15

Parallelization, Special Hardware and Post-Newtonian Dynamics in Direct N-Body Simulations

Rainer Spurzem[1,5], Ingo Berentzen[1,5], Peter Berczik[1,5], David Merritt[2], Pau Amaro-Seoane[3], Stefan Harfst[4,2,5] and Alessia Gualandris[2,4]

[1]Astronomisches Rechen-Institut, Zentr. Astron. Univ. Heidelberg (ZAH), Mönchhofstrasse 12-14, 69120 Heidelberg, Germany
[2]College of Science, Dept. of Physics, Rochester Instute of Technology, 85 Lomb Memorial Drive, Rochester, NY 14623-5603, USA
[3]Max-Planck Institut für Gravitationsphysik (Albert-Einstein-Institut), Am Mühlenberg 1, D-14476 Potsdam, Germany
[4]Astronomical Institute *Anton Pannekoek* and Section Computational Science, University of Amsterdam, The Netherlands
[5]The Rhine Stellar Dynamical Network
spurzem@ari.uni-heidelberg.de

15.1 Introduction

The formation and evolution of supermassive black hole (SMBH) binaries during and after galaxy mergers is an important ingredient for our understanding of galaxy formation and evolution in a cosmological context, e.g. for predictions of cosmic star formation histories or of SMBH demographics (to predict events that emit gravitational waves). If galaxies merge in the course of their evolution, there should be either many binary or even multiple black holes, or we have to find out what happens to black hole multiples in galactic nuclei, e.g. whether they come sufficiently close to merge resulting from emission of gravitational waves, or whether they eject each other in gravitational slingshot interactions.

According to the standard theory, the subsequent evolution of the black holes is divided in three successive stages (Begelman, Blandford & Rees 1980). 1. Dynamical friction causes a transfer of the black holes' kinetic energy to the surrounding field stars, and the black holes spiral to the centre where they form a binary. 2. While hardening, the effect of dynamical friction reduces and the evolution is dominated by superelastic scattering processes, that is, the interaction with field stars closely encountering or intersecting the binaries' orbit, thereby increasing the binding energy. 3. Finally, the black holes coalesce through the emission of gravitational radiation, potentially detectable by the planned space-based gravitational wave antennae LISA. For a more detailed

Spurzem, R. et al.: *Parallelization, Special Hardware and Post-Newtonian Dynamics in Direct N-Body Simulations*. Lect. Notes Phys. **760**, 377–389 (2008)
DOI 10.1007/978-1-4020-8431-7_15

account of the state of research in this field, see Milosavljević & Merritt (2001, 2003); Makino & Funato (2004); Berczik, Merritt & Spurzem (2005). In our context the problem will be used as an example, where relativistic dynamics becomes important during the evolution of an otherwise classical Newtonian N-body system.

15.2 Relativistic Dynamics of Black Holes in Galactic Nuclei

Relativistic stellar dynamics is of paramount importance for the study of a number of subjects. For instance, if we want to have a better understanding of what the constraints on alternatives to supermassive black holes are, in order to explore the possibility of ruling out stellar clusters, one must do detailed analysis of the dynamics of relativistic clusters. Furthermore, the dynamics of compact objects around an SMBH or multiple SMBHs in galactic nuclei requires the inclusion of relativistic effects. Our current work deals with the evolution of two SMBHs, in bound orbit, and looks at the phase when they get close enough to each other that relativistic corrections to Newtonian dynamics become important, which ultimately leads to gravitational radiation losses and coalescence.

Efforts to understand the dynamical evolution of a stellar cluster in which relativistic effects may be important have already been made by Lee (1987), Quinlan & Shapiro (1989, 1990) and Lee (1993). In the earlier work, $1\mathcal{PN}$ and $2\mathcal{PN}$ terms were neglected (Lee 1993) and the orbit-averaged formalism (Peters 1964) used. We describe here a method to deal with deviations from Newtonian dynamics more rigorously than in most existing literature (but compare Mikkola & Merritt (2007); Aarseth (2007), which are on the same level of $\mathcal{PN}$ accuracy). We modified the NBODY6++ code to allow for post-Newtonian ($\mathcal{PN}$) effects of two particles getting very close to each other, implementing the $1\mathcal{PN}$, $2\mathcal{PN}$ and $2.5\mathcal{PN}$ corrections fully from Soffel (1989) and Kupi, Amaro-Seoane & Spurzem (2006).

