Mapping the Arnold web with a graphic processing unit

A. Seibert, a) S. Denisov, A. V. Ponomarev, and P. Hänggi
Institute of Physics, University of Augsburg, Universitätsstr. 1, D-86159 Augsburg, Germany

(Received 25 July 2011; accepted 18 October 2011; published online 14 November 2011)

The Arnold diffusion constitutes a dynamical phenomenon which may occur in the phase space of a non-integrable Hamiltonian system whenever the number of the system degrees of freedom is $M \geq 3$. The diffusion is mediated by a web-like structure of resonance channels, which penetrates the phase space and allows the system to explore the whole energy shell. The Arnold diffusion is a slow process; consequently, the mapping of the web presents a very time-consuming task. We demonstrate that the exploration of the Arnold web by use of a graphic processing unit-supercomputer can result in distinct speedups of two orders of magnitude as compared with standard CPU-based simulations.

Several archetype results of dynamical chaos theory can unambiguously be attributed to unforeseen outcomes of numerical experiments. Two such examples are the absence of thermalization in the Fermi-Pasta-Ulam (FPU) chain and the discovery of the Lorenz attractor. In some cases, the use of numerical simulations presents the only method to obtain insight into the behavior and evolution of the system of interest. There are several branches of modern physics, which are computational by their own nature, with Computational Cosmology being such a paradigm. Nowadays, computational cosmologists perform simulations with more than $10^{10}$ particles. The main driving force underpinning this advance is the high parallelization on the scale of much smaller systems. For example, the averaging over many realizations of a stochastic force or of quenched disorder to arrive at scaling exponents, the Monte-Carlo sampling etc., present typical routines in many areas of computational physics. It is evident that by running $N$ different realizations on $N$ processors in parallel, one can speed up the statistical collection process by a factor of $N$. One of the possibilities is then to use a computational cluster by sending the program to run on many CPUs simultaneously, collecting the output data and finally sample them. This “sending-collecting-sampling” process is cumbersome work, which, however, could be avoided by using special designed scripts. The advent of graphics processing units (GPUs) has brought such numerical experiments into a new level of performance. Originally used as hardware chips, designed as graphic data-pipelines, GPUs were soon recognized as an additional beneficial tool: researchers from such areas as medical imaging, computational electromagnetics, and hydrodynamics, have successfully implemented them for data processing. Nowadays, computational physics is marked by an impressive boost of the “General Purpose Computing on GPU” (GPCC GPU) ideology. With this work, we attempt to demonstrate how the use of a GPU-supercomputer can provide useful insight for complex problems of nonlinear dynamics such as Arnold diffusion.

I. INTRODUCTION

The Arnold diffusion is not that kind of phenomenon that involves the time evolution of billions of particles. In fact, it can appear in the phase space dynamics of a Hamiltonian system whenever the number of system degrees of freedom is $M \geq 3$. The Arnold diffusion is relevant in celestial mechanics and astronomy, plasma dynamics in stellarators and tokamaks. It also influences the evolution of a Rydberg atom placed in crossed magnetic and electric fields or it might explain experimentally observed effects of emittance growth in the Tevatron colliders. Moreover, Arnold diffusion may be a mechanism of the ultraviolet cut-off in statistical mechanics. Except some specially designed models, where the presence of the diffusion can be rigorously proved and the diffusion timescales can theoretically be estimated, only little work is generally analytically possible. Therefore, numerical experiments play an prominent role in the studies of Arnold diffusion. However, conclusive numerical output then requires extremely large time scales; because the actual rate of the process is not known ab initio. With this study, we demonstrate that the presence of the Arnold diffusion in the dynamics of a model Hamiltonian system can be visualized by scanning the system phase space with a giant ensemble of trajectories. This has been realized within the compute unified device architecture (CUDA) framework with its benchmarks performed on a NVIDIA Tesla graphic processing unit (GPU). We detect the Arnold web, i.e., a chaotic network which can carry the system over the energy shell, in the limit of extremely weak perturbations. We also resolve a rich fractal structure of the Arnold web in the regime of moderate non-integrability, when the clusters of high-order resonances start to contribute to diffusion.

a)Author to whom correspondence should be addressed. Electronic mail: armin.seibert@physik.uni-augsburg.de.
II. MODEL FOR ARNOLD DIFFUSION

One of the typical models to study the Arnold diffusion is a system of coupled standard maps, see, for example, Refs. 20. Such systems are easy to handle because they only need to be iterated and do not demand sophisticated integration schemes. However, they represent a class of driven Hamiltonian systems, and their exposition to a train of delta-kicks leads to an energy pumping and an unlimited diffusion in the momentum subspace. Therefore, time evolution of such systems is not restricted to a compact manifold in the corresponding phase space.

