Natural Parallelization Paradigm Implementation for Grid Technology

Alexander Bogdanov
Institute for High-Performance Computing and Integrated Systems
Saint-Petersburg, Russia
e-mail: bogdanov@csa.ru

Elena Stankova
Saint-Petersburg State University
Institute for High-Performance Computing and Integrated Systems
Saint-Petersburg, Russia
e-mail: lena@csa.ru

ABSTRACT
We discuss the possibility to apply an approach, elaborated for derivation of population's kinetics equations, for effective calculation of a wide range of computational problems using modern Grid technology. The main idea of the algorithm is to propose functional transformation of variables on the base of symmetry properties of pertinent physical system to make it quasi-diagonal. The approach is illustrated with the help of three examples: computations of quantum systems evolution, calculation of the chemical kinetics for many component systems and solution of non integrable nonlinear equations of wave evolution. The use of proposed approach shows substantial speed-up over the standard algorithms and seems to be more effective with the increase of the size of the problem. We discuss also a possible strategy for implementing a grid-based approach to realize the immense computational resources required to compute the physical problems under consideration.

Keywords
Computer science, Grid technology, computational algorithms, kinetic equations, nonlinear equations

1. INTRODUCTION
There are a lot of scientific problems that represent a real challenge for computational technologies. Among them there are: computations of quantum system evolution, numerical solution of chemical kinetic equations for the many component systems including evaluation of chemical reaction cross-sections and rate-constants and solution of non integrable nonlinear equations of wave evolution. The main difficulties are connected with the multi-dimension and multi-scale character of the problems. Both the accuracy and the detailed computations are limited by the availability of computational resources and by the efficiency of the computational algorithm used in their calculation. Even comprehensive multi-processor systems (supercomputers and clusters) do not satisfy the huge computational requirements. In our opinion it does not seem so strange since with the increase of number of parallel computational processes the price, which we pay for exchange of data between processes, is becoming more and heavier load, that makes the increase of number of nodes ineffective. That is why it is very important in deriving the algorithms to minimize the data exchange between computational processes. This leads to the necessity for seeking new approaches to satisfying the seemingly insatiable demand of such calculations for computational resources. We examine here a possible solution to this problem by proposing the algorithm that will take into account the symmetries of the physical problem while mapping it to corresponding algorithm and computer architecture. It is easy to do it for evolution operator and so we shall illustrate our main ideas on the examples, where the physical meaning of the proposed procedure is quite transparent. At the same time the proposed approach really makes it possible to get on large-scale problems substantial speed-ups over the standard algorithms. Moreover it can be effectively used in Grid technologies for effective calculations.

2. GENERAL DISCRIPTION OF THE ALGORITHM
It is easy to understand bases of the proposed algorithm on the example of the evolution equation of the type

$$iDu/Dt = \mathrm{P}u \qquad (1)$$

with P being some operator, which we will suppose to be Hermitian. Any standard approach normally will transform it to the linear system ordinary differential equations of the type:

$$i\,d\mathbf{v}/dt = \boldsymbol{P}\mathbf{v} \qquad (2)$$

with **v** being the large vector and ***P*** being symmetric matrix. The problem is not difficult for any size of **v**, if ***P*** is almost diagonal, but for many important situations it is not the case. From the point of view of the theory of dynamic systems [1] large non diagonal members in ***P*** mean bad choice of representation for Eq. (1), although often such representation is forced by physical considerations.

Usually it is not difficult to find the transformation, which will make P quasi diagonal, some of the beautiful approaches in the theory of nonlinear equations [2] can have even natural physical background. At the same time some problems may cause the reverse transformation to original variables. Here we shall study the problem, for which it is not necessary to make reverse transformation since the solution of Eq. (1) is distribution function, used for computation of average values, which can be as effectively done in new variables, as in old ones.

3. SOME PARTICULAR CASES
3.1. Algorithm for quantum system evolution
Here we consider the problem of calculating scattering cross sections and rate constants for one of the processes shown in

the schematic diagram below, (m, n and p are initial or final diatomic quantum numbers).

$$A+(BC)_m \rightarrow \begin{cases} A+(BC)_n \\ (AB)_n + C \\ (AC)_p + B \\ A+B+C \\ (ABC)^* \rightarrow \begin{cases} A+(BC)_n \\ (AB)_n + C \\ (AC)_p + B \\ A+B+C \end{cases} \end{cases}$$

