С 173(04)
С-73

Экз. чит. зала

COMPUTATIONAL MODELLING AND COMPUTING IN PHYSICS



Editors

E.P.Zhidkov

A.V.Fedorov

R.R.Shahbaghyan

JOINT INSTITUTE FOR NUCLEAR RESEARCH

Proceedings of the 9th International Conference

COMPUTATIONAL MODELLING AND COMPUTING IN PHYSICS

September 16 – 21, 1996,
Dubna, Russia

Editors
E.P.Zhidkov
A.V.Fedorov
R.R.Shahbaghyan

Dubna 1997

Preface

The traditional International Conference on Computational Modelling and Computing in Physics organized by Laboratory of Computing Techniques and Automation was held in Dubna (Russia) on September 16-21 1996. The Conference follows the rich traditions of the previous conferences on Numerical Methods and Applications that have been held in Dubna since 1964. The present Conference was organized in honour of M.G.Meshcheryakov the famous scientist and the founder of the LCTA.

More than 160 scientists from 15 countries attended the Conference and of about 130 reports and short communications were presented. The topics cover the wide scope of problems including computational simulations and models, numerical methods (finite difference methods, finite and boundary elements methods, numerical linear algebra, parallel algorithms and their applications), inverse and ill-posed problems, computer algebra, neural networks, wavelets, data processing, bifurcations and chaos, computing and networking for physical research, computer education.

This book contains refereed and edited versions of the papers presented at the Conference.

I am grateful to all the authors for their contributions. I appreciate the cooperation of Prof R.Pose (Director of LCTA, member of the Organizing Committee), Prof I.V.Puzynin (Deputy Director of LCTA, member of the Organizing Committee), Dr V.V.Korenkov (Deputy Director of LCTA, Co Chairman), Academician A.A.Samarski (Moscow, Honorary Member), Prof V.P.Shirikov (LCTA, member of the Organizing Committee), Dr S.P.Ivanova (LCTA, member of the Organizing Committee) and Dr L.A.Sevastyanov (Moscow, member of the Organizing Committee). The success of the Conference has been ensured in many respects by the efficiency of the corresponding offices of JINR. Special thanks to Dr A.P.Korol (JINR International Department, member of the Organizing Committee). The tough daily work related to the technical and scientific organization has been shared by my colleagues, members of the Organizing Committee: Dr E.A.Airian, Dr A.V.Fedorov (Scientific Secretary), Dr S.O.Lukianov (Deputy Director of LCTA), Dr R.R.Shahbaghyan and Dr T.A.Strizh. Without their efforts it would have been impossible to organize this conference. The financial support of the Russian Foundation for Basic Research, Russian Ministry of Science and Technical Policy and Russian University of People's Friendship is gratefully acknowledged.

E.P.Zhidkov
Chairman of CMCP'96

ISBN 5-85165472-4
© Объединенный институт ядерных исследований. Дубна, 1997

Contents

Preface

On numerical-analytic investigation of scalar scattering by ideal prolate spheroids and triaxial ellipsoids 9
A.A.Abramov, A.L.Dyshko, N.B.Konyukhova and T.V.Levitina

A universal program to uncouple linear systems 16
S.A.Abramov and E.V.Zima

Numerical method for soving the inverse problem of quantum scattering theory 27
R.G.Airapetyan, J.Busha and E.P.Zhidkov

Some application of the continuation method for the numerical solving of the three-dimensional polaron problem 32
P.G.Akishin I.V.Puzynin and Yu.S.Smirnov

On some problems of numerical investigation of eigenvalue problems in momentum space 40
I.V.Amirkhanov, E.V.Zemlyanaya, I.V.Puzynin, T.P.Puzynina and T.A.Strizh

Numerical investigation of quantumfield model of strong-coupling binucleon 48
I.V.Amirkhanov, E.V.Zemlyanaya, I.V.Puzynin, T.P.Puzynina, T.A.Strizh and V.D.Lakhno

Nonlinear dynamics of particle beams with space charge 55
S.Andrianov

Dynamic modeling paradigm and computer algebra 60
S.Andrianov

Calculations of contrast structures in singulary perturbated equations and systems 65
V.A.Anikeeva, A.B.Vasil'eva and A.A.Plotnikov

Taking into account the finite statistics of Monte Carlo simulation for parametrs estimation from experimental data 70
V.B.Anykeyev, A.V.Popov and V.P.Zhigunov

The method of consideration of a strong scattering anisotrophy in transport equation 77
E.N.Aristova and V.Ya.Gol'din

Spline approximation and smoothing using the adaptive projective filters 82
M.V.Avdeev and N.D.Dikoussar

Soliton-like solutions of linear differential equations 91
V.S.Barashenkov and L.G.Sapogin

On the numerical solution of Shrödinger equation 95
B.Batgerel and T.Zhanlav

Energetic particle fluxes data base of the spacecraft 100
A.Beliaev, A.Dmitriev, S.Kuznetsov, O.Nechaev and M.Panasyuk

Calculation of decay rates of the $HeH\mu$ systems 105
V.B.Belyaev, O.I.Kartavtsev, V.I.Kochkin and E.A.Kolganova

Physics, accelerator and software data bases at the WWW/Oracle server of JINR 110
V.A.Berezhnoj, B.S.Volkov, Yu.N.Isaev, A.P.Sokolov, V.F.Borisovskij, G.Musulmanbekov, V.V.Koren'kov, E.G.Nikonov, M.S.Plyashkevich, A.V.Ribalkin and V.P.Shirikov

The $U(1)$ gauge-invariant Heisenberg antiferromagnet model and multidimensional topological solitons 115
I.L.Bogoljubsky and A.A.Bogoljubskaya

The Hall effect in two-dimensional MHD problems with transverse magnetic field 121
K.V.Brushlinsky and T.A.Ratnikova

JAVA and WWW for NEMO experiment 126
A.D.Budnik and T.V.Filipova

Using cellular automata to model cave stuctures 130
M. Casolino and P. Picozza

Calculation of canonical normal distributions on the group $SO(3)$ and their application 137
E.A.Davidzhan, T.I.Savyolova and T.M.Ivanova

Using "TG-GSS" dialogue-graphics platform for processing the cosmic gamma-ray 142

spectra in CORONAS-I space experiment
A.V.Dmitriev, S.N.Kuznetsov, S.P.Ryumin, P.N.Zaikin, T.B.Podosenova, I.I.Petsilyak, Yu.D.Kotov and A.S.Glyanenko

Visualization of relativistic string dynamics 146
V.V.Dyachin, S.V.Klimenko and I.N.Nikitin

On a variant of domain decomposition method for solution of mathematical physics problems 153
A.A.Egorov

Approximate formulae for functional integrals based on interpolation of measures 156
A.D.Egorov and A.V.Zherelo

Two-dimensional solidifying and melting of needle-like crystals in the presence of hydrodynamic flow 159
Yu.E.Hohlov, S.Howison and K.Kornev

Calculations of entropy and critical phenomena by Monte Carlo methods 165
C.-K.Hu

Fast multilevel iterative solver for harmonic boundary value problem on polygonal domain 172
B.N.Khoromskij and E.G.Nikonov

Three-body resonances in framework of the Faddeev configuration space approach 177
E.A.Kolganova and A.K.Motovilov

Transitions from periodic to glass-like structures in foam saturated capillaries 181
K.Kornev, V.Mourzenko and R.Dautov

Modelling of toroidal plasma equilibrium evolution with account of non-ohmic currents 186
D.P.Kostomarov, E.S.Zaitsev, V.V.Nefedov, A.G.Shishkin, D.C.Robinson, M.Cox and M.R.O'Brien

On numerical solving of the problem of preassigned motions formation of charged particles in magnetic field 193
E.D.Kotina

Cellular automation and elastic net for event reconstruction in NEMO-2 experiment 198
V.Kovalenko (for NEMO collaboration)

The new alalytic tools of the space R^n. 203
Yu.I.Kuznetsov

Codes package for the RFQ dynamics and optimization 210
A.Loukianova

Deformable templates for circle recognition 214
L.Muresan, R.Muresan, G.Ososkov and Yu.Panebratsev

Transformation theory of multiple hypergeometric series and computer aided symbolic calculations 219
A.W.Niukkanen

On numerical solving of the direct regulator problem 224
A.D.Ovsyannikov

Algorithms for selecting the form of approximation expressions and splines 229
B.A.Popov

Simulation of a boundary layer between the high density non-ideal plasma and condenced medium (part I) 234
P.V.Poritsky, A.Z.Kononov, P.D.Starchyk and L.M.Voitenko

Simulation of a boundary layer between the high density non-ideal plasma and condenced medium (part II) 237
P.V.Poritsky, A.Z.Kononov, P.D.Starchyk and L.M.Voitenko

High accuracy Newton iteration scheme for solving the multichannel eigenvalue problem of some exotic few-body systems 240
I.V.Puzynin, T.P.Puzynina, A.Yu.Tyukhtyaev and S.I.Vinitsky

Analytical study of the Gary scheme on PC. Displacement of overlapping grids gives the stability 246
S.Serdyukova and M.Thuné

Grid approximation of the solution and diffusion flux for singularly perturbed parabolic equations with discontinuous coefficients and concentrated factors 253
G.I.Shishkin

Pixel decimation motion estimation using wavelet transform 259
T.M.Y.Wong and J.F.Zhang

Execution time analysis for least squares problems on massively parallel distributed memory computers 265
T.Yang

Linear operator interpolation in a space of functional matrices 282
L.A.Yanovich and F.M.Manyuk

