 1 機見つかった。この衛星は、視野内の移動速度が CHINASAT 5 の軌道傾斜角からはじき出される速度におおよそ一致したので，CHINASAT 5 と断定した．また，GEO リストによると，東経 115 ° 近傍に CHINASAT 5 以外の動く静止衛星は存在しない．したがって，最初に捕捉した飛翔体はデブリ (以下，D と呼ぶ) であると判断した．

2 月 14 日は，CINASAT 5 と併せて約 30 分間隔で撮像し，夜明けで観測が不可能になるまで通算 9 時間半追尾した．その結果 (経度・緯度) を図 2に示す．また，経度の時間推移を図 3に示す．図 3を見ると，時間経過によって D の経度はほぼ直線的に変化している．経度が完全に時間に比例して変化しているのならば，D の軌道は真円軌道である．この場合，D は任務を終えた静止衛星の可能性がある．運用中の衛星の邪魔にならないよう，静止軌道より数百 km 程度高い位置へ軌道離脱 (de orbit) したため，地球局から見ると少しずつ西側へ移動して見える．グラフでは，僅かに三角関数状の膨らみがあるものの，おおよそ直線であり軌道は円軌道で近似可能であると判断した．経度 ($y[^\circ]$) の時間推移 (t[h]) を最小二乗法でフィッティングした直線は $y = -0.1341t + 116.5$ で与えられた．よって，次の日の同じ時刻には，西に $0.1341 \times 24 = 3.2^\circ$ ずれた場所に現れるはずである．

15 日は曇天の為，観測を延期した．16 日に上記の考え方のもと，21 時 50 分 (JST) に東経 108 ° 付近の赤道上空を通過するであろう，と予測をたて，21 時 30 分 (JST) 頃から望遠鏡を向けて待ち構えた．しかし，予定時刻になっても現れなかった．付近一帯を捜索したところ，東経 110 ° にそれらしきものが 1 機見つかった．図 4 がその CCD 画像である．本当に D かどうか一見しただけでは判別できないため，追跡し軌道決定をする．その間にも捜索を続けたが，結局この 1 機以外見つけ出すことはできなかった．その結果 (経度・緯度) を図 2に，経度の時間推移を図 3に重ねた．緯度・経度で表した軌跡は，14 日の D とほとんど同じ傾向を示し，図 3の近似曲線の傾きも 14 日とほぼ同じである．よっ

Table 1: Orbital elements of the debris found by us.

sample dates		t (UT)	a[km]	e	i[°]	Ω[°]	ω[°]	M[°]
(1)	14	14th, 0:00	42350.424	0.00151015	0.9630	88.4345	230.0255	301.0312
(2)	16	16th, 0:00	42351.658	0.00144334	0.9676	88.7131	230.2970	297.8603
(3)	17	17th, 0:00	42346.753	0.00150786	0.9689	88.8095	230.0681	296.6275
(4)	14,16	14th, 0:00	42345.822	0.00160761	0.9634	88.4332	229.9252	301.0944
(5)	14,16,17	14th, 0:00	42345.818	0.00160583	0.9634	88.4280	229.9453	301.0794

て, この時点で同一のデブリを追跡したと判断した.

17 日は, 14 日と 16 日の 48 時間で, 赤道上空を通過した経度がどれだけずれているかを算出し, 2 で割って 24 時間のずれを出して, 16 日の赤道通過経度から差し引いた. これによると, 1 日分のずれは 2.3 ° で, 翌 17 日には東経 107.90 ° の赤道上空を 21 時 42 分 (JST) に通過する. 誤差の分も含め予定時刻の 1 時間程前から観測を開始した. 観測結果を図 2, 3に示す. 赤道通過経度と時刻はおおよそ予測と一致し, 観測 3 日目にして D の動きを把握することに成功した.

これら 3 日間分の観測データを元に, 軌道決定を行なった. 軌道の決定とは, ある時刻 t における軌道要素 — 軌道長半径 (a), 離心率 (e), 軌道面傾斜角 (i), 昇降点赤経 (Ω), 近地点引数 (ω), 平均近点離角 (M) — を決めることである. 軌道決定の原理は, 最初にモデル的な値を仮定し, 摂動の影響を加味して観測値との差の二乗平均値が小さくなるよう, 反復計算を繰り返すものである.

最小二乗法により決定された軌道要素を表 1に示す。 1 日分の観測データのみを使って軌道決定をしたときは, 各々の日で a, e に無視できない大きさのばらつきが生じていることがわかる (表中 (1),(2),(3) を参照のこと). また, 基準時刻 t が 2 月 14 日 0 時 UT の軌道要素を見ると, 1 日分と 2 日分の観測データから求めた a, e には大きなずれを生じている (表中 (1),(4) を参照のこと). しかし, 同じ t = 2 月 14 日 0 時 UT の軌道要素でも, 2 日分と 3 日分の観測データから求めた a, e にはほとんど差がない. (表中 (4),(5) を参照のこと). ゆえに, 光学観測装置を用いて衛星, デブリの軌道決定を行う場合は, 2 日間データを取れば, かなりの精度で軌道が決まることがわかった. 最後に, 追跡二日目の観測において D が予測位置に現れなかったのは, 僅かな離心率 ($e \sim 1/1000$) の影響であることがわかった.

今後の課題に一つに, デブリを自動検出するアルゴリズムの先行開発 [3] を参考にしながら, 独自の機器開発を進め, 静止軌道近傍のデブリサーベイを進めることが挙げられる.

謝辞

システム開発に携わった株式会社エイ・イー・エス, 澤田史武氏 (通信総合研究所), 吉川真氏 (宇宙科学研究所) に感謝します.

参考文献

[1] Guide Star Catalog - Ver.1.1, *Space Telescope Science Institute*, 1992.

[2] GEO list, *North American Air Defence Command*, http://www.spacecom.af.mil/norad/

[3] Schildknecht, T., Hugentobler, U., Verdun, A.; Optical observations of space debris with the zimmerwald 1-meter telescope, *Advances in Space Research*, **19**[2], 221-228, 1997.

Clustering control of geostationary satellites using artificial interactions

Hiroaki UMEHARA
Kashima Space Research Center, Communications Research Laboratory
Hirai, Kashima, Ibaraki 314-0012, Japan (ume@crl.go.jp, http://www.crl.go.jp/ka/)

Abstract. The collective motion of geostationary satellites is searched for with artificial interactions of the Gaussian potential. These interactions are treated as a time-continuous maneuver. The potential parameters are found so that the five satellites maintain a cluster configuration in the tolerance region of the geostationary orbit.

1 Introduction

The number of geostationary satellites is greatly increasing although the available region is limited around the geostationary orbit. Therefore, station-keeping maneuvers are more complicated to prevent collisions, near-misses, etc. The linear theory has been used to develop optimal collision-avoidance strategies in a few satellite systems (Härting *et al.*, 1988; Eckstein *et al.*, 1989). Our aim is to search for an optimal procedure in a cluster system by using the mechanism of collective motion found in many-body systems.

We apply external forces with attractive interactions on five geostationary sattelites closely located to each other. These satellites are considered to be massless particles which are affected by the terrestrial gravitation and the solar radiation pressure. It is assumed that the external force is given by the Gaussian-potential function of mutual distances with several parameters. We will search for the parameters where the particles form a cluster in the tolerance window of the geostationary orbit.

The perturbations (the lunar-solar gravity and the earth non-sphericity) with the exception of the solar radiation are ignored since these forces accelerate the neighboring satellites in the same way and the perturbed orbits would be parallel to each other. The solar radiation pressures act on the satellites with various magnitudes causing complicated maneuvers since the effective cross-sections are different from each other.

Here, we investigate only the condition of spontaneously collective motion. In particular, we study the effect of attractive force on the motion. Our future work will be to search for impulsive collision-avoidance maneuvers which minimize the overall fuel consumption as well as operational efforts on the ground.

2 Simulation model

Let us describe the relative motion of each satellite with the nominal stationary position denoted by S_0. The positions of the five satellites are denoted by S_i, $i = 1,2,\ldots,5$. The rotational coordinate system with the vector from the earth to S_0 is introduced. The x-axis is directed radially outward from the earth. The y-axis is along the direction of motion and the z-axis is normal to the orbit plane of S_0. Let $\mathbf{q}_i = (x_i, y_i, z_i)^T$ be $\overrightarrow{S_0S_i}$. The equations of motion are given by:

$$\ddot{\mathbf{q}}_i = \begin{bmatrix} 0 & 2\omega & 0 \\ -2\omega & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \dot{\mathbf{q}}_i + \begin{bmatrix} 3\omega^2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -\omega^2 \end{bmatrix} \mathbf{q}_i + \mathbf{f}_i(t) - \frac{\partial U}{\partial \mathbf{q}_i}, \tag{1}$$

$$\mathbf{f}_i(t) = \sigma_i P \begin{bmatrix} -\cos\Omega t\ \cos\omega t - \cos\delta_0\ \sin\Omega t\ \sin\omega t \\ +\cos\Omega t\ \sin\omega t - \cos\delta_0\ \sin\Omega t\ \cos\omega t \\ -\sin\delta_0\ \sin\Omega t \end{bmatrix}, \tag{2}$$

$$U = -c\sum_{i}^{n-1}\sum_{j>i}^{n} a\exp(-a^2 r_{ij}^2), \quad r_{ij} = \overline{S_i S_j}, \tag{3}$$

where ω is the earth rotation angular velocity, Ω is the earth revolution angular velocity, δ_0 is the maximum declination of the sun, σ_i is the effective area/mass ratio of a satellite S_i, and P is the solar pressure constant (4.5×10^{-5}[N/m^2]). The time is initialized when the x-axis is directed to the vernal equinox. The perturbation of the solar radiation pressure is expressed in eq. (2). The last term of eq. (1) corresponds to the external force, which is considered as a maneuver and is given by the Gaussian-potential function U (see eq. (3)). It is noted that the linear part of eq. (1) corresponds to the Hill equation in the earth-S_0-S_i system. The analytic solution of eq. (1) without the last term is derived by Kawase (1989).

The parameters of individual satellites are fixed as $\sigma_1 = 0.20$, $\sigma_2 = 0.10$, $\sigma_3 = 0.05$, $\sigma_4 = 0.02$, and $\sigma_5 = 0.01$. The initial points stand at $S_1(0, 0.2L, 0)$, $S_2(0, 0.1L, 0)$, S_3(0,0,0), $S_4(0, -0.1L, 0)$, and $S_5(0, -0.2L, 0)$, where $L = 74$ km. The tolerance window is regulated as $\delta y = \delta z = 2L$. We assume the x-tolerance as $\delta x = 1.5L$.

3 Numerical survey

Figure 1(a) shows five-satellite orbits without control ($c = 0$). The three-dimensional orbits are projected onto the (x, y) plane. About two weeks later, S_1 is ejected from the tolerance window. Figure 1(b) represents the time-dependent moment of inertia $I(t)$ where the masses of the respective satellites are considered as a unit. The maximal cluster size increases with increasing time.

Next, the parameters (a, c) of the artificial force function which maintains the cluster configuration in the five-satellite system are searched for in Fig. 2. The duration of the cluster is long in the range:

$$c > 1, \qquad a \sim 1. \tag{4}$$

The above parameter region is suitable for the clustering control.

Our model is the time-dependent Hamiltonian system. The time-independent part of the Hamiltonian ($\mathbf{f_i} = \mathbf{0}$) is given by:

$$H = \sum_{i}^{n}\left[\frac{1}{2}p_i^2 + \omega(y_i u_i - x_i v_i) - \frac{\omega^2}{2}(3x_i^2 - q_i^2)\right] + U. \tag{5}$$

Here, $\mathbf{p}_i = (u_i, v_i, w_i)^T$ is the canonically conjugate momentum of $\mathbf{q}_i$. We find that the $|H|$-distribution is similar to the duration distribution in the (a, c) space.

Acknowledgement

The author would like to thank Dr. Seiichirou Kawase, Kazuhiro Kimura, and Fumitake Sawada for their helpful discussions.

References

Eckstein, M. C., Rajasingh, C. K., and Blumer, P., Colocation strategy and collision avoidance for the geostationary satellites at 19 degrees west, *The CNES International Symposium on Space Dynamics*, Toulouse (France), 6-10 Nov. 1989.

Härting, A., Eckstein, M. C., Leibold, A., Srinivasamurthy, K. N., On the collision hazard of colocated geostationary satellites, DFVLR FB 88-02, 1988.

Kawase, S., Real-time relative motion monitoring for co-located geostationary satellites, Journal of the Communications Research Laboratory, **36**, 125-135, 1989.

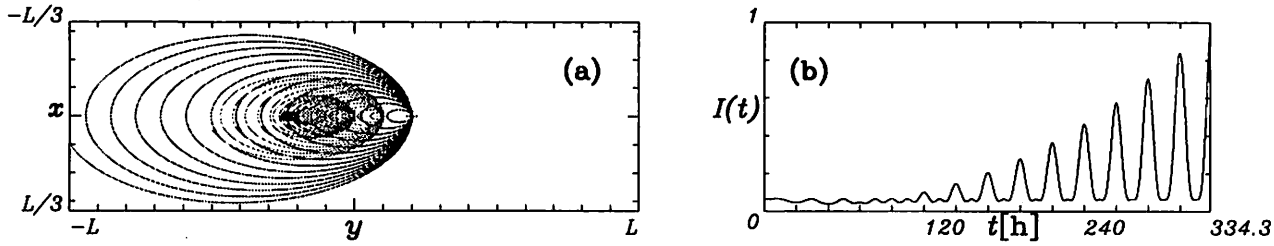


Figure 1: Five-satellite orbits without control (a) and time-dependent moment of inertia of the system (b).