The quantum mechanical Hamiltonian in the curvilinear coordinate system may be written in the form [3].The multi-channel scattering problem in the three-body system is described by the Schrödinger equation

$$\left(\hat{H}_S - E\right)\Psi = 0$$

The full wave function should obey boundary conditions

$$\lim_{x^0 \to -\infty} \Psi^{(+)}(m, x_S) = \Psi_{in}(m, x_S) + \sum_{n \neq m} R_{nm}\Psi_{in}(n, x_S)$$

$$\lim_{x^0 \to +\infty} \Psi^{(+)}(m, x_S) = \sum_n S_{nm}\Psi_{out}(n, x_S)$$

The main idea of the approach is to make in functional or path representation of quantum observable the transformation of variables. We will apply the functional transformation to so called interaction coordinates:

$$A(i, f') = \langle G(i,t)G(t, f')\rangle \qquad (3)$$

That will allow to reduce Hamiltonian function of the problem to actually responsible for described process, to reduce integration interval to finite one, corresponding to interaction region, to transform asymptotically boundary conditions to standard ones, to make it possible to get the expression directly to scattering amplitude as an average over Green's functions of the problem but in mixed representation and in interaction coordinates

The average is taken over the coordinate space of the problem with Green's functions determined in terms of path integrals over the phase space with the weights of the type:

$$\exp(-i + \int XdP + iXP - i\int Hdt) \qquad (4)$$

Since we use Green's functions only for computation of averages (3) we can make any phase coordinates transformations of the type:

$$H(P, X) \rightarrow H\left(X, \frac{\partial F}{\partial X}\right) + \frac{\partial F}{\partial t}$$

With F being the Generator of the transformation. It is convenient to choose F as a solution of certain equation that guarantees the possibility of evaluation of path integral with the weight (4). In that case instead of computation of path integral we have to solve four times the equation for F, that is more convenient since it is partial differential equation of the first order.

The sequence of the transformation are as follows:

$$(\mathrm{P},\mathrm{X}) \rightarrow (\mathrm{QY})$$

$$\mathrm{H}(\mathrm{P},\mathrm{X}) \rightarrow \mathrm{H}(\mathrm{Q},\mathrm{Y}) = \mathrm{H}(\mathrm{X},\ \partial \mathrm{F1}/\partial \mathrm{t}) + \frac{\partial \mathrm{F_1}}{\partial t}|_{\mathrm{Q,Y}}$$

$$\mathrm{S} = -\mathrm{P}_*\mathrm{X}_* + \int \mathrm{dF_1} + \int QdY$$

$$\mathrm{G}(\mathrm{P}_*\mathrm{X}) = \exp\left[-\mathrm{i}/\mathrm{h}\ \mathrm{P}_*\mathrm{X}_* + \mathrm{i}/\mathrm{hF_1}\,|_* + \mathrm{i}/\mathrm{hQ}_*(\mathrm{Y}_0 - \mathrm{Y}_*)\right]$$

$$\mathrm{A}(\mathrm{P}_i \rightarrow \mathrm{P}_f) = \int \mathrm{dX_0C\ d}(\mathrm{X_0P_0})\exp\left[\mathrm{i}/\mathrm{h}\left(\mathrm{P}_f\mathrm{X}_f - \mathrm{P}_i\mathrm{X}_i\right)\right]$$

$$+\mathrm{i}/\mathrm{hF1}|_* + \mathrm{i}/\mathrm{hF1}\,|_* + \mathrm{i}/\mathrm{hQ}_i(\mathrm{Y_0} - \mathrm{Y}_i) + \mathrm{i}/\mathrm{hQ}_f(\mathrm{Y}_f\text{-}\mathrm{Y_0})$$

$$\mathrm{F1:}\ \ \mathrm{H}(\mathrm{X},\ \partial \mathrm{F1}/\partial \mathrm{x}) = \mathrm{E}$$

$$(\mathrm{P},\mathrm{X}) \rightarrow (\mathrm{QY})$$

The resulting amplitude representation is multidimensional integral:

$$T(P_i \rightarrow P_f) = \int dX_0 C\delta(X_0, P_0)\exp\left\{\frac{i}{\hbar}(P_f X_f - P_i X_i) + \right.$$

$$\left. + \frac{i}{\hbar}F_1|_i^0 + \frac{i}{\hbar}F_1|_0^f + \frac{i}{\hbar}Q_i(Y_0 - Y_i) + \frac{i}{\hbar}Q_f(Y_f - Y_0)\right\}\ .$$

As a result we will be able to achieve the following results:

- the phase of path integral becomes quadric functional
- the new representation for observable is reduced to standard multidimensional integral
- the solution, for every point in coordinate space of that integral, will be of the first order partial differential equation system.

It is clear, that proposed procedure leads to absolutely parallel algorithm, since the equations for generators should be solved independently for every coordinate of the integral.