"TG-GSS" manager to generate applied software for analysis of data like set of time series and ill-posed inverse problems 288
P.N.Zaikin and T.B.Podosenova

Методы решения некоторых многопараметрических спектральных задач математической физики 293
А.А.Абрамов, А.Л.Дышко, Н.Б.Конюхова, Т.В.Левитина, В.И.Ульянова

Поведение решений сингулярно возмущенного дифференциального уравнения бесконечного порядка 300
И.В.Амирханов, С.А.Васильев, Е.П.Жидков, И.Е.Жидкова

О некоторых нелинейных сингулярных задачах в моделях релятивистской космологии 304
Н.А.Воронов, А.Л.Дышко, Н.Б.Конюхова, И.Б.Староверова

Модификация метода К.Ланцоша выявления скрытых периодичностей 311
Е.А.Гребеников, С.В.Миронов

Численное моделирование некоторых магнитных систем соленоидального типа 314
Е.П.Жидков, В.П.Овсянников, Р.В.Полякова, И.П.Юдин

Повышение точности приближенных решений краевой задачи на полупрямой 321
Е.П.Жидков, А.Г.Соловьев

Применение теплицевых матриц для решения задач квантовой механики 327
И.В.Ибрагимов, Е.Е.Тыртышников

Методика преобразования координат локальных детекторов в единую систему 332
З.М.Иванченко, И.М.Иванченко, Н.Н.Карпенко

Параллельная система обработки экспериментальных данных (ПАРИС-96) 337
И.М.Иванченко, В.В.Кореньков, Ю.А.Кретов

Распространение радиоволн в неоднородном пространстве с нерегулярной границей раздела между двумя средами 341
А.С.Ильинский Т.Н.Галишникова

Монотонные разностные схемы бегущего счета на неравномерных сетках 347
П.П.Матус

Гидродинамический подход к моделированию динамики морских волн, создаваемых ураганом 352
В.Н.Масленникова, М.Е.Боговский

Maple-пакет для построения решений линейных обыкновенных дифференциальных уравнений в виде степенных рядов 360
А.А.Рябенко

Операторно-разностные схемы для задач математической физики 363
А.А.Самарский, П.Н.Вабищевич

Итерационное решение ретроспективной обратной задачи теплопроводности 372
А.А.Самарский, П.Н.Вабищевич, В.И.Васильев

Author Index

On numerical–analytic investigation of scalar scattering by ideal prolate spheroids and triaxial ellipsoids [1]

A. A. Abramov, A. L. Dyshko, N. B. Konyukhova, and T. V. Levitina

Computing Center, Russian Academy of Sciences.
Vavilova 40, 117967 Moscow GSP-1, Russia
e-mail: alalabr@ccas.ru, nadja@ccas.ru, lev@ccas.ru

Abstract. The brief survey of papers [1-6] is given. The exterior Dirichlet and Neumann problems for the Helmholtz equation are considered for scatterers of a shape pointed out in the title. The problems are solved by the Fourier method via separation of variables in either spheroidal or ellipsoidal coordinate system. In so doing, complicated special functions arise, which are calculated by methods elaborated by the authors. The methods allow one to carry out calculations for the listed above classical scattering problems in a wide range of parameter variations. Calculations of distinct terms in the Fourier series are independent of each other, therefore they can be done in parallel. The calculations for a triaxial ellipsoid have been carried out on a multiprocessor computer based on a net of transputers.

1. The plane monochromatic wave $u^i(\vec{r}) = \exp\left[ik(\vec{l}, \vec{r})\right]$ propagates through a homogeneous isotropic space and hits a finite obstacle D with surface S being either a spheroid or a triaxial ellipsoid. Here the time dependence factor $\exp(-i\omega_0 t)$ is omitted, ω_0 is the incident wave frequency, $k = \omega_0/c_0$ is the wave number, c_0 is the sound speed, $\vec{l} = (l_x, l_y, l_z)$ is the unit vector specifying the direction of the plane wave propagation $\vec{r} = (x, y, z)$ is the radius vector of the observation point. Outside the obstacle both the incident $u^i(\vec{r})$ and scattered $u^s(\vec{r})$ waves satisfy the Helmholtz equation

$$\Delta u + k^2 u = 0, \quad \vec{r} \in \mathbf{R}^3/D, \tag{1}$$

where $\Delta \equiv \partial^2/\partial x^2 + \partial^2/\partial y^2 + \partial^2/\partial z^2$ is the Laplace operator. We suppose as well that on S either a Dirichlet or Neumann boundary condition is imposed:

$$u^s(\vec{r})|_s = -u^i(\vec{r})|_s \quad \text{or} \quad \partial u^s(\vec{r})/\partial n|_s = -\partial u^i(\vec{r})/\partial n|_s. \tag{2}$$

Here $\partial/\partial n$ is the derivative in the direction of the exterior normal to the surface S.

Far from the scattering object the scattered field is close to a spherical wave

$$u^s(r, \theta, \varphi) = F(\theta, \varphi)\frac{\exp(ikr)}{kr} + \mathrm{O}\Big[\frac{1}{r^2}\Big], \quad r \to \infty, \tag{3}$$

uniformly with respect to (θ, φ), where (r, θ, φ) are spherical coordinates. For the solutions of the Helmholtz equation relation (3) is equivalent to the Sommerfeld radiation condition

$$\lim_{r\to\infty} r\left[\frac{\partial u^s}{\partial r} - ik\,u^s\right] = 0 \tag{4}$$

[1]This work was supported by the Russian Foundation for Basic Research under Grant No. 96-01-00951.

uniformly with respect to (θ, φ).

The solution of (1)&(2)&(4) yields both the far field amplitude $F(\theta, \varphi)$ from the asymptotic representation (3), and other principal scattering characteristics of the ellipsoid S such as the differential and total cross-section. For a more detailed discussion of the problem see, e.g, [1, 2].

Since the scattering surface may be identified with a coordinate surface of an orthogonal curvilinear coordinate system which enables separation of the variables in the Helmholtz equation, this problem may be solved 'exactly' via the Fourier method.

2. In case of a prolate spheroid the solution of the above stated problem via the Fourier method has been considered carefully in [1] and the literature cited therein. For methods for computing the spheroidal functions one should be referred to [4,5].

Further we will consider briefly the case of an ellipsoid. The ellipsoidal coordinate system associated with the ellipsoid S of principal semiaxes $a > b > c > 0$ is defined by the equation $x^2/(a^2+\eta) + y^2/(b^2+\eta) + z^2/(c^2+\eta) = 1$. Going over to dimensionless variables $\xi = (\eta + a^2)/(a^2 - b^2)$, one writes

$$\frac{x^2}{(a^2-b^2)\xi} + \frac{y^2}{(a^2-b^2)(\xi-1)} + \frac{z^2}{(a^2-b^2)(\xi-\rho^2)} = 1.$$

Here we have set $\rho^2 = (a^2-c^2)/(a^2-b^2)$, i.e. $\rho^2 > 1$. This equation for each point (x, y, z), $xyz \neq 0$, has exactly three different roots: $\xi_1 \in I_1 = (0, 1)$, $\xi_2 \in I_2 = (1, \rho^2)$, and $\xi_3 \in I_3 = (\rho^2, \infty)$. If $\xi = \xi_1$, the surface defined above is a hyperboloid of two sheets, for $\xi = \xi_2$ it is a hyperboloid of one sheet and for $\xi = \xi_3$ it is an ellipsoid (e.g., the equation $\xi_3 = \xi_3^* = a^2/(a^2-b^2)$ defines S).

The above three families of confocal surfaces form an orthogonal curvilinear coordinate system related to Cartesian coordinates within each octant (e. g., $x \geq 0$, $y \geq 0$, $z \geq 0$) by the one–to–one correspondence.

On separating variables, we arrive at a system of three identical equations known as the Lamé wave equations

$$\sqrt{f(\xi_i)}\frac{d}{d\xi_i}\left\{\sqrt{f(\xi_i)}\frac{d\Lambda_i}{d\xi_i}\right\} + \frac{1}{4}q(\xi_i, h, l, \omega)\Lambda_i = 0, \quad \xi_i \in I_i, \quad i = 1, 2, 3, \tag{5}$$

where h, l are the separation constants, $f(\xi) = \xi(\xi-1)(\xi-\rho^2)$, $q(\xi, h, l, \omega) = h\rho^2 - l\rho^2\xi + \omega^2\xi^2$, $\omega^2 = k^2(a^2-b^2)$.

Owing to the symmetry of the problem, a particular solution must be either odd or even with respect to each coordinate plane. Since coordinate planes correspond to the ends of the intervals of ellipsoidal coordinate variations, we have

$$\begin{aligned}
&\lim_{\xi_1 \to +0} \Lambda_1(\xi_1) = 0 \quad \text{or} \quad \lim_{\xi_1 \to +0}\left(\sqrt{f(\xi_1)}\Lambda'(\xi_1)\right) = 0,\\
&\lim_{\xi_1 \to 1-0} \Lambda_1(\xi_1) = 0 \quad \text{or} \quad \lim_{\xi_1 \to 1-0}\left(\sqrt{f(\xi_1)}\Lambda_1'(\xi_1)\right) = 0,\\
&\lim_{\xi_2 \to 1+0} \Lambda_2(\xi_2) = 0 \quad \text{or} \quad \lim_{\xi_2 \to 1+0}\left(\sqrt{-f(\xi_2)}\Lambda_2'(\xi_2)\right) = 0,\\
&\lim_{\xi_2 \to \rho^2-0} \Lambda_2(\xi_2) = 0 \quad \text{or} \quad \lim_{\xi_2 \to \rho^2-0}\left(\sqrt{-f(\xi_2)}\Lambda_2'(\xi_2)\right) = 0,
\end{aligned} \tag{6}$$

besides the boundary conditions for $\xi_1 \to 1-0$ and for $\xi_2 \to 1+0$ must be identical as well as the conditions for $\xi_2 \to \rho^2 - 0$ and $\xi_3 \to \rho^2 + 0$.