Relativistic corrections to the Newtonian forces are expressed by expanding the relative acceleration between two bodies in a power series of $1/c$ in the following way (Damour & Dereulle 1987; Soffel 1989),

$$\underline{a} = \underbrace{\underline{a}_0}_{\text{Newt.}} + \underbrace{\underbrace{c^{-2}\underline{a}_2}_{1\mathcal{PN}} + \underbrace{c^{-4}\underline{a}_4}_{2\mathcal{PN}}}_{\text{periastron shift}} + \underbrace{\underbrace{c^{-5}\underline{a}_5}_{2.5\mathcal{PN}}}_{\text{grav. rad.}} + \mathcal{O}(c^{-6}), \tag{15.1}$$

where $\underline{a}$ is the acceleration of particle 1, $\underline{a}_0 = -Gm_2\underline{n}/r^2$ is the Newtonian acceleration, G is the gravitation constant, m_1 and m_2 are the masses of the two particles, r is the distance of the particles, $\underline{n}$ is the unit vector pointing from particle 2 to particle 1, and the $1\mathcal{PN}$, $2\mathcal{PN}$ and $2.5\mathcal{PN}$ are post-Newtonian corrections to the standard acceleration, responsible for the pericentre shift

($1\mathcal{PN}$, $2\mathcal{PN}$) and the quadrupole gravitational radiation ($2.5\mathcal{PN}$), correspondingly, as shown in (15.1). The expressions for the accelerations are

$$\underline{a}_2 = \frac{Gm_2}{r^2} \cdot \left\{ \underline{n} \left[-v_1^2 - 2v_2^2 + 4v_1 v_2 + \frac{3}{2}(nv_2)^2 + 5\left(\frac{Gm_1}{r}\right) + 4\left(\frac{Gm_2}{r}\right) \right] + (\underline{v}_1 - \underline{v}_2)\left[4nv_1 - 3nv_2\right] \right\}, \tag{15.2}$$

$$\begin{aligned}
\underline{a}_4 = \frac{Gm_2}{r^2} \cdot \Bigg\{ \underline{n} \Bigg[& -2v_2^4 + 4v_2^2(v_1 v_2) - 2(v_1 v_2)^2 + \frac{3}{2}v_1^2(nv_2)^2 \\
& + \frac{9}{2}v_2^2(nv_2)^2 - 6(v_1 v_2)(nv_2)^2 - \frac{15}{8}(nv_2)^4 \\
& + \frac{Gm_2}{r} \cdot \left(\; 4v_2^2 - 8v_1 v_2 + 2(nv_1)^2 - 4(nv_1)(nv_2) - 6(nv_2)^2 \right) \\
& + \frac{Gm_1}{r} \cdot \left(-\frac{15}{4}v_1^2 + \frac{5}{4}v_2^2 - \frac{5}{2}v_1 v_2 + \frac{39}{2}(nv_1)^2 - 39(nv_1)(nv_2) + \frac{17}{2}(nv_2)^2 \right) \Bigg] \\
& + (\underline{v}_1 - \underline{v}_2) \Bigg[\; v_1^2(nv_2) + 4v_2^2(nv_1) - 5v_2^2(nv_2) - 4(v_1 v_2)(nv_1) \\
& + 4(v_1 v_2)(nv_2) - 6(nv_1)(nv_2)^2 + \frac{9}{2}(nv_2)^3 \\
& + \frac{Gm_1}{r} \cdot \left(-\frac{63}{4}nv_1 + \frac{55}{4}nv_2 \right) + \frac{Gm_2}{r} \cdot \left(\; -2nv_1 - 2nv_2 \right) \Bigg] \Bigg\} \\
& + \frac{G^3 m_2}{r^4} \cdot \underline{n} \left[-\frac{57}{4}m_1^2 - 9m_2^2 - \frac{69}{2}m_1 m_2 \right],
\end{aligned} \tag{15.3}$$

$$\begin{aligned}
\underline{a}_5 = \frac{4}{5}\frac{G^2 m_1 m_2}{r^3} \Bigg\{ & (\underline{v}_1 - \underline{v}_2) \left[-(\underline{v}_1 - \underline{v}_2)^2 + 2\left(\frac{Gm_1}{r}\right) - 8\left(\frac{Gm_2}{r}\right) \right] \\
& + \underline{n}(nv_1 - nv_2) \left[3(\underline{v}_1 - \underline{v}_2)^2 - 6\left(\frac{Gm_1}{r}\right) + \frac{52}{3}\left(\frac{Gm_2}{r}\right) \right] \Bigg\}.
\end{aligned} \tag{15.4}$$