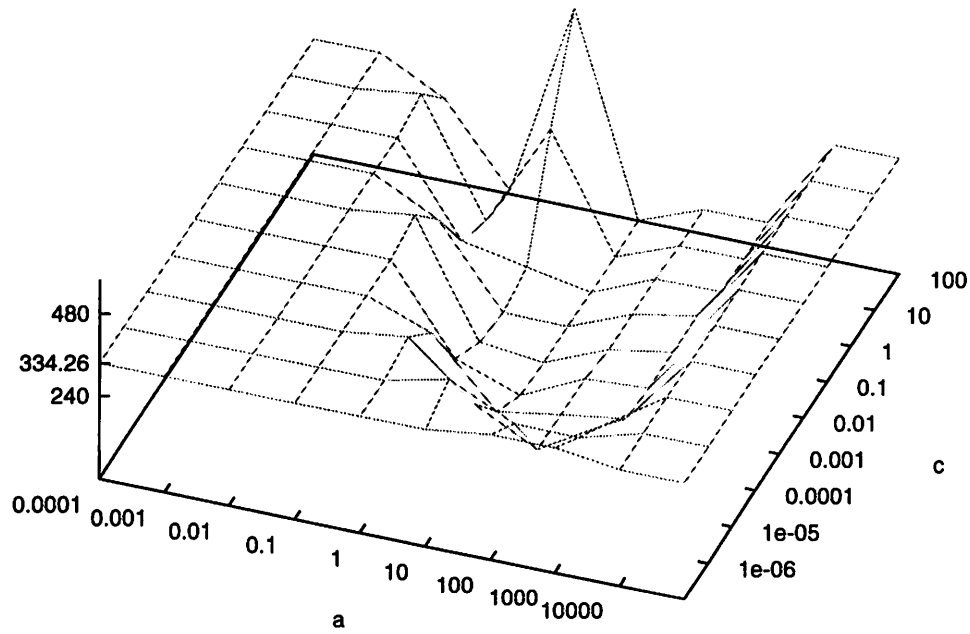


Figure 2: Duration of the cluster in the tolerance window.

静止衛星の軌道解析に関する研究－Ⅴ.
－観測距離と計算距離の残差と気象要素の関係－

河合雅司*

A Study on the Orbital Analysis of Geosynchronous Satellites -V.
- The Relation between Post-fit Range Residuals and Meteorological Elements -

Masashi KAWAI*

Abstract

It was tried to investigate the relations between meteorological elements and post-fit range residuals for discussing the causes of the 12-hour and 24-hour period fluctuations in post-fit range residuals in orbital analysis of geosynchronous satellites. Meteorological elements such as air pressure, air temperature and relative humidity were observed at meteorological stations around the tracking station. The results of the investigation are as follows.

(1) There is a correlation between 24-hour period fluctuations and relative humidity.
(2) In case that 12-hour period fluctuation appears in a clear shape, relative humidity doesn't change almost.

There is a correlation between the 12-hour period fluctuations and tidal generating forces, therefore it seems that the fluctuations are caused by tidal generating forces. It was verified that the 12-hour period fluctuations are explained by only thinking that the tracking station moves horizontally with earth crust. It is considered that there are horizontal flows of asthenosphere which are generated by tidal generating forces, and the earth crusts moves horizontally with the flows in the mantle such as ocean currents and tidal currents in ocean.

It is considered that 24-hour period fluctuations in post-fit range residuals are due to relative humidity and the observed range to the geosynchronous satellite CS-3b increases about one meter by tropospheric refraction of a radio wave as relative humidity increases 20 percents. These results are reported in this paper.

1．緒言

電波で観測された静止衛星のデータを用いて軌道解析を行うと軌道決定後の観測距離と計算距離の残差（O-C）に12時間周期又は24時間周期の変動が生じる。12時間周期の変動については、潮汐力と強い相関があることから、潮汐力によって衛星観測局が主として水平方向に変動していると考えることにより説明可能である。[1] しかし、O-Cの24時間周期の変動については潮汐力とは無関係であり、その原因についてはよく分かっていない。そこで、O-Cの周期変動の原因ついて調べるために、O-Cの周期変動と衛星観測局周辺の気象要素（気圧，相対湿度，気温，蒸気圧）との関係について調べた。その結果、O-Cに12時間周期の変動がきれいな形で表れる場合は相対湿度がほとんど変化していないこと、及びO-Cの24時間周期の変動と相対湿度の変動との間に相関があることが明らかになった。尚、筆者の論文”静止衛星の軌道解析に関する研究－Ⅳ”[2]における潮汐力による地殻水平変動モデルの式に誤りがあることが分かったので、本論文では訂正させて頂いた。これらの結果について以下に報告する。

2．軌道計算システムの概要

人工衛星軌道計算システムのアルゴリズムについては、他の文献に詳しく報告されている[3]ので、ここでは、本研究で使用された軌道解析システム TOCS(Toyama Orbit Computation System) の衛星力学モデル等について簡単に紹介する。

[1] 衛星力学モデル

衛星運動方程式はJ2000.0赤道面座標系（慣性座標系）において次の力を考慮して積分した。

(1)地球万有引力による加速度
(2)地球重力乱れポテンシャルによる加速度（重力モデル：GEM-10B[4]　4次で打ち切り）
(3)月，太陽，惑星の引力による加速度（天体歴：DE200/LE200 JPL ephemerides [5]）
(4)太陽輻射圧による加速度
(5)地球潮汐による加速度（ラプモデル[6]，Lag Angleは2.5°，ラプ定数は0.25）
(6)一般相対性理論による加速度[7]
(7)未知の力による加速度[8]（衛星速度ベクトルの方向に加速度を仮定）

[2] 電波伝搬補正モデル

電波で衛星までの距離を測定する時、電波が大気中を進む時にその速度が変化したり、屈折したりするために、観測距離が実際の距離よりも長くなる。これを補正するのが大気補正である。本研究では、電離層及び対流圏における補正を行った。電離層における補正については土屋モデル[9]を用い、対流圏における補正については GSFC NONAME式[10]を用いた。又、測角データについては角度の変化量が小さい（方位変化は0.05°未満、高度変化は0.1°未満）ので補正量は一定と考えて定誤差として処理し、大気屈折補正は行わなかった。本研究で用いた対流圏補正モデル（GSFC NONAME式）を次に示す。

$$\Delta R = \frac{8432.336 \cdot N_S}{0.026+\sin(E)} \ \mathrm{m} \qquad \text{ただし、} \Delta R_{MAX}=101.5\mathrm{m} \qquad (1)$$

$N_S=N_D+N_W$

ΔR：距離の補正量，　N_S：地表の屈折率（月毎の定数）
N_D：N_Sの乾燥分，　N_W：N_Sの湿潤分，　E：衛星仰角

(1)式におけるN_Sは、月ごとに定められた定数であり、衛星までの距離を観測した月が同じであれば、対流圏電波伝搬補正量ΔRは衛星仰角Eにのみ依存する。静止衛星の場合、衛星仰角はほぼ一定であるので、対流圏電波伝搬補正量は一定値になる。

Table 1　Refractivity on the ground (N_D:Dry Components, N_W:Wet components)

Jan.	$N_D=0.2851\times10^{-3}$, $N_W=0.02427\times10^{-3}$	Jul.	$N_D=0.2644\times10^{-3}$, $N_W=0.10950\times10^{-3}$
Feb.	$N_D=0.2845\times10^{-3}$, $N_W=0.02669\times10^{-3}$	Aug.	$N_D=0.2636\times10^{-3}$, $N_W=0.11310\times10^{-3}$
Mar.	$N_D=0.2817\times10^{-3}$, $N_W=0.03040\times10^{-3}$	Sep.	$N_D=0.2672\times10^{-3}$, $N_W=0.09391\times10^{-3}$
Apr.	$N_D=0.2765\times10^{-3}$, $N_W=0.04572\times10^{-3}$	Oct.	$N_D=0.2735\times10^{-3}$, $N_W=0.06696\times10^{-3}$
May	$N_D=0.2721\times10^{-3}$, $N_W=0.06421\times10^{-3}$	Nov.	$N_D=0.2784\times10^{-3}$, $N_W=0.04768\times10^{-3}$
Jun.	$N_D=0.2680\times10^{-3}$, $N_W=0.08468\times10^{-3}$	Dec.	$N_D=0.2829\times10^{-3}$, $N_W=0.03106\times10^{-3}$

[3] 地球潮汐による観測局位置の変動

地球潮汐による観測局位置の変動については、IERS Standards(1992)に紹介されているWahrの理論によるモデル（２次の潮汐まで考慮）を用いて補正を行った。

[4] 潮汐力による地殻の水平変動

月や太陽等の潮汐力により地殻が水平方向に変動していると考えられる。この地殻の水平方向の変動について本研究では次の様に考えた。

(1) 日本周辺における地殻の水平変動方向（θ_D）は315°。
(2) 水平変動の振幅（$D(T_F)$）は月と太陽の潮汐力（T_F）に比例している。

(3) 水平変動の周期は約12時間であり、その変動の位相は観測局における潮汐力の位相に対して、約２～４時間（角度で約30°～60°）遅れている。
(4) 地殻の周期変動の振幅は大潮の場合（潮汐力が大きい場合）で2m程度、小潮の場合（潮汐力が小さい場合）で1m程度である。

以上を考慮し、Table 2.に示された衛星観測データの解析において O-Cにおける12時間周期の変動がうまく除去出来る様にモデルを作った。（静止衛星の軌道解析に関する研究－Ⅳ．で示したモデルには誤りがあったのでここではそれを訂正した。）そのモデルを以下に示す。

$$\Delta X=A\cdot\sin\theta_D\ ,\quad \Delta Y=A\cdot\cos\theta_D \qquad (2)$$

$$A=D(T_F)\cdot\cos[2\cdot(L_O-L_D-\alpha_M)]$$

$$D(T_F)=C\cdot T_F=1.60\cdot T_F$$

$$L_D=18^\circ\cdot T_F+25^\circ$$

$$T_F=T_{FM}+\cos(2\cdot\theta_{MES})\cdot T_{FS}$$

$$T_{FM}=GM_M[1/(D_M-A_E)^2-1/D_M^2]\times10^6$$

$$T_{FS}=GM_S[1/(D_S-A_E)^2-1/D_S^2]\times10^6$$

ΔX:地殻変動量の東西方向成分（東が＋，西が－）[単位:m]
ΔY:地殻変動量の南北方向成分（北が＋，南が－）[単位:m]
A:地殻変動量[単位:m]
θ_D:地殻変動の方向（θ_D：北を0°として時計回りに360°まで測る。地域により値は異なり、日本周辺ではθ_D=315°と仮定）
L_O:観測局の経度
L_D:地殻変動の潮汐力に対する位相遅れ角度
θ_{MES}:地球から月へのベクトルと地球から太陽へのベクトルがなす角
$D(T_F)$:地殻変動の振幅[単位:m]（係数C は地域により異なった値をとると考えられる。本研究では日本周辺での値としてC=1.60を用いる。）
α_M:月の地位の経度（地球固定座標系における月位置の経度）
T_{FM}:月による潮汐力に10^6を乗じた値[単位:m/s²]
T_{FS}:太陽による潮汐力に10^6を乗じた値[単位:m/s²]
T_F:月と太陽による潮汐力の大きさを表わすパラメタ[単位:m/s²]
G:万有引力定数，　M_M:月の質量，　M_S:太陽の質量
A_E:地球赤道半径，　D_M:地球と月の距離，　D_S:地球と太陽の距離

観測局の位置は地殻の周期変動により東西方向に ΔX、南北方向に ΔYだけ偏位しているので、上記の地殻変動モデルを用いて補正する。

３．解析データ

君津衛星管制センターで、1993年 6月～ 7月及び1994年 6月に観測された通信衛星(CS-3b) のデータを用いて軌道解析を行った。それらのデータをTable 2.に示す。Table 2.においてθは、観測期間における潮汐力の大きさを表す指標でθが0°又は180°に近い時は潮汐力が大きく、θが 90°に近い時は潮汐力が小さい。（θは観測期間のほぼ中央で太陽と月が地球においてなす角である。）　距離データの精度については、軌道決定後の O-Cの１時間毎の標準偏差で判断して、約0.4mである。角度データの精度は約0.002°（距離換算で約1000m）　である。各衛星の観測に使用された観測局アンテナ位置をTable 3.に示し、各観測局に月が正中する時刻（日本標準時）をTable 4.に示す。

又、本研究で用いた対流圏補正モデルでは、静止衛星の場合電波伝搬遅延量がほぼ一定となっている。しかし、実際には気象条件によって、対流圏における伝搬遅延量が相当変化している可能性が強いと思われるので、これを確かめるために君津衛星管制センター付近の気象データ（館山，勝浦，銚子，横浜の各観測所における気象庁地上気象観測データ）を使用した。

Table 2. CS-3b Tracking Data Observed at Kimitu
θ is the angle between the sun and the moon at the earth.

データ	データ観測期間 (JST)	θ
Cb1	1993年06月05日08時～1993年06月07日07時	162°
Cb2	1993年06月12日08時～1993年06月14日07時	82°
Cb3	1993年06月19日08時～1993年06月21日07時	2°
Cb4	1993年06月26日08時～1993年06月28日07時	90°
Cb5	1993年07月03日08時～1993年07月05日07時	177°
Cb6	1993年07月10日08時～1993年07月12日07時	101°
Cb7	1993年07月17日08時～1993年07月19日07時	20°
Cb8	1994年06月04日08時～1994年06月06日07時	48°
Cb9	1994年06月11日08時～1994年06月13日07時	30°
Cb10	1994年06月18日08時～1994年06月20日07時	118°

Table 3. The Locations of Tracking Stations
TSCJ-C2 is station for observing ranges of CS-3b at Kimitu.
TSCJ-ANG is station for observing angles of CS-3b at kimitu.