No reliable methods for the evaluation of ellipsoidal wave functions existed until papers [5,6] appeared and therefore the application of these functions had been delayed (see paper *A. A. Abramov, A. L. Dyshko, N. B. Konyukhova, T. V. Levitina, V. I. Ul'yanova* of this volume). This is because the solutions to (5) defined on I_1 and I_2 — the angle ellipsoidal wave functions — compose the eigenfunctions of a two–parameter self–adjoint eigenvalue problem. That is, one must find such a couple of values for the separation constants, that (5) has on both I_1 and I_2 non-trivial solutions satisfying the corresponding boundary conditions. The eigenvalue of the problem is such a couple of values, and the product of the corresponding solutions to the system gives its eigenfunction (see, e.g., [7]).

Eigenfunctions associated with different eigenvalues are orthogonal on $\Pi = I_1 \times I_2$ with respect to the weight $\sigma(\xi_1, \xi_2) = (\xi_2 - \xi_1)\,[-f(\xi_1)f(\xi_2)]^{-1/2}$. Since an eigenfunction is defined only on a part of S enclosed in an octant, one must to prolong it to the hole space in accordance with its parity, i.e., with the boundary conditions chosen from (6). We will call such a prolonged function a surface ellipsoidal function $\Psi(\xi_1, \xi_2)$. Normalized by

$$\iint\limits_{\Pi} \left[\Psi(\xi_1, \xi_2)\right]^2 \sigma(\xi_1, \xi_2)\, d\xi_1 d\xi_2 = 1/8,$$

all the functions for all possible boundary conditions constitute an orthonormal basis in $\mathcal{L}^2(S, \sigma)$.

To solve the problem under consideration we also need the radial ellipsoidal functions, i.e., the solutions to (5) defined on a semi–infinite interval I_3 and satisfying at infinity the radiation boundary condition

$$\Lambda^{(3)}(\xi_3) = \frac{1}{\sqrt{\omega\xi_3}} \exp\{i(\omega\sqrt{\xi_3} + \delta(h,l)\} + O(\frac{1}{\xi_3}), \quad \xi_3 \to \infty.$$

Here the phase shift $\delta(h,l)$ is determined (up to a summand $n\pi$) by the requirement that $\Lambda^{(1)}(\xi_3) = \mathrm{Re}\Lambda^{(3)}(\xi_3)$ is subject to (6). Such a nontrivial solution exists for any h and l.

The methods proposed in [4–6] provide an efficient and economic way to calculate the wave functions for spheroids and ellipsoids at a rather wide range of parameter variations, they are involved in the survey [8]. .

Thus, when separating the variables in terms of ellipsoidal (spheroidal) coordinates, particular solutions of the Helmholtz equation are formed as products of the surface and radial ellipsoidal (spheroidal) wave functions. For comparison, in case of a sphere the analogue of a surface ellipsoidal wave function is a spherical function, while the radial ones are the Bessel and Hankel functions.

3. Here we do not dwell on the expanding of the incident and scattered wave as well as the far field amplitude. For details we refer to [1–3]. We will only present the ultimate expressions for the case of an ellipsoid.

For the plane wave propagating in the direction $\vec{l}$ specifeid with spherical coordinates θ' and φ' (the angle θ' is counted off the axis x) we have the expansion in ellipsoidal wave functions at an observation point $\vec{r} = (\xi_1, \xi_2, \xi_3)$

$$e^{-ik(\vec{l},\vec{r})} = -\frac{16\pi i}{\sqrt{\omega}} \sum_n e^{i\delta_n} \Psi_n(\xi_1, \xi_2) \Psi_n(\xi_1', \xi_2') \Lambda_n^{(1)}(\xi_3). \tag{7}$$

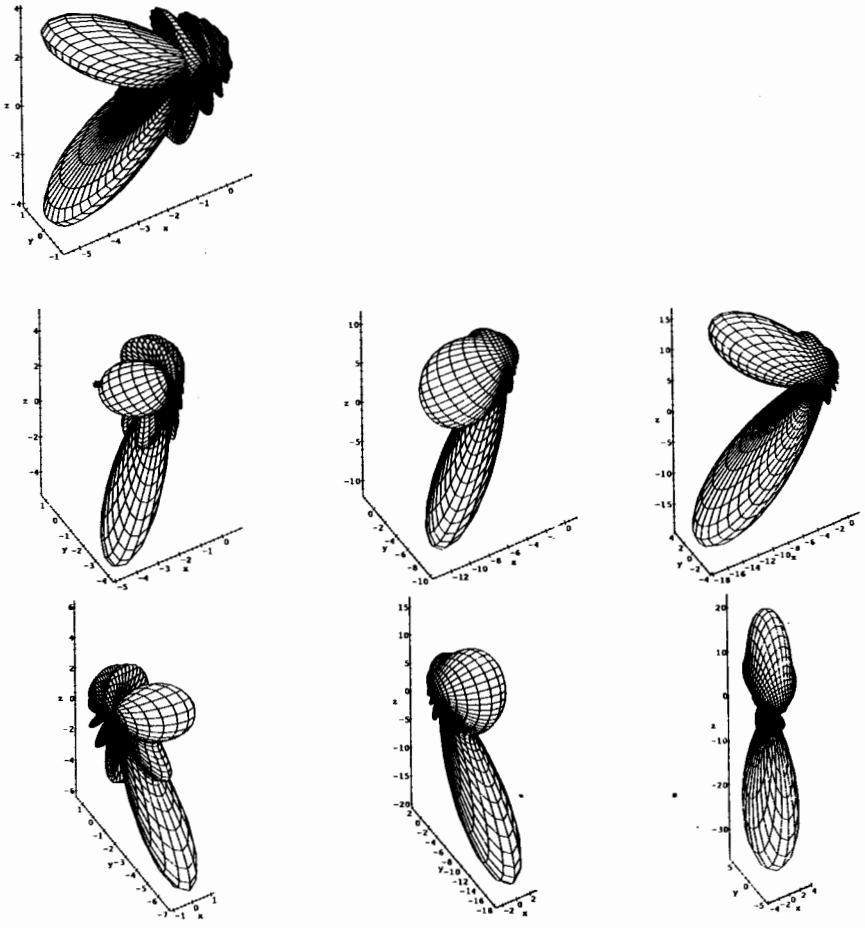


Figure 1.

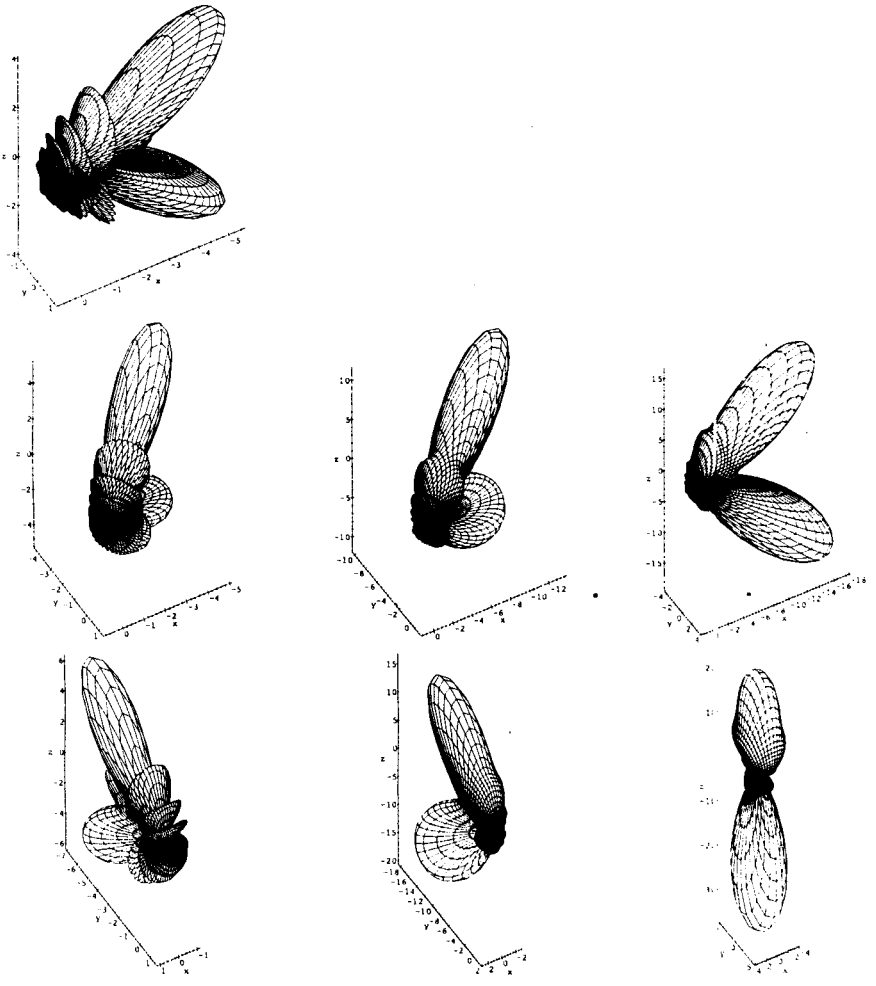


Figure 2.