In the last expressions $\underline{v}_1$ and $\underline{v}_2$ are the velocities of the particles. For simplification, we have denoted the vector product of two vectors, $\underline{x}_1$ and $\underline{x}_2$, as $x_1 x_2$. The basis of direct NBODY4 and NBODY6++ codes relies on an improved Hermite integration scheme (Makino & Aarseth 1992; Aarseth 1999) for which we need not only the accelerations but also their time derivatives. These derivatives are not included here for succinctness. We include our correction terms in the KS *regularisation* scheme (Kustanheimo & Stiefel 1965) as perturbations, similarly to what is done to account for passing stars influencing a KS pair. Note that formally the perturbing force in the KS equations does not need to be small compared to the two-body force (Mikkola 1997). If the internal KS time step is properly adjusted, the method works even for relativistic terms becoming comparable to the Newtonian force component.

15.3 Example of Application to Galactic Nuclei

In Fig. 15.1 the importance of relativistic, post-Newtonian dynamics for the separation of the binary black holes in our simulations is seen. The curve deviates from the Newtonian results when gravitational radiation losses set in and causes a sudden coalescence ($1/a \to \infty$) at a finite time. Gravitational radiation losses are enhanced by the high eccentricity of the SMBH binary. It is interesting to note that the inclusion or exclusion of the conservative $1\mathcal{PN}$ and $2\mathcal{PN}$ terms changes the coalescence time considerably. Details of these results will be published in a larger parameter study (Berentzen et al. 2008, in preparation). Note that Aarseth (2003a) presents two models very similar to those discussed here, which agree qualitatively with our work regarding the relativistic merger time and the eccentricity of the SMBH binary.

Once the SMBH binary starts to lose binding energy dramatically due to gravitational radiation, its orbital period drops from a few thousand years to less than a year very quickly (time-scale much shorter than the dynamical time-scale in the galactic centre, which defines our time unit). Then the SMBH binary will enter the LISA band, i.e. its gravitational radiation will be detectable by LISA. The Laser Interferometer Space Antenna is a system of three space probes with laser interferometers to measure gravitational waves, see e.g. `http://lisa.esa.int/`. Once the SMBH binary decouples from the rest of the system we just follow its relativistic two-body evolution, starting

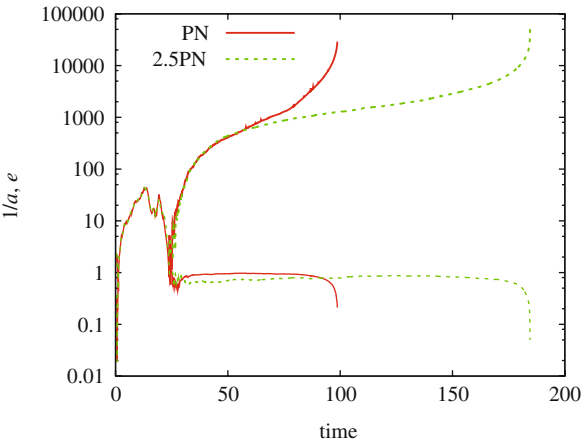

Fig. 15.1. Effect of post-Newtonian (PN) relativistic corrections on the dynamics of black hole binaries in galactic nuclei. Plotted are inverse semi-major axis and eccentricity as a function of time. The solid line uses the full set of PN corrections, while the dashed line has been obtained by artificially only using the dissipative 2.5PN terms. Note that the coalescence time in the latter case has changed significantly. Further details will be published elsewhere (Berentzen et al. 2008, in preparation)

with exactly the orbital parameters (including eccentricity) as they were extracted from the N-body model. It is then possible to predict the gravitational radiation of the SMBH binary relative to the LISA sensitivity curve (Preto et al. 2008, in preparation). For some values of the eccentricity our simulated SMBH binaries indeed enter the LISA sensitivity regime; for a circular orbit the $n = 2$ harmonic of the gravitational radiation is dominant, while for eccentric orbits higher harmonics are stronger (Peters & Mathews 1963; Peters 1964).

15.4 *N*-Body Algorithms and Parallelization

Numerical algorithms for solving the gravitational N-body problem (Aarseth 2003) have evolved along two main lines in recent years. Direct-summation codes compute the complete set of N^2 interparticle forces at each time step. These codes are designed for systems in which the finite-N graininess of the potential is important or in which binary- or multiple-star systems form, and until recently, were limited by their $\mathcal{O}(N^2)$ scaling to moderate ($N < 10^5$) particle numbers. The best-known examples are the `NBODY` series of codes (Aarseth 1999) and the `Starlab` environment developed by McMillan, Hut and collaborators (e.g. Portegies Zwart et al. 2001).

