

I MOTI QUASI PERIODICI E LA STABILITÀ DEL SISTEMA SOLARE.

II: Dai tori di Kolmogorov alla stabilità esponenziale.

ANTONIO GIORGILLI

*Università degli Studi di Milano, Dipartimento di Matematica,
Via Saldini 50, 20133 — Milano, Italy.*

1. Premessa

Nella prima parte di questo lavoro ho illustrato alcuni problemi classici della meccanica celeste, soffermandomi in particolare su tre punti: la descrizione della dinamica planetaria come sovrapposizione di moti periodici, che si riconduce essenzialmente al metodo classico degli epicicli; il problema delle risonanze, che si manifesta negli sviluppi perturbativi sotto la forma dei piccoli divisori; la scoperta, ad opera di Poincaré, del fenomeno del punto omoclino e della dinamica caotica che ne consegue.

Dedicherò la seconda parte di questo lavoro alla discussione di alcuni sviluppi delle nostre conoscenze che hanno segnato la seconda metà del novecento, e che continuano nei nostri giorni. Vorrei però premettere due brevi osservazioni che mostrano come gli sviluppi teorici recenti costituiscano una continuazione ideale di temi che già sono presenti nei *Méthodes Nouvelles*.

La prima osservazione riguarda il problema, già discusso nella prima parte del lavoro, della convergenza degli sviluppi classici e della conseguente rappresentabilità della dinamica planetaria come quasi periodica. Ho già accennato in particolare al fatto che il teorema di Kolmogorov, di cui parlerò più diffusamente nel paragrafo 2, mostra in modo indiretto la convergenza delle serie di Lindstedt, almeno per valori particolari delle frequenze. La grande rilevanza del teorema di Kolmogorov per il problema della stabilità delle orbite planetarie è messa in grande evidenza da Arnold, che così scrive all’inizio di un lavoro ad esso dedicato:^[1]

“ One of the most remarkable of A.N. Kolmogorov’s mathematical achievements is his work on classical mechanics of 1954. A simple and novel idea, the combination of very classical and essentially modern methods, the solution of a 200 year-old problem, a clear geometric picture and a breadth of outlook — these are the merits of the work. ”

La seconda osservazione, su cui mi dilungherò un poco, riguarda l’utilità degli sviluppi perturbativi, indipendentemente dalla convergenza. A questo proposito, vorrei citare una

pagina che si trova proprio all'inizio del secondo volume dei *Méthodes Nouvelles*, nel capitolo VIII.[†]

“ Il y a entre les géomètres et les astronomes une sorte de malentendu au sujet de la signification du mot *convergence*. Les géomètres, préoccupés de la parfaite rigueur et souvent trop indifférents à la longueur de calculs inextricables dont ils conçoivent la possibilité, sans songer à les entreprendre effectivement, disent qu' une série est convergente quand la somme des termes tend vers une limite déterminée, quand même les premiers termes diminueraient très lentement. Les astronomes, au contraire, ont coutume de dire qu' une série converge quand les vingt premiers termes, par exemple, diminuent très rapidement, quand même les termes suivants devraient croître indéfiniment.

Ainsi, pour prendre un exemple simple, considérons les deux séries qui ont pour terme général

$$\frac{1000^n}{1 \cdot 2 \cdot 3 \dots n} \quad \text{et} \quad \frac{1 \cdot 2 \cdot 3 \dots n}{1000^n} .$$

† “ Tra i geometri (oggi si direbbe piuttosto gli analisti, *n.d.a.*) e gli astronomi c'è una sorta di malinteso sul significato del termine *convergenza*. I geometri, interessati al rigore assoluto e spesso troppo indifferenti alla lunghezza dei calcoli inestricabili di cui concepiscono la possibilità, senza per questo sognarsi di intraprenderli davvero, dicono che una serie è convergente se la somma dei termini tende ad un limite ben definito, e ciò anche se i primi termini decrescono in modo estremamente lento. Gli astronomi, al contrario, usano affermare che una serie converge se, diciamo, i primi venti termini diminuiscono molto rapidamente, anche se i termini successivi crescono indefinitamente.

Così, per fare un semplice esempio, consideriamo le due serie che hanno come termine generale

$$\frac{1000^n}{1 \cdot 2 \cdot 3 \dots n} \quad \text{e} \quad \frac{1 \cdot 2 \cdot 3 \dots n}{1000^n} .$$

I geometri diranno che la prima serie converge, e pure rapidamente, perché il milionesimo termine è ben più piccolo del 999 999^o; essi classificheranno invece la seconda serie come divergente, perché il termine generale cresce oltre ogni limite.

Al contrario, gli astronomi considereranno la prima serie come divergente, perché i primi 1000 termini sono crescenti, mentre classificheranno la seconda serie come convergente perché i primi 1000 termini decrescono, e all'inizio la diminuzione è molto rapida.

Ambedue le regole sono legittime: la prima nella ricerca teorica, la seconda nelle applicazioni numeriche. Ambedue devono regnare, ma in domini separati da confini che dovremmo conoscere in modo ben preciso.

.....

Il primo esempio che abbia messo in evidenza la legittimità di certi sviluppi divergenti è quello classico della serie di Stirling. Cauchy ha mostrato che i termini di quella serie prima decrescono, e poi iniziano a crescere, sicché la serie diverge; ma arrestando il calcolo al termine più piccolo si rappresenta la funzione Euleriana con un'approssimazione sempre migliore al crescere dell'argomento.“

Les géomètres diront que la première série converge, et même qu'elle converge rapidement, parce que le millionième terme est beaucoup plus petit que le $999\,999^e$; mais ils regarderont la seconde comme divergente, parce que le terme général peut croître au delà de toute limite.

Les astronomes, au contraire, regarderont la première série comme divergente, parce que les 1000 premiers termes vont en croissant; et la seconde comme convergente, parce que les 1000 premiers termes vont en décroissant et que cette décroissance est d'abord très rapide.

Les deux règles sont légitimes: la première, dans les recherches théoriques; la seconde, dans les applications numériques. Toutes deux doivent régner, mais dans deux domaines séparés et dont il importe de bien connaître les frontières.

.....

Le premier exemple qui a montré clairement la légitimité de certains développements divergents est l'exemple classique de la série de Stirling. Cauchy a montré que les termes de cette série vont d'abord en décroissant, puis en croissant, de sorte que la série diverge; mais si l'on s'arrête au terme le plus petit, on représente la fonction eulérienne avec une approximation d'autant plus grande que l'argument est plus grand. ”

Lo sviluppo di queste considerazioni conduce alla teoria della stabilità esponenziale e superesponenziale, di cui parlerò nel paragrafo 4.

Nel discutere il problema della dinamica del sistema solare e della sua stabilità non si può neppure passare sotto silenzio il contributo della simulazione numerica. I metodi classici di calcolo delle effemeridi sono stati sostituiti, nella maggior parte dei casi, da metodi che si riconducono in qualche modo all'integrazione numerica di equazioni differenziali. Ciò è, tutto sommato, una scelta naturale: la velocità dei calcolatori elettronici consente oggi di svolgere una mole di calcoli impensabili per i calcolatori (intesi come persone che applicavano gli algoritmi di calcolo) degli osservatori astronomici dell'ottocento o della prima metà del novecento. Ma il calcolo numerico non si è fermato qui: già all'inizio degli anni 50 del secolo ormai trascorso fisici ed astronomi iniziarono a far uso del calcolatore per lo studio della dinamica di sistemi nonlineari. Posso citare come esempi significativi: il rapporto dei laboratori di Los Alamos N. LA-1940, ad opera di Fermi, Pasta ed Ulam^[13], dal titolo significativo “Studies of nonlinear problems” (1955); i primi calcoli di orbite galattiche svolti da Per-Olof Lindblad e Contopoulos[‡] a partire dal 1956; il celeberrimo lavoro di Hénon e Heiles^[18] del 1964 in cui si mostra con evidenza come il comportamento caotico si presenti in modelli dall'apparenza estremamente semplice.

È del tutto spontaneo, ed altrettanto lecito, chiedersi quale sia la validità delle simulazioni numeriche, soprattutto quando vengono svolte facendo uso, ad esempio, di integrazioni di equazioni differenziali su intervalli temporali di lunghezza tale che una stima dell'errore globale dovuto alle inevitabili approssimazioni degli algoritmi numerici diventa

[‡] Si veda [11], §1.4.

di fatto impossibile. È opinione personale dell'autore di queste note che l'atteggiamento corretto nei confronti dell'uso del calcolatore nella simulazione dinamica sia quello assunto da M. Hénon in risposta ad una domanda specifica sull'attendibilità dei risultati dell'esplorazione numerica:^b

“ Il est certain qu'on ne peut rien prouver, au sens mathématique du terme, par des calculs de ce genre. Ce travail est a ranger dans la catégorie des expériences numériques, qui à mon avis ont la même valeur, ni plus ni moins, que des expériences de physique. ”

2. L'esistenza di moti quasi periodici: il teorema di Kolmogorov

Al Congresso Internazionale dei Matematici di Amsterdam del 1954 Kolmogorov presentò una relazione dal titolo *On the general theory of dynamical systems and Classical Mechanics*, nella quale espose il suo teorema sulla persistenza di tori invarianti in sistemi Hamiltoniani integrabili soggetti ad una piccola perturbazione. Il teorema fu pubblicato in una breve nota^[21] nella quale è delineata in modo sintetico ma completo la linea della dimostrazione. Un teorema analogo riferito al caso di mappe che conservano l'area venne pubblicato da Moser^[35] nel 1962. Un anno più tardi Arnold pubblicò la dimostrazione completa del teorema di Kolmogorov in due corposi lavori^{[1][2]}. Da qui prese le mosse quella che oggi viene chiamata *teoria KAM*.