Here, we want to address the case of autonomous Hamiltonian systems. Hamiltonians of such systems represent integrals of motion, and their time evolution is restricted to manifolds of constant energy.

An autonomous Hamiltonian system with one degree of freedom, \( M = 1 \), is always integrable since it possesses the integral of motion, \( E = H(x,p) \). Hamiltonian systems with two degrees of freedom, \( M = 2 \), evolve in a four-dimensional phase space, \( \Omega = \{(x,y,p_x,p_y)\} \), but the system evolution is confined to the energy shell of energy \( E = H(x(0),y(0),p_x(0),p_y(0)) \), determined by the initial conditions. Regular two-dimensional invariant manifolds, tori, separate the system energy shell into different regions. Different regions exhibit different dynamics, chaotic or regular ones, but each region is perfectly isolated from the remaining ones. This is so because the two-dimensional tori provide a complete partition of the three-dimensional energy shell. This topological argument does not work anymore in higher dimensions: the \( M \)-dimensional tori cannot partition the \( (2M - 1) \)-dimensional shell whenever the number of the degrees of freedom \( M \geq 3 \). Therefore, there could be trajectories that slip between regular tori and thus are able in exploring the whole energy shell.

To be more specific, we consider a three-dimensional system with a Hamiltonian

\[
H(P,X) = \frac{P^2}{2} + eH_p(X) = H_0(P) + eH_p(X),
\]

where \( P = (p_x,p_y,p_z) \) and \( X = (x,y,z) \) denote the momentum and coordinate vectors. For \( e = 0 \), the system is completely integrable, and the vector \( P \) is constant along any trajectory of the system, \( X(t) = X(0) + Pt \). For a given energy \( E \), the energy shell forms a sphere in the momentum subspace, \( S: P^2 = 2E \), and system trajectories are represented by fixed points on the sphere. Consider now the regime of weak perturbation, \( e \ll 1 \). This implies that the system phase space \( \Omega \) remains almost completely filled with invariant tori, which form the set \( \Omega_{\text{tori}} \). The motion is regular on the manifold \( \Omega_{\text{tori}} \), but there also appears a tiny manifold, the relative complement of \( \Omega_{\text{tori}} \) in \( \Omega \), which constitutes the Arnold web, \( \Omega_{\text{web}} = \Omega \setminus \Omega_{\text{tori}} \). The Arnold web covers the resonance lines \( \Omega_{\text{res}} = \{ P \text{im} \cdot P = 0 \} \), where \( m = (m_x,m_y,m_z) \) is a triplet of coprimes, by narrow chaotic channels. A typical channel width depends on the order of the corresponding resonance, \( m = \text{max} m \), and it usually decreases with the increase of \( m \). The total volume occupied by the Arnold web is expected to scale as \( \sqrt{e} \).

The appearance of the Arnold diffusion may take a while. The motion along the web can be detected only when the change \( \Delta P(t) = \| P(t) - P(0) \| \) assumes a noticeable value. The Nekhoroshev theorem predicts that such a change can be observed after a time \( t_{\alpha} \), which scales at least exponentially with \( 1/e^{2} \). Moreover, only a tiny fraction of the initial conditions, which entered into the manifold \( \Omega_{\text{web}} \), can diffuse. Therefore, even the numerical detection of the Arnold diffusion is a rather difficult and time-consuming task.