A second class of N-body algorithms replaces the direct summation of forces from distant particles by an approximation scheme. Examples are the Barnes–Hut tree code (Barnes & Hut 1986), which reduces the number of force calculations by subdividing particles into an oct-tree, and fast multipole algorithms that represent the large-scale potential via a truncated basis-set expansion (van Albada & van Gorkom 1977; Greengard & Rokhlin 1987). Such algorithms have a milder $\mathcal{O}(N \log N)$ or even $\mathcal{O}(N)$ scaling for the force calculations and can handle much larger particle numbers, although their accuracy are substantially lower than that of the direct-summation codes (Spurzem 1999). The efficiency of both sorts of algorithm can be considerably increased by the use of individual time steps for advancing particle positions (Aarseth 2003).

A natural way to increase both the speed and the particle number in an N-body simulation is to parallelize (Dubinski 1996; Pearce & Couchman 1997). Parallelization on general-purpose supercomputers is difficult, however, because the calculation cost is often dominated by a small number of particles in a single dense region, e.g. the nucleus of a simulated galaxy. Communication latency becomes the bottleneck; the time to communicate particle positions between processors can exceed the time spent computing the forces. The best such schemes use systolic algorithms (in which the particles are successively passed around a ring of processors) coupled with non-blocking communication between the processors to reduce the latency (Makino 2002; Dorband, Hemsendorf & Merritt 2003).

A major breakthrough in direct-summation N-body simulations came in the late 1990s with the development of the GRAPE series of special-purpose computers (Makino & Taiji 1998), which achieve spectacular speed-ups by implementing the entire force calculation in hardware and placing many force pipelines on a single chip. The GRAPE-6, in its standard implementation (32 chips, 192 pipelines), can achieve sustained speeds of about 1 Tflops at a cost of just $\sim$ \$50 K. In a standard setup, the GRAPE-6 is attached to a single host workstation, in much the same way that a floating-point or graphics accelerator card is used. Advancement of particle positions [$\mathcal{O}(N)$] is carried out on the host computer, while coordinate and velocity predictions and inter-particle forces [$\mathcal{O}(N^2)$] are computed on the GRAPE. More recently, "mini-GRAPEs" (GRAPE-6A) (Fukushige, Makino & Kawai 2005) have become available, which are designed to be incorporated into the nodes of a parallel computer. The mini-GRAPEs have four processor chips on a single PCI card and deliver a theoretical peak performance of $\sim$ 131 Gflops for systems of up to 128 K particles, at a cost of about \$6 K. By incorporating mini-GRAPEs into a cluster, both large (10^6) particle numbers and high (1 Tflops) speeds can be achieved.

In the following we describe the performance of direct-summation N-body algorithms on two computer clusters that incorporate GRAPE hardware.

15.5 Special Hardware, GRAPE and GRACE Cluster

The GRAPE-6A board (Fig. 15.2, top panel) is a standard PCI short card on which a processor, an interface unit and a power supply are integrated. The processor is a module consisting of four GRAPE-6 processor chips, eight SSRAM chips and one FPGA chip. The processor chips each contain six force calculation pipelines, a predictor pipeline, a memory interface, a control unit and I/O ports (Makino et al. 2003). The SSRAM chips store the particle data. The four GRAPE chips can calculate forces, their time derivatives and the scalar gravitational potential simultaneously for a maximum of 48 particles at a time; this limit is set by the number of pipelines (six force calculation pipelines each of which serves as eight virtual multiple pipelines). There is also a facility to calculate neighbour lists from predefined neighbour search radii; this feature is not used in the algorithms presented below. The forces computed by the processor chips are summed in an FPGA chip and sent to the host computer. A maximum of 131 072 (2^{17}) particles can be held in the GRAPE-6A memory. The peak speed of the GRAPE-6A is 131.3 Gflops (when computing forces and their derivatives) and 87.5 Gflops (forces only), assuming 57 and 38 floating-point operations, respectively, per force calculation (Fukushige, Makino & Kawai 2005). The interface to the host computer is via a standard 32-bit/33 MHz PCI bus. The FPGA chip (Altera EP1K100FC256) realizes a 4-input, 1-output reduction when transferring data from the GRAPE-6 processor chip to the host computer. The complete

Fig. 15.2. *Top*: interior of a node showing a GRAPE-6A card (note the large black fan) and an Infiniband card. *Bottom*: the GRACE cluster at ARI. The head node and the 14 Tbyte raid array are visible on the central rack. The other four racks hold a total of 32 compute nodes, each equipped with a GRAPE-6A card and MPRACE cards