Here angle ellipsoidal coordinates ξ_1' and ξ_2' specifyi ıg the incident direction are related to the angles θ' and φ' by the equations

$$\begin{aligned} \xi_1'\xi_2' &= \rho^2\cos^2\theta', \\ (\xi_1'-1)(\xi_2'-1) &= -(\rho^2-1)\sin^2\theta'\cos^2\varphi', \\ (\xi_1'-\rho^2)(\xi_2'-\rho^2) &= \rho^2(\rho^2-1)\sin^2\theta'\sin^2\varphi'. \end{aligned} \tag{8}$$

The scattered wave may be represented as a series

$$u^s(\vec{r}) = \sum_n u_n \Psi_n(\xi_1,\xi_2)\Psi_n(\xi_1',\xi_2')\Lambda_n^{(3)}(\xi_3). \tag{9}$$

where

$$u_n = 16\pi i \frac{1}{\sqrt{\omega}} e^{i\delta_n} \begin{cases} \Lambda_n^{(1)}(\xi_3^*)/\Lambda_n^{(3)}(\xi_3^*) & \text{for the Dirichlet problem,} \\ \dfrac{d}{d\xi_3}\Lambda_n^{(1)}(\xi_3^*)\Big/\dfrac{d}{d\xi_3}\Lambda_n^{(3)}(\xi_3^*) & \text{for the Neumann problem.} \end{cases}$$

Finally the far field amplitude of the ellipsoid S is given by

$$F(\theta,\varphi,\theta',\varphi') = \sqrt{\omega}\sum_n u_n e^{i\delta_n}\Psi_n(\xi_1,\xi_2)\Psi_n(\xi_1',\xi_2'), \tag{10}$$

where the angle ellipsoidal coordinates ξ_1, ξ_2 are connected with the angle spherical coordinates θ and φ via the same relations as connect ξ_1', ξ_2' with θ' and φ' (see (8)). The expressions (7), (9) and (10) have been first obtained by M. V. Fedoryuk in [9].

These expressions are simplified, if the direction of the plane wave incidence coincides with one of the coordinate axes. Therefore in [2,3] th e consideration has been restricted to the case of incidence in the direction parallel to the axis OX, $u^i(\vec{r}) = \exp(-ikx)$, still these calculations require considerable time. That is why the parallel version of the programs evaluating a plane wave field at a given point and a far field amplitude for a triaxial ellipsoid have been devised in "Parsytec C" languag e for a net of transputers. The idea of parallel realization is natural for the Fourier metho d, as distinct harmonics in the series are calculated independently of each other, and the efore can be calculated in parallel.

As a basis of parallel programs the so-called 'farm–processor' concept exploits: there is 'master' processor, which sends tasks (i.e., the number n of an eigenvalue and a set of boundary conditions to 'worker' processors, receives the replay (the array of values of the corresponding ellipsoidal wave functions $\Lambda_n(\xi_1)$, $\Lambda_n(\xi_2)$, $\Lambda_n^{(3)}(\xi_3)$ at given points) and adds the new term to the series. The time required to calculate any term of the above series, is much more than that of data exchange among processors, therefore regardless of the configuration of a transputer net the speeding up may be attained up to nearly p times, where p is the amount of occupied processors. (Here we assume that $p \ll N$, where N is a total number of calculated harmonics.)

In case of an arbitrary incident direction the total amount of calculations becomes four times more, therefore the possibility to carry out computations in parallel is of special significance.

Here we present the results of several calculations done by means of the above programs. Fig. 1 and Fig. 2 show the same far field amplitudes of ideally rigid ellipsoids (the Neumann problem) observed from opposite points of view. All the calculations have been done for $a^2k^2 = 105.0$, $b^2k^2 = 55.0$ and $c^2k^2 = 5.0$. For both pictures the incident directions specified by spherical angles φ', θ' are given by the table:

$\varphi' = 0 \quad \theta' = 0$		
$\varphi' = 0 \quad \theta' = \pi/4$	$\varphi' = \pi/4 \quad \theta' = \pi/4$	$\varphi' = \pi/2 \quad \theta' = \pi/4$
$\varphi' = 0 \quad \theta' = \pi/2$	$\varphi' = \pi/4 \quad \theta' = \pi/2$	$\varphi' = \pi/2 \quad \theta' = \pi/2$

Calculation of each series continued till all the series terms corresponding to the main index n became sufficiently small. Besides, the particular sums of the plane wave expansion were compared with its exact value, in order to test ellipsoidal wave functions and to estimate the number of harmonics required. To verify the reliability of the program we have compared the results of calculations for nearly prolate spheroid with the calculations obtained for a corresponding spheroid in [2]. We have found that these calculations are in a good correspondence with each other. Note that the far field amplitudes for the case of an arbitrary incident direcition are published for the first time.

References

[1] A. A. Abramov, N. B. Konyukhova, and T. V. Levitina, Differentsial'nye Uravneniya **29** (8), 1347 (1993).

[2] A. A. Abramov, A. L. Dyshko, N. B. Konyukhova, and T. V. Levitina, Zh. Vychisl. Matem. i Matem. Fiz. **35** (9), 1374 (1995).

[3] A. A. Abramov, N. B. Konyukhova, and T. V. Levitina, Design Engineering Technical Conferences (ASME, New–York) **3**, Part B, 437 (1995).

[4] A. A. Abramov, A. L. Dyshko, N. B. Konyukhova, T. V. Pak, and E. S. Pariiskii, Zh. Vychisl. Matem. i Matem. Fiz. **24** (1), 3 (1984).

[5] A. A. Abramov, A. L. Dyshko, N. B. Konyukhova, and T. V. Levitina, Zh. Vychisl. Matem. i Matem. Fiz. **31** (2), 212 (1991).

[6] A. A. Abramov, A. L. Dyshko, N. B. Konyukhova, and T. V. Levitina, Zh. Vychisl. Matem. i Matem. Fiz. **29** (6), 813 (1989).

[7] H. Volkmer, Multiparameter Eigenvalue Problems and Expansion Theorems, (Springer–Verlag, Berlin–Heidelberg) (1988).

[8] D. W. Lozier, and F. W. Olver, Proc. of Symposia in Appl. Math. (ed. W. Gautschi, AMS, Providence) **48**, 79 (1994).

[9] M. V. Fedoryuk, Differentsial'nye Uravneniya **25** (11), 1990 (1989).

A Universal Program to Uncouple Linear Systems *

Sergei A. Abramov
Computer Center of
the Russian Academy of Science,
Vavilova 40, Moscow 117967, Russia.
abramov@ccas.ru
sabramov@cs.msu.su

Eugene V. Zima
Comput. Math. & Cybernetics dept.,
Moscow State University,
Moscow 119899,
Russia.
zima@cs.msu.su

Abstract. We describe a program to uncouple a linear system of equations. An algorithm is described in terms of an arbitrary Ore polynomial ring. It allows to adjust the program for the differential, difference, q-difference and many other cases.

1 Introduction

The search for solutions of a system of linear ordinary differential equations can be reduced to the search for solutions of scalar (i.e., with one unique unknown function) linear equations. Such a reduction is possible for systems of linear difference and q-difference equations as well.

Theoretically, this reduction is not hard for any kind of system. Algorithms to uncouple linear differential and difference systems are based on very similar ideas. There is a mathematical abstraction connecting, in particular, linear ordinary differential, difference and q-difference equations. This is Ore polynomial rings [1, 2, 3]. In this paper we describe a **Maple** implementation of a universal (based on Ore polynomials) algorithm to uncouple a linear system to scalar equations. The implementation assumes that coefficients are polynomials (in the usual sense). If coefficients are rational functions then preliminary clearing of denominators is needed.

The algorithm has been described by the authors in [4]. It is a generalization to an arbitrary Ore polynomial ring of an algorithm to uncouple a system of linear ordinary differential equations [5].

In Section 2 we recap basic notions connected with Ore polynomial rings. Then, in Section 3 we describe shortly the uncoupling algorithm. Section 4 discusses a **Maple** implementation. A few examples of uncoupling can be found in Section 5.

*Work reported herein was supported in part by INTAS-RFBR under Grant 95-412 and State program "Informatization of Russia".

2 Basic notions of Ore polynomial rings

Let k be a field of characteristic zero, θ an indeterminate over k, σ an automorphism of k, and $\delta : k \to k$ a map satisfying

$$\delta(a+b) = \delta a + \delta b, \quad \delta(ab) = \sigma(a)\,\delta b + \delta a\, b \tag{1}$$

for any $a, b \in k$. The Ore polynomial ring $k[\theta; \sigma, \delta]$ given by σ and δ is the ring of polynomials in θ over k with the usual addition. Multiplication in $k[\theta; \sigma, \delta]$ is defined by the following relation

$$\theta a = \sigma(a)\theta + \delta a \tag{2}$$

for any $a \in k$.

Let V be a linear space over k (if k is a functional field then usually V is a wider set of functions). Let the application of θ to elements of V be defined in such a way that

$$\theta(u+v) = \theta u + \theta v, \quad \theta(au) = \sigma(a)\theta u + \delta a\, u \tag{3}$$

for $u, v \in V, a \in k$. Then we can consider the application of Ore polynomials $A(\theta), B(\theta), \ldots$ to elements of V, with $A(\theta)B(\theta)v = A(\theta)(B(\theta)v), (A(\theta)+B(\theta))v = A(\theta)v + B(\theta)v$ and so on. It allows to consider V as a module over $k[\theta; \sigma, \delta]$. From here on $\theta a, \theta u, \ldots$ we denote the results of applying θ to $a, u, \ldots$ rather than products of the elements of $k[\theta; \sigma, \delta]$ as in (2).