Senza entrare nei dettagli tecnici, può essere utile almeno mettere in evidenza le due idee nuove introdotte da Kolmogorov, che si rivelarono cruciali. A tal fine occorre riflettere per un momento su quello che Poincaré chiamava *il problema generale della dinamica*: studiare un sistema canonico con Hamiltoniana

$$(1) \quad H(p, q) = h(p) + \varepsilon f(p, q; \varepsilon)$$

dove $q \in \mathbb{T}^n$ sono variabili angolari, $p \in \mathcal{G} \subset \mathbb{R}^n$ sono variabili d'azione definite in un aperto \mathcal{G} , ed ε è un (piccolo) parametro che controlla la taglia della perturbazione. Si assume senz'altro che le funzioni siano analitiche in tutte le variabili, e che in particolare $f(p, q; \varepsilon)$ sia sviluppabile in serie di potenze di ε convergente in un intorno di $\varepsilon = 0$.

La dinamica del sistema imperturbato, per $\varepsilon = 0$, è relativamente semplice. Scrivendo le equazioni di Hamilton

$$\dot{q} = \frac{\partial h}{\partial p} =: \omega(p), \quad \dot{p} = 0$$

si vede subito che lo spazio delle fasi $\mathcal{G} \times \mathbb{T}^n$ ammette una foliazione in tori invarianti parametrizzati dalle azioni p , e il moto su ciascun toro è quasi periodico con frequenze

^b “È certo che calcoli del genere non dimostrano nulla nel senso matematico del termine. Questo lavoro si deve collocare nella categoria delle esperienze numeriche, che a mio parere hanno lo stesso grado di validità, né più né meno, delle esperienze di fisica.” La citazione è tratta dalla discussione riportata alla fine del lavoro [41].

$\omega(p) = \frac{\partial h}{\partial p}(p)$. La dinamica del sistema solare in approssimazione Kepleriana è sostanzialmente di questo tipo, con la differenza non banale che le frequenze invece di essere $3N$, pari al numero di gradi di libertà del sistema, sono solo N . Le $2N$ frequenze mancanti vengono reintrodotta dalla teoria di Lagrange e Laplace.

La domanda che si pone è cosa accada ai tori invarianti quando si ponga $\varepsilon \neq 0$. La risposta data da Kolmogorov si può sintetizzare brevemente, senza eccessivi dettagli tecnici, come segue: *se l'Hamiltoniana imperturbata $h(p)$ è non degenere (nel senso che il suo hessiano non si annulla), allora per ε sufficientemente piccolo i tori invarianti caratterizzati da frequenze fortemente non risonanti persistono sotto la perturbazione.*

Per “frequenze fortemente non risonanti” si intendono quelle che soddisfano la cosiddetta *condizione diofantea*

$$(2) \quad |\langle k, \omega \rangle| \geq \frac{\gamma}{|k|^\tau}, \quad \gamma > 0, \quad \tau \geq n - 1, \quad 0 \neq k \in \mathbb{Z}^n .$$

È noto dalla teoria dei numeri che quasi tutte le frequenze soddisfano questa condizione (nel senso della misura).[‡]

La prima delle idee cruciali cui ho accennato sopra si basa sul fatto elementare seguente: si consideri l'Hamiltoniana

$$(3) \quad H(p, q) = \sum_j \lambda_j p_j + \mathcal{R}(p, q), \quad \mathcal{R}(p, q) = O(p^2),$$

dove $\lambda \in \mathbb{R}^n$ è il vettore delle frequenze, e $\mathcal{R}(p, q)$ è almeno quadratico nelle azioni p . Scrivendo le equazioni di Hamilton

$$\dot{q}_j = \omega + \frac{\partial \mathcal{R}}{\partial p_j}, \quad \dot{p}_j = -\frac{\partial \mathcal{R}}{\partial q_j}$$

si vede subito che il toro $p = 0$ è invariante, e che le orbite sul quel toro sono quasi periodiche con frequenze $\lambda_1, \dots, \lambda_n$: basta porre come dati iniziali $p(0) = 0$ e $q(0)$ arbitrario. Naturalmente, nulla possiamo sapere di ciò che accade ad orbite con dato iniziale $p(0) \neq 0$.

La tecnica dimostrativa si riconduce al procedimento classico di riduzione del sistema di equazioni ad una *forma normale*, in un senso che richiede qualche ulteriore considerazione. Molti dei procedimenti classici fanno uso della cosiddetta *forma normale di Birkhoff*. Si intende con questo che si cerca una trasformazione canonica di variabili prossima all'identità, ossia della forma

$$q = q' + \varepsilon \varphi_1(q', p') + \varepsilon^2 \varphi_2(q', p') + \dots, \quad p = p' + \varepsilon \psi_1(q', p') + \varepsilon^2 \psi_2(q', p') + \dots$$

che trasformi l'Hamiltoniana (1) in una funzione $H'(p')$ delle sole azioni. Se ciò fosse possibile, si concluderebbe immediatamente che tutte le soluzioni del sistema con Hamiltoniana

[‡] La dimostrazione è essenzialmente una generalizzazione in dimensione n dello schema con cui si dimostra che i razionali del segmento $[0, 1]$ hanno misura di Lebesgue nulla. Il lettore interessato potrà trovare una dimostrazione completa in [2] o in [16], p. 130.

della forma (1) sono quasi periodiche con n frequenze fondamentali $\omega'(p') = \frac{\partial H'}{\partial p'}(p')$. In altre parole, tutti i moti planetari sarebbero descrivibili mediante epicicli. La trasformazione viene costruita tipicamente in modo iterativo come composizione di una successione di infinite trasformazioni, oppure come sviluppo in serie di potenze del parametro ε in cui i coefficienti delle potenze successive del parametro ε vengono determinati in modo ricorsivo.

Che tale programma sia irrealizzabile in generale era ben noto a Poincaré: si tratta in buona sostanza del teorema di non esistenza di integrali primi già contenuto nella memoria presentata per il premio del Re di Svezia, e riesposto in dettaglio nel capitolo VIII dei *Méthodes Nouvelles*. Il fallimento è dovuto proprio all'esistenza di risonanze, che si concretizzano nei divisori nulli, o molto piccoli, che vengono introdotti ad ogni iterazione, ed il cui accumulo provoca la divergenza delle iterazioni o della serie risultante.

Il suggerimento di Kolmogorov consiste nel sostituire la forma normale (3) a quella di Birkhoff. Più in dettaglio, si inizia col determinare un toro p^* imperturbato (ovvero per $\varepsilon = 0$) del sistema (1) con frequenze $\omega(p^*) = \lambda$ di tipo diofanteo, e si sviluppa l'Hamiltoniana in serie di Taylor nell'intorno di p^* . Applicando anche la traslazione $p' = p - p^*$, e sopprimendo gli apici per semplificare la scrittura, si ottiene per l'Hamiltoniana un'espressione della forma

$$(4) \quad H(p, q) = \sum_j \lambda_j p_j + \frac{1}{2} \sum_{j,k} C_{j,k}(q) p_j p_k + A(q) + \sum_j B_j(q) p_j + O(p^3),$$

dove $C_{jk}(q)$ è una matrice simmetrica, e $A(q)$ e $B_j(q)$ sono piccoli con ε , perché derivano dallo sviluppo della perturbazione $\varepsilon f(p, q)$. Il problema è eliminare dall'Hamiltoniana proprio i contributi $A(q)$ e $B_j(q)$. Anche in questo caso il procedimento di eliminazione introduce i piccoli divisori che sono ormai una vecchia conoscenza, ma qui viene in aiuto la condizione diofantea sulle frequenze, associata alla seconda idea introdotta da Kolmogorov.

Descrivere in breve la seconda idea è impresa ardua: si tratta di una questione squisitamente tecnica, e quindi difficile da apprezzare per chi non abbia sufficiente familiarità con i metodi della teoria classica delle perturbazioni. Mi limiterò quindi ad accennare al fatto che la tecnica dimostrativa usata da Kolmogorov rinuncia agli sviluppi in serie di potenze del parametro perturbativo ε , facendo invece uso di procedimento di approssimazioni successive, che Kolmogorov stesso chiama “generalized Newton method”, la cui rapidità di convergenza ricorda appunto quella del metodo di Newton per la ricerca degli zeri di una funzione. In effetti, la chiave della dimostrazione di Kolmogorov sta proprio nel fatto che la superconvergenza riesce a controllare l'azione per così dire esplosiva dell'accumulo dei piccoli divisori. Purtroppo, quest'ultima affermazione si capisce, credo, solo ripercorrendo in dettaglio la dimostrazione.

Non è infrequente trovare nella letteratura degli anni 60–70 l'affermazione che il teorema di Kolmogorov risponde positivamente alla domanda se il sistema solare sia stabile, ed in particolare se i moti planetari siano rappresentabili come funzioni quasi periodiche del tempo, ovvero mediante epicicli (si veda ad esempio [36], Cap. I, § 2). Vedremo più avanti che tali affermazioni peccavano di eccessivo ottimismo.

In ogni caso potremmo parlare di stabilità delle orbite solo in senso probabilistico:

la maggior parte dei dati iniziali dà origine ad orbite quasi periodiche descritte come moto su tori invarianti della teoria KAM. Ciò che rende il teorema di Kolmogorov meno attraente dal punto di vista dell'applicazione ad un sistema reale è l'aspetto topologico: il complemento dei tori invarianti è aperto e denso. D'altra parte, proprio nella regione che costituisce il complemento dei tori invarianti si deve celare il comportamento caotico delle orbite previsto da Poincaré.