GRAPE-6A unit is roughly 11 cm × 19 cm × 7 cm in size. Note that 5.8 cm of the height is taken up by a rather bulky combination of cooling body and fan, which may block other slots on the main board. Possible ways to deal with this include the use of even taller boxes for the nodes (e.g. 5U) together with a PCI riser of up to 6 cm, which would allow the use of slots for interface cards beneath the GRAPE fan, or the adoption of the more recent, flatter designs such as that of the GRAPE6-BL series. The reader interested in more technical details should seek information from the GRAPE (`http://astrograpе.org`) and Hamamatsu Metrix (`http:/www.metrix.co.jp`) websites.

A computer cluster incorporating GRAPE-6A boards became fully operational at the Rochester Institute of Technology (RIT) in February 2005. This cluster, named "gravitySimulator," consists of 32 compute nodes plus one head node, each containing dual 3 GHz-Xeon processors. In addition to a standard Gbit-ethernet, the nodes are connected via a low-latency Infiniband network with a transfer rate of 10 Gbits. The typical latency for an Infiniband network is of the order of 10^{-6} seconds, or a factor ~ 100 better than the Gbit-Ethernet. A total of 14 Tbyte of disc space is available on a level 5 RAID array. The disc space is equivalent to 2.5×10^5 N-body data sets each with 10^6 particles. The discs are accessed via a fast Ultra320 SCSI host adapter from the head node or via NFS from the compute nodes, which in addition are each fitted with an 80 Gbyte hard disc. Each compute node also contains a GRAPE-6A PCI card (Fig. 15.2, top panel). The total, theoretical peak performance is approximately 4 Tflops if the GRAPE boards are fully utilized. Total cost was about $ 450 000, roughly half of which was used to purchase the GRAPE boards.

Some special considerations were required in order to incorporate the GRAPE cards into the cluster. Since our GRAPE-6A's use the relatively old PCI interface standard (32 bit/33 MHz), only one motherboard was available, the SuperMicro X5DPL-iGM, that could accept both the GRAPE-6A and the Infiniband card. (A newer version of the GRAPE-6A which uses the faster PCI-X technology is now available.) The PC case itself has to be tall enough (4U) to accept the GRAPE-6A card and must also allow good air flow for cooling since the GRAPE card is a substantial heat source. The cluster has a total power consumption of 17 kW when the GRAPEs are fully loaded. Cluster cooling was achieved at minimal cost by redirecting the air conditioning from a large room toward the air-intake side of the cluster. Temperatures measured in the PC case and at the two CPUs remain below 30°C and 50°C, respectively.

A similar cluster, called "GRACE" (GRAPE + MPRACE), has been installed in the Astronomisches Rechen-Institut (ARI) at the University of Heidelberg (Fig. 15.2, bottom panel). There are two major differences between the RIT and ARI clusters. (1) Each node of the ARI cluster incorporates a reconfigurable FPGA card (called "MPRACE") in addition to to the GRAPE board. MPRACE is optimized to compute neighbour forces and other non-Newtonian forces between particles, in order to accelerate calculations of

molecular dynamics, smoothed-particle hydrodynamics, etc. (2) The newer main board SuperMicro X6DAE-G2 was used, which supports Pentium Xeon chips with 64-bit technology (EM64T) and the PCIe (PCI express) bus. This made it possible to use dual-port Infiniband interconnects via the PCI express Infiniband ×8 host interface card, used in the ×16 Infiniband slot of the board (it has another ×4 Infiniband slot, which is reserved for the MPRACE-2 Infiniband card). As discussed below, the use of the PCIe bus substantially reduces communication overhead. The benchmark results presented here for the ARI cluster were obtained from algorithms that do not access the FPGA cards.

15.6 Performance Tests

Initial conditions for the performance tests were produced by generating Monte-Carlo positions and velocities from self-consistent models of stellar systems. Each of these systems is spherical and is completely described by a steady-state phase-space distribution function $f(E)$ and its self-consistent potential $\Psi(r)$, where $E = v^2/2 + \Psi$ is the particle energy and r is the distance from the centre. The models were a Plummer sphere, two King models with different concentrations and two Dehnen models (Dehnen 1993) with different central density slopes. The Plummer model has a low central concentration and a finite central density; it does not represent any class of stellar system accurately, but is a common test case. King models are defined by a single dimensionless parameter W_0 characterizing the central concentration (e.g. ratio of central to mean density); we used $W_0 = 9$ and $W_0 = 12$, which are appropriate for globular star clusters. Dehnen models have a divergent inner density profile, $\rho \propto r^{-\gamma}$. We took $\gamma = 0.5$ and $\gamma = 1.5$, which correspond approximately to the inner density profiles of bright and faint elliptical galaxies.