Station	Coordinate	Latitude	Longitude	Height
TSCJ-C2	GEM-10B	35° 11' 35.795"N	140° 04' 29.486"E	189.742m
TSCJ-ANG	GEM-10B	35° 11' 36.822"N	140° 04' 31.221"E	200.836m

Table 4. The Time of Transit of Moon across the Kimitu Meridian

Data	Time of Meridian Transit of the Moon (JST)
Cb1	1993. 6. 6. 1^H, 1993. 6. 7. 2^H
Cb2	1993. 6.13. 6^H, 1993. 6.14. 7^H
Cb3	1993. 6.19.11^H, 1993. 6.20.12^H
Cb4	1993. 6.26.17^H, 1993. 6.27.18^H
Cb5	1993. 7. 3.23^H, 1993. 7. 5. 0^H
Cb6	1993. 7.11. 5^H, 1993. 7.12. 5^H
Cb7	1993. 7.17.10^H, 1993. 7.18.11^H
Cb8	1994. 6. 4. 8^H, 1994. 6. 5. 8^H
Cb9	1994. 6.11.13^H, 1994. 6.12.14^H
Cb10	1994. 6.18.19^H, 1994. 6.19.20^H

4．解析結果

Table 2.に示された48時間アークのデータを用いて静止衛星の軌道を決定し、軌道決定後の観測距離と計算距離の残差 (O-C) を求めた。

距離データの重みを１、角度データを距離に換算した量の重みを0.0001として重み付き最小自乗法によりパラメタを推定した。同時に推定されたパラメータは次の通りである。

(1) 衛星６軌道要素 (X, Y, Z, V_x, V_y, V_z)
(2) 太陽輻射圧における反射係数
(3) 慣性座標系における衛星速度ベクトル方向の加速度
この加速度の符号は、速度ベクトルの方向が正、その逆方向が負。
(4) 定誤差
角度データ（高度，方位角）の定誤差。

観測距離とは観測局から衛星に搭載されたトランスポンダーまでを電波が往復するのに要した時間に光速度を乗じた値の 1/2であり、計算距離とは軌道解析理論に基づいて計算

された衛星の重心と観測局の距離である。トランスポンダーの位置と衛星重心とはずれているので観測距離に対してその補正を行う必要があるが、本研究ではこれを無視している。Table 2.に示された衛星観測データの解析結果をFig.1～Fig.6に示す。これらの図の横軸は日本標準時、縦軸は軌道決定後の O-Cの１時間毎の平均である。

Fig.1は、軌道決定後の O-Cに12時間周期の変動がきれいな形で表れた場合の一例である。これは、データ Cb3と Cb6の解析結果であり、 Cb3は潮汐力が大（大潮）の時のデータであり、 Cb6は潮汐力が小（小潮）の時のデータである。この図から月が観測局に正中してから2～4時間後に O-Cが最大になっており、12時間周期の変動の振幅については、潮汐力が大きい時は大きく、潮汐力が小さい時は小さくなっており、潮汐力が大きい時の振幅は潮汐力が小さい時の振幅の約２倍になっていることが分かる。更に、潮汐力による地殻水平変動モデルを用いて地殻水平変動の振幅を調べた所、潮汐力が大きい場合は2m程度、潮汐力が小さい場合（データCb6 の場合）は、約1mであった。

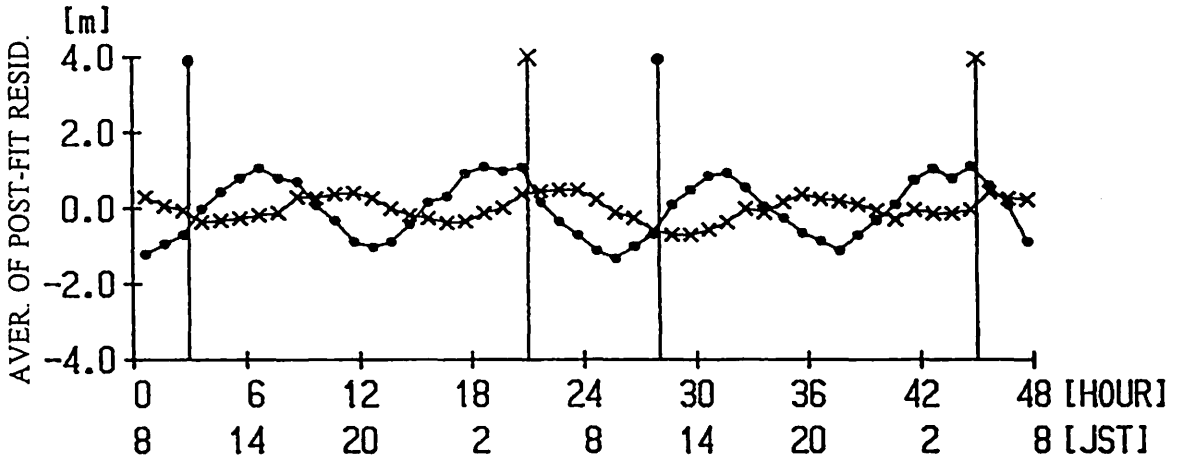


Fig.1 The averages of post-fit residuals of range data for every hour. Vertical line shows time of transit of the moon across the station meridian. The effect on O-C by horizontal movement of the earth crust is remained. θ is the angle between the sun and the moon at the earth.
・ : Data Cb3,RMS=0.9m, θ =2° ・× : Data Cb6,RMS=0.5m, θ =101°

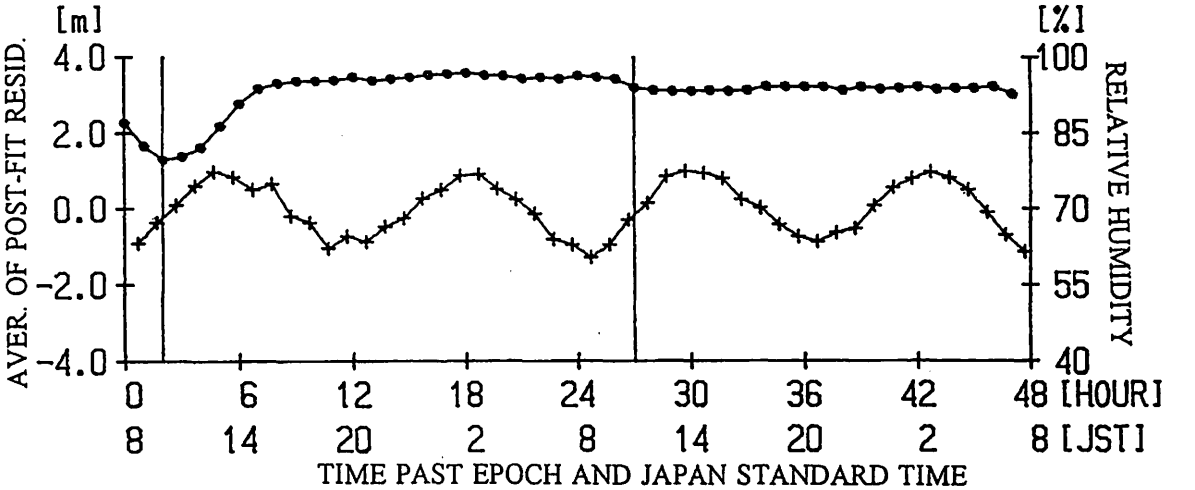


Fig.2 The averages of post-fit residuals of range data Cb7 for every hour. Vertical line shows time of transit of the moon across the station meridian. The effect on O-C by horizontal movement of the earth crust is remained. θ is the angle between the sun and the moon at the earth.
+ : Data Cb7,RMS=0.8m, θ = 20°
・ : Mean relative humidity around Kimitu Satellite Control Center

Fig.2～Fig.4は、O-Cと相対湿度の関係を示したものである。気象データ（相対湿度）については、君津衛星管制センターの周辺にある館山，勝浦，銚子，横浜の各観測所で1時間毎に観測された気象庁地上気象観測データ（相対湿度）を時間毎に平均したものである。Fig.2とFig.3は、12時間周期の変動がきれいな形で表れた場合（データ Cb7, Cb3）のO-Cであり、相対湿度はほとんど変化していない。又Fig.4はデータCb10 に関するものであり最初の18時間は相対湿度が約30% 変化しており、O-Cの周期の変動の形が崩れているが、18時間以後は相対湿度が安定しているため O-Cに12時間周期の変動がきれいな形で表れていることが確認出来る。データCb3, Cb6, Cb7 については、相対湿度の変化が少ないために、O-Cの12時間周期の変動をきれいな形で検出することが出来た。

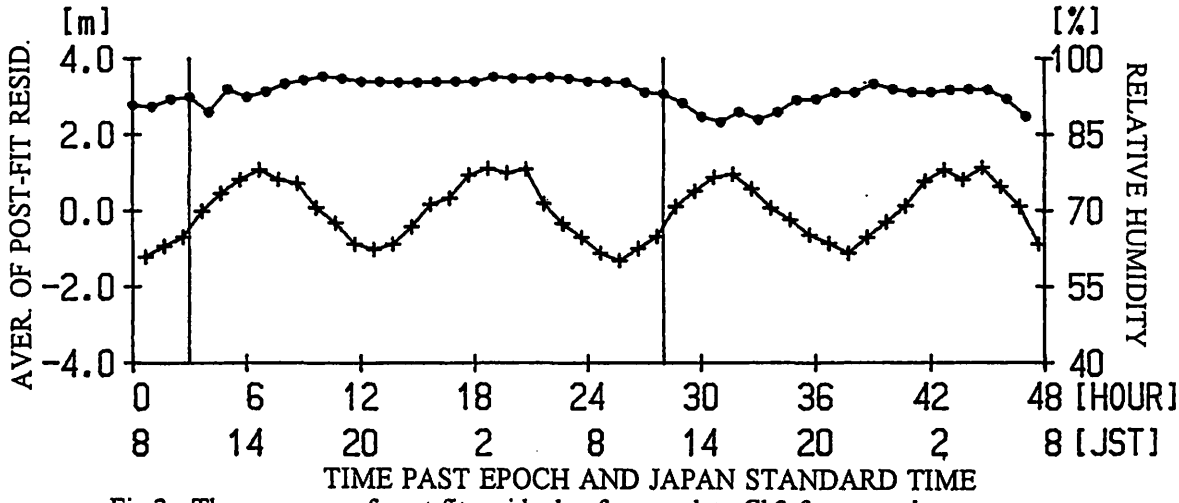


Fig.3 The averages of post-fit residuals of range data Cb3 for every hour. Vertical line shows time of transit of the moon across the station meridian. The effect on O-C by horizontal movement of the earth crust is remained. θ is the angle between the sun and the moon at the earth.
+ : Data Cb3,RMS=0.9m, $\theta = 2^{\circ}$
・ : Mean relative humidity around Kimitu Satellite Control Center

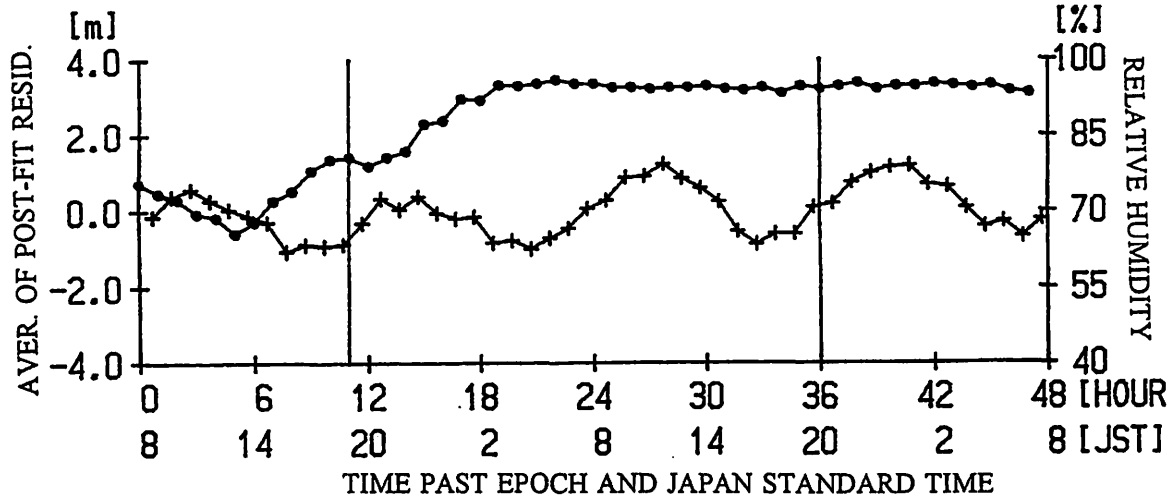


Fig.4 The averages of post-fit residuals of range data Cb10 for every hour. Vertical line shows time of transit of the moon across the station meridian. The effect on O-C by horizontal movement of the earth crust is remained. θ is the angle between the sun and the moon at the earth.
+ : Data Cb10,RMS=0.8m, $\theta = 118^{\circ}$
・ : Mean relative humidity around Kimitu Satellite Control Center

Fig.5, Fig.6は、 O-Cの12時間周期の変動と気圧及び気温（相対湿度の場合と同様に4カ所の気象庁観測所で観測された海面気圧及び気温の時間毎の平均）との関係を示したものである。これらの図から O-Cの12時間周期の変動と気圧及び気温の変動との間に相関はなく両者が無関係であることが確認出来る。