Here, we consider a perturbation Hamiltonian of the explicit form

\[
H_p(x,y,z) = \cos(x) \cos(y)[1 + \cos(2z)].
\]

The total Hamiltonian \( H(P,X) \) in Eqs. (1) and (2) cannot be separated into several independent two- or one-dimensional Hamiltonians; therefore, the system is manifestly three-dimensional. The system in Eqs. (1) and (2) falls within the range of Nekhoroshev theorem since the Hamiltonian \( H_0 \) is convex and the function \( H_p(X) \) is analytic. Moreover, the perturbation potential (2) is bounded, \( |H_p(X)| \leq 2 \), and the total system energy, \( E = H(P,X) \), effectively controls the strength of perturbation. Therefore, we set \( e = 1 \) and introduce an effective perturbation parameter

\[
e_{\text{eff}} = \max |H_p(X)| / H(P,X) = 2/E.
\]

In the momentum subspace, the system evolves within the sphere \( S \). Namely, the system dynamics is confined to a thin layer enclosing the surface of the sphere \( S \), \( L_{\text{eff}}(P) \); \( 2(E - e_{\text{eff}}) \leq P^2 \leq 2(E + e_{\text{eff}}) \). Resonance lines \( m \cdot P = 0 \) form a set of circles on the surface of the sphere. In the limit \( e_{\text{eff}} \ll 1 \), the Arnold web represents a sparse net on \( S \). Upon increasing the perturbation strength, \( e_{\text{eff}} \), the Arnold web tends to become “thicker”: the invariant tori, which were located far outside of the low-order resonances (the distance is defined by the Diophantine resonance conditions), and were unaffected by the weak perturbation, become now modified. The increasing perturbation involves more and more resonances into the growing Arnold web. This all results in the appearance of a complex, fine-structured network braiding whole areas on the surface of the sphere.

III. COMPUTATIONAL METHOD

Numerical studies of Arnold diffusion demand long runs. Therefore, one should resort to the integrators which consistently respect the symplectic properties of the original systems. A symplectic numerical scheme replaces the original, continuous-time Hamiltonian by its discretized version, and the numerical propagation produces the exact time evolution of the corresponding Hamiltonian map. The so discretized system dynamics, thus, is a slightly perturbed version of the original one, so that there is no secular growth of the system integrals of motions, which is the total system energy \( E \) in our case. The smaller the time step of the corresponding Hamiltonian map is, the closer both systems, original and discretized version stay close to each other. For the propagation of the Hamiltonian in Eqs. (1) and (2), we used the sixth-order symplectic integrator from Ref. 25.
time step $h = 10^{-2}T$, where the timescale is determined by the characteristic frequency of the particle oscillations at the bottom of the potential well, $T = 2\pi$, the absolute error in the system energy, $\Delta E(t) = |E(t) - E(0)|$, did remain below $10^{-8}$ during the whole integration time.

All our calculations have been performed on a NVIDIA TESLA M2050 GPU, with 448 processing units on board, and by using CUDA and double-precision accuracy. We employed the complete parallelization, which allowed us in total to run $N = 10^7$ realizations simultaneously. The performance gain is shown in Figure 1. While the calculation time on a single stack-structured CPU grows linearly with $N$, at the same time, up to $N = 7168$ realizations can be calculated on a GPU simultaneously, due to the distribution procedure performed by CUDA. This number is a multiple of the number of GPU-cores, 7168 = 16 × 448. A further increase of the number of realizations causes a re-distribution of threads between the fixed number of cores and slow down the calculations speed, note the steps on the corresponding curve in Figure 1.

IV. MAIN RESULTS OF THE SIMULATIONS

In order to visualize the Arnold web, we employed the following procedure. After having the initial conditions distributed over a certain region of the system phase space the ensemble of trajectories was launched. The location and the shape of the distribution are tunable, so that we can scan different regions of the phase space at will. For every trajectory from the ensemble, $\{X(t), i = 1, \ldots, N\}$, we calculate the vector of the averaged finite-time velocity, $V(t) = (V_x(t), V_y(t), V_z(t))$, with the components reading: $V_x(t) = |X(t + \Delta t) - X(t)|/\Delta t$, wherein the averaging interval is given by $\Delta t$. In other words, it corresponds to a point on $S$. By running many such realizations in parallel, we collected the statistics, and finally projected a probability distribution function, $P_m(V; t)$, on the surface of the sphere $S$. The distribution is a nonstationary function, in the sense that its shape depends on $t$, so that the distribution profile reflects the dynamics of the ensemble.