In what follows we adopt standard N-body units $G = M = -4E = 1$, where G is the gravitational constant, M the total mass and E the total energy of the system. In some of the models, the initial time step for some particles was smaller than the minimum time step $t_{\min}$ set to 2^{-23}. These models were then rescaled to change the minimum time step to a large enough value. Since the rescaling does not influence the performance results, we will present all results in the standard N-body units.

We realized each of the five models with 11 different particle numbers, $N = 2^k$, $k = [10, 11, \ldots, 20]$, i.e. $N = [1\,\mathrm{K}, 2\,\mathrm{K}, \ldots, 1\,\mathrm{M}]$.[1] We also tested Plummer models with $N = 2\,\mathrm{M}$ and $N = 4\,\mathrm{M}$; the latter value is the maximum N-value allowed by filling the memory of all 32 GRAPE cards. Thus, a total of 57 test models were used in the timing runs.

Two-body relaxation, i.e. exchange of energy between particles due to gravitational scattering, induces a slow change in the characteristics of the

[1] Henceforth, we use K to denote a factor of $2^{10} = 1024$ and M to denote a factor of $2^{20} = 1,048,576$.

models. In order to minimize the effects of these changes on the timing runs, we integrated the models for only one time unit. The standard softening ϵ was set to zero for the Plummer models and to 10^{-4} for the Dehnen and King models. For the time step parameters used see Harfst et al. (2007).

We analyzed the performance of the hybrid scheme as a function of particle number and also as a function of number of nodes, using $p = 1, 2, 4, 8, 16$, and 32 nodes. The compute time w for a total of almost 350 test runs was measured using `MPI_Wtime()`. The timing was started after all particles had finished their initial time step and ended when the model had been evolved for one time unit. No data evaluation was made during the timing interval.

The top panel of Fig. 15.3 shows wallclock times $w_{N,p}$ from all integrations on the ARI cluster. For any p, the clock time increases with N, roughly as N^2 for large N. However, when N is small, communication dominates the total clock time, and w *increases* with increasing number of processors. This behaviour changes as N is increased; for $N > 10\,\mathrm{K}$ (the precise value depends on the model), the clock time is found to be a decreasing function of p, indicating that the total time is dominated by force computations.

The speedup for selected test runs is shown in the bottom panel of Fig. 15.3. The speedup s is defined as

$$s_{N,\,p} = \frac{w_{N,\,1}}{w_{N,\,p}}. \tag{15.5}$$

The ideal speedup (optimal load distribution, zero communication and latency) is $s_{N,p} = p$. For particle numbers $N \geq 128\,\mathrm{K}$ the wallclock time $w_{N,1}$ on one processor is undefined as N exceeds the memory of the GRAPE card. In that case we used $w_{N,1} = w_{128\,\mathrm{K},1}(N/128\,\mathrm{K})^2$, assuming a simple N^2-scaling. In general, the speedup for any given particle number is roughly proportional to p for small p, then reaches a maximum before reducing at large p. The number of processors at maximum speedup is "optimum" in the sense that it provides the fastest possible integration of a given problem. The optimum p is roughly the value at which the sum of the communication and latency times equals the force computation time; in the zero-latency case, $p_{opt} \propto N$ (Dorband, Hemsendorf & Merritt 2003). Figure 15.3 (bottom panel) shows that for $N \geq 128\,\mathrm{K}$, $p_{opt} \geq 32$ for all the tested models. The reader interested in more details is referred to Harfst et al. (2007).

15.7 Outlook and Ahmad–Cohen Neighbour Scheme

At present there exist only the relatively simple parallel N-body code described above and in Harfst et al. (2007), which uses GRAPE special hardware in parallel, but always computes full forces for every particle at every step. This code, sometimes dubbed p-GRAPE (sources are freely available, see link in the cited paper) also does not include any special few-body treatments (regularisations), as in the N-body codes of Aarseth (1999, 2003).