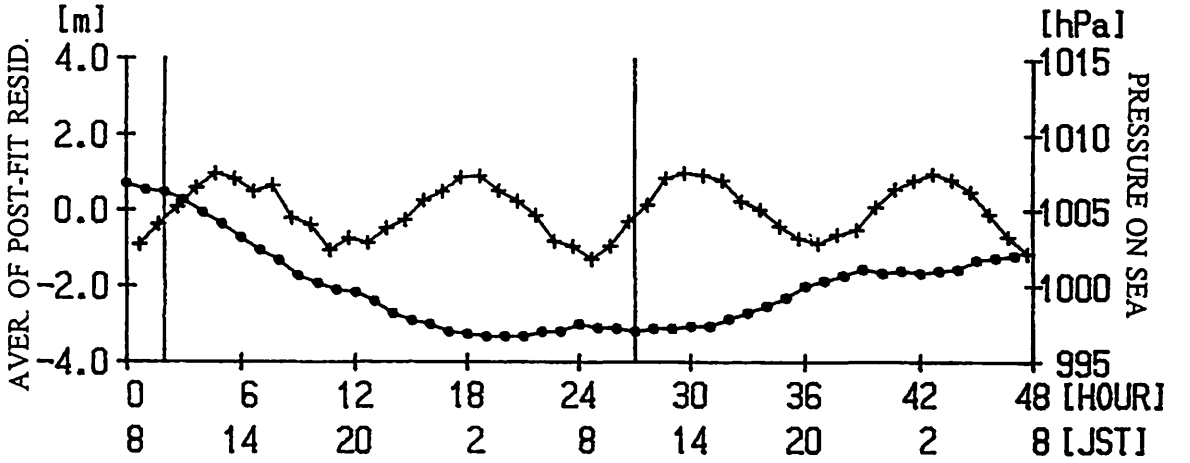


Fig.5 The averages of post-fit residuals of range data Cb7 for every hour. Vertical line shows time of transit of the moon across the station meridian. The effect on O-C by horizontal movement of the earth crust is remained. θ is the angle between the sun and the moon at the earth.
+ : Data Cb7,RMS=0.8m, θ =20°
・ : Mean pressure on sea surface around Kimitu Satellite Control Center

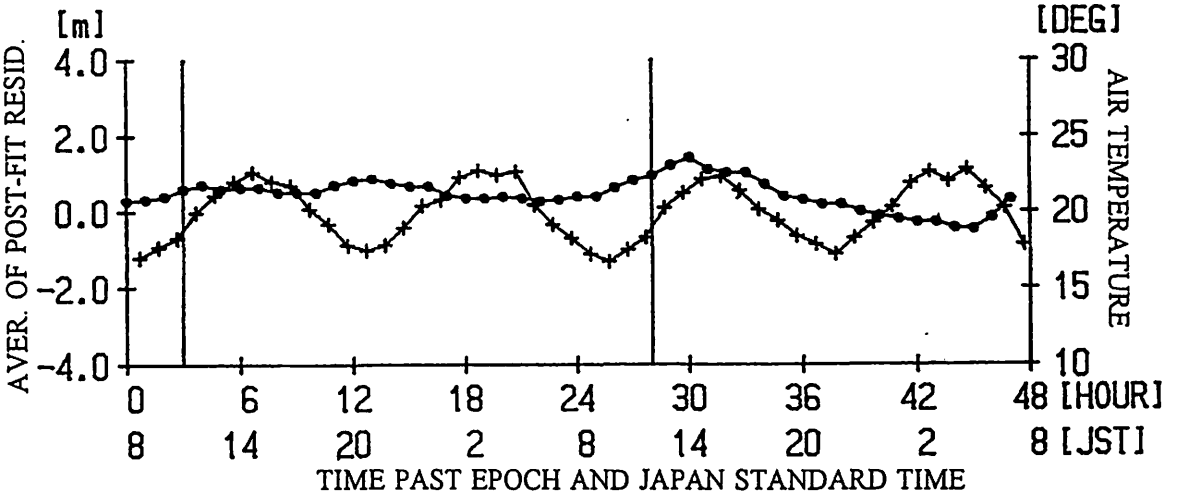


Fig.6 The averages of post-fit residuals of range data Cb3 for every hour. Vertical line shows time of transit of the moon across the station meridian. The effect on O-C by horizontal movement of the earth crust is remained. θ is the angle between the sun and the moon at the earth.
+ : Data Cb3,RMS=0.9m, θ = 2°
・ : Mean air temperature around Kimitu Satellite Control Center

Fig.7, Fig.8は、潮汐力による地殻水平変動モデルを用いて O-Cに存在する12時間周期の変動を除去した後に24時間周期の変動が存在する場合（データCb1, Cb2, Cb4, Cb5, Cb8 ）について O-Cと相対湿度及び気温との関係を示したものである。これらの図における相対湿度及び気温は、各観測所で観測された値を時間毎に平均し、更に各データ期間における値を時間毎に平均したものである。Fig.7から、潮汐力による地殻水平変動が O-Cに与える影響を除去した後の O-Cと相対湿度の間には相関があり、両者とも日本標準時の 2～ 3時頃最大となり、14～15時頃最小となっており、通信衛星CS-3b（衛星仰角約49°）の場合、相対湿度が約20%増加すると衛星までの観測距離が電波伝搬遅延量の増加により約1m長くなっていることが分かる。次に Fig.8から、 O-Cの変動と気温の変動との間には相関があり、気温が 10°増加すると電波伝搬遅延量が1m減少していると考えることも出来る

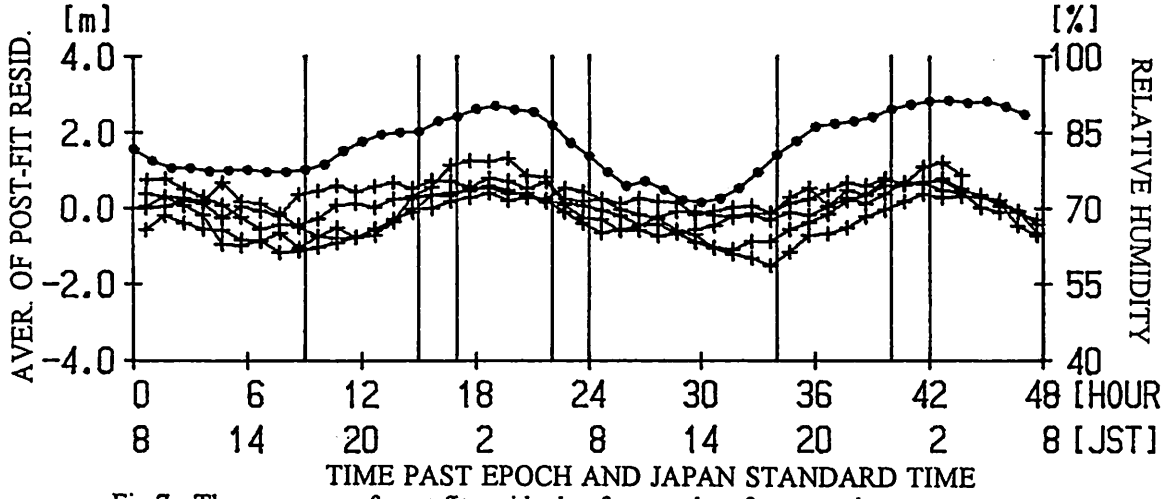


Fig.7 The averages of post-fit residuals of range data for every hour. Vertical line shows time of transit of the moon across the station meridian. The effect on O-C by horizontal movement of the earth crust is removed. θ is the angle between the sun and the moon at the earth.
+ : Data Cb1,Cb2,Cb4,Cb5,Cb8
• : Mean relative humidity around Kimitu Satellite Control Center

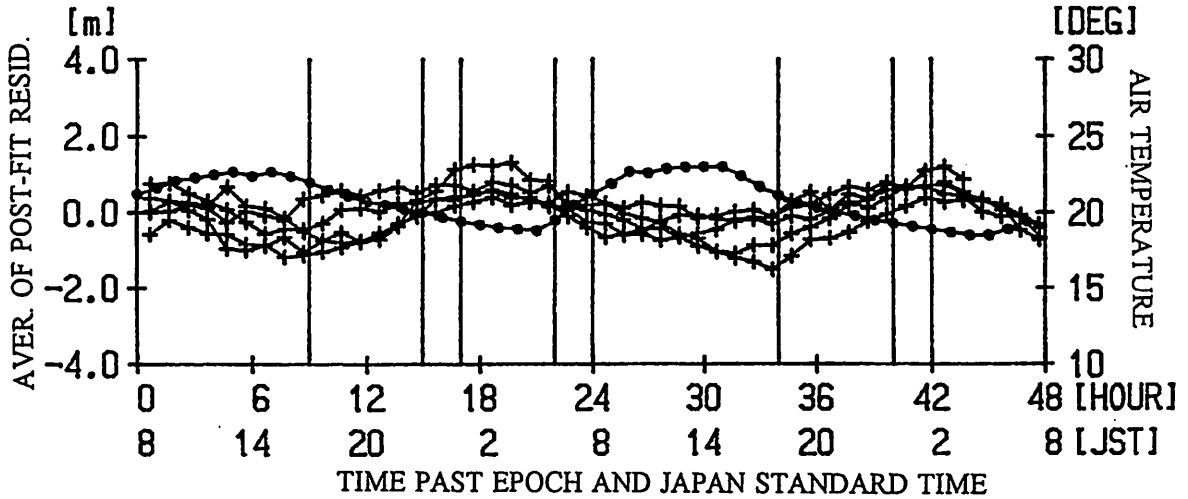


Fig.8 The averages of post-fit residuals of range data for every hour. Vertical line shows time of transit of the moon across the station meridian. The effect on O-C by horizontal movement of the earth crust is removed. θ is the angle between the sun and the moon at the earth.
+ : Data Cb1,Cb2,Cb4,Cb5,Cb8
• : Mean air temperature around Kimitu Satellite Control Center

が、本研究で使用している対流圏補正モデルでは、気温の低い冬よりも気温の高い夏のほうが電波伝搬遅延量が大きくなっており、気温が増加すると電波伝搬遅延量が減少すると考えると矛盾が生じる。又、 Fig.6から、気温が 10°以上変化しても O-Cにほとんど影響を与えていないことが分かっており、Fig.7及びFig.8における O-Cの24時間周期の変動の原因は、相対湿度の変化による電波伝搬遅延量の変化であると考えるのが合理的である。

本研究では、トランスポンダーの位置と衛星重心がずれていることに対する補正を行っていないので、衛星が重心を中心に揺れている場合は、その揺れが O-Cの変動となって表れることになる。しかし、衛星が潮汐力の大きさや位相に合わせて揺れたり、又は相対湿度の変化に合わせて揺れることは物理的に考えられない。又、 O-Cの12時間周期の変動は湿度，気圧，気温等の気象要素とは無関係であること、及び南東方向にある静止衛星（ETS-Ⅴ）の場合は周期変動が大きく[11]、南西方向にある静止衛星（BS-3a）の場合は、12時間周期の変動が存在しない[12]ことから、この周期変動は衛星観測局が水平方向（北西～南東）に周期的に動くことによって生じていると考えるのが合理的である。

5．結言

本研究における解析結果をまとめると次の様になる。

(1) 12時間周期の変動が表れた全ての場合について、その変動の位相は潮汐力に対して２時間～４時間遅れている。
(2) O-Cの12時間周期の変動は、相対湿度がほとんど変化しない場合にきれいな形で表れる。
(3) O-Cの12時間周期の変動と気圧，相対湿度，気温等の気象要素の変動との間に相関は無い。
(4) O-Cの24時間周期の変動と相対湿度の変動との間には相関がある。
(5) 衛星仰角が約49°の静止衛星の場合、相対湿度が約20%増加すると衛星までの観測距離が約1m長くなる。（対流圏における電波伝搬遅延量が1m増加する。）

又、以前の研究から南東方向にある静止衛星（ETS-Ⅴ）の場合は周期変動が大きく、南西方向にある静止衛星（BS-3a）の場合は12時間周期の変動が存在しない[11][12]ことが分かっており、この様な結果を矛盾なく説明するためには、衛星観測局の位置が、北西～南東方向に潮汐力に対して2～4時間の位相遅れで周期的に変動しており、その変動の振幅は、潮汐力が小さい時は1m程度、潮汐力が大きい時は2m程度であると考える必要がある。

観測局の位置が水平方向に周期的に変動する原因については、海洋に存在する潮流や海流の様な水平方向の流れがマントル内部にも存在（アセノスフェアの流動）し、この様な流れとともにその上に浮いているリソスフェアすなわち地殻が水平方向に動き、その結果観測局の位置が変動すると考えることが出来る。又、自転している地球に月や太陽の潮汐力が作用することにより生じるマントル内部の流動（アセノスフェアの流動）を考えることにより、極運動や大陸移動等の地殻の水平変動についても説明することが可能である。更に、この様な地球内部の流動（アセノスフェアの流動等）は地球磁場の発生にも深く係わっている様に思われる。

O-Cの24時間周期の変動については、本研究で用いられた対流圏補正モデル（GSFC NONAME式）は標準的な大気を仮定して電波の伝搬補正を行っており、大気に含まれる水蒸気量が標準的な状態からずれることにより補正誤差が生じ、それが O-Cの24時間周期の変動として表れていると考えることが出来る。