Let $k \subset V$ (this inclusion is natural in the functional case). Thanks to (3),

$$\theta a = \sigma(a)\theta 1 + \delta(a) \tag{4}$$

for $a \in k$. Therefore, in order to define the result of application of θ to an element of k it is enough (in addition to σ and δ) to know $\theta 1$, i.e. the result of application of θ to the unit of k. The following table shows that linear ordinary differential, recurrent, difference and q-analogs of recurrent and differential equations can be considered in the framework of the theory of Ore polynomial rings:

case	θ	σ	δ	$\theta 1$
differential	$\theta f(x) = \frac{d}{dx} f(x)$	1_k	$\frac{d}{dx}$	)
recurrent	$\theta f(x) = Ef(x) = f(x+1)$	E	0	1
difference	$\theta f(x) = \Delta f(x) = f(x+1) - f(x)$	E	Δ	)
q-recurrent	$\theta f(x) = Qf(x) = f(qx)$	Q	0	1
q-differential	$\theta f(x) = D_q f(x) = \frac{f(qx)-f(x)}{qx-x}$	Q	D_q	)

The list above is not exhaustive. It can be extended easily by special cases like and so on.

a nonstandard	$\theta f(x) = x\frac{d}{dx}f(x)$	1_k	$x\frac{d}{dx}$	0
q-difference	$\theta f(x) = \Delta_q f(x) = f(qx) - f(x)$	Q	Δ_q	0

3 The algorithm

Consider a linear system (in general, inhomogeneous)

$$f_1\theta y_1 = a_{11}y_1 + a_{12}y_2 + ... + a_{1n}y_n + r_1,$$
$$...$$
$$f_n\theta y_n = a_{n1}y_1 + a_{n2}y_2 + ... + a_{nn}y_n + r_n,$$

$a_{ij} \in k$, $r_i, y_i \in V$, where y_i are unknown. The result of the application of θ is defined in accordance with (3).

The application of the algorithm starts with the following transformations, which give a scalar equation in y_1.

Consider the first equation. If $a_{12} = ... = a_{1n} = 0$ then we already have an equation in y_1 alone. Let $a_{12} \neq 0$ (if $a_{12} = 0$, but $a_{1i} \neq 0, i > 2$, we can re-enumerate unknowns). Introduce new unknown z_2 (instead of y_2):

$$z_2 = a_{12}y_2 + ... + a_{1n}y_n. \tag{5}$$

The first equation can be rewritten as

$$f_1\theta y_1 = a_{11}y_1 + z_2.$$

In all the remaining equations we can eliminate y_2 w th the help of (5). After this elimination the second equation will have the form

$$g_2\theta y_2 = b_{21}y_1 + b_{22}z_2 + ... + b_{2n}y_n + s_2. \tag{6}$$

Application of θ to both the sides of (5) gives us

$$\theta z_2 = \sigma(a_{12})\theta y_2 + \delta(a_{12})y_2 + \ldots + \sigma(a_{1n})\theta y_n + \delta(a_{1n})y_n. \tag{7}$$

We can use (6), (5) together with

$$g_3\theta y_3 = b_{31}y_1 + b_{32}z_2 + b_{33}y_3 + ... + b_{3n}y_n + s_3,$$
$$...$$
$$g_n\theta y_n = b_{n1}y_1 + b_{n2}z_2 + b_{n3}y_3 + .. + b_{nn}y_n + s_n$$

in order to eliminate $y_2, \theta y_2, \theta y_3, \ldots, \theta y_n$ from (7). After that (7) contains only $\theta z_2, y_1, y_3, \ldots, y_n$:

$$h_2\theta z_2 = c_{21}y_1 + c_{22}z_2 + c_{23}y_3 + ... + c_{2n}y_n + t_2.$$

Hence we have as the result of these actions the system

$$\begin{aligned}
&f_1\theta y_1 = a_{11}y_1 + z_2 + r_1,\\
&h_2\theta z_2 = c_{21}y_1 + c_{22}z_2 + c_{23}y_3 + \ldots + c_{2n}y_n + t_2,\\
&g_3\theta y_3 = b_{31}y_1 + b_{32}z_2 + b_{33}y_3 + \ldots + a_{3n}y_n + s_3,\\
&\ldots\\
&g_n\theta y_n = b_{n1}y_1 + b_{n2}z_2 + b_{n3}y_3 + \ldots + a_{nn}y_n + s_n.
\end{aligned} \tag{8}$$

If $c_{23} = c_{24} = \ldots = c_{2n} = 0$ we can obtain a second order equation in y_1 from the first two equations of (8). The first equation allows to express z_2 via $y_1, \theta y_1$. Applying θ to the first equation we can express θz_2 via $y_1, \theta y_1, \theta^2 y_1$. Substituting in the second equation we have the desired result.

If at least one of $c_{23}, c_{24}, \ldots, c_{2n}$ (for example c_{23}) is nonzero, then we introduce the new unknown z_3 (instead of y_3)

$$z_3 = c_{23}y_3 + \ldots + c_{2n}y_n, \tag{9}$$

and so on until for a value $l \leq n$ we obtain a system

$$\begin{aligned}
&\alpha_1\theta y_1 = \beta_{11}y_1 + z_2 + r_1,\\
&\alpha_2\theta z_2 = \beta_{21}y_1 + \beta_{22}z_2 + z_3 + t_2,\\
&\ldots\\
&\alpha_{l-1}\theta z_{l-1} = \beta_{l-1,1}y_1 + \beta_{l-1,2}z_2 + \ldots + \beta_{l-1,l-1}z_{l-1} + z_l + u_{l-1},\\
&\alpha_l\theta z_l = \beta_{l,1}y_1 + \beta_{l,2}z_2 + \ldots + \beta_{l,l-1}z_{l-1} + \beta_{l,l}z_l + u_l.
\end{aligned} \tag{10}$$

We can get an equation of degree l in y_1 with the help of these equations.

In the rest of the initial system we consider $y_1, z_2, \ldots, z_l$ as knowns and perform the same transformations. It gives us an equation of order $m, m \leq n - l$, for y_{l+1} and so on. Finally, we will have several scalar equations (in y_1, y_{l+1} etc). The sum of their orders will be equal to n. Relations, analogous to (10), allow to express all z_i via $y_1, y_{l+1}, \ldots$. Formulas (5), (9) and so on form a triangle system of linear algebraic equations. Hence, we can express all remaining y_i via solutions of scalar equations.

Let the initial conditions for equations of the initial system be given. Then it is easy to get initial conditions for split scalar equations. For example for an l-order equation in y_1 one can find linear expressions of $\theta y_1, \ldots, \theta^{l-1}y_1$ via $y_1, \ldots, y_n$. If $y_1, \ldots, y_n$ are functions and the values $y_1(0), \ldots, y_n(0)$ are known, then these linear expressions allow to compute $\theta y_1(0), \ldots, \theta^{l-1}y_1(0)$. In order to construct the linear expressions mentioned above, we apply θ to the first equation of the initial system and then eliminate in the right-hand side all θy_i by means of equations of the initial system. This gives a linear expression for $\theta^2 y_1$. Repeating these steps we obtain expressions for $\theta^3 y_1, \theta^4 y_1, \ldots$ Analogous expressions can be obtained for the scalar equation in y_{l+1} and so on.

4 The program

The algorithm from the previous section has been implemented in **Maple 5.3** with several improvements like

- preliminary splitting of scalar equations on each step,

- reducing coefficients by their gcd,
- choosing the most suitable unknown for excluding on each step

and others.

The user is provided with the three main procedures **set_Ore_ring**, **uncpl_sys** and **get_sol**. Procedure **set_Ore_ring(indvar, case)** sets the program to a concrete $k[\theta;\sigma,\delta]$ by selecting the independent variable and concrete δ, σ and $\theta 1$. It takes unassigned names as parameters. The value of the first parameter stands for the name of the independent variable, the value of the second parameter selects a concrete $k[\theta;\sigma,\delta]$. In the standard cases (**differential**, **difference**, **recurrent**, **q-recurrent**, **q-difference**, **q-differential**) adjustment of the program is hidden from the user; for example, the call

```
set_Ore_ring(x,differential)
```

sets the program to the differential case with x as independent variable.

If the value of the second parameter is **nonstandard**, the program prompts the user for input of definitions of δ and σ (in procedural form) and the value of $\theta 1$. Below we show a Maple session setting the program to the nonstandard case:

```
> set_Ore_ring(x, 'nonstandard');
   Enter definition of automorphism sigma
> proc(y) y end;
   Enter definition of map delta
> proc(y) x^2*diff(y,x) end;
   What is the value of theta_1 ?
> 0
   program is set to the  nonstandard case
```

Procedure **uncpl_sys(sys,unklst)** takes a system of linear equations as the first parameter, a list of unknowns as the second parameter and returns the result of uncoupling. The result is a list of scalar equations, expressions of auxiliary unknowns in terms of solutions of these equations and linear expressions of the remaining unknowns via solutions of the scalar equations and the auxiliary unknowns. There is the optional third parameter **ch_flag** of type Boolean, which is set **true** by default. If **ch_flag=false** then the choice of the most suitable unknown for excluding is switched off.

Procedure **get_sol(unres, unkl)** takes the result of uncoupling and the list of unknowns as parameters and finds a solution of the initial system by means of standard Maple tools. In the differential case the procedure **dsolve** is used for this purpose. If the optional third papameter has the value **rat**, then Bronstein's procedure **ratlode** (from the Maple share library) to find rational solutions is used instead of **dsolve**.