3. Il ruolo delle risonanze

Il quadro che emerge dalla discussione svolta fin qui induce a pensare che la stabilità del sistema solare possa dipendere dall'esistenza o meno di risonanze tra le frequenze dei moti planetari.[†] In effetti, non si tratta di una faccenda del tutto nuova: l'idea che la sorte delle orbite planetarie potesse dipendere dalla razionalità o irrazionalità dei rapporti tra le frequenze era già stata avanzata nell'ottocento, alla luce della constatazione che i procedimenti perturbativi formali entravano sistematicamente in difficoltà in caso di risonanza. Ma non era facile far convivere un'idea del genere con l'apparente regolarità dei moti planetari, che sembravano non preoccuparsi troppo delle difficoltà formali che tanto assillavano i matematici. La difficoltà nell'accettare certe bizzarrie è ben riassunta da un commento di Weierstrass, riportato da Moser:^[36]

“ Thus we see that the stability cannot depend on rationality or irrationality of certain quantities. Precisely this error has been committed by Biot. Ever since this mistake has been made it is spread in big lectures over the organization of the world. He said, a slight change in the distance between Saturn and Jupiter would be sufficient that the strangest of all planets in our system would escape our system forever. However, one forgot to mention that also Jupiter could escape, and this would indeed simplify the work of astronomers considerably, since this is precisely the planet causing the largest perturbation. ”

Ma procediamo con la teoria matematica. Il teorema di Kolmogorov non dice nulla sulla dinamica nel complemento dell'insieme dei tori invarianti, che è aperto e denso, e per più di due gradi di libertà è connesso. Non si può escludere che un'orbita possa visitarlo tutto nel corso del tempo, seguendo, come già aveva intuito Poincaré, le connessioni eterocline tra orbite periodiche. Il primo esempio che mette in evidenza un tal fenomeno è

[†] Ad esempio, è noto dai tempi di Lagrange e Laplace che il rapporto tra le frequenze di Giove e Saturno è vicino a $2/5$, e che questa risonanza ha effetti osservabili sulle orbite. Più recentemente Moltchanov ha avanzato l'ipotesi che i piccoli effetti dissipativi presenti in qualunque sistema fisico reale facciano evolvere il sistema solare verso stati caratterizzati da risonanze, ed ha anche compilato un paio di tabelle delle risonanze del sistema solare.^[29] Ma, osservava Moser, dati 9 numeri reali (in questo caso le frequenze del moto ellittico dei nove pianeti) è sempre possibile trovare un certo numero di relazioni di risonanza che siano soddisfatte entro certi limiti di accuratezza.

stato proposto da Arnold^[3]; da qui il termine *diffusione di Arnold*. Il modello appare però alquanto artificiale: se la diffusione di Arnold sia un fatto generico in sistemi hamiltoniani è stato oggetto di dibattito negli ultimi decenni, e la discussione non può ancora dirsi chiusa.

Supponiamo ora che tale fenomeno esista, il che, allo stato attuale delle nostre conoscenze, può ben considerarsi del tutto plausibile. Nel caso del sistema solare questo significa, ad esempio, che l'eccentricità dell'orbita di un pianeta potrebbe crescere sistematicamente finché l'orbita arrivi ad intersecare quella di altri pianeti, rendendo così possibile una collisione o almeno un incontro ravvicinato in grado di modificare in modo consistente le orbite. Fenomeni analoghi potrebbero interessare i semiassi maggiori delle orbite, con la possibilità che un pianeta finisca col lasciare il sistema solare o precipitare sul sole, o che la luna si sganci dalla terra, e così via. Una simile sorte potrebbe essere riservata alle inclinazioni delle orbite rispetto al piano medio: un fenomeno all'apparenza non preoccupante, se non fosse che le orbite ad alta inclinazione sono instabili.

3.1 *Il caos nelle orbite planetarie*

I primi dubbi sull'applicabilità del teorema di Kolmogorov al sistema solare furono sollevati da Hénon. Egli osservò ben presto che le conclusioni che si traevano dal teorema potevano essere ottimistiche ed un po' affrettate. In effetti, pur senza entrare in dettagli tecnici complicati sulla verifica di tutte le ipotesi, resta l'osservazione non banale che il teorema si applica a condizione che le mutue perturbazioni dei pianeti siano sufficientemente piccole. Ora, se si fa uso delle stime contenute nei lavori di Moser ed Arnold, si conclude grosso modo che la condizione di applicabilità del teorema di Kolmogorov al sistema solare è che la massa di Giove sia largamente inferiore a quella del protone. Naturalmente, le stime analitiche sono sempre pessimistiche, ma l'osservazione di Hénon lascia almeno intendere che il problema è tutt'altro che risolto.

In mancanza di strumenti teorici, alcuni astronomi iniziarono una serie di simulazioni della dinamica del sistema solare su tempi sempre più lunghi. Tra i primi studi numerici significativi possiamo senz'altro annoverare quello di Milani, Nobili e Carpino^{[28][9]}, che portarono a termine un'integrazione sull'arco di qualche centinaio di milioni di anni mettendo in evidenza le variazioni secolari dei semiassi maggiori dei pianeti. Negli anni successivi il considerevole incremento della potenza di calcolo degli elaboratori (senza trascurare la drastica riduzione dei costi) ed il raffinamento delle tecniche analitiche consentirono calcoli che coprivano un arco di tempo paragonabile all'età stessa del sistema solare: ricorderò tra questi i lavori di Sussman e Wisdom^[42], Laskar^{[22][23]} e Murray e Holman^[37].

La base comune a tutti questi calcoli è l'integrazione numerica delle equazioni di Newton per il sistema planetario, o almeno per una parte di esso. Gli strumenti usati, oltre all'evoluzione temporale di semiassi maggiori, eccentricità ed inclinazioni delle orbite, sono l'analisi delle frequenze dei moti planetari condotta sia mediante i metodi classici delle trasformate di Fourier con le varianti tradizionali, sia con metodi che permettono una determinazione più precisa delle frequenze stesse, quali ad esempio il metodo denominato

da Laskar “analisi in frequenza”, sia mediante il calcolo degli esponenti di Lyapounov.[‡] Accennerò qui ai risultati di Laskar, dato che il suo lavoro ha avuto maggior risonanza rispetto ad altri.

La tesi di Laskar, esplicitamente enunciata in [22], è che la dinamica del sistema solare sia caotica, e non quasi periodica. L’affermazione si fonda sulla constatazione che le indicazioni della simulazione numerica danno un esponente di Lyapounov dell’ordine di $1/(5 \times 10^6)$. Sulla base di queste indicazioni si conclude, ad esempio, che un errore di 15 metri nel determinare la posizione della terra sulla sua orbita si potrebbe amplificare fino a 150 milioni di chilometri (ovvero la distanza attuale della terra dal sole) su un arco di tempo di 100 milioni di anni, decisamente breve rispetto ai 5 miliardi di anni che rappresentano l’età stimata del sistema solare.

Di fronte ad una previsione così catastrofica vien da chiedersi come possa ancora esistere il sistema solare. In realtà occorre riflettere sul fatto che gli esponenti di Lyapounov misurano solo la velocità di allontanamento *locale* delle orbite, ma non danno nessuna indicazione su quanto sia estesa la regione che esse possono visitare, e che potrebbe essere piccola. In altre parole, l’esistenza di un moto caotico non è a priori incompatibile con la possibilità che quantità critiche quali i semiassi maggiori, le eccentricità e le inclinazioni dei pianeti abbiano variazioni piccole, perché il caos potrebbe interessare prevalentemente gli angoli e restringersi ad una zona piccola nelle azioni; ciò che avviene, infatti.

Lo stesso Laskar ha introdotto anche un’idea notevole: esaminare come evolvono le frequenze delle orbite. L’interesse del metodo, denominato *analisi in frequenza*, sta nel fatto che in tal modo si rivelano le deviazioni dal moto quasi periodico sotto forma di variazione nel tempo delle frequenze — un’informazione molto più dettagliata rispetto al mero calcolo di un indicatore globale, e tutto sommato grossolano, come il massimo esponente di Lyapounov. Le conclusioni possono così riassumersi. I quattro pianeti maggiori — Giove, Saturno, Urano e Nettuno — hanno frequenze ben determinate e praticamente costanti, sicché il loro moto è compatibile con la teoria di Kolmogorov. Non così accade invece per i pianeti minori, inclusa la Terra: le loro orbite presentano variazioni caotiche delle frequenze che sono incompatibili coi moti quasi periodici previsti dal teorema di Kolmogorov. L’immagine tradizionale del sistema solare come orologio perfetto ne esce perlomeno compromessa.

[‡] Un cenno al significato degli esponenti di Lyapounov è d’obbligo. In termini semplici, si valuta come evolve nel tempo la distanza $d(t)$ tra le due orbite che abbiano dati iniziali vicini. Nel caso di moti quasi periodici la piccola differenza tra le fasi provoca un incremento lineare nel tempo della differenza tra le fasi, sicché si ha ad esempio $d(t) \sim t$. Nel caso di moto caotico la distanza cresce come $\exp(\chi t)$, e χ viene detto esponente di Lyapounov dell’orbita. L’inverso di χ viene detto tempo di Lyapounov, e viene considerato come un’indicazione del tempo necessario perché si perda praticamente la memoria dei dati iniziali. Il lettore interessato ad una discussione più approfondita degli esponenti di Lyapounov e dei metodi per calcolarli potrà consultare, ad esempio, [4], e [5].