衛星力学モデルで仮定した未知の力の原因について調べること、及び対流圏電波伝搬補正モデルの改良については今後の課題である。

最後に通信衛星の貴重な観測データを提供して頂いた通信・放送機構に心から感謝の意を表する。更に、研究全般に渡って貴重な助言をして頂いた群馬天文台長の古在由秀氏、国立天文台の木下宙氏、磯部琇三氏、東京商船大学の安田明生氏及び軌道解析プログラムの作成において色々指導して頂いた海上保安庁水路部の仙石新氏、国立天文台の福島登志夫氏、宇宙科学研究所の吉川真氏、又計算機の使用等において指導して頂いた国立天文台の平山智啓氏、富山商船高等専門学校の門村英城氏に心から感謝の意を表する。

参考文献

(1)河合雅司:潮汐力による地殻の水平変動について,第30回天体力学研究会集録,pp.121-130,1998
(2)河合雅司:静止衛星の軌道解析に関する研究－Ⅳ －潮汐力による観測局位置の補正モデル－,第29回天体力学研究会集録,pp.134-143,1997
(3)Minoru Sasaki:STUDY OF THE EARTH'S DYNAMICS BY MEANS OF SATELLITE LASER RANGING TECHNIQUES, Report of Hydrographic Researches,No.26,pp.99-187,1990.
(4)Francis J.Lerch,Barbara H.Putney,Carl A.Wagner and Steven M.Klosko:Goddard Earth Models for Oceanographic Applications(GEM10B and 10C),Marine Geodesy,Vol.5,No.2,pp.145-187,1981.
(5)E.M.Standish,Jr.:Orientation of the JPL Ephemerides,DE200/LE200,to the Dynamical Equinox of J2000,Astronomy and Astrophysics,114,pp.297-302,1982.
(6)古在由秀:人工衛星による地球・海洋潮汐の検出,月刊海洋科学,地球と海洋潮汐,106号,pp.640-643,1978.
(7)Dennis D.McCarthy:IERS STANDARDS 1992,International Earth Rotation Service,1992.
(8)河合雅司:静止衛星の軌道解析に関する研究－Ⅱ,第27回天体力学研究会集録,pp.19-28,1995
(9)宇宙開発事業団:レンジとレンジレートの補正,宇宙開発事業団データベース,ABS No.2612.C.
(10)J.H.Berbert and H.C.Parker:GEOS Satellite Corrections for Refraction in the Troposphere,International Symposium on Electromagnetic Distance Measurement and Atmospheric Refraction,Boulder Colorado,1969.
(11)河合雅司:静止衛星ETS-Vの軌道解析に関する一考察,日本航海学会論文集,第89号,pp.253-260,1993.
(12)河合雅司:静止衛星の軌道解析に関する研究－Ⅰ. －通信衛星と放送衛星の観測データの解析－,日本航海学会論文集,第91号,pp.153-159,1994.

飛翔体の軌道と性能を推定する

仙石　新（海上保安庁水路部）

Estimation of a flying object trajectory and its major specifications

Arata Sengoku (Hydrographic Department)

Abstract

This paper describes a method to estimate a flying object trajectory and its major specifications. It is generally hard to know the trajectory or performance of the flying object without knowledge how the object is controlled. However, we can roughly estimate them with several assumptions. This method is useful when time of launch, burn-out and fall is known as well as distance between the lunch site and fall points.

ロケット、ミサイルなどの飛翔体の軌道は、燃料の噴射速度など飛翔体の推進装置の性能による制限があるものの、どのような軌道にするかは制御する側で自由にコントロールすることが可能である。しかし、飛翔体の打ち上げ、各段の点火・切り離し・落下等の時刻や場所がわかると、ある程度推進装置などの性能や飛翔体の軌道について推定することが可能となる。

1．飛翔体の軌道

燃料が燃焼している段階では、飛翔体の軌道は、推進力、燃焼時間、排気速度、燃料の重量比といった推進装置の性能によってある程度規定される。しかし、各瞬間の飛翔体の向きや噴射速度を調整するなどの制御を行えば、飛翔体の軌道はかなり自由に変えることができる。いわば、制御する側に無限の自由度があるといえる。

一方、推進力が無く free fall している段階では、飛翔体の軌道は決定論的に論ずることが可能である。簡単のため大気抵抗を無視すると、飛翔体の軌道は、ケプラー運動となり、初期位置と速度のみで決定される。

1．1．加速時の飛翔体の運動

ここでは、重力は無視し、飛翔体は直線的に運動するものと仮定する。飛翔体の質量を m、速度を v、有効排気速度を u'、とすると（冨田、1993）、

$$m\frac{dv}{dt} = -u'\frac{dm}{dt}$$

ここで、排気量と有効排気速度が一定と仮定すると、上式は解析的に解くことができて、

$$v(t) = v_0 - u'\ln(\frac{m(t)}{m_0}) = v_0 - u'\ln(1-\alpha t)$$

$$v_f = v_0 - u'\ln(1-\mu)$$

$$l(t) = l_0 + v_0 t + \frac{u'}{\alpha}[(1-\alpha t)\ln(1-\alpha t) + \alpha t]$$

ただし、vo は初期速度、vf は最終速度、μ は燃料の質量比、l は初期位置からの移動距離、α

は質量の変化速度をそれぞれ表す。

$$m = m_0(1-\alpha t)$$

1．2．非加速時の飛翔体の運動

非加速時の飛翔体の運動はケプラー運動で記述できる。簡単のため、軌道面は加速時と同一平面であると仮定する。また、大気抵抗は無視する。

ケプラー運動であるから、元期の位置・速度を与えると軌道が決まる（平面問題であるから自由度４）が、元期の位置・落下地点・飛行時間の４つの量を与えても軌道が決まる。初期位置と落下位置の地心角距離をΦとすると（Figure１参照)、地球は球であるものと仮定して、

$$\Phi = 2\tan^{-1}\left(\frac{b+\sqrt{b^2+ac}}{a}\right)$$

ただし、a、b、cは以下で定義されるものとする。

$$a = 2r_e(1+\tan^2\gamma_0)-(r_0+r_e)\nu_0$$
$$b = \nu_0 r_e \tan\gamma_0$$
$$c = (r_0-r_e)\nu_0$$
$$\nu_0 = r_0 v_0^2/GM$$

ここで、re、$r0$、$\gamma 0$はそれぞれ地球半径、初期の地心距離、速度の仰角である。飛行時間Tは、

$$T = t-t_0 = \frac{E-E_0-e(\sin E-\sin E_0)}{n}$$

ただし、Eは離心近点角、nは平均運動で、真近点角f、離心率eなどと以下の関係にある。

$$\tan\frac{f}{2} = \sqrt{\frac{1+e}{1-e}}\tan\frac{E}{2}$$
$$f = f_0+\Phi$$
$$\cos f_0 = \frac{(p/r_0-1)}{e}$$
$$n = \sqrt{GM(1-e^2)^3/p^3}$$
$$e = \sqrt{1+\nu_0(\nu_0-2)\cos^2\gamma_0}$$
$$p = a(1-e^2) = r_0\nu_0\cos^2\gamma_0$$

以上から、$r0$、$f0$、T、Φを与えて、$\gamma 0$、$\nu 0$を求め、これらからケプラー要素を計算することができる。上式は$\gamma 0$、$\nu 0$について初等的に解くことができないので、数値的に解を求めることになる。

2．三段ロケットのevent time sequenceから飛翔体の性能と軌道を推定する

飛翔体は、一般的に多段ロケットの構造となっている。これは、打ち上げの効率を向上させるために必要となるものである。ここでは、三段ロケットの場合を考える。Figure２は三段ロケットの打ち上げスキームである。一般的に、第二段ロケットの燃焼終了後、先端のカバーをはずし、人工衛星を軌道に投入するための準備が行われた後に第三段が切り離される。

ここで、三段ロケットの打ち上げに関して、各段の点火・切り離し・落下時刻と各段の落下位置が既知である場合のロケットの軌道について考察する。

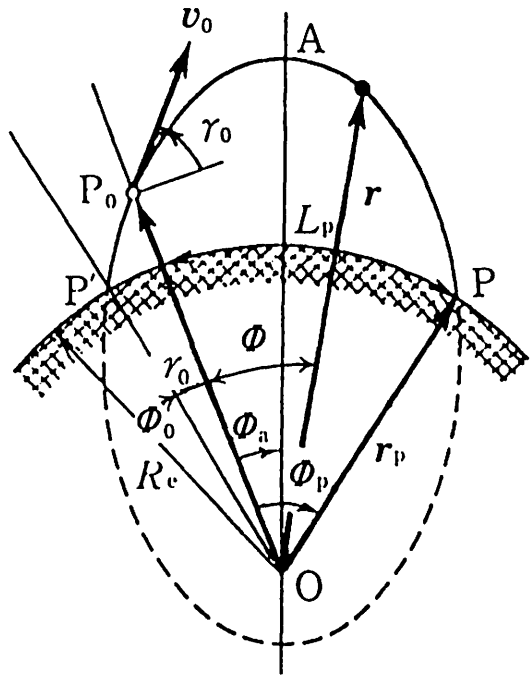


Figure 1. Trajectory of a flying object. Po and P stand for burn-out/separation and fall points, respectively (after Tomita, 1993).

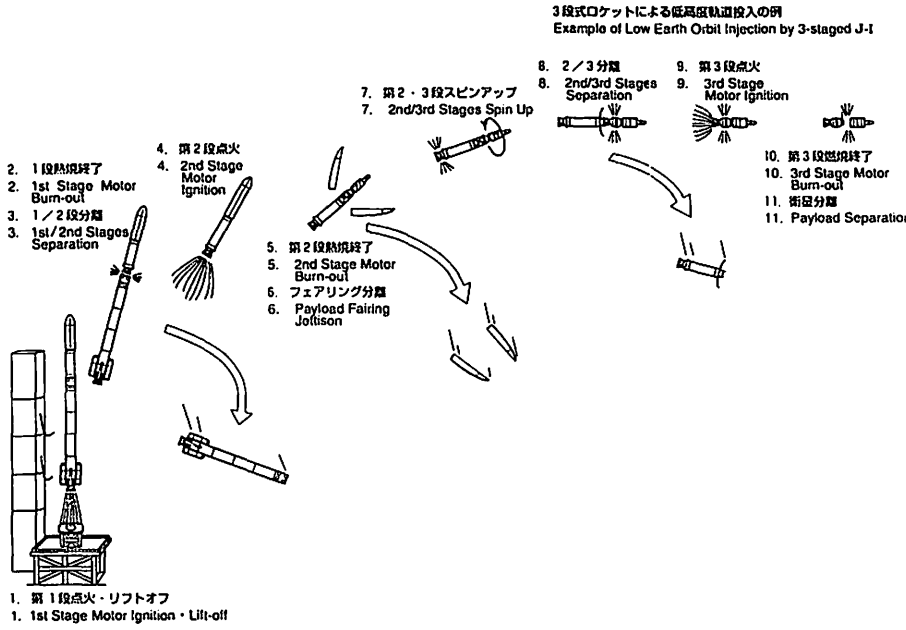


Figure 2. Launch scheme of a flying object with three stages (from a brochure of J1 rocket by National Space Development Agency of Japan).

1節にも述べたように、加速時の軌跡については直線運動を仮定する。しかし、これだけの仮定では不十分であるため、以下の仮定をさらに加える。

・各段毎に直線運動を仮定する。加速時の軌跡は折れ線となる。
・各段の軌跡の仰角（地平線からの高度）は、点火時と切り離し時の平均値である。
・各段の有効排気速度は等しい。
・燃料が燃え尽きると同時に切り離される。
・切り離しと同時に点火される。

また、飛翔体はどちらの方向にも打ち上げられる可能性があるが、簡単のため真東に打ち上げられるものとする。

自由落下時の軌道は、ケプラー運動とし、１．２に述べた公式を用いるものとする。

例題として、以下のような場合を考えてみよう。

これは、1998年8月31日に打ち上げられたテポドンに関する新聞情報をまとめたものである。

Table 1. Time of separation/fall and distance between the launch site and the fall points

	Time since the launch	distance
Separation (first stage)	1m24s	—
Separation (cover)	2m10s	—
Separation (second stage)	4m24s	—
Fall (first stage)	5m00s	410km
Fall (cover)	9m20s	1090km
Fall (second stage)	10m04s	1620km

２．１．第一段ロケットの軌跡と性能

第一段ロケットははじめ垂直に打ち上げられるものとし（$\gamma_1=90$deg)、第一段ロケットの燃料の重量比を0.6と仮定する。飛翔体の有効排気速度がわかっていれば重量比の代わりにこちらを仮定しても良い。第一段ロケットの燃焼が終わって切り離された時の仰角（γ_2）と有効排気速度（u'）は変数とする。第一段ロケットの切り離し地点の水平位置と高度は、有効排気速度などによって変わってくる。第一段ロケットは燃え尽きた後、仰角（γ_2）で放り投げ上げられて自由落下するものとする。打ち上げ地点と落下地点間の距離は地球固定系で測定されているので地球の自転を考慮して慣性系で見た地心角Φを計算しておく必要がある。

u'とγ_2を変数にして、落下地点までの距離と落下時刻が合う（γ_2, u'）の組み合わせを探す（Figure３）と、以下の値となる。

$u'=2.1$km/sec
$\gamma_2=17$deg

有効排気速度は、ロケットの基本的な性能のひとつで、一般的に 2～4.5km/sec 程度の値をとる。

２．２．カバーと第二段ロケットの軌跡

第二段ロケットの有効排気速度は第一段ロケットと等しいと仮定する。初期の仰角（γ_1）は17degとする。燃料の重量比とカバーが切り離された時の仰角（γ_2）は変数である。