3.2. The Equations of Nonequilibrium Kinetics

The system of kinetic equations, describing the relaxation in molecular systems [4], consists in realistic cases of thousands of equations of the type

$$Dc(\mathbf{j},t)/Dt = I(\mathbf{j}-\mathbf{1},t) - I(\mathbf{j},t) \qquad (5)$$

with j being the multi-index, describing molecular states and I being the molecular current in index space of the form $I(\mathbf{j},t) = K(\mathbf{j},\mathbf{j}+\mathbf{1})c(\mathbf{j},t) - K(\mathbf{j}+\mathbf{1},\mathbf{j})c(\mathbf{j}+\mathbf{1},t)$, where K is a rate constant for molecule-central particle collision.

The problems with Eq.(5) actually come from two factors – there are large nondiagonal members, corresponding to important physical transitions, and values of $I(\mathbf{j},t)$ are very large with difference between them in r.h.s. of Eq. (5) is relatively small. The situation becomes dramatic if you start integration with thermal equilibrium, when all I's are equal.
To overcome those difficulties it is useful to introduce new variables [5]

$$f(\mathbf{j},t) = c(\mathbf{j}+\mathbf{1},t)/c(\mathbf{j},t)a(\mathbf{j},t)$$

with $a(\mathbf{j},t)$ being the ratio of to rate constants $K(\mathbf{j},\mathbf{j}+\mathbf{1})$ and $K(\mathbf{j}+\mathbf{1},\mathbf{j})$. f's are so called slow variables, which become constants at equilibrium conditions. The equations for f's are [5]

$$df(\mathbf{j},t)/dt = \tilde{R}(f(\mathbf{j},t)) + H(\mathbf{j},t)f(\mathbf{j},t) + S(\mathbf{j})f(\mathbf{j},t) \qquad (6)$$

with $\tilde{R}$ being the quadric relaxational term, that is diagonal in j, H is the source term, proportional to hydrodynamic gradients, and S is the source of population change, the only term, that is nondiagonal in $f(\mathbf{j},t)$.

The main advantage of the last equation is, that not only the sum of three terms in r.h.s. it is small, but they are small separately and it is easy to determine their relative values beforehand. Moreover, it is important, that the major contribution in r.h.s. of the equation is diagonal that opens interesting opportunities for parallel algorithm.

3.3. Solution of nonlinear equations of wave evolution

The solution of non integrable nonlinear equations is very difficult even numerically and practically impossible by standard analytical technique. Although they appear in many applications and there is a lot of work done on their analysis it is hard to quote any result, that can be called conclusive. The only favorable exception is the systems not far from completely integrable. We can cast up perturbation theory, starting from nonlinear integrable system, that can give a direct way to build the unique solution at least for finite time intervals. Some beautiful examples were given in [6,7] on the base of different methods of solution of basic nonlinear integrable problem.

We shall give some examples of such analysis on the base of nonlinear waves evolution study in multi phase media with chemical reaction [8]. It was shown, that for one dimensional gas dynamic problem, described by Navier-Stockes equations, equation of state, and simple linear relaxational equation after expansion up to the second order near the equilibrium state one gets for the velocity nonlinear evolution equation of the form

$$v_t + vv_x + \alpha v_{xx} + \beta v_{xxx} = \gamma I(v)$$

where α — the measure of dissipational effects, β — the measure of dispersion and expressed via transport coefficients and relaxation times, γ — the measure of interphase interactions and is expressed via integral brackets and relaxational times

Integral operator, that to the first approximation is linear:

$$I(v) = -\int_0^t G(t,\tau) v_\tau d\tau$$

where G — for different models of interaction being exponential or inverse power function.

We are interested only in such features of the wave, that change the form of soliton, i.e. tails and pilot waves. We can give here simple quasiclassical conditions for pilot wave appearance. It can be shown, that change of amplitude in our approximation is given by integral:

$$\Phi(\kappa) = 2\int_0^\infty \exp(-\kappa\theta')\int_0^\infty \cosh^{-2}(\theta+\gamma\theta')\times$$
$$\times \cosh^{-3}\theta \sinh\theta d\theta' d\theta$$

and the evolution of the amplitude is determined by the equation:

$$\frac{da}{d\tau} = -\Phi(\kappa)a$$

with κ being the combination of parameters of soliton and perturbation. Pilot wave can appear only if

$$\Phi(\kappa) < 0$$

The important possibility, which is the most interesting for us, is the use of quasiclassical approximation within nonlinear change of variables for our system. The best known example is Whiner-Hopf transformation, that transforms Burgers equation to linear problem. For our equation the corresponding transformation is given by [9]. Following the method one can give a beautiful quasiclassical description of the equation, even uniform one, so in this approach the only serious problem is to return to initial variables. After the described transformation we can apply to Hamiltonian form of the solution the same approach as described in 3.1.

Using this quasiclassical approximation within nonlinear change of variables we can also effectively extract the qualitative behavior of the solution.