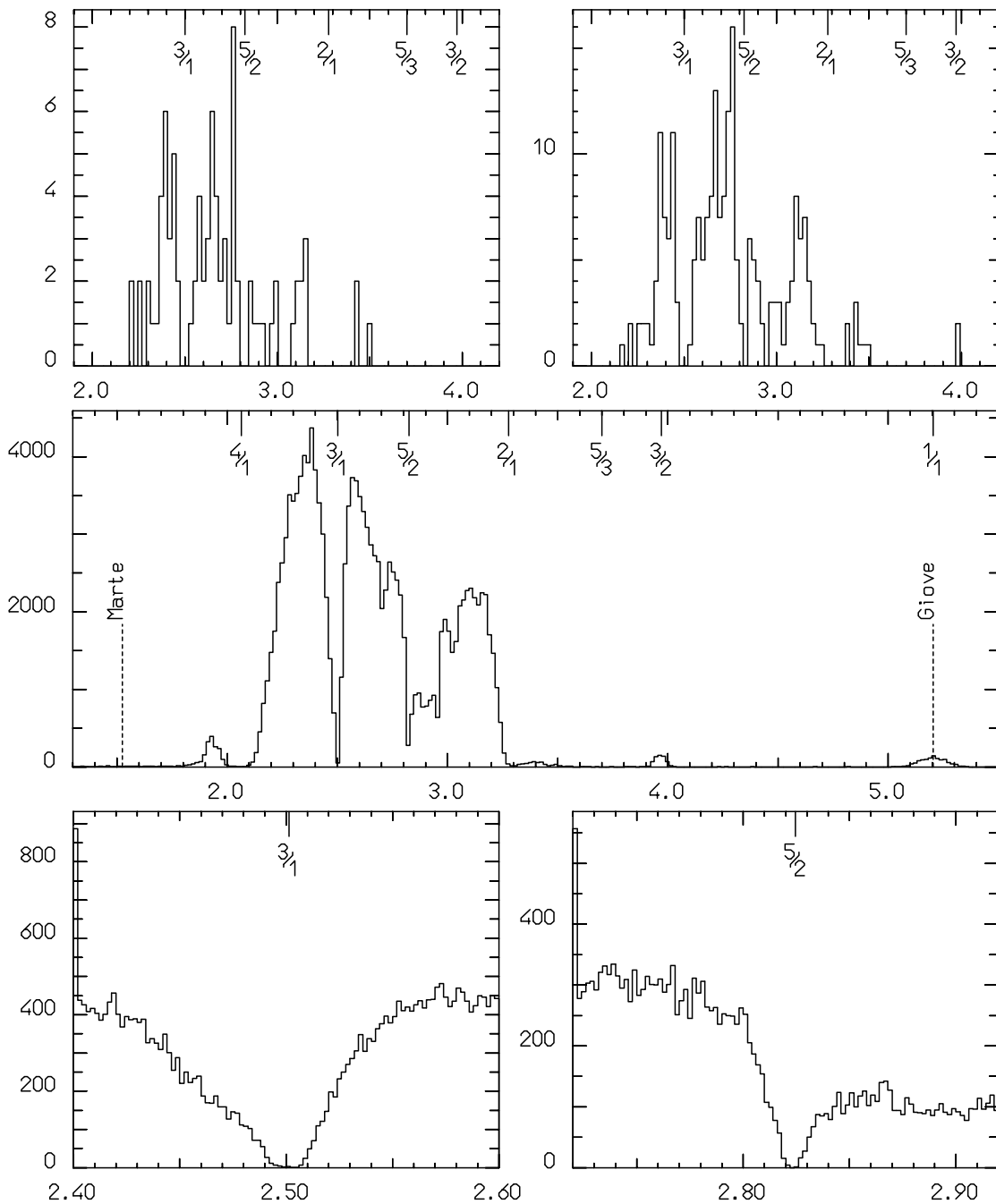


Figura 1. In alto: a sinistra la distribuzione degli asteroidi scoperti fino al 1866, anno della prima pubblicazione di Kirkwood; a destra quelli scoperti fino al 1884, anno della lettera di Kirkwood citata nel testo. Al centro: distribuzione degli asteroidi scoperti fino al dicembre 2006 (oltre 145 000). In basso: ingrandimento della regione intorno alle risonanze $3/1$ (a sinistra) e $5/2$ (a destra). In ascissa il semiasse maggiore dell'orbita, in unità astronomiche; in ordinata il numero di asteroidi in ciascun intervallo. I dati sono tratti dal catalogo del Minor Planet Center al 13 dicembre 2006.

3.2 *Gli asteroidi e le risonanze*

Come esempio illustrativo dell'azione delle risonanze possiamo considerare il caso delle *lacune di Kirkwood*.

L'ipotesi che il sistema solare tragga la sua origine dalla condensazione di una nube di polvere primordiale attorno al sole condurrebbe spontaneamente ad attendersi una distribuzione uniforme dei numerosi corpi celesti che si trovano nella fascia tra Marte e Giove. La scoperta che le cose stanno in modo ben diverso risale al 1866, quando Kirkwood notò che in corrispondenza ai valori dei semiassi maggiori per cui il moto medio dell'asteroide, calcolato grazie alla seconda legge di Keplero, è in risonanza con quello di Giove vi erano delle zone vuote.^[19] L'osservazione di Kirkwood fu ripresa da un altro astronomo, Proctor, che la considerò plausibile pur accennando alla possibilità che la posizione delle lacune fosse un fatto accidentale, considerato che gli asteroidi noti erano meno di un centinaio. In una lettera del 1884 Kirkwood osservò che le nuove scoperte intevente nel frattempo – 120 nuovi asteroidi – corroboravano la sua tesi.^[20] Il lettore potrà farsi un'idea della situazione osservando i due riquadri in alto della figura 1, ove è riportata in forma di istogramma la distribuzione dei semiassi maggiori degli asteroidi noti alla fine del 1866 (a sinistra) e del 1884 (a destra).

Sulle cause di un fatto dall'apparenza così strana lo stesso Kirkwood scrive:

“ In those parts of the zone of the minor planets where a simple relation of commensurability would obtain between the period of an asteroid and that of Jupiter the original planetary matter was liable to great perturbations. The result of such disturbance by the powerful mass of Jupiter was the necessary formation of gaps in the asteroid zone. ”

In altre parole, la risonanza avrebbe un effetto destabilizzante, capace di scacciare gli asteroidi dalle loro orbite originarie.

Per confronto, può essere interessante osservare nella fascia centrale della figura 1 che lo stesso istogramma, tracciato con i dati disponibili oggi (oltre 145 000 asteroidi), sembra confermare la tesi di Kirkwood. All'apparenza alcune lacune, quali quelle corrispondenti alle risonanze 3/1 e 5/2, sembrano non essere completamente vuote. Ma un istogramma localizzato intorno a quelle zone, quale quello riportato in basso nella figura 1, mostra che le lacune esistono davvero. Nel contempo però si osserva che vi sono dei gruppi di asteroidi in corrispondenza alle risonanze 3/2 e 1/1. Il primo di questi viene detto gruppo di *Hilda*, dal nome del primo asteroide scoperto. Il secondo gruppo è detto dei *Troiani* e dei *Greci*, perché ai primi asteroidi scoperti sono stati assegnati i nomi degli eroi dell'Iliade. Il semiasse maggiore di questi ultimi asteroidi è molto vicino a quello di Giove, il che significa che si trovano praticamente sulla stessa orbita. Si deve osservare che l'esistenza dei Troiani non era nota a Kirkwood: il primo di essi, Achille, fu scoperto solo nel 1906. La scoperta di Hilda è del 1875, e quindi esso era noto a Kirkwood al tempo della sua lettera del 1884, assieme ad un secondo asteroide (Ismene). Ma due asteroidi sono troppo pochi per pensare ad un gruppo associato alla risonanza.

L'esistenza di asteroidi in zone risonanti mette in discussione la convinzione di Kirk-

wood che la risonanza abbia necessariamente un effetto destabilizzante. In effetti, nel paragrafo 5 della prima parte di questa nota ho già messo in evidenza, discutendo la standard map, come la risonanza possa creare orbite periodiche instabili col loro corredo di separatrici, punti omoclini e conseguente comportamento caotico, ma anche orbite periodiche stabili che formano le “isole di moto ordinato nel mare di orbite caotiche”, per riprendere l’immagine suggestiva di Hénon.

La spiegazione delle lacune di Kirkwood è stata perseguita a lungo nell’ambito del *problema ristretto dei tre corpi*. Si considerano anzitutto due corpi, detti *primari*, in moto su un’orbita Kepleriana (eventualmente circolare) attorno al baricentro comune. Si aggiunge poi un terzo corpo, detto *planetoide*, di massa trascurabile e se ne studia il movimento sotto l’azione dei due primari.

Era già noto a Lagrange che il problema ristretto, circolare e piano ammette cinque soluzioni distinte tali che la posizione reciproca dei tre corpi resti invariata nel tempo. Se, grazie alla semplificazione del caso circolare, ci si pone in un riferimento che ruota uniformemente con i primari questo significa che i tre corpi occupano posizioni fisse, e per questo si parla di *equilibri lagrangiani*. Tre di queste soluzioni, dette *collineari*, vedono il planetoido in posizione fissa sulla retta congiungente i primari. Negli altri due casi i tre corpi sono disposti ai vertici di un triangolo equilatero, e si parla appunto di soluzioni *triangolari*. Lo studio della stabilità di queste ultime mediante approssimazione lineare delle equazioni porta a concludere che esse sono stabili a condizione che uno dei primari abbia massa abbastanza piccola rispetto all’altro, condizione ampiamente soddisfatta sia nel caso del sistema Sole–Giove che del sistema Terra–Luna.