以上から切り離し地点を計算し、落下地点と落下時刻が合う（γ_2、μ）を探す（Figure４）と、以下の値となる。

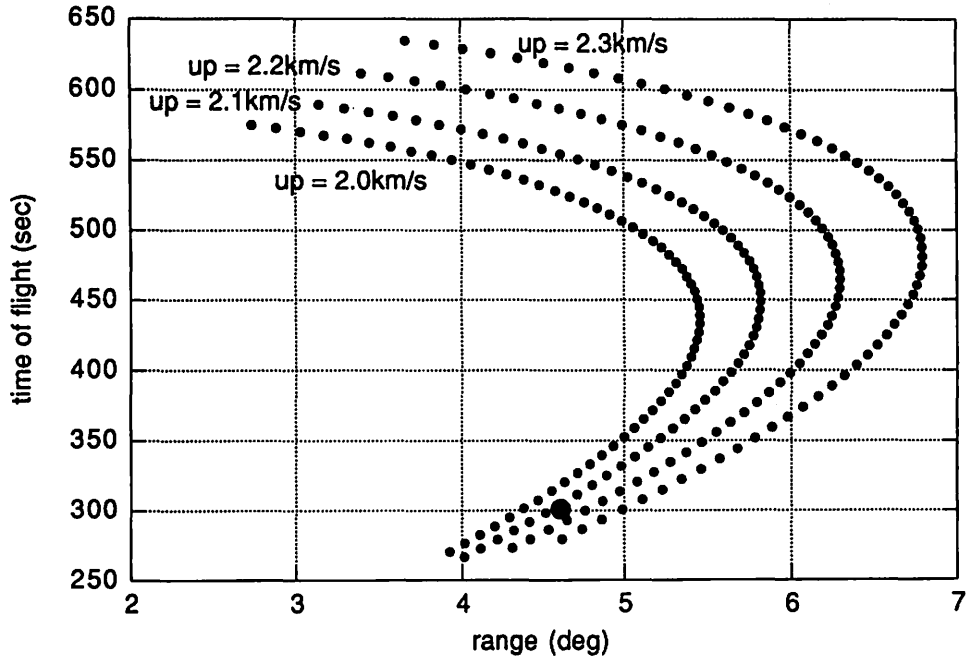


Figure 3. Time of flight vs. range as a function of u' (first stage). A large circle stands for the observation.

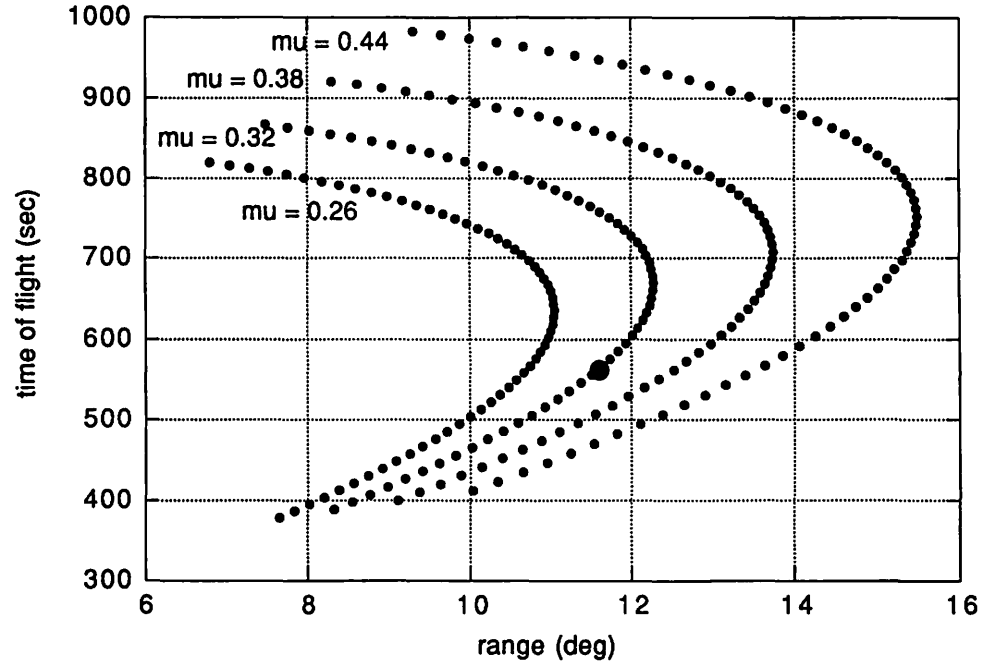


Figure 4. Time of flight vs. range as a function of μ (cover). A large circle stands for the observation.

μ ＝0.32

γ_2＝28deg

この場合、燃料の重量比が小さすぎるため、カバーを切り離した後もさらに加速したと考えるべきであろう。第二段ロケットについて、燃料の重量比と切り離し時の仰角を変数にして、落下地点と落下時刻が合致する（γ_2、 μ）の組み合わせを探すと、

μ ＝0.48

γ_2＝3deg

となる。ただし、燃料の消費速度はカバー切り離し前と同一である仮定した（2:10～3:18 にロケット燃焼)。カバーがはずれる前と合わせると、第二段ロケットに搭載されていた燃料の重量比は0.8になる。

以上から、推定された各段の free fall 時の軌道要素は、Table２のとおりとなる。

Table 2. Estimated orbit characteristics during the free fall

	First stage	cover	second stage
Velocity at separation	2.3km/sec	3.1km/sec	4.5km/sec
Altitude at separation	62km	120km	260km
Height at separation	17deg	28deg	3deg
*p=a(1-e*e)*	500km	800km	2200km
Eccentricity	0.92	0.88	0.67
Semi-major axis	3400km	3500km	4000km
Apogee height	90km	250km	270km
Displacement during Accelerating stage	70km	120km	250km

２．３．第三段について

第二段ロケットが切り離されたと考えられる高度260km で円運動をするためには、飛翔体の速さは7.7km/sec でなければならない。このためには、第三段の燃料の重量比μが0.8以上であることが必要である。この場合、第三段からごく小さな人工衛星を軌道に投入することが可能である。

３．第三段が推進力を持たない場合、第三段はどこまで飛ぶか？

ある高度からある速度で飛翔体を free fall させた場合、飛翔体を最も遠くまで飛ばすためには仰角を以下のようにすればよい。１．２節で述べたように、free fall の初期位置から落下地点までの地心から見た離角は、

$$\Phi = 2\tan^{-1}\left(\frac{b+\sqrt{b^2+ac}}{a}\right)$$

ただし、a、b、c は初期の高度、速さ、速度の仰角（γ_0）の関数である。高度と速さが一定であるとして、γ_0で上式を偏微分し、これを０と置くと最も遠くまで飛ばすことができる仰角γ_0が求められる。これを maximum range trajectory と呼ぶ。このようなγ_0は以下を満たす。

$$\tan^2\gamma_0 = \frac{-B \pm \sqrt{B^2 - 4AC}}{2A}$$

ここで、A、B、Cは以下のとおりである。

$$A = 4\{2r_e(r_0 - r_e) + \nu_0 r_e^{\,2}\}$$
$$B = 4(r_0 - r_e)\{2r_e - \nu_0(r_0 + r_e)\}$$
$$C = -\nu_0\{2r_e - \nu_0(r_0 + r_e)\}^2$$

Figure 5 に高度 260km で飛翔体を free fall させた場合の maximum range trajectory を示す。maximum range trajectory では、初期の仰角（γ_0）、落下地点までの地心角距離、飛行時間を、初速度の関数として表すことができる。宇宙速度を越えると maximum range は無限大となる。初速が 0 または宇宙速度に近い場合、飛翔体を水平に打ち出すのが最も効率的になる。中間的な速度の場合は、初期の高度が 260km の場合は 40 度弱にするのが効率的である。ただし、これは初期の高度の関数になる。

初速度として Table 2 の第 2 段ロケットの最終速度（4.5km/sec）をとると、仰角を調整することにより、最大 2900km まで飛ばすことができることがわかる。すなわち、Table 1 で与えられる飛翔体は、第 3 段目が推進力を持たない場合、最大で 2900km まで飛ばすことができることがわかる。

4．まとめ

飛翔体の軌道について推定する簡便な手法を見いだした。この手法は、飛翔体について限られた情報しか存在しない場合に有効である。

Table 1 の飛翔体情報に基づいて推定を行うと、有効排気速度は 2.1km/sec となる。また、第三段目が推進力を持たない場合、第三段は最大で 2900km の距離まで飛翔体を飛ばすことが可能である。第三段目の燃料重量比が 0.8 以上であれば、ごく小さな人工衛星を軌道に投入することも可能である。

本稿では、大気抵抗力と加速時の重力の影響を無視している。今後これらを考慮することが必要である。これらの影響を考慮した場合、有効排気速度はより大きな値となる必要がある。また、加速時の軌跡を直線で近似したが、より現実的な仮定をおくべきかもしれない。

参考文献

冨田　信之（1993）：宇宙システム入門，東京大学出版会、74 ページ.

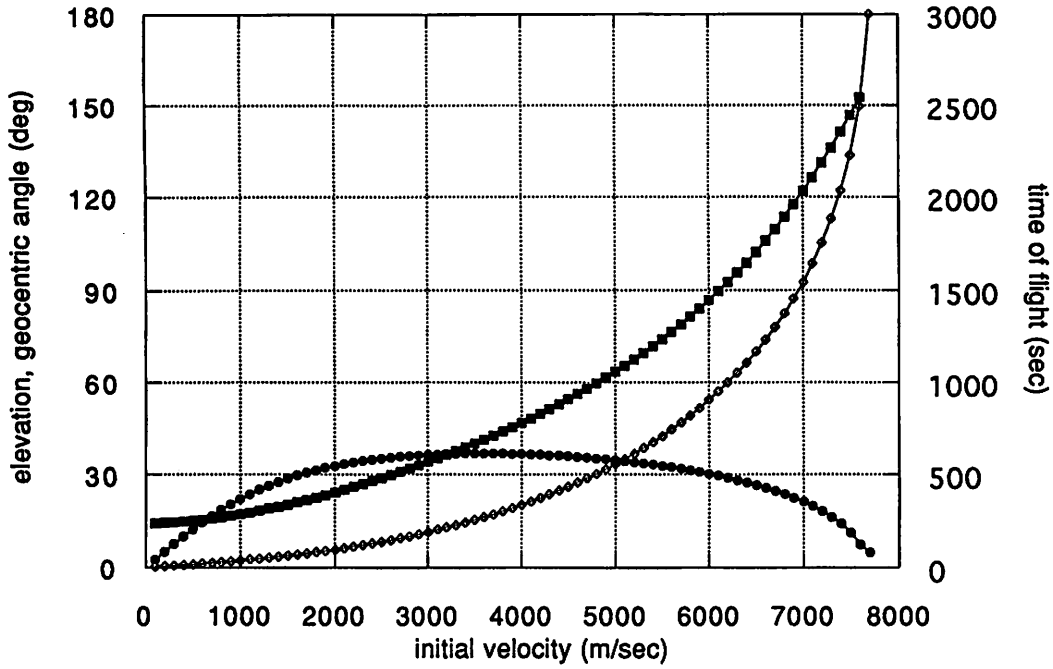


Figure 5. Maximum range trajectory (initial height=260km). Black circles, white squares and black squares stand for initial elevation, geocentric angle between the launch site and the fall point, and time of flight, respectively.

Visualization of Satellite Orbital Motions by Virtual Reality

Kazuhiro Kimura

Kashima Space Research Center
Communications Research Laboratory (CRL)
Hirai 893-1, Kashima, Ibaraki, 314-0012, Japan
E-mail: kimura@crl.go.jp

Abstract

A system for visualization of satellite orbital motions by virtual reality has been developed. The system can display orbital calculation results of the satellites as moving computer graphics images. Three-dimensional orbital motions of a number of satellites can be simulated with the viewpoint and the viewing direction of selected and controlled arbitrarily. The merit of this system is that complex orbital motions of satellites can be intuitively understood by visualization. Not only the motion of each satellite, but the variations of constellation, coverage pattern, visibility, and relative positions of approaching satellites also can be clearly displayed.

1. Introduction

In general, orbital motions of satellites are analyzed using analytical methods or numerical calculations based on the equations of orbital dynamics and orbital elements of all satellites. Although exact positions and velocities of all satellites at arbitrary times can be calculated in these analyses, it is not easy to understand the motions of the satellites in the three dimensional space. Especially, it is difficult to get a whole image of a system that consists of a lot of satellites. Even if there are some trivial mistakes in the equations or calculation program, it is difficult to find them only by looking at the calculated values.

In order to simulate and visualize satellite orbital motions, an orbital motion analysis system using virtual reality has been developed. The results of orbital calculation of the satellites can be displayed as moving computer graphics images using this system. The system projects three-dimensional orbital motion of a number of satellites. The viewpoint and the viewing direction of can be moved arbitrarily. This system is useful not only in orbit analysis and design but also in presentation of satellite orbital motions for such non-specialists of orbital dynamics as mission operators and users of satellite systems.

When the initial system which can display only the motions of planets and asteroids in the solar system was being developed, the outline of the system has already been introduced.[1] The functions of simulating orbital motions of artificial satellites have been newly installed.

This paper describes in outline the hardware and software construction of this system, features of this system, and applications. Some examples of images generated by this system are also shown.

2. Outline of the system

2.1. Hardware

The system is installed in a large capacity data storing and processing system, which exists for research and development of satellite communications technologies with very high data rate for high quality multimedia communications, and which can generate virtual reality images as the data sources for satellite communications experiments.