5 Examples of uncoupling

In this section we give several examples of using the program from the previous section together with timings[1].

[1] All the reported timings were obtained on 66Mhz IBM PC 486DX.

Example 1. Consider the well known combinatorial problem: n pairs of inimical knights are invited to dinner. How many ways $y_1(n)$ are there to seat them at a round table so that no enemies are beside each other? Let $y_2(n)$ be the number of ways to seat them so that exactly one pair of enemies is beside each other; $y_3(n)$ be the number of ways to seat them so that exactly two pairs of enemies are beside each other. It is possible to show [6], that for $n \geq 2$

$$\begin{aligned} Ey_1(n) &= 2(2n-1)ny_1(n) + 2(2n-1)y_2(n) + 2y_3(n), \\ Ey_2(n) &= 4n(n+1)y_1(n) + 2(n+1)y_2(n), \\ Ey_3(n) &= 2n(n+1)y_1(n) + 2(n+1)y_2(n), \end{aligned} \tag{11}$$

$y_1(2) = 2, y_2(2) = 0, y_3(2) = 4$. In the session below we first set the program to the recurrent case and then (after initial assignments) call the procedure uncpl_sys:

```
> set_Ore_ring(n, 'recurrent'):
> eq1:= theta(y1(n),n) =
>       2*(2*n-1)*n*y1(n) + 2*(2*n-1)*y2(n) + 2*y3(n):
> eq2:= theta(y2(n),n) =
>       2*(n+1)*2*n*y1(n) + 2*(n+1)*y2(n):
> eq3:= theta(y3(n),n) =
>       2*(n+1)*n*y1(n) + 2*(n+1)*y2(n):
> eq:=[eq1,eq2,eq3]:
> varl:=[y1(n), y2(n), y3(n)]:
> rez:=uncpl_sys(eq, varl);
```

$$\begin{aligned} rez := \Big[& [(64n^2+32n+8n^4+40n^3)\,y1(n) + (-1-n)\,\theta(\theta(\theta(y1(n)))) \\ & + (20n^2+20+34n+4n^3)\,\theta(\theta(y1(n))) \\ & + (40+108n+108n^2+8n^4+48n^3)\,\theta(y1(n)) = 0, \\ & y3_z(n) = \theta(y1(n)) - (4n^2-2n)\,y1(n), y2_z(n) = \%1 - 6\,\theta(y1(n))\,n \\ & - 2\,\theta(y1(n)) - 4\,\theta(y1(n))\,n^2 - (28n^2+12n+16n^3)\,y1(n)], \\ & \left[y2(n) = \frac{y2_z(n)}{16n+8+8n^2}, y3(n) = \frac{1}{2}\,y3_z(n) - \frac{1}{2}\,\frac{(4n-2)\,y2_z(n)}{16n+8+8n^2} \right] \Big] \\ & \%1 := \theta(\theta(y1(n))) \end{aligned}$$

Here (in 2 seconds) we get as the result one scalar equation in $y1(n)$, expressions of auxiliary functions via $y1(n)$ and linear expressions of $y2(n), y3(n)$ via $y1(n)$ and auxiliary functions. Using the remark at the end of Section 3, it is possible to get the initial conditions for the scalar equation $y1(2) = 2$, $y1(3) = 32$, $y1(4) = 1488$. The scalar equation has the following solution with these initial conditions

$$y_1(n) = \sum_{k=0}^{n} (-1)^k \binom{n}{k} 2^k (2n-k-1)!\,.$$

Example 2. In order to demonstrate the work of the program in the q-recurrent case we consider the system (11) using Q instead of E.

```
> set_Ore_ring(n, 'q-recurrent'):
> eq1:= theta(y1(n),n) =
>       2*(2*n-1)*n*y1(n) + 2*(2*n-1)*y2(n) + 2*y3(n):
> eq2:= theta(y2(n),n) =
>       2*(n+1)*2*n*y1(n) + 2*(n+1)*y2(n):
> eq3:= theta(y3(n),n) =
>       2*(n+1)*n*y1(n) + 2*(n+1)*y2(n):
> eq:=[eq1,eq2,eq3]:
> varl:=[y1(n), y2(n), y3(n)]:
> rez:=uncpl_sys(eq, varl);
```

$$
\begin{aligned}
rez := \Big[&[(8n^2q + 8n^3q^2 + 8n^2q^2 + 8nq)\,y1(n) + (2q + 4n^2q^4)\,\theta(\theta(y1(n)))\\
&+ (12n^2q^3 + 8n^3q^4 + 4nq^2 - 4n^2q^2 - 4nq)\,\theta(y1(n))\\
&- \theta(\theta(\theta(y1(n)))) = 0,\\
&y3_z(n) = \theta(y1(n)) - (4n^2 - 2n)\,y1(n), y2_z(n) = \%1\\
&- 4n^2q^2\,\theta(y1(n)) + 2nq\,\theta(y1(n))\\
&- (-4n^2 - 4n + 16n^3q + 16n^2q)\,y1(n)]\\
&\Big[y2(n) = \frac{y2_z(n)}{8n^2q + 8nq}, y3(n) = \frac{1}{2}\,y3_z(n) - \frac{1}{2}\,\frac{(4n-2)\,y2_z(n)}{8n^2q + 8nq}\Big]\Big]\\
&\%1 := \theta(\theta(y1(n)))
\end{aligned}
$$

Here (in 4 seconds) we again get one scalar equation in $y1(n)$, and expressions for the remaining unknowns via $y1(n)$.

Example 3. In both the examples above we get one scalar equation. Now we consider a differential linear system [8] which will be split into several scalar equations:

$$
\begin{aligned}
x^2y_1' &= (1+4x)y_1 - 5xy_2 + 7xy_3 - 8xy_4 + 8xy_5 - 6xy_6,\\
x^2y_2' &= -10xy_1 + (9x+1)y_2 - 14xy_3 + 16xy_4 - 16xy_5 + 12xy_6,\\
x^2y_3' &= -5xy_1 + 5xy_2 + (1-8x)y_3 + 8xy_4 - 8xy_5 + 6xy_6,\\
x^2y_4' &= 10xy_1 - 10xy_2 + 14xy_3 + (1-17x)y_4 + 16xy_5 - 12xy_6,\\
x^2y_5' &= 5xy_1 - 5xy_2 + 7xy_3 - 8xy_4 + (1+7x)y_5 - 6xy_6,\\
x^2y_6' &= -5xy_1 + 5xy_2 - 7xy_3 + 8xy_4 - 8xy_5 + (1+5x)y_6.
\end{aligned}
\tag{12}
$$

In the session below we first set the program to the differential case and then (after initial assignments) call the procedure **uncpl_sys**:

```
> set_Ore_ring(x, 'differential'):
> eq1:= theta(y1(x),x)*x^2 =
>       (1+4*x)*y1(x) + (-5*x)*y2(x) + (7*x)*y3(x) +
>       (-8*x)*y4(x)  + (8*x)*y5(x)  + (-6*x)*y6(x):
> eq2:= theta(y2(x),x)*x^2 =
>       (-10*x)*y1(x) + (9*x+1)*y2(x) + (-14*x)*y3(x) +
>       (16*x)*y4(x)  + (-16*x)*y5(x)  + (12*x)*y6(x):
> eq3:= theta(y3(x),x)*x^2 =
```

```
>       (-5*x)*y1(x) + (5*x)*y2(x) + (-8*x+1)*y3(x) +
>       (8*x)*y4(x)  + (-8*x)*y5(x)  + (6*x)*y6(x):
> eq4:= theta(y4(x),x)*x^2 =
>       (10*x)*y1(x) + (-10*x)*y2(x) + (14*x)*y3(x) +
>       (1-17*x)*y4(x)  + (16*x)*y5(x)  + (-12*x)*y6(x):
> eq5:= theta(y5(x),x)*x^2 =
>       (5*x)*y1(x) + (-5*x)*y2(x) + (7*x)*y3(x) +
>       (-8*x)*y4(x)  + (1+7*x)*y5(x)  + (-6*x)*y6(x):
> eq6:= theta(y6(x),x)*x^2 =
>       (-5*x)*y1(x) + (5*x)*y2(x) + (-7*x)*y3(x) +
>       (8*x)*y4(x)  + (-8*x)*y5(x)  + (1+5*x)*y6(x):
> eq:=[eq1,eq2,eq3,eq4,eq5,eq6]:
> varl:=[y1(x), y2(x), y3(x), y4(x), y5(x), y6(x)]:
> rez:=uncpl_sys(eq, varl);
```

$$
\begin{aligned}
rez := \Big[&[-x^4\,\theta(\mathrm{y1}(x)) + 3x^3\,\theta(\mathrm{y1}(x)) + 2x^2\,\theta(\mathrm{y1}(x)) + 5x^2\,\mathrm{y1}(x) \\
&- 5x\,\mathrm{y1}(x) - \mathrm{y1}(x) = 0, \mathrm{y2_z}(x) = x^2\,\theta(\mathrm{y1}(x)) - (1+4x)\,\mathrm{y1}(x)], [\\
&-x^2\,\theta(\mathrm{y3}(x)) = 5x\,\mathrm{y1}(x) + \mathrm{y2_z}(x) + (x-1)\,\mathrm{y3}(x), \\
&-x^2\,\theta(\mathrm{y4}(x)) = -10x\,\mathrm{y1}(x) - 2\,\mathrm{y2_z}(x) + (x-1)\,\mathrm{y4}(x), \\
&-x^2\,\theta(\mathrm{y5}(x)) = -5x\,\mathrm{y1}(x) - \mathrm{y2_z}(x) + (x-1)\,\mathrm{y5}(x), \\
&-x^2\,\theta(\mathrm{y6}(x)) = 5x\,\mathrm{y1}(x) + \mathrm{y2_z}(x) + (x-1)\,\mathrm{y6}(x)], \\
&\left[\mathrm{y2}(x) = -\frac{1}{5}\frac{\mathrm{y2_z}(x)}{x} + \frac{7}{5}\mathrm{y3}(x) - \frac{8}{5}\mathrm{y4}(x) + \frac{8}{5}\mathrm{y5}(x) - \frac{6}{5}\mathrm{y6}(x)\right]\Big]
\end{aligned}
$$