L’esistenza degli equilibri triangolari getta un po’ di luce sulla dinamica degli asteroidi troiani. Se si considerano come primari il Sole e Giove, allora gli asteroidi di questo gruppo sono in librazione attorno a quegli equilibri, la cui stabilità lineare appare come una buona giustificazione del fatto che possano tranquillamente muoversi sulla stessa orbita di Giove senza esserne catturati.

Per le altre risonanze il problema è alquanto più complesso. In effetti, i numerosi tentativi di spiegare le lacune di Kirkwood e, allo stesso tempo, il gruppo di Hilda sono più o meno miseramente falliti. Uno studio svolto da Morbidelli e dall’autore mostra come le principali risonanze delle lacune di Kirkwood siano circondate da una zona fortemente stabile, dalla quale un asteroide non potrebbe sfuggire se non in tempi estremamente lunghi, e conclude che non esiste una spiegazione delle lacune di Kirkwood nell’ambito del problema dei tre corpi ristretto, anche tenendo conto del caso ellittico spaziale.^{[30][31]}

L’idea nuova, che ha cominciato a prender forma negli anni ’80 ed è stata ampiamente sviluppata nel corso del decennio successivo, è che non basti considerare solo le risonanze tra le frequenze kepleriane dei moti di pianeti e satelliti, ovvero, nel linguaggio degli astronomi, quelle dovute al *moto medio*. Occorre invece tener conto delle cosiddette *risonanze secolari*. In termini un poco più precisi: le orbite planetarie sono soggette alla precessione sia dei perielii che dei nodi — che sono gli angoli coniugati rispettivamente ad eccentricità ed inclinazioni. Le frequenze corrispondenti sono molto inferiori a quelle del moto medio: da qui l’aggettivo secolare loro attribuito. Il fatto è che anche queste

frequenze possono entrare in risonanza.

Nel caso degli asteroidi si è compreso che si deve tener conto delle risonanze secolari considerando, accanto all'azione di Giove, anche di quella di Saturno. Una mole considerevole di studi svolti negli ultimi vent'anni ha mostrato che la distribuzione attuale degli asteroidi è stata letteralmente scolpita da queste risonanze, la cui azione lenta (come se volessimo mettere in movimento un'altalena dandole un colpo al secolo), ma ripetuta per milioni o centinaia di milioni di anni, ha progressivamente eliminato gli asteroidi dalle regioni instabili. Ad esempio, nel caso della lacuna di Kirkwood corrispondente alla risonanza 3/1 l'effetto delle risonanze secolari è una variazione decisamente ampia dell'eccentricità dell'orbita dell'asteroide, che può raggiungere valori prossimi a 1. In tali condizioni l'orbita dell'asteroide arriva ad intersecare quella dei pianeti interni, rendendo possibile, se non l'evento catastrofico di una collisione, almeno quello di un incontro ravvicinato. In tali condizioni l'orbita dell'asteroide subisce modifiche consistenti, sicché esso può essere espulso dalla regione di risonanza, e magari cadere sul sole.

Una discussione approfondita di tutti i fenomeni qui accennati non può trovare spazio sufficiente in una breve nota come questa. Inoltre, negli anni più recenti la disponibilità di telescopi sempre più potenti ha permesso di osservare anche un gran numero di oggetti che hanno orbite esterne a quella di Nettuno, ed una dinamica in qualche modo simile a quella degli asteroidi. Il lettore interessato potrà consultare il libro di Morbidelli [33].

A conclusione di questo paragrafo possiamo ben affermare che le simulazioni numeriche riaprono la discussione che il teorema di Kolmogorov sembrava aver chiuso con pieno successo. Ma tale questione non può risolversi con pieno rigore ricorrendo ai soli metodi numerici. Mi si permetta dunque di tornare ad esporre alcuni strumenti e risultati teorici.

4. La stabilità per tempi lunghi

La discussione in cui vorrei addentrarmi ora richiede, per così dire, un cambio di prospettiva rispetto a quanto ho detto fin qui. Si tratta di rinunciare a cercare informazioni sulle orbite di un sistema perturbato per tempo infinito, e di accontentarsi di informazioni su tempi finiti, ma lunghi. Qui svolge un ruolo centrale il carattere asintotico delle serie perturbative di cui si parla nella pagina di Poincaré che ho riportato nel paragrafo 1.

Restando sempre nell'ambito del problema della stabilità del sistema solare potremmo formulare la domanda seguente: *possiamo garantire che i semiassi maggiori, le eccentricità e le inclinazioni delle orbite planetarie resteranno vicini a quelli attuali per tempi paragonabili all'età del sistema solare stesso, o all'età dell'universo?*

Il modo stesso in cui ho formulato la domanda mette in evidenza il fatto che per un sistema fisico sembra ragionevolmente semplice stabilire cosa significhi "tempo lungo": basta far riferimento alla durata stessa del sistema. Così, per il sistema solare possiamo prendere come riferimento l'età stimata dell'universo, che corrisponde a circa 10^9 rivoluzioni di Giove; per un acceleratore di particelle basta considerare pochi giorni, corrispondenti a

circa 10^9 rivoluzioni del fascio; per una galassia l'età dell'universo corrisponde a qualche centinaio di rivoluzioni delle stelle che la formano, il che, per inciso, significa che una galassia può ben esistere anche se la sua dinamica non è molto stabile: un piccolo vortice di polvere che in breve si dissolve nello spazio infinito.

Dal punto di vista matematico la questione si presenta ben più difficoltosa, non essendovi alcun criterio plausibile per stabilire se un tempo sia lungo o no. Per semplificare il problema facciamo riferimento, ad esempio, ad un punto di equilibrio di un sistema di equazioni differenziali. Possiamo porci il problema seguente: *sia Δ_ρ un intorno di raggio ρ del punto di equilibrio; dimostrare che qualunque orbita con dato iniziale $x(0) \in \Delta_\rho$ soddisfa $x(t) \in \Delta_{2\rho}$ per un tempo $|t| < T(\rho)$ che tenda all'infinito per $\rho \rightarrow 0$.*

Nel caso del sistema solare, o del problema generale della dinamica, si può porre la domanda in termini simili: *mostrare che le azioni p del sistema soddisfano $|p(t) - p(0)| \simeq \varepsilon^b$ (con $b > 0$) per un tempo $T(\varepsilon)$ che tenda all'infinito per $\varepsilon \rightarrow 0$.*

L'obiezione naturale è che una tale formulazione sembra ricondursi alla continuità delle soluzioni rispetto ai dati iniziali. Ma l'esempio elementare dell'equazione $\dot{x} = x$ mostra immediatamente che la mera continuità è una proprietà troppo debole per essere utile. In effetti, si vede subito che in quest'ultimo caso la distanza di un qualunque punto $x \neq 0$ dall'equilibrio $x = 0$ raddoppia in un tempo $T = \log 2$.

Un modo per caratterizzare la stabilità su tempi finiti consiste proprio nell'esaminare quale forma assuma $T(\varepsilon)$. Lo sviluppo storico della teoria conduce ad identificare cinque possibilità che sono state effettivamente esplorate.

- i. $T(\varepsilon) \simeq 1/\varepsilon$: è la teoria degli *invarianti adiabatici*, ampiamente usata in fisica, ed alla quale si riconduce in sostanza la teoria di Laplace e Lagrange per il sistema solare.
- ii. $T(\varepsilon) \simeq 1/\varepsilon^r$ con $r > 1$: è la teoria della *stabilità completa*, introdotta da Birkhoff proprio nel caso di punti di equilibrio di sistemi Hamiltoniani ([8], cap. IV, §2 e §4).
- iii. $T(\varepsilon) \simeq \exp(1/\varepsilon^a)$ con $0 < a \leq 1$: è la teoria della *stabilità esponenziale*, anticipata da Moser^[34] e Littlewood^{[24][25]}, e formulata in modo completo da Nekhoroshev^{[38][39]}.
- iv. $T(\varepsilon) \simeq \exp(\exp(1/\varepsilon^a))$ con $0 < a \leq 1$: è la *stabilità superesponenziale*, messa in luce da Morbidelli e dall'autore^[32].
- v. $T(\varepsilon) = \infty$: è la *stabilità perpetua*, alla quale Lyapounov ha dedicato lunghi studi ed a cui si fa solitamente riferimento in matematica.

Alla luce della discussione condotta fin qui, che in pratica esclude per il sistema solare la possibilità della stabilità perpetua, discuterò brevemente i primi tre casi. Ometto la discussione della stabilità superesponenziale perché richiede una familiarità con gli aspetti tecnici degli altri risultati che non è possibile trasmettere in queste poche pagine.

4.1 La teoria adiabatica e la stabilità completa alla Birkhoff

Per illustrare il procedimento possiamo considerare il caso semplificato di un sistema la cui Hamiltoniana abbia la forma

$$(5) \quad H(p, q) = \sum_{j=1}^n \omega_j p_j + \varepsilon f(p, q) ,$$

dove $p \in \mathcal{G} \subset \mathbb{R}^n$ sono le variabili d'azione, $q \in \mathbb{T}^n$ sono le variabili angolari, ed ε è un parametro che si suppone piccolo. L'Hamiltoniana è simile a quella del problema generale della dinamica, salvo che per $\varepsilon = 0$ si ha un sistema completamente isocrono. Il problema appare alquanto artificiale, ma costituisce la base per lo studio della dinamica sia nell'intorno di un punto di equilibrio ellittico, sia nell'intorno di un toro KAM.