Figure 1 shows the configuration of the hardware. There are five computers in this system connected with a local area network. The main computer is a Power ONYX Reality Engine 2, which gives excellent performances on image processing. It can output 3-pipe computer graphics images simultaneously to three CRT monitors. Two graphical workstations (Indigo 2) are also used to analyze and simulate orbital motions. The data and images generated by the computers can be stored in the file servers. The main computer can produce oblong images with the screen width equivalent to three CRT monitors just arranged side by side. The images are divided into three parts, and output simultaneously from three terminals of the computer. Output image signals from each terminal are distributed to the corresponding CRT monitor and projector. Three projectors reproduce original wide images onto the wide arch screen with a height of 2 meters and a width of 8 meters.

Stereographic images can also be seen using special glasses with synchronized liquid crystal shutters. The main computer generates the images for the left eye and right eye alternately with high frequency such that the viewer cannot recognize the switching of the images. The computer sends synchronization signals to the 3-D controller coincident with each transmission of images. The synchronization signal emitter converts the signals into infrared signals that can be received by the 3-D glasses. When the images for the left eye are projected onto the screen, the glasses open the liquid crystal shutter for the left eye and close that for the right eye; as a result, only the left glass becomes transparent, and one can see the images only through the appropriate one's left eye. As the images for each eye can be seen with the appropriate eye, one can see three-dimensional images by using the glasses.

The main computer is equipped with a joystick controller as one of the input devices.

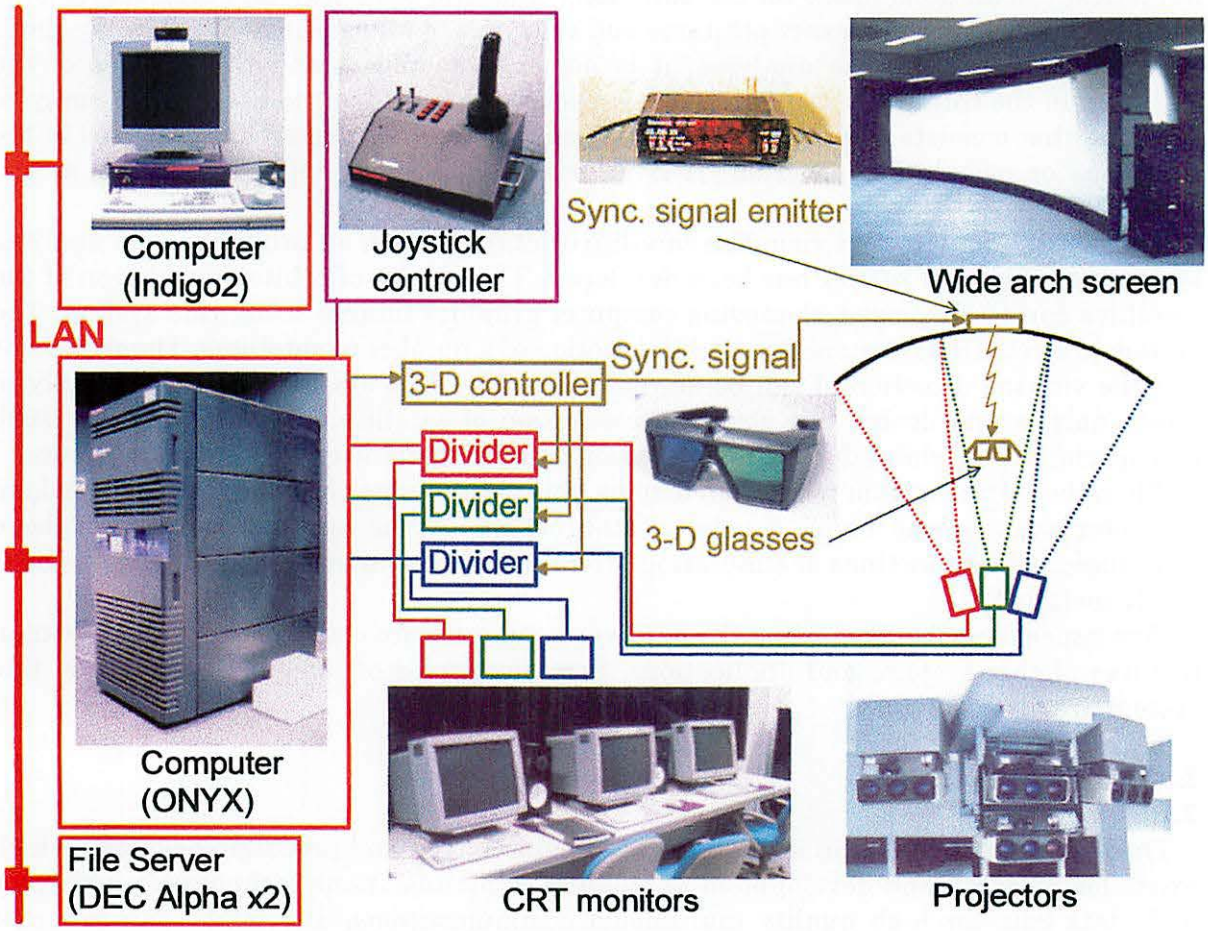


Figure 1: Hardware configuration

When satellite orbital motions are simulated with this system, the viewpoint and the viewing direction can be moved manually by using the device. Consequently, the hardware constituting this system enables us to see rather large images in a three-dimensional way with an arbitrary view, and really to experience the world of satellite orbital motions.

2.2. Software

Virtual reality software to analyze and display orbital motions of satellites has been developed and installed on the computers. This software is based on ready-made software of virtual reality for general use. The functions of the software have been enhanced by adding special functions to display orbital motions, such as, displaying lines representing satellite orbits, character strings, and dots for small objects. Data conversion programs have also been developed. The programs make data files to be loaded to the virtual reality software, which requires a proper data format in accordance

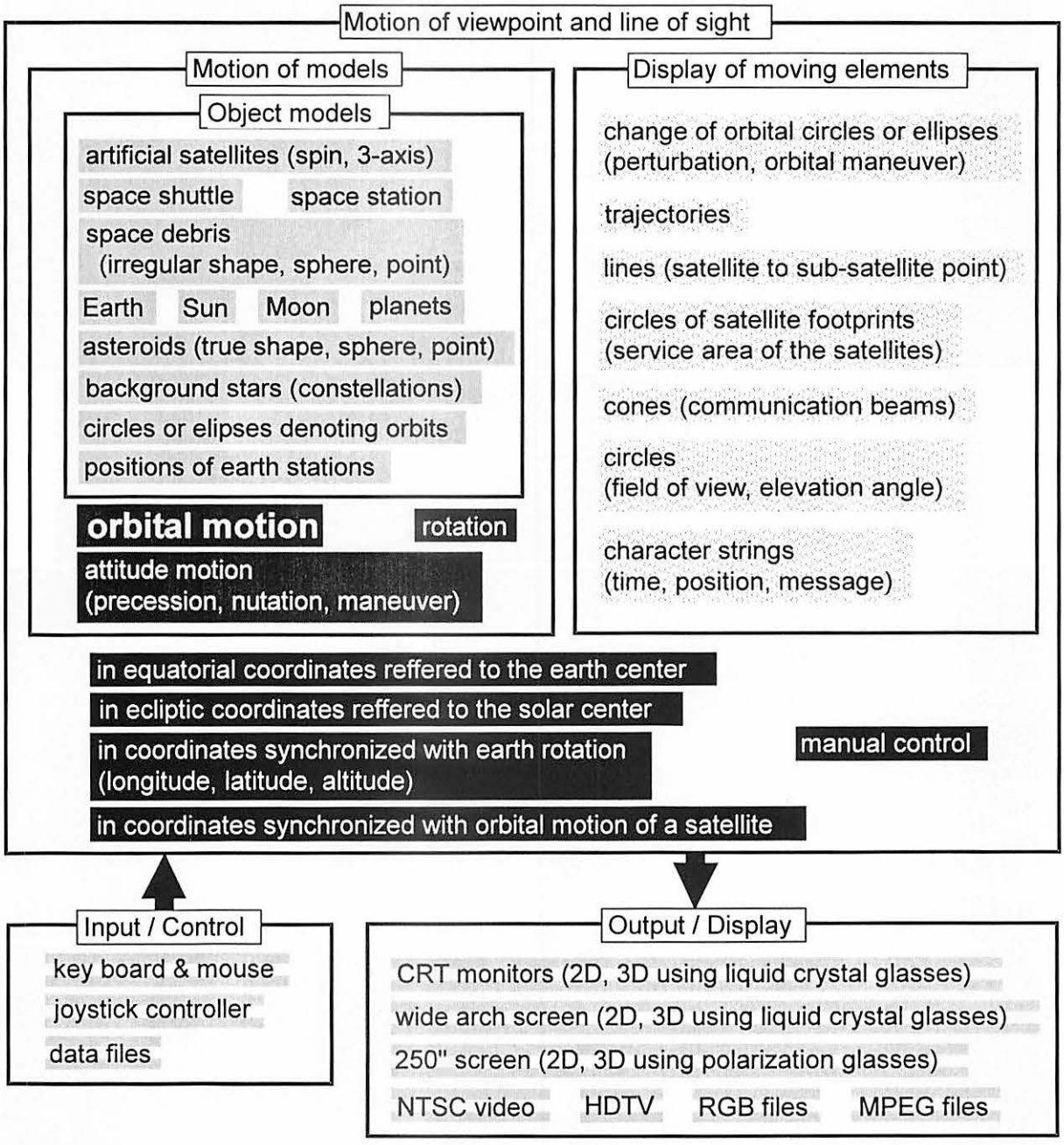


Figure 2: Functions of the simulation software

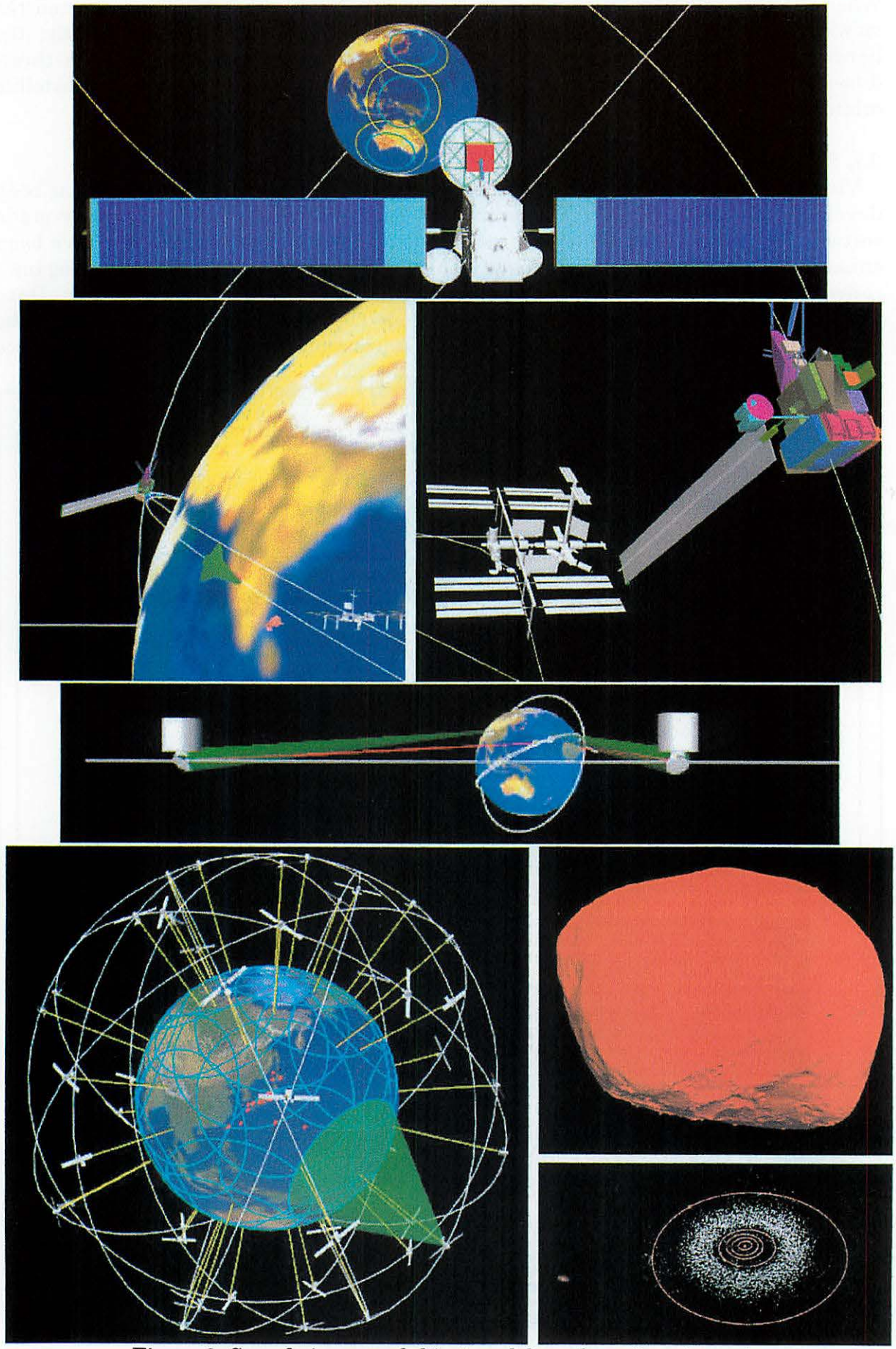


Figure 3: Sample images of object models and moving elements

with the proper coordinates for virtual reality, from the technical data such as orbital elements, positions of earth stations, and change of view.