If one of the ensemble trajectories was launched from a region filled with regular trajectories, it remains at the close vicinity of the initial momentum vector. There is no mixing in regular regions. Therefore, the projection of the corresponding part of the initial probability density distribution onto the momentum sphere remains invariant. When one of the ensemble trajectories entered the thin chaotic layer along a resonance $m$, it stays within the channel for some time ("sticking event"), and during this time the system moves with the velocity vector, $P$, which nearly exactly obeys the resonance condition, $m \cdot P \approx 0$. All the realizations within the chaotic layer contribute to the probability distribution, $P_m(V; t)$, concentrated on the corresponding resonance. The resonance appears as a bright line enclosed by an empty "dark zone," and the width of the dark zone corresponds to the width of the web around the resonance. An increase of the averaging time will induce further localization of the distribution at the corresponding resonance lines, but the widths of channels, i.e., dark zones on the momentum sphere, are fixed and do not depend on $\Delta t$. Therefore, the color representation of $P_m(V; t)$ allows one to clearly identify the location of the Arnold channels.

At this point, we may refer to the two-dimensional limit of the Hamiltonian (1 - 2). The momentum sphere represents a circle in this limit, and alternating chaotic and regular zones provide a full partition of the circle into a set of sectors. A trajectory, being placed initially into one of the zones, stays forever within the corresponding sector. Each chaotic zone can be characterized by the corresponding asymptotic velocity, $V = (V_x, V_y)$. When the averaging time $\Delta t$ approaches infinity, the distribution function inside the corresponding chaotic sector shrinks to the point $V$. Therefore, in the asymptotic limit sectors, corresponding to chaotic zones look like dark regions with bright points at their centers.

The averaging time interval $\Delta t$ is tunable and there exists no a priori best choice for it. Namely, every resonance channel $m$ can be characterized by a probability distribution of the corresponding sticking times, $\psi_m(t_{\text{stick}})$. The distribution always possesses a finite first moment, $20$ a mean sticking time $t_{\text{stick}}(m) = \int_0^\infty \psi_m(t_{\text{stick}}) dt$. If the averaging time interval is much larger than this mean sticking time, the corresponding resonance channel cannot be resolved. If, on the contrary, $\Delta t \ll t_{\text{stick}}(m)$, the oscillations of the momentum vector $P(m)$ along the high-order resonance, with $m \gg 1$, will smear the probability density over a broad region thus preventing a resolution of the fine structure of the web. In the following we used the value $\Delta t = 50T$.

The high degree of parallelism capability of the GPU offers a massive speedup for the ensemble propagation, see in Figure 1. With our Figures 2 and 3 we show the results of our numerical experiment with $N \approx 10^9$ independent realizations and the propagation time has been set at $T_p = \Delta t$. The overall simulation time of each experiment assumed 45 minutes. For a very long time evolution, see in Figure 4, we propagated an ensemble of $N = 3.2 \cdot 10^9$ particles up to a time $T_p = 10^5T$. The simulation time in this case was 24 h. The standard CPU-based run of a numerical experiment of the same scale would require (i) approximately one year of calculations on a standard desktop PC or (ii) about five months on a more advanced CPU of the Augsburg University computational cluster.

FIG. 1. (Color online) Performance of a central processing unit (CPU) of the Augsburg University computational cluster (Intel Xeon Processor 2.93 GHz Quad-Core X5570 Gainedstown) versus the GPU (NVIDIA TESLA M2050) for the propagation of an ensemble of Hamiltonian systems (1, 2): the overall calculation time is depicted as a function of the number of realizations $N$ (see text for more details). The propagation time of a single realization was $t = 5000T$. 

Downloaded 14 Nov 2011 to 193.174.246.171. Redistribution subject to AIP license or copyright; see http://chaos.aip.org/about/rights_and_permissions
The results corresponding to the weak perturbative regime at a total energy \( E = 400 \), corresponding to \( \epsilon_{\text{eff}} = 0.005 \), are shown in Figure 2. We used an ensemble with the initial conditions uniformly distributed over the sphere \( S \), and over the torus \([0,2\pi] \times [0,2\pi] \times [0,2\pi]\) in the coordinate space \( \Omega_X \). The most of the sphere surface is occupied by regular trajectories, therefore, the initial uniform distribution remains invariant almost everywhere, representing the uniform dark orange area on Figure 2(a). The brightest narrow stripes which braid the sphere correspond to the lower order, \( m \leq 2 \), resonances. We also scanned the energy shell by using an ensemble of trajectories with the initial conditions uniformly distributed within a small segment of the sphere (enclosed by the two squares in Fig. 2(a), with the aim to resolve the presence of the Arnold web around some higher-order resonances. The results, depicted in Figs. 2(b) and (c), show that the structures of the resonance intersections are topologically similar.