Si ricorre alla forma normale di Birkhoff, di cui ho già detto nel paragrafo 2. Si cerca una trasformazione canonica della forma

$$(6) \quad q = q' + \varepsilon\varphi(p', q') , \quad p = p' + \varepsilon\psi(p', q')$$

che ponga l'Hamiltoniana nella forma

$$(7) \quad H'(p', q') = \sum_{j=1}^n \omega_j p_j + \varepsilon Z(p') + \varepsilon^2 f'(p', q') ,$$

simile alla (5) da cui siamo partiti, ma con la caratteristica rilevante che la dipendenza dagli angoli è diventata di ordine ε^2 . L'operazione è possibile a condizione che le frequenze siano non risonanti: la costruzione della generatrice della trasformazione canonica rimette in circolo i piccoli divisori che ormai sono diventati il ritornello di questa nota.

La teoria della stabilità alla Birkhoff si fonda sull'iterazione del procedimento che ho appena illustrato: con una successione di trasformazioni canoniche della forma (6), ma con $\varepsilon^2, \varepsilon^3, \dots$ anziché ε , si eliminano progressivamente gli angoli dalla forma normale, sempre assumendo che le frequenze siano non risonanti. Supponendo di aver compiuto r iterazioni si arriverà ad un'Hamiltoniana in forma normale di Birkhoff fino all'ordine r

$$(8) \quad H^{(r)} = \sum_j \omega_j p'_j + \varepsilon Z_1(p') + \dots + \varepsilon^r Z_r(p') + \varepsilon^{r+1} f^{(r)}(p', q') .$$

Qui si applica l'argomento di Birkhoff. Per $r \rightarrow \infty$ il procedimento diverge, in generale, ma se ci si ferma dopo r passi il problema della convergenza non si pone, e si può ben affermare che *esiste un numero positivo B_r tale che valga $|f^{(r)}(p, q)| < B_r \varepsilon^{r+1}$* . Vediamone le conseguenze. Le nuove variabili p', q' differiscono dalle p, q originarie per una quantità ε , per la (6). Esse obbediscono alle equazioni

$$(9) \quad \dot{q}'_j = \omega_j + \varepsilon \frac{\partial Z_1}{\partial p'_j} + \dots + \varepsilon^r \frac{\partial Z_r}{\partial p'_j} + \varepsilon^{r+1} \frac{\partial f^{(r)}}{\partial p'_j} , \quad \dot{p}'_j = -\varepsilon^{r+1} \frac{\partial f^{(r)}}{\partial q'_j} .$$

In prima battuta, rimuoviamo il termine di ordine ε^{r+1} . Allora le equazioni assumono la forma

$$\dot{q}'_j = \Omega_j(p') , \quad \dot{p}'_j = 0 ,$$

dove $\Omega_j(p') = \omega_j + \varepsilon \frac{\partial Z_1}{\partial p'_j} + \dots + \varepsilon^r \frac{\partial Z_r}{\partial p'_j}$. Dunque, le nuove azioni p' restano costanti nel tempo, e gli angoli q' evolvono linearmente con frequenze che dipendono dal dato iniziale: in questa approssimazione il moto è quasi periodico. Se prestiamo attenzione alle azioni p che compaiono nell'Hamiltoniana da cui siamo partiti osserviamo che esse non sono

più costanti, ma la loro variazione temporale resta limitata, perché la differenza $p - p'$ è una quantità di ordine ε , e sono comunque funzioni quasi periodiche del tempo, perché la trasformazione che le lega alle nuove azioni p' è funzione periodica dei nuovi angoli q' . Si dice comunemente che questa variazione è frutto di una *deformazione* dei tori invarianti.[†]

Se ora consideriamo l'intero sistema (9), vediamo che all'effetto di deformazione si sovrappone una variazione molto lenta, la cui derivata è di ordine ε^{r+1} , che diventa visibile (ad esempio confrontabile con la deformazione) solo dopo tempi dell'ordine di $1/\varepsilon^r$. A questo secondo fenomeno si dà comunemente il nome di *diffusione*.

Possiamo riassumere la discussione dicendo che *esistono delle coordinate p', q' che differiscono dalle p, q di un quantità di ordine ε tali che le nuove azioni p' possono variare al più di una quantità di ordine ε per tempi dell'ordine di $1/\varepsilon^r$* . Questa è la teoria della stabilità completa alla Birkhoff, che per $r = 1$ si riduce alla teoria adiabatica.

4.2 La stabilità esponenziale e il teorema di Nekhoroshev

L'argomento alla Birkhoff si ferma necessariamente al risultato che abbiamo appena visto finché non siamo in grado di dare qualche informazione più dettagliata sulla costante B_r che entra nella stima della diffusione.

Ciò che si riesce a mostrare è che $B_r < C^r (r!)^a$, con una costante $a > 1$, tipicamente $a \simeq n$, il numero di gradi di libertà del sistema. Ottenere questa stima è faccenda tecnica, e risparmierei al lettore il calcolo che occuperebbe diverse pagine. Vorrei invece richiamare l'attenzione sul fatto che le serie della teoria delle perturbazioni si comportano come quella suggerita da Poincaré nella pagina che ho citato: il paradigma di riferimento è la serie è $\sum_{k \geq 0} k! \varepsilon^k$, ossia quella che un analista considera divergente ed un astronomo convergente, almeno per piccoli valori di ε .

Tornando alle equazioni (9), osserviamo che nella seconda abbiamo $\dot{p}'_j \simeq r! \varepsilon^{r+1}$. Se ora fissiamo ε , al crescere di r il membro di destra raggiunge un minimo per $r \simeq 1/\varepsilon$. Per avere un risultato ottimale occorre fermarsi a questo valore di r . Aggiungiamo una stima elementare, facendo uso della formula di Stirling:

$$r! \varepsilon^r \simeq e^{-r} r^r \varepsilon^r \simeq e^{1/\varepsilon},$$

ossia la derivata temporale di p' diventa esponenzialmente piccola con l'inverso del parametro di perturbazione. Si conclude che il tempo necessario perché si manifesti l'effetto della diffusione è dell'ordine di $e^{1/\varepsilon}$, quanto affermato dalla teoria della stabilità esponenziale. Littlewood commentava: "if not eternity, this is a considerable slice of it".

La proprietà di stabilità esponenziale messa in evidenza qui è locale: si applica solo all'intorno di un equilibrio o di un toro invariante di Kolmogorov, ove la parte imperturbata dell'Hamiltoniana assume la forma lineare che vediamo nella (5). Nel caso del problema

[†] Il lettore osserverà che il procedimento appare molto simile a quello di Kolmogorov. La differenza cruciale è che qui *tutte* le orbite sono quasi periodiche, anche quelle con frequenze risonanti, mentre il teorema di Kolmogorov assicura la persistenza dei soli tori fortemente non risonanti.

generale della dinamica la tecnica perturbativa diventa alquanto più complessa, perché occorre tener conto anche delle regioni ove le risonanze tra le frequenze generano un moto caotico.

La formulazione generale del teorema di Nekhoroshev, può sintetizzarsi come segue: se la parte imperturbata $h(p)$ dell'Hamiltoniana (1) soddisfa sul dominio $\mathcal{G} \in \mathbb{R}^n$ la condizione di convessità

$$|\langle C(p)v, v \rangle| \geq m\|v\|^2 \quad \forall v \in \mathbb{R}^n, \quad C_{jk}(p) = \frac{\partial^2 h}{\partial p_j \partial p_k}(p),$$

e se ε è abbastanza piccolo, allora per ogni dato iniziale in \mathcal{G} vale $|p(t) - p(0)| < \varepsilon^b$ per tutti i tempi

$$|t| \leq T(\varepsilon) = T_* \exp\left(\frac{1}{\varepsilon^a}\right),$$

con costanti positive $a < 1$ e $b < 1$. Nella formulazione originale di Nekhoroshev la condizione di convessità viene rimpiazzata da una condizione più debole, detta *steepness*.[†]

È interessante descrivere brevemente il quadro della dinamica che emerge dal teorema di Nekhoroshev, e che la sinteticità dell'enunciato non riesce ad esprimere: per farsene un'idea occorrerebbe ripercorrere la dimostrazione. Vorrei fare un tentativo, ricorrendo ad un'immagine suggestiva proposta dallo stesso Nekhoroshev.

Immaginiamo di seguire un'orbita nello spazio delle fasi, e precisamente nel complemento dell'insieme dei tori invarianti la cui esistenza è garantita dal teorema di Kolmogorov. Lasciando che gli angoli evolvano secondo le loro frequenze, possiamo immaginare la regione descritta dalla azioni come una foresta, in cui gli alberi svolgono il ruolo dei tori invarianti di Kolmogorov. In questa foresta possono ben esserci delle radure — le regioni dominate dalle risonanze. Alcune sono abbastanza grandi: quelle corrispondenti a risonanze di basso ordine (gli interi coinvolti nella relazione di risonanza $\langle k, \omega \rangle = 0$ sono piccoli); altre sono molto piccole, perché la larghezza delle regioni risonanti diminuisce esponenzialmente con l'ordine della risonanza stessa. Nelle regioni prive di risonanze di basso ordine gli alberi sono molto fitti — hanno misura relativa prossima a uno — e lasciano solo piccolissimi varchi. Il nostro cammino è inevitabilmente lento quando gli alberi sono fitti, e ci è permessa una breve accelerazione solo quando arriviamo ad una radura che possiamo attraversare di buon passo, per poi tornare ad addentrarci nel fitto degli alberi. Il tempo necessario per attraversare la foresta sarà inevitabilmente lungo. Anzi, ci dice Nekhoroshev, esponenzialmente lungo.

[†] Oltre che nei lavori originali di Nekhoroshev [38], [39] si può trovare un'esposizione semplificata della dimostrazione in [6],[7]. Il caso di una perturbazione dipendente dal tempo in modo non periodico, purché analitico, è discusso in [14]. Le note [16] contengono una dimostrazione completa, scritta con intenti didattici. È interessante anche segnalare il lavoro [27], che propone uno schema di dimostrazione che si discosta da quello originale, essendo fondato su argomenti aritmetici di approssimazione in teoria dei numeri.