Main functions of the software are shown in Figure 2. The software arranges selected objects in virtual reality space and displays them. Models such as satellites, earth, and planets can be processed in the space with orbital and attitude motions including rotation, if necessary. Not only object models without transformation but also moving elements whose shapes are altered at every moment can be displayed. By using this function, it is possible to simulate changing orbits by maneuvers or orbital perturbations. Motions of cones that express communication beams and change of satellite coverage denoted by circles on the earth also can be displayed. Sample images of object models and moving elements are shown in Figure 3. Although motions of viewpoints and line of sight must be described in virtual reality coordinates primarily, the conversion programs help to change the data based on the coordinates listed in Figure 2. It is also possible to make the data only giving a table describing a time sequence of view control. Each line of the table consists of time, object of observer, and target object, where 'object' means a satellite, an earth station, or the center of the earth. Using this function, images of the earth or target satellite viewed from a satellite or earth station are acquired very easily. Manual control of view using joystick controller or mouse is also supported. Although this software normally outputs moving images on CRT monitors or screen, recording the images to an NTSC video recorder or saving still images to RGB files can also be performed.

3. Simulation procedure

The procedure of simulation by this system is illustrated in Figure 4. At first, it is necessary to make four files: an orbital element file in which the Kepler elements of all satellites are written, an earth station file containing the positions (longitude, latitude, and altitude) of all stations and reference points in coordinates rotating synchronously with the earth, a moving element definition file, and a view file. Shapes of the circles under the satellites, both ends of cones, and lines to the sub-satellite points including colors of all these elements are described in the moving element definition file. Time variations of viewpoint and line of sight are defined in the view file.

After these files are prepared, the conversion program is executed to produce data files for the virtual reality software. Next, the files read by the virtual reality software are edited. After making a scenario file to control models appearing and disappearing, change of scenes, and so on, simulation can be done by executing the virtual reality software. After the simulation, if it is necessary to modify parameters or conditions, the procedure mentioned above is repeated.

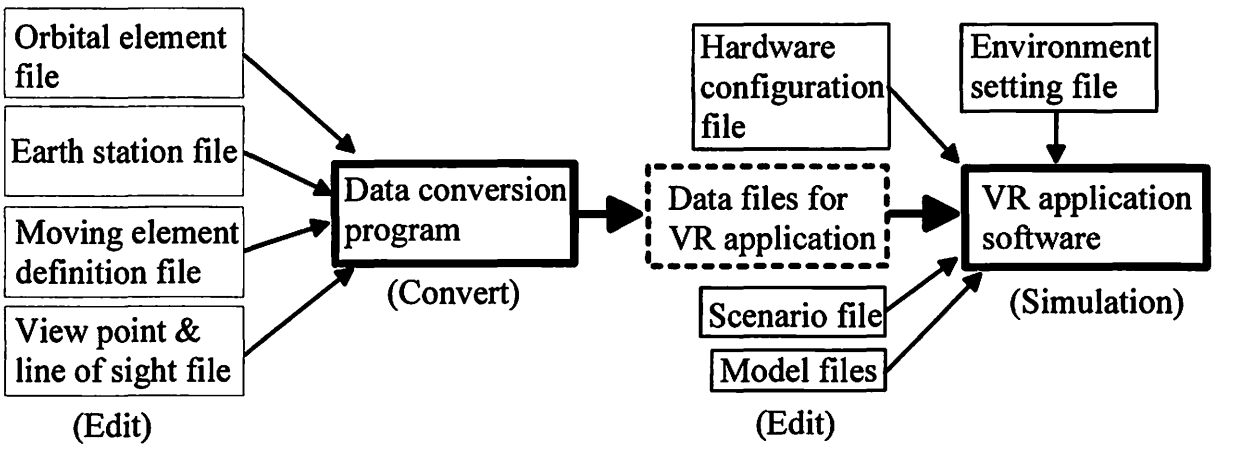


Figure 4: Procedure of simulation

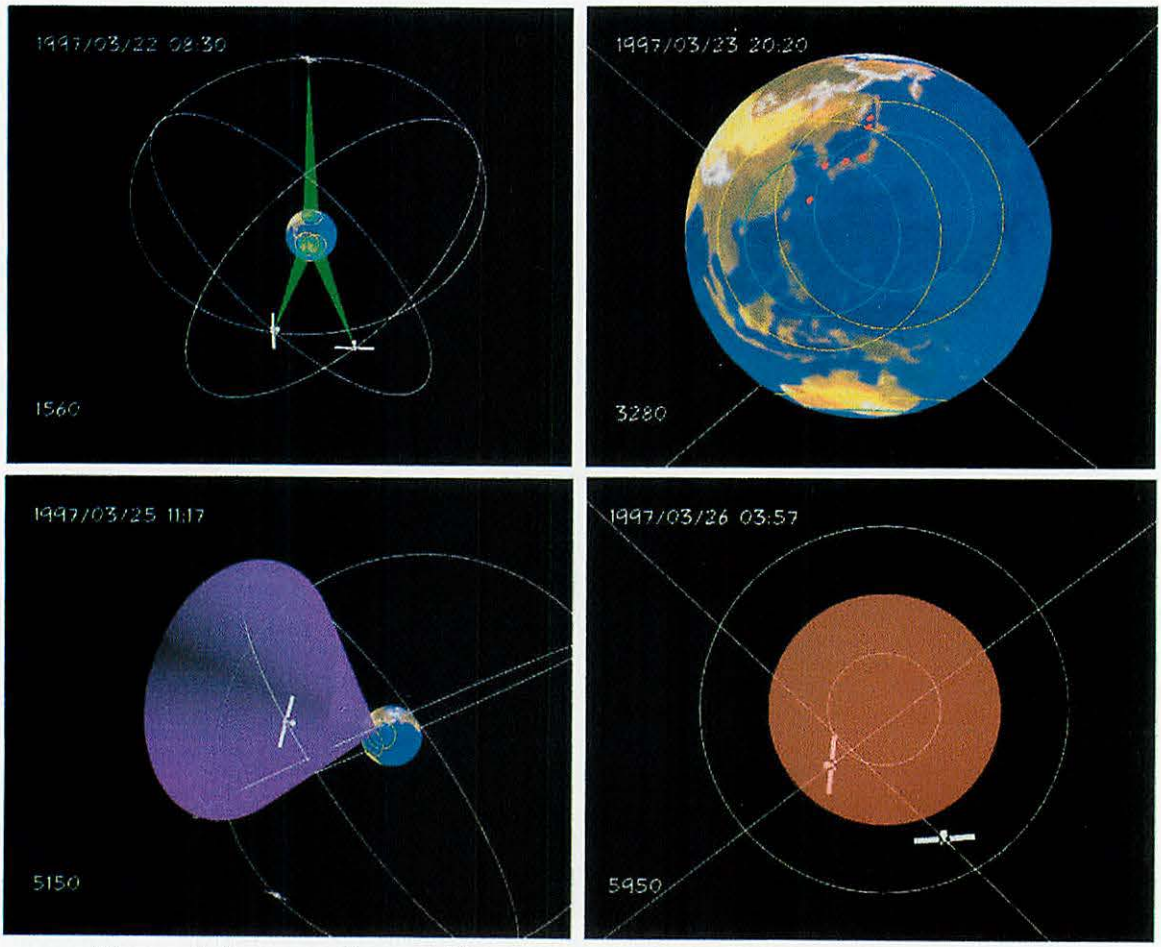


Figure 5: Images of a satellite system using inclined synchronous orbits

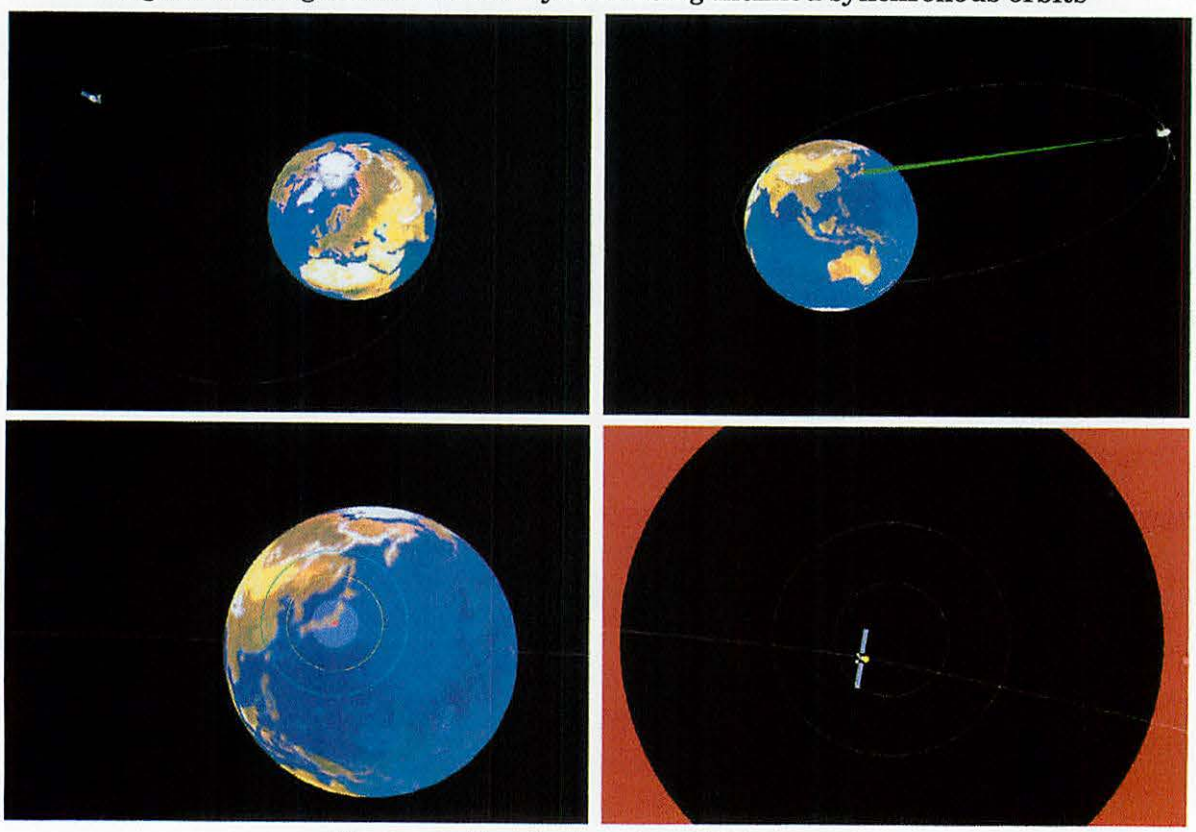


Figure 6: Simulation of COMETS orbit

4. Simulation examples

Not only the motion of each satellite, but the variations of constellation, coverage pattern, visibility, and relative positions of approaching satellites also can be clearly displayed. Visualization of the results of orbital motion analysis by numerical calculation is also useful for the verification of the algorithm and computing software. Accordingly, designing of satellite constellation, planning of station keeping operation, mission analysis of rendezvous, and so on, may be efficiently performed by this system. It may be useful to simulate the orbital motions of deep space missions such as explorations of planets and asteroids.

Actually, this system is applied to design satellite constellations of non-geostationary satellites in our laboratory. The orbital motions of a satellite communications system using inclined synchronous circular orbits are being analyzed. Some scenes produced by this virtual reality system are shown in Figure 5. The upper left image of the figure is the satellite constellation in inertial coordinates. Communication beams from the satellites are also displayed. The upper right image shows the coverage area of the satellites where each satellite can be seen with elevation angles more than 60 (within outer circles) and 70 (within inner circles) degrees. The lower left image indicates that more than one of the three satellites always exists in the communication beam pointed to the zenith with a divergence angle of 20 degrees. The variation of elevation angles of satellites viewed from the surface of the earth can be seen by the images like the lower right one. Thus this system is useful in order to optimize orbital parameters of the satellites and to visualize various properties concerned with satellite orbital motions. The images made by this system are very convenient to introduce the concepts of investigating satellite systems to the peoples with insufficient knowledge of orbital dynamics.

Orbital motions of COMETS (communications and broadcasting engineering test satellite) are also visualized using this system. The satellite was launched on 21 February, 1998. However, due to a failure of second stage engine of the H-II rocket, it was failed to be injected into the geostationary orbit. The orbit of the satellite was changed by seven orbital maneuvers in order to conduct communications experiments. Finally, the satellite was injected into a sub-recurrent elliptical orbit on which the satellite revolve nine times in every two days. Images of the COMETS orbit simulated by this system are shown in Figure 6. The upper left image is the orbit viewed from the direction of the orbit normal. The shape of the elliptical orbit with an eccentricity about 0.56 can be grasped at a glance. Experiments are conducted while the satellite are orbiting around the apogee. A geometry at the timing is shown in the upper left image. An image of the earth viewed from the satellite with an attitude bias for antenna pointing to the earth station located at Kashima can be seen in the lower right image. The lower left image shows a satellite motion viewed from the earth station. It is useful to know conditions of experiments in advance.

5. Summary

A system for visualizing satellite orbital motions by virtual reality has been developed. The greatest merit of this system is that complex orbital motions of satellites can be intuitively understood by visualizing the results of orbital analysis. It is also useful for various presentations for non-specialists of the orbital dynamics.

An advanced simulation software is being developed. In this system, parameters such as the orbital elements of each satellite can be easily changed by graphical user interface, and the consequent orbital motions are immediately displayed.

Reference

[1] M. Yoshikawa, "Visualization of Celestial Mechanics --Orbital Motion Analysis Software by Virtual Reality --", Proceedings 28th Symposium on Celestial Mechanics, 137-142, 1996.

Author Index and Participants List

(氏名 Alphabet 順, ページ数は筆頭著者のみ, 所属は 1999 年 3 月現在)

Kashima Space Research Center

http://www.crl.go.jp/ka