Here we get (in 3 seconds) as the result a scalar equation of the second order in $y1(x)$, an expression of auxiliary function $y2_z(x)$ via $y1(x)$, a list of four scalar first-order equations and a linear expression of $y2(x)$ via solutions of the scalar equations and the auxiliary function. Now we can find the solution of the system using procedure **get_sol**:

```
> get_sol(rez,varl);
```

$$
\left[\frac{e^{(-\frac{1}{x})}\,(_C1 + _C2\,x^6)}{x}, -\frac{2}{5}\frac{e^{(-\frac{1}{x})}\,(-3_C1 + 5_C2\,x^6)}{x}, -\frac{e^{(-\frac{1}{x})}\,(_C2\,x^6 - _C1)}{x},\right.
$$
$$
\left.\frac{e^{(-\frac{1}{x})}\,(2_C2\,x^6 + _C1)}{x}, \frac{e^{(-\frac{1}{x})}\,(_C1 + _C2\,x^6)}{x}, -\frac{e^{(-\frac{1}{x})}\,(_C2\,x^6 - _C1)}{x}\right]
$$

Example 4. Consider the previous example in the nonstandard case, taking $\delta = x^2\frac{d}{dx}$:

```
> set_Ore_ring(x, 'nonstandard');
   Enter definition of automorphism sigma
> proc(y) y end;
   Enter definition of map delta
> proc(y) x^2*diff(y,x) end;
```

```
   What is the value of theta_1 ?
> 0
   program is set to the  nonstandard case
> eq1:= theta(y1(x),x) =
>       (1+4*x)*y1(x) + (-5*x)*y2(x) + (7*x)*y3(x) +
>       (-8*x)*y4(x)  + (8*x)*y5(x)  + (-6*x)*y6(x):
> eq2:= theta(y2(x),x) =
>       (-10*x)*y1(x) + (9*x+1)*y2(x) + (-14*x)*y3(x) +
>       (16*x)*y4(x)  + (-16*x)*y5(x)  + (12*x)*y6(x):
> eq3:= theta(y3(x),x) =
>       (-5*x)*y1(x) + (5*x)*y2(x) + (-8*x+1)*y3(x) +
>       (8*x)*y4(x)  + (-8*x)*y5(x)  + (6*x)*y6(x):
> eq4:= theta(y4(x),x) =
>       (10*x)*y1(x) + (-10*x)*y2(x) + (14*x)*y3(x) +
>       (1-17*x)*y4(x)  + (16*x)*y5(x)  + (-12*x)*y6(x):
> eq5:= theta(y5(x),x) =
>       (5*x)*y1(x) + (-5*x)*y2(x) + (7*x)*y3(x) +
>       (-8*x)*y4(x)  + (1+7*x)*y5(x)  + (-6*x)*y6(x):
> eq6:= theta(y6(x),x) =
>       (-5*x)*y1(x) + (5*x)*y2(x) + (-7*x)*y3(x) +
>       (8*x)*y4(x)  + (-8*x)*y5(x)  + (1+5*x)*y6(x):
> eq:=[eq1,eq2,eq3,eq4,eq5,eq6]:
> varl:=[y1(x), y2(x), y3(x), y4(x), y5(x), y6(x)]:
> rez:=uncpl_sys(eq, varl);
```

$$
\begin{aligned}
rez := \Big[&[-\theta(\theta(\mathrm{y1}(x))) + 2\,\theta(\mathrm{y1}(x)) + 5\,x\,\theta(\mathrm{y1}(x)) + 5\,x^2\,\mathrm{y1}(x) - 5\,x\,\mathrm{y1}(x) \\
&- \mathrm{y1}(x) = 0, \mathrm{y2_z}(x) = \theta(\mathrm{y1}(x)) - (1+4\,x)\,\mathrm{y1}(x)], [\\
&-\theta(\mathrm{y3}(x)) = 5\,x\,\mathrm{y1}(x) + \mathrm{y2_z}(x) + (x-1)\,\mathrm{y3}(x), \\
&-\theta(\mathrm{y4}(x)) = -10\,x\,\mathrm{y1}(x) - 2\,\mathrm{y2_z}(x) + (x-1)\,\mathrm{y4}(x), \\
&-\theta(\mathrm{y5}(x)) = -5\,x\,\mathrm{y1}(x) - \mathrm{y2_z}(x) + (x-1)\,\mathrm{y5}(x), \\
&-\theta(\mathrm{y6}(x)) = 5\,x\,\mathrm{y1}(x) + \mathrm{y2_z}(x) + (x-1)\,\mathrm{y6}(x)], \\
&\left[\mathrm{y2}(x) = -\frac{1}{5}\frac{\mathrm{y2_z}(x)}{x} + \frac{7}{5}\mathrm{y3}(x) - \frac{8}{5}\mathrm{y4}(x) + \frac{8}{5}\mathrm{y5}(x) - \frac{6}{5}\mathrm{y6}(x)\right]\Big]
\end{aligned}
$$

Here the result of uncoupling (obtained in 2 seconds) looks slightly different from the previous one. But again we have a second-order scalar equation in $y1(x)$ and four split first-order scalar equations. Using procedure **get_sol** we obtain the same solution as in the standard case:

```
> get_sol(rez,varl);
```

$$
\left[\frac{e^{(-\frac{1}{x})}(_C1 + _C2\,x^6)}{x}, -\frac{2}{5}\frac{e^{(-\frac{1}{x})}(-3\,_C1 + 5\,_C2\,x^6)}{x}, -\frac{e^{(-\frac{1}{x})}(_C2\,x^6 - _C1)}{x},\right.
$$
$$
\left.\frac{e^{(-\frac{1}{x})}(2\,_C2\,x^6 + _C1)}{x}, \frac{e^{(-\frac{1}{x})}(_C1 + _C2\,x^6)}{x}, -\frac{e^{(-\frac{1}{x})}(_C2\,x^6 - _C1)}{x}\right]
$$

Example 5. Now we demonstrate how to use the program in order to get rational solutions of linear systems. Consider the system [7]

$$
\begin{aligned}
y_1' &= y_2,\\
y_2' &= y_3 + y_4,\\
y_3' &= y_5,\\
y_4' &= 2y_1 + 2xy_2 + y_5,\\
y_5' &= x^2y_1 + 2xy_3 + y_6,\\
y_6' &= x^2y_2 - 2y_3.
\end{aligned}
$$

```
> set_Ore_ring(x, 'differential'):
> eq1:= theta(y1(x),x) =   y2(x):
> eq2:= theta(y2(x),x) =   y3(x) + y4(x):
> eq3:= theta(y3(x),x) =   y5(x):
> eq4:= theta(y4(x),x) =   2*y1(x) + (2*x)*y2(x) + y5(x):
> eq5:= theta(y5(x),x) =   x^2*y1(x) + 2*x*y3(x) + y6(x):
> eq6:= theta(y6(x),x) =   x^2*y2(x) -2*y3(x):
> eq:=[eq1,eq2,eq3,eq4,eq5,eq6]:
> varl:=[y1(x), y2(x), y3(x), y4(x), y5(x), y6(x)]:
> rez:=uncpl_sys(eq, varl);
```

$$
\begin{aligned}
rez := \Big[&[-\theta(\theta(\%1))) + 6\,\%1 + 4\,x\,\theta(\%1) = 0, y2_z(x) = \theta(y1(x))],\\
&y3_z(x) = \%1, y5_z(x) = \theta(\%1) - 2\,y1(x) - 2\,x\,\theta(y1(x))),\\
&y6_z(x) = \theta(\theta(\%1))) - 4\,\theta(y1(x))) - 6\,x\,\%1 - 2\,x^2\,y1(x)],\\
&[2\,\theta(y4(x))) = 4\,y1(x) + 4\,x\,y2_z(x) + y5_z(x)], \Big[\\
&y6(x) = \frac{1}{2}\,y6_z(x) + 2\,x\,y4(x), y5(x) = \frac{1}{2}\,y5_z(x),\\
&y3(x) = y3_z(x) - y4(x), y2(x) = y2_z(x)\Big]\Big]\\
&\%1 := \theta(\theta(y1(x)))
\end{aligned}
$$

This result was obtained in 4 seconds and consists of a fifth-order scalar equation in $y1(x)$, a first order equation in $y4(x)$ and linear expressions for the remaining unknowns via solutions of the scalar equations. We can use the procedure **get_sol** with the third parameter set to **rat** in order to find all rational solutions of the initial system:

```
> get_sol(rez, varl, rat);
```

$$
\begin{aligned}
\Big[&_C_0 + _C_1\,x, _C_1, -_C_2 - _C_2\,x - _C_2\,x^2, _C_2 + _C_2\,x + _C_2\,x^2, -_C_0 - 2\,_C_1\,x,\\
&(-_C_1 + 2\,_C_2)\,x^3 + (-_C_0 + 2\,_C_2)\,x^2 + 2\,_C_2\,x - 2\,_C_1\Big]
\end{aligned}
$$

6 The uncoupling program family

Another universal algorithm to uncouple linear systems is proposed in [7]. Additionally the paper gives a review and an analysis of known methods for uncoupling systems of linear ordinary differential equations. It was observed that only very few of them can be generalized to arbitrary Ore polynomial rings. Once those methods are generalized, different methods give very different uncoupled equations. The generalization of the successive elimination algorithm which has been done in our publication [4] was considered in this context. The conclusion is that the generalization "tends to yield equations with fewer extra singularities than others, which improves the efficiency of the solving process". This remark gives a hope that the program which we describe can be useful.