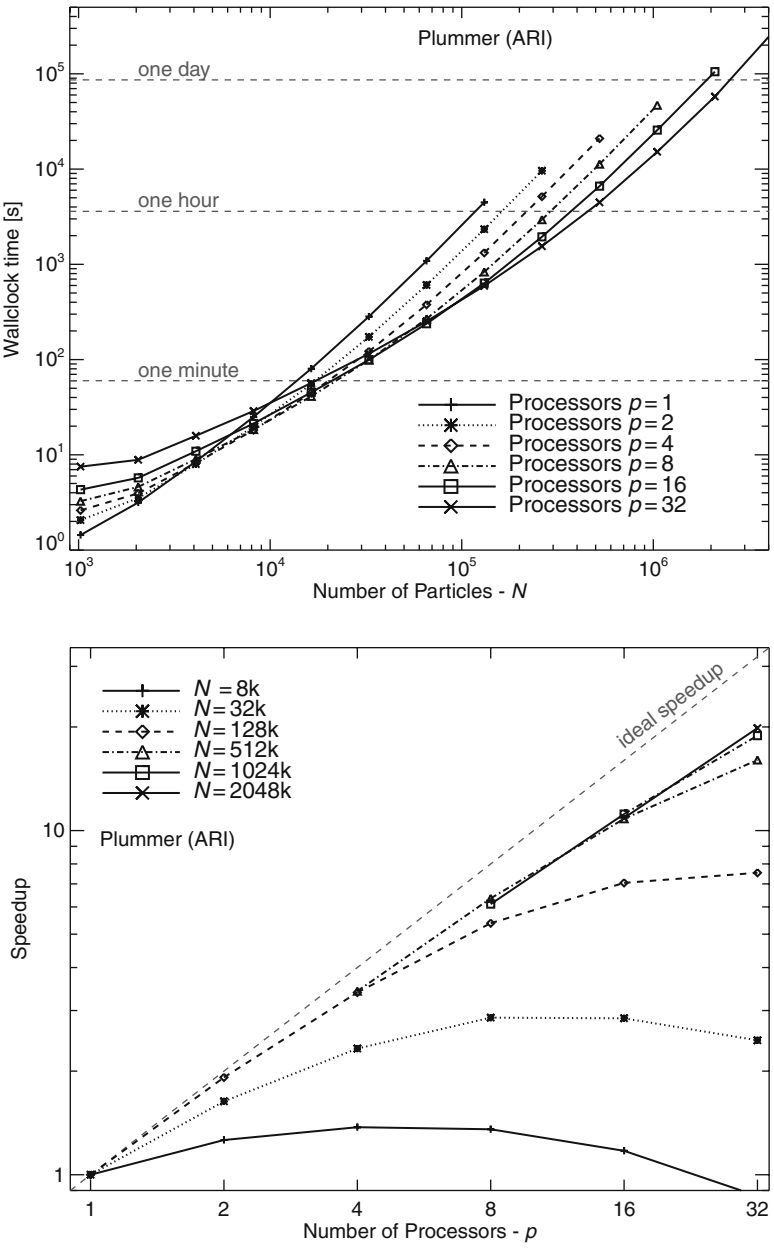

Fig. 15.3. *Top*: wallclock time w versus particle number N for different numbers of processors p. *Bottom*: speedup s versus processor number p for different N. Both the plots show the results obtained for a Plummer model on the ARI cluster

There is the already mentioned parallel N-body code NBODY6++, which includes all regularizations and the use of the Ahmad-Cohen neighbour scheme (Ahmad & Cohen 1973) as in the standard NBODY6 code. However, the publicly provided source code (`ftp://ftp.ari.uni-heidelberg.de/pub/staff/spurzem/nb6mpi/`) is not yet able to make parallel use of special hardware. It parallelizes very efficiently over the regular *and* irregular force loops (cf. Spurzem 1999; Khalisi et al. 2003), but current work is in progress on an implementation of NBODY6++ for special-purpose hardware (such as GRAPE, MPRACE or graphical processing units GPU) as well as on an efficient parallel treatment of many regularized perturbed binaries (see first results in Maalej et al. 2005). New results in these topics will be published early at the wiki of NBODY6++ developers and users at `http://nb6mpi.pbwiki.com/`. Last but not least, a nice visualization interface, specially developed for NBODY6++ , is hosted by FZ Jülich, see `http://www.fz-juelich.de/jsc/xnbody/`.

Similar to the GRAPE development nearly two decades ago, the recent introduction of GPUs and other new hardware devices (such as FPGA or MPRACE cards in the GRACE project,

`http://www.ari.uni-heidelberg.de/grace/`) is inspiring a new interest in improving and developing efficient N-body algorithms. It is expected that very soon the use of most advanced special hardware and software (such as NBODY6 and NBODY6++) will not mutually exclude each other any more.