5. Applicazione al sistema solare

Il problema dell'applicabilità del teorema di Kolmogorov al sistema solare si estende anche al teorema di Nekhoroshev. La domanda resta la stessa: se la soglia di applicabilità dei due teoremi sia abbastanza grande da includere il caso dei pianeti.

Pensare di rispondere a questa domanda facendo uso solo di strumenti analitici appare impresa disperata, se non impossibile. Per questo motivo si è fatto ricorso a metodi che chiamerò *semianalitici*: si ricorre alla manipolazione algebrica al fine di migliorare le stime della soglia di applicabilità del teorema.

L'osservazione che sta alla base è tutto sommato semplice: se con un calcolo esplicito di forma normale si riduce una perturbazione di intensità ε ad un resto di ordine ε^r con $r > 1$ la condizione di applicabilità del teorema diventa del tipo $\varepsilon^r < \varepsilon^*$, il che migliora sensibilmente la verifica delle condizioni di applicabilità del teorema ad un sistema assegnato.

Il calcolo esplicito della forma normale può svolgersi al calcolatore con metodi di manipolazione algebrica. Per comprenderne il significato in termini semplici si pensi a funzioni rappresentabili in serie di potenze di cui si dà un'espressione troncata ad un ordine finito, sicché si ha a che fare di fatto con polinomi di un grado fissato. In tal caso il polinomio è noto una volta che siano stati determinati i coefficienti, e le consuete operazioni algebriche di somma, prodotto e derivata vengono ricondotte ad operazioni sui coefficienti. Per svolgere il calcolo basta dunque avere a disposizione un programma che rappresenti in modo efficiente i polinomi e sia in grado di svolgere le operazioni appena menzionate. Per le serie trigonometriche si ha una situazione del tutto analoga: si fa ricorso ai polinomi trigonometrici. Resta naturalmente il problema che i coefficienti possono rappresentarsi solo con la precisione finita del calcolatore, ed in questo senso intendo l'aggettivo *semianalitici* che ho usato poco sopra sopra.[†] I risultati possono rendersi rigorosi ricorrendo alla cosiddetta

[†] A mia conoscenza, il primo programma di questo tipo è stato realizzato all'inizio degli anni '60 del secolo ormai trascorso da Contopoulos. Di pochi anni successivo è il lavoro di Gustavson^[17]. La tecnica è diventata poi ampiamente usata, e non è possibile dare una bibliografia completa anche perché programmi di questo genere non vengono quasi mai pubblicati. In stretta connessione con la discussione qui svolta posso citare il pacchetto sviluppato da Laskar, e da lui usato per trasformare le equazioni in una forma adatta ai fini dell'integrazione numerica di cui ho detto, e quello sviluppato dall'autore, utilizzato per i risultati di cui si parla poco più avanti. Ad una diversa categoria appartengono invece i manipolatori algebrici in grado di svolgere calcoli in forma simbolica. Tra i primi programmi di una certa consistenza possiamo ricordare *Macsyma* (spesso usato in Meccanica Celeste ed ormai fuori commercio, ma sostituito da una versione disponibile nell'ambito del free software sotto il nome leggermente modificato di *Maxima*) *Mathematica* e *Maple*. L'elenco, naturalmente, ben si guarda dall'essere esaustivo (e soprattutto non deve intendersi come una forma di pubblicità a favore dei pacchetti menzionati). Si deve osservare però che programmi di quest'ultimo tipo, se usati per calcoli di tipo perturbativo, risultano essere alquanto meno efficienti delle versioni sviluppate *ad hoc*, che svolgono calcoli in forma semianalitica.

aritmetica degli intervalli: invece di un coefficiente valutato a prezzo di errori di arrotondamento si determina un intervallo entro il quale il coefficiente è contenuto. Questo metodo è alla base delle tecniche di *dimostrazione assistita dal calcolatore* (orrenda traduzione del *computer assisted* inglese, ma ormai entrata nell'uso).

I metodi semianalitici sono stati ampiamente usati nel calcolo esplicito degli sviluppi perturbativi. Le applicazioni connesse con i teoremi di Kolmogorov e di Nekhoroshev sono ancora poche, ma alcuni risultati sono abbastanza promettenti.

Per quanto riguarda l'applicabilità del teorema di Kolmogorov si è osservato un miglioramento costante delle stime di applicabilità nel corso degli anni. Ricorderò qui solo i due risultati più recenti: il primo, dovuto a Celletti e Chierchia^[10], in cui si considera l'asteroide *Veritas* nell'ambito del problema ristretto dei tre corpi; il secondo, dovuto a Locatelli ed all'autore^[26], in cui si considera il problema generale dei tre corpi nel caso Sole–Giove–Saturno. L'interesse di questi lavori sta nel fatto che si dimostra l'esistenza di tori invarianti per i parametri (masse) ed i dati iniziali (elementi orbitali) effettivamente osservati, mentre in diversi lavori precedenti si era costretti a considerare masse più piccole o parametri orbitali modificati in modo da allontanare le orbite, diminuendo così l'azione reciproca dei pianeti. Resta, naturalmente, l'approssimazione del modello: considerare l'intero sistema solare è ancora fuori dalla portata di calcoli di questo genere.

L'applicabilità del teorema di Nekhoroshev è un problema ancor più complesso. Allo stato attuale si può dire che il problema realistico più significativo tra quelli studiati sia quello degli asteroidi Troiani. In questo caso si riesce a determinare una regione di dimensioni ragionevolmente ampie (circa 1/8 della distanza del Sole da Giove), caratterizzata da stabilità su tempi pari all'età dell'universo, ma la geometria non banale del problema consente solo di assicurare che in quella regione è contenuta una piccola frazione degli asteroidi effettivamente osservati^[15]. Facendo ricorso all'approssimazione di una mappa, metodo meno rigoroso ma efficace, Efthimiopoulos e Sándor sono riusciti a mostrare l'applicabilità della teoria a circa metà degli asteroidi noti^[12]. D'altra parte alcune ricerche numeriche svolte da Robutel, Gabern e Jorba^[40] tendono a mostrare come l'azione delle risonanze possa ben causare una lenta fuga di asteroidi dalla regione di librazione attorno ai punti lagrangiani: basta aspettare, appunto, un tempo dell'ordine di miliardi di anni, spesso ben superiore all'età dell'universo.

6. Finale

Non posso affermare che i risultati che ho discusso fin qui coprano tutto lo spettro delle ricerche nel campo della dinamica del sistema solare. Del resto, anche una discussione completa (che richiederebbe un numero di pagine ben maggiore) sarebbe inevitabilmente destinata ad una rapida obsolescenza: l'attività di ricerca nel campo dei sistemi dinamici, ed in particolare della dinamica del sistema planetario, è in una fase di grande sviluppo.

Credo però che pur non essendo in grado di rispondere in modo definitivo alla domanda se il sistema solare sia stabile si possano almeno chiarire alcuni aspetti che in passato sono

stati discussi a lungo, come ho documentato almeno in parte. Mi riferisco in particolare ai due problemi che ho ripetutamente richiamato in questa nota: la possibilità di descrivere il moto planetario con funzioni quasi periodiche (ovvero con gli epicicli) ed il ruolo delle risonanze nel generare comportamenti caotici.

Ho cercato di descrivere come questi due problemi conducano a visioni contrastanti: da un lato il sistema solare come orologio perfetto; dall'altro una dinamica dominata dall'imprevedibilità, se non dal caos.

Il teorema di Kolmogorov sembra appoggiare decisamente l'immagine dell'orologio quasi periodico perfetto, almeno in senso probabilistico. Ma la sua applicabilità è alquanto dubbia, a meno che non si consideri solo una parte alquanto ristretta del sistema. È ben vero che immaginare che un piccolo asteroide o addirittura un meteorite possa perturbare il moto di Giove e del Sole sembra essere ipotesi ridicola. Ma il principio di azione e reazione non lascia scampo: l'asteroide esercita su Giove una forza eguale e contraria a quella che Giove esercita sull'asteroide. Vedere gli effetti di questa debolissima azione sull'orbita di Giove è solo questione di tempo: se avessimo la pazienza di attendere qualche miliardo di anni ce ne accorgeremmo.

L'osservazione che ho appena fatto apre una crepa nella descrizione della dinamica che emerge dal teorema di Kolmogorov. Se consideriamo, ad esempio, il sistema formato dai pianeti maggiori, allora possiamo immaginare di condensare l'azione di asteroidi, comete e quant'altro in una debole dipendenza temporale dell'Hamiltoniana che descrive la dinamica del sistema solare. Ma una tale dipendenza, per quanto piccola, distruggerebbe immediatamente il quadro di Kolmogorov: le risonanze sono dense, ed una piccola variazione di un'orbita che giace su un toro invariante può ben accendere l'azione della risonanza.

Il teorema di Nekhoroshev appare più robusto: il risultato sopravvive anche ad una debole dipendenza temporale nell'Hamiltoniana^[14]. C'è un prezzo da pagare: si rinuncia ad avere una descrizione delle orbite su un tempo infinito. C'è un guadagno: il teorema si applica a condizioni iniziali in un aperto, e ciò è compatibile con gli inevitabili errori di osservazione.