The results for the regime of moderate perturbation with total energy \( E = 15 \); i.e., \( \epsilon_{\text{eff}} = 2/15 \), are depicted with Fig. 3. Although the perturbation is much stronger now, it is not yet strong enough so as to destroy all the resonance tori and bring the system into the regime of the global Chirikov diffusion. There are several patches of the resonance web, which form leaf-like clusters on the surface of the sphere. The zooms into clusters reveal a fine fractal structure, see Figs. 3(b) and (c).

Finally, we computed the diffusive spreading of the ensemble initially injected at the point of intersection of three low-order resonances, \( W = \Omega_{X(1,-1,0)} \cap \Omega_{X(1,-1,-2)} \cap \Omega_{X(1,-1,2)} = (\sqrt{E}, \sqrt{E}, 0) \), which is at the center of a well-developed chaotic region. The evolution dynamics proceeds in a step-wise manner. In a first stage, we detect a fast spreading over the chaotic region, which encloses the injection point. Then the spreading slows down and eventually enters a second stage of slow diffusion through the Arnold web. Both stages are illustrated on Figure 4. The system phase space is not
Although the radius of all spheres in the momentum space is the same, the probability density structure of the web with a relatively small ensemble we averaged the probability over the whole time span between consecutive plots. Therefore, the evolution of ultracold atoms is Hamiltonian. The instantaneous momentum distribution can be measured by using the standard time-of-flight measurement technique. Finally, a tunable effective Planck constant allows to probe both the semiclassical and the deep quantum limits.

Another intriguing application that comes to mind is the implementation of the GPU-based mapping procedure to find the footprints of the Arnold diffusion in the emission patterns of three-dimensional optical resonators.

Our work here illustrates the advantages, provided by general purpose computing (GPC)–GPU ideology, for non-linear dynamics studies by using a particular example. Yet the potential of this approach reaches far beyond: it has already been put to work to explore (i) the noisy phase dynamics in a Josephson junction and the noisy Kuramoto model, (ii) the long-time evolution of nonlinear lattices, and (iii) the functioning of inertial Brownian motors that are driven by colored noise. We thus believe that the use of GPU computing in nonlinear science is only at a beginning: in the immediate future its great potential likely will provide many unforeseen findings and possibly even unravel manifest new phenomena.

**ACKNOWLEDGMENTS**

This work has been supported by the DFG Grant No. HA1517/31-2 (S.D. and P.H.).

---

3. See, for example, Computational Cosmology Center, US (https://c3.lbl.gov/) and Institute for Computational Cosmology, Durham University (http://www.icc.dur.ac.uk/).
5. For many examples, see http://www.nvidia.com/object/cuda_showcase.html.
6. A Nvidia Tesla C2050 GPU, having the size of a standard PC tower, is able to perform double precision computations at the rate of 0.5 TFLOPS, while the first supercomputer which crossed the 1 TFLOPS limit fifteen years ago, was a huge ASCI Red that occupied about 230 m² and required 850 kW of power (not including air conditioning).


24Numerical evolution is “exact” up to the round-off error, which is $10^{-32}$ for double-precision calculations.


26More precisely, it corresponds to a point within the thin layer beneath the sphere, $L_b(P)$. However, the layer width is small for the range of the energies we will consider here, so finally we use the central projection of the point $V$ on the surface of the sphere.


34Numerical evolution is “exact” up to the round-off error, which is $10^{-32}$ for double-precision calculations.


36More precisely, it corresponds to a point within the thin layer beneath the sphere, $L_b(P)$. However, the layer width is small for the range of the energies we will consider here, so finally we use the central projection of the point $V$ on the surface of the sphere.


44Numerical evolution is “exact” up to the round-off error, which is $10^{-32}$ for double-precision calculations.


46More precisely, it corresponds to a point within the thin layer beneath the sphere, $L_b(P)$. However, the layer width is small for the range of the energies we will consider here, so finally we use the central projection of the point $V$ on the surface of the sphere.


54Numerical evolution is “exact” up to the round-off error, which is $10^{-32}$ for double-precision calculations.