Acknowledgement

Computing time at NIC Jülich on the IBM Jump is acknowledged. Financial support comes partly from Volkswagenstiftung (I/80 041-043), German Science Foundation (DFG) via SFB439 at the University of Heidelberg and Schwerpunktprogramm 1177 (Project ID Sp 345/17-1) 'Black Holes Witnesses of Cosmic History'. It is a pleasure to acknowledge many enlightening discussions with and support by Sverre Aarseth, and very useful interactions about relativistic dynamics with A. Gopakumar and G. Schäfer.

References

Aarseth S. J., 1999, PASP, 111, 1333 379, 381, 386
Aarseth S. J., 2003a, ApSS, 285, 367 380
Aarseth S. J., 2003, Gravitational N-Body Simulations. Cambridge University Press, Cambridge 381, 386
Aarseth S. J., 2007, MNRAS, 378, 285 378
Ahmad A., Cohen L., 1973, J. Comput. Phys., 12, 349
Barnes J., Hut P., 1986, Nature, 324, 446 381
Begelman M. C., Blandford R. D., Rees M. J., 1980, Nature, 287, 307 377
Berczik P., Merritt D., Spurzem R., 2005, ApJ, 633, 680 378

Berczik P., Merritt D., Spurzem R., Bischof H.-P., 2006, ApJ, 642, L21
Berentzen I., Preto M., Berczik P., Merritt D., Spurzem R., 2008, to be submitted 380
Damour T., Dereulle N., 1987, Phys. Lett., 87, 81 378
Dehnen W., 1993, MNRAS, 265, 250 385
Dorband E. N., Hemsendorf, M., Merritt, D., 2003, J. Comput. Phys., 185, 484 381, 386
Dubinski J., 1996, New Astron., 1, 133 381
Fukushige T., Makino J., Kawai A., 2005, PASJ, 57, 1009 382
Greengard L., Rokhlin V., 1987, J. Comput. Phys., 73, 325 381
Harfst S., Gualandris A., Merritt D., Spurzem R., Portegies Zwart S., Berczik P., 2007, New Astron., 12, 357 386
Khalisi E., Omarov C. T., Spurzem R., Giersz M., Lin D. N. C., 2003, in Krause E., Jaeger W., Resch M., eds, Performance Computing in Science and Engineering. Springer Verlag, p. 71 388
Kupi G., Amaro-Seoane P., Spurzem R., 2006, MNRAS, 371, L45 378
Kustaanheimo P., Stiefel E., Journ. für die reine und angew. Math., 1965, 218, 204
Lagoute C., Longaretti P. -Y., 1996, A&A, 308, 441
Lee H. M., 1987, ApJ, 319, 801 378
Lee M. H., 1993, ApJ, 418, 147 378
Maalej K. P., Boily C., David R., Spurzem R., 2005, in Casoli F., Contini T., Hameury J. M., Pagani L., eds, SF2A-2005: Semaine de l'Astrophysique Francaise. EdP-Sciences, Conference Series, p. 629 388
Makino J., 2002, New Astron., 7, 373 381
Makino J., Aarseth S. J., 1992, PASJ, 44, 141 379
Makino J., Fukushige T., Koga M., Namura K., 2003, PASJ, 55, 1163 382
Makino J., Funato Y., 2004, ApJ, 602, 93 378
Makino J., Taiji M., 1998, Scientific Simulations with Special-Purpose Computers — the GRAPE systems. Wiley 382
Mikkola S., 1997, Celes. Mech. Dyn. Ast., 68, 87 379
Mikkola S., Merritt D., 2007, ArXiv e-prints 709, arXiv:0709.3367 378
Milosavljević M., Merritt D., 2001, ApJ, 563, 34 378
Milosavljević M., Merritt D., 2003, ApJ, 596, 860 378
Pearce F. R., Couchman H. M. P., 1997, New Astron., 2, 411 381
Peters P. C., 1964, Phys. Rev., 136, B1224 378, 381
Peters P. C., Mathews J., 1963, Phys. Rev., 131, 435 381
Portegies Zwart S. F., McMillan S. L. W., Hut P., Makino J., 2001, MNRAS, 321, 199 381
Preto M., Berentzen I., Berczik P., Spurzem R., 2008, in preparation 381
Quinlan G. D., Shapiro S. L., 1989, ApJ, 343, 725 378
Quinlan G. D., Shapiro S. L., 1990, ApJ, 356, 483 378
Soffel M. H., 1989, Relativity in Astrometry, Celestial Mechanics and Geodesy. Springer-Verlag 378
Spurzem R., 1999, J. Comput. Applied Math., 109, 407 381, 388
van Albada T. S., van Gorkom J. H., 1977, A&A, 54, 121 381