4. POSSIBLE USE OF GRID TECHNOLOGY

Grid technology is an approach that provides a mechanism for handling dynamically changing environments of geographically distributed computers and data resources. The Grid facilitates the integration of the efforts of scientists, engineers, institutions, and enterprises by creating virtual organizations in which resources are shared under certain rules. Scientific problems, described above make great demand on computational and data resources both for calculations and for data management. The Grid technology provides possibilities for solving such computationally intensive problems.

The basic principles of the complex scientific calculations and can be stated as follows:

1. Present computer and data resources as a single virtual environment by developing a web portal on the Grid technology.
2. Build an easy-to-use user web interface for providing access to these resources.
3. Facilitate the sharing of results of research.
4. Organize archiving of input, output, and intermediate data.

As the basic software for computational Grid of such kind the Globus code will be used because of the following distinguishing advantages [10]. It is the most widely used package used for the practical implementation of the Grid concept. It provides a standards-based secure model, a modular structure, a decentralized approach for control and deployment of software, a clear model for maintaining of local control of resources and open source code. As the scheduling system we plan to use the Nimrod/G tool [11, 12], which is designed to manage the computational process including the transfer of input data and of the results of the calculations.

The basic principles of the Grid for chemical reaction calculation were discussed in [13]. All the principles, presented in the works can be used with some modification for applying Grid facilitates for kinetic and wave evolution equations calculations.

5. CONCLUSIONS

In this paper we have discussed a possible strategy for solving complex computational problems using Grid technology. We believe that this approach offers the best possibility for the practical realization of the immense computational requirements needed for their solution. Unfortunately, as in many other cases the effective use of grid possibilities makes it necessary to build new problem solving environment for each particular class of problems and substantially change applied codes and sometimes even pertinent algorithms.

We have shown that the use of some physical considerations makes it possible to derive some new algorithms for solution of the evolution equations for physical variables like distribution function. With those algorithms we can reduce the needed computer time orders of magnitude, go to substantially larger number of processor and work out approximate methods, which can be used for mass computations in technical applications. Using Grid technology we simulated on the loosely coupled cluster the solution of the complex

problem. The results showed that our approach essentially reduce computational time in comparison with the parallel approach. In future we intend to port the application on the Grid in full volume.

REFERENCES

[1] W. Ebeling, J. Freund, F.Schweitzer, "Komplexe Strukturen: Entropic und Information. " B.G. Teubner, Stuttgart, Leipzig (1998)

[2] V.E. Zakharov, E.A. Kuznetzov, "Hamiltonian formalism for nonlinear waves". In: *Russian Uspekhi Fizicheskhich Nauk*, Vol. 167. Nauka, Moscow 1137–1167, 1997

[3] A.V. Bogdanov, Yu.E. Gorbachev, V.M. Strelchenya, "Theory of vibrational and rotational excitation of polyatomic molecules", *Physics Reports*, vol. 181, N 3, P. 121-206. 1989

[4] A.L. Itkin, E.G. Kolesnichenko, "Microscopic Theory of Condensation in Gases and Plasma. " *World Scientific*, Singapore New Jersey London Hong Kong , 1998

[5] A.V. Bogdanov, Yu.E. Gorbachev, V.M. Strelchenya,: "Relaxation and condensation processes influence on flow dynamics. " in *Proccedings of XV Intern. Symp. Rarefied Gas Dyn.* Grado (Italy) 407-408, 1986

[6] V.I., Karpman E.M. Maslov *JETP*, , v.2, No. 8, p.537--548, 1977 (In Russian).

[7] J.R. Keener, D.W. Mclaughlin *Phys. Rev.*, v. A16, No. 2, p. 777—790, 1977.

[8] A.V. Bogdanov, S.A. Vakulenko, V.M. Strelchenya In: *Tchislennye Metody Mechaniky Sploshnoy Sredy*, v. 11, No. 3, p. 18—26, 1980 (In Russian).

[9] G.B. Witham, Linear and nonlinear waves. John Wiley & Sons, New~York, London, 1974.

[10] W. E. Johnston, The NASA IPG Engineering Team, and The DOE Science Grid Team Implementing production Grid

[11] D. Abramson, J. Giddy, L. Kotler,: "High Performance Parametric Modeling with Nimrod/G: Killer Application for the Global Grid? " *International Parallel and Distributed Processing Symposium (IPDPS)*, Cancun, Mexico (May 2000) 520- 528; http://www.csse.monash.edu.au/~davida/nimrod/

[12] Shoshmina, I., Bogdanov, A.V. and Abramson D. "Whither the Grid?" *Proceedings of the International Conference Distributed Computing and Grid Technologies in Science and Education*. July 2004, Dubna, Russia.

[13] A.V. Bogdanov, E. N. Stankova, A.Goncharuk "Example of a Potential Grid Technology Application for Effective Kinetic Equation Calculations" *Proceedings of the Third International Conference "Distributed Computing and Grid-Technologies in Science and Education"* Dubna, June 30 - July 4, 2008