Accettare di limitare la nostra conoscenza a tempi finiti ma lunghi sembra essere un modo efficace per risolvere l'apparente conflitto tra epicicli e risonanze. Se mi è permesso ricorrere a qualche immagine tratta dalla nostra esperienza, pensiamo alla lenta azione dei torrenti che nel corso dei secoli hanno modellato le valli delle nostre montagne scavando la roccia, o a quella delle onde del mare che hanno levigato i piccoli sassi del fondo marino formando la sabbia delle nostre spiagge, o ancora a quella del vento che ha sminuzzato le pietre fino a trasformarle nella sabbia dei deserti — “minuscoli frammenti della fatica della natura”, cantava De André. L'azione delle risonanze è qualcosa di simile, solo incomparabilmente più lenta. Agendo pazientemente per milioni e milioni di anni esse hanno scolpito la distribuzione degli asteroidi. Agendo per miliardi e miliardi di anni esse potrebbero ben modificare anche le orbite dei pianeti. L'osservazione un po' ironica di Weierstrass che esprimeva il suo scetticismo sul fatto che le risonanze potessero avere effetti catastrofici si dissolve di fronte a tempi così lunghi. La rappresentazione tramite epicicli, tanto cara a Weierstrass come a molti matematici dell'ottocento, può ben aver successo, e continua

ad essere uno strumento prezioso, se siamo interessati a periodi relativamente brevi — centinaia o migliaia di anni — ma può anche entrare in difetto su tempi che siano una frazione significativa dell'età del sistema solare.

L'idea emergente è che forse ponendo il problema della stabilità del sistema solare sulla base dei nostri modelli matematici che pretendono risultati su tempo infinito abbiamo sbagliato la domanda. Perché il sistema solare possa esistere basta che la sua dinamica ne assicuri la stabilità per il tempo necessario — di fatto, la sua durata.

Se accettiamo la visione che ho delineato, il ricorso a forme deboli di stabilità, quale quella di Nekhoroshev, sembra essere inevitabile. Significa che siamo disposti ad accettare che il miglior modello planetario che abbiamo saputo costruire — la gravitazione Newtoniana eventualmente corredata da correzioni relativistiche — contenga in sé anche la possibilità del disfacimento del sistema solare che esso descrive. Ciò non ci impedirà di stupirci di fronte al rosso di un tramonto, né di alzare il nostro sguardo alle stelle in segno di speranza.

*Oritur sol, et occidit, et ad locum suum revertitur,
 ibique renascens gyrat per meridiem, et flectitur ad aquilonem . . .
 . . . et proposui in animo meo quaerere et investigare sapienter
 de omnibus quae fiunt sub sole.
 Hanc occupationem pessimam dedit Deus filiis hominum,
 ut occuparentur in ea.*

Bibliografia

- [1] Arnold, V. I.: *Proof of a theorem of A. N. Kolmogorov on the invariance of quasi-periodic motions under small perturbations of the Hamiltonian*, Usp. Mat. Nauk, **18**, 13 (1963); Russ. Math. Surv., **18**, 9 (1963).
- [2] V.I. Arnold: *Small denominators and problems of stability of motion in classical and celestial mechanics*, Usp. Math. Nauk **18** N.6, 91 (1963); Russ. Math. Surv. **18** N.6, 85 (1963).
- [3] V.I. Arnold: *Instability of dynamical systems with several degrees of freedom*, Sov. Math. Dokl. **5** N.1, 581–585 (1964).
- [4] G. Benettin, L. Galgani, A. Giorgilli and J. M. Strelcyn: *Tous les nombres caractéristiques de Ljapunov sont effectivement calculables*, C. R. Acad. Sc. Paris **268** A, 431–433 (1978).
- [5] G. Benettin, L. Galgani, A. Giorgilli and J.M. Strelcyn: *Lyapunov characteristic exponents for smooth dynamical systems and for Hamiltonian systems; a method for computing all of them, part 1: theory*, Meccanica, 9–20 (1980); *Part 2: numerical applications*, Meccanica, 21–30 (1980).
- [6] G. Benettin, L. Galgani, A. Giorgilli: *A proof of Nekhoroshev's theorem for the stability times in nearly integrable Hamiltonian systems*. Cel. Mech., **37**, 1–25 (1985).

- [7] G. Benettin, G. Gallavotti: *Stability of motions near resonances in quasi-integrable Hamiltonian systems*. Journ. Stat. Phys., **44**, 293 (1986).
- [8] G.D. Birkhoff: *Dynamical systems*, New York (1927).
- [9] M. Carpino, A. Milani, A. Nobili: *Long term numerical integration and synthetic theories for the motion of outer planets*, Astronomy and Astrophysics **181**, 182–194 (1987).
- [10] A. Celletti, L. Chierchia: *KAM stability and Celestial Mechanics*, Memoirs of AMS, in stampa (2005).
- [11] G. Contopoulos: *Order and chaos in dynamical Astronomy*, Springer–Verlag (2002).
- [12] C. Efthimiopoulos, Z. Sándor: *Optimized Nekhoroshev stability estimates for the Trojan asteroids with a symplectic mapping model of co-orbital motion*, Mon. Not. R. Astron. Soc., (2005).
- [13] Fermi, E., Pasta, J. and Ulam, S.: *Studies of nonlinear problems*, Los Alamos document LA–1940 (1955).
- [14] A. Giorgilli, E. Zehnder: *Exponential stability for time dependent potentials*, ZAMP (1992).
- [15] A. Giorgilli, Ch. Skokos: *On the stability of the Trojan asteroids*, Astron. Astroph. **317**, 254–261 (1997).
- [16] A. Giorgilli: *Notes on exponential stability of Hamiltonian systems*, in *Dynamical Systems, Part I: Hamiltonian systems and Celestial Mechanics* Pubblicazioni del Centro di Ricerca Matematica Ennio De Giorgi, Pisa, 87–198 (2003).
- [17] F.G. Gustavson: *On constructing formal integrals of a Hamiltonian system near an equilibrium point*, Astron. J. **71** (1966), 670–686.
- [18] M. Hénon, C. Heiles: *The applicability of the third integral of motion: some numerical experiments*, Astron. J. **69**, 73–79 (1964).
- [19] D. Kirkwood: *Proceedings of the American Association for the Advancement of Science for 1866*.
- [20] D. Kirkwood: *The zone of asteroids and the ring of Saturn*, Astronomical Register, **22** 243–247 (1884).
- [21] A.N. Kolmogorov: *Preservation of conditionally periodic movements with small change in the Hamilton function*, Dokl. Akad. Nauk SSSR, **98**, 527 (1954).
- [22] J. Laskar: *A numerical experiment on the chaotic behaviour of the solar system*, Nature, **338**, 237–238 (1989).
- [23] J. Laskar: *Large scale chaos in the solar system*, Astron. Astroph. **287** (1994).
- [24] J.E. Littlewood: *On the equilateral configuration in the restricted problem of three bodies*, Proc. London Math. Soc.(3) **9**, 343–372 (1959).
- [25] J.E. Littlewood: *The Lagrange configuration in celestial mechanics*, Proc. London Math. Soc.(3) **9**, 525–543 (1959).
- [26] U. Locatelli, A. Giorgilli: *Invariant tori in the Sun–Jupiter–Saturn system*, DCDS-B **7**, 377 - 398 (2007).
- [27] P. Lochak: *Canonical perturbation theory via simultaneous approximations*, Usp. Math. Nauk. **47**, 59–140 (1992). English transl in Russ. Math. Surv.

- [28] A. Milani, A. Nobili, M. Carpino: *Secular variations of the semimajor axes: theory and experiments*, Astronomy and Astrophysics **172**, 265–269 (1987).
- [29] A. Moltchanov: *The resonant structure of the solar system*, Icarus **8**, 203–215 (1968).
- [30] A. Morbidelli e A. Giorgilli: *On the dynamics in the asteroids' belt. Part I: general theory*, Cel. Mech. **47**, 145–172 (1990).
- [31] A. Morbidelli e A. Giorgilli: *On the dynamics in the asteroids' belt. Part II: detailed study of the main resonances*, Cel. Mech. **47**, 173–204 (1990).
- [32] A. Morbidelli, A. Giorgilli: *Superexponential stability of KAM tori*, J. Stat. Phys. **78**, 1607–1617 (1995).
- [33] A. Morbidelli: *Modern celestial Mechanics. Aspects of solar system dynamics*, Taylor & Francis, London (2002).
- [34] J. Moser: *Stabilitätsverhalten kanonischer differentialgleichungssysteme*, Nachr. Akad. Wiss. Göttingen, Math. Phys. Kl IIa, nr.6, 87–120 (1955).
- [35] J. Moser: *On invariant curves of area-preserving mappings of an annulus*, Nachr. Akad. Wiss. Göttingen, Math. Phys. Kl II, 1–20 (1962).
- [36] Moser, J.: *Stable and random motions in dynamical systems*, Princeton University press, Princeton (1973).
- [37] N. Murray, M. Holman: *The Origin of Chaos in Outer Solar System*, Science, **283**, Iss. 5409, 1877 (1999).
- [38] N.N. Nekhoroshev: *Exponential estimates of the stability time of near-integrable Hamiltonian systems*. Russ. Math. Surveys, **32**, 1 (1977).
- [39] N.N. Nekhoroshev: *Exponential estimates of the stability time of near-integrable Hamiltonian systems, 2*. Trudy Sem. Petrovs., **5**, 5 (1979).
- [40] P. Robutel, F. Gabern, A. Jorba: *The observed Trojans and the global dynamics around the Lagrangian points of the Sun–Jupiter system*, Celest. Mech. Dyn. Astr. **92**, 53–69 (2005).
- [41] J. Roels e M. Hénon: *Recherche des courbes invariantes d'une transformation ponctuelle plane conservant les aire*, Bull. Astron. **32** 267–285 (1967).
- [42] G.J. Sussman, J. Wisdom: *Numerical evidence that the motion of Pluto is chaotic*, Science, **241**, 433–437 (1988).