Acknowledgement

We would like to thank Francis J. Wright who provided us with useful comments on earlier draft.

References

[1] Ore O. (1933): *Theory of non-commutative polynomials,* Annals of Mathematics **34**, 480–508.

[2] M. Bronstein & M. Petkovšek (1994): *On Ore rings, linear operators and factorisation,* Programmirovanie **1**, 27–45.

[3] M. Bronstein, M. Petkovšek (1996): An introduction to pseudo-linear algebra, *Theoretical Computer Science* **157**, 3–33.

[4] Abramov S.A., Zima E.V. (1995): *Ore polynomial rings and linear systems reduction (in Russian),* Vestnik MGU, Ser.15, Computat.,Math. and Cybernetics **3**, 50–56.

[5] Murray F.J., Miller K.S. (1954): *Existence theorems for ordinary differential equations.* New York Univ. Press, Intersciences, New York.

[6] Vilenkin N.Ja. (1969) *Combinatorics. (in Russian)* Nauka, Moscow.

[7] Bronstein M., Zürcher B. (1996): *Uncoupling Algorithms for Pseudo-Linear Systems,* (submitted to AAECC).

[8] Barkatou M.A. *An algorithm for computing a companion block diagonal form for a system of linear differential equations.* Applicable Algebra in Engineering, Communication and Computing, 1993, **4**, pp. 185-195.

NUMERICAL METHOD FOR SOLVING THE INVERSE PROBLEM OF QUANTUM SCATTERING THEORY

R.G.Airapetyan*), J.Busha**), E.P.Zhidkov*)

*) *Joint Institute for Nuclear Research, Dubna, Russia*
**) *Technical University, Kosice, Slovak Republic*
e-mail:airapet@cv.jinr.ru

Abstract. The numerical method for the inverse problem of quantum scattering theory for a set of coupled Schrödinger radial equations with zero orbital momentum is suggested. It is based on the analysis of an asymptotic behaviour of S-matrix and Jost function for high energies and on the corresponding asymptotic regularizations.

1 INTRODUCTION

The problem of a numerical reconstruction of a potential by scattering data is well known and important from the mathematical point of view and for such physical applications as the analysis of a nuclear interaction potential by experimental data. The main approaches at the theoretical investigations of the problem are well-known Gelfand-Levitan, Marchenko and Krein methods reducing the problem to the families of integral equations. At the same time the development of the corresponding numerical methods is sufficiently complicated by the reason of ill-posedness of the mentioned inverse problems (as it has shown by A.M. Denisov, their solutions are nonstable with respect to perturbations of input data). On the other hand the algorithms based on the Tikhonov method of regularization of incorrect problems are considerably more effective. There is a series of the papers devoted to the applications of the Tikhonov method to the numerical solving of the inverse problems.

In this report rather different regularization is suggested. It is based on the analysis of the asymptotic of the spectral properties of the problem for high energies. By means of the obtained asymptotic characteristics of the spectral data, the regularizing algorithms are constructed. The essential steps of the algorithms are the regularization of the Fourier transformation and the special inversion of "almost" Toeplitz, block Toeplitz and Hankel matrices. The last one permits to determine the asymptotical behaviour of the inverse matrix with respect to the discretization step and, by this way, to avoid the unstable procedure of a numerical differentiation. The testing of the programs based on this algorithms shows their high accuracy and stability with respect to small perturbations of input data.

[1]This work was supported in part by RFFI under Grant No. 95-01-01467a

2 STATEMENT OF THE PROBLEM

The following Cauchy problem for a set of coupled Schrödinger radial equation with zero orbital momentum and $k \in [0, +\infty)$ is considered:

$$\left(\frac{d^2}{dr^2} + k^2\right) \Psi(r,k) = V(r)\Psi(r,k), \tag{1}$$

$$\Psi(0,k) = 0, \quad \Psi'(0,k) = kI, \tag{2}$$

where $\Psi(r,k)$ is a n-dimensional vector function defined on $R^+ \times R^+$, $V(r)$ is a real symmetric $n \times n$ matrix function, and I is an identity matrix.

It is well-known that, for the potentials satisfying the condition

$$\int_0^\infty r|V_{\alpha\beta}(r)|dr < \infty, \quad \alpha,\beta = 1,\ldots,n, \tag{3}$$

the wave function has the following asymptotic behaviour:

$$\Psi(r,k) = \frac{1}{2i}(G(k)e^{ikr} - G(-k)e^{-ikr}) + o(1), \quad r \mapsto \infty, \tag{4}$$

or

$$\Psi(r,k) = \frac{1}{2i}(S(k)e^{ikr} - e^{-ikr})G(-k) + o(1), \quad r \mapsto \infty, \tag{5}$$

where $G(k)$ is Jost function and $S(k)$ is S-matrix:

$$S(k) = G(k)G^{-1}(-k), \quad S(-k) = S^{-1}(k). \tag{6}$$

The inverse problem of the Quantum Scattering Theory is the problem of the reconstruction of unknown potential $V(r)$ by the known S-matrix $S(k)$. The problem is uniquely solvable if there are no bound states. In this case

$$\arg\det S(+\infty) = \arg\det S(-\infty). \tag{7}$$

So let us assume that this condition holds.

3 THE KREIN APPROACH

In the framework of the Krein's approach at the first step we determine the matrix function $F(t)$ from the following relations:

$$F(t) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-ikt}[S(k) - I_n]\,dk = \frac{1}{\pi}\int_0^{+\infty} \mathrm{Re}\,\{[S(k) - I_n]e^{-ikt}\}\,dk =$$

$$= \frac{1}{\pi}\int_0^{+\infty} [(\mathrm{Re}\,S(k) - I_n)\cos kt + \mathrm{Im}\,S(k)\sin kt]\,dk. \tag{8}$$

The last integral is conditionally convergent. So, for the numerical calculation of $F(t)$ one needs some regularization procedure.

At the second step we solve the integral equation:

$$K(t) + \int_0^{+\infty} K(s)F(s-t)\,ds = -F(-t), \qquad 0 \le t < +\infty. \tag{9}$$

Then we determine the inverse matrix-function to the Jost function $G(k)$ and the matrix-functions $H(t)$ and $H'(t)$

$$G^{-1}(k) = I_n + \int_0^{+\infty} K(t)e^{-ikt}\,dt, \tag{10}$$

$$H(t) = \sqrt{\frac{2}{\pi}} \int_0^{+\infty} \frac{[G^T(-k)G(k)]^{-1} - I_n}{\sqrt{2\pi}} \cos kt\,dk, \tag{11}$$

$$H'(t) = -\sqrt{\frac{2}{\pi}} \int_0^{+\infty} k\frac{[G^T(-k)G(k)]^{-1} - I_n}{\sqrt{2\pi}} \sin kt\,dk, \tag{12}$$

Further we solve the one-parametric family of integral equations:

$$X_{2r}(t) + \int_0^{2r} H(t-s)X_{2r}(s)\,ds = H(t), \qquad 0 \le t \le 2r, \tag{13}$$

and finally reconstruct the potential by means of the formula:

$$V(r) = 2\frac{d}{dr}[X_{2r}(2r) - X_{2r}(0)]. \tag{14}$$

4 ASYMPTOTIC BEHAVIOUR OF S-MATRIX AND OF JOST FUNCTION

Let $\mathcal{M}_n$ be the set of real symmetric matrices of order n. Denote by $L_m^1(\bar{R}_+, \mathcal{M}_n)$, $m > 1$, the set of $\mathcal{M}_n$-value functions $V(r)$ belonging to $L^1([0,+\infty), \mathcal{M}_n)$ with all derivatives $V^{(j)}(r)$, $j \le m$.

Lemma 1. If $V(r) \in L_2^1(\bar{R}_+, \mathcal{M}_n) \cap C^1(\bar{R}_+, \mathcal{M}_n)$, then

a) the Jost function has the following asymptotic

$$G(k) = I_n + V_2k^{-2} + iV_0k^{-1} + \mathcal{O}(k^{-3}), \tag{15}$$

b) S-matrix has the asymptotic

$$S(k) = I_n - 2V_0^2k^{-2} + 2iV_0k^{-1} + \mathcal{O}(k^{-3}), \tag{16}$$

c) the "module" of the Jost function has the asymptotic

$$[G^T(-k)G(k)]^{-1} - I_n = -\frac{V(0)}{2}k^{-2} + \mathcal{O}(k^{-3}), \tag{17}$$

where

$$V_0 = -\frac{1}{2}\int_0^{+\infty} V(r)\,dr, \tag{18}$$

$$V_2 = \frac{1}{4}\left[V(0) - \int_0^{+\infty} V(r)W(r)\,dr\right], \quad W(r) = \int_0^r V(\xi)\,d\xi. \tag{19}$$

5 Algorithm

I. From (16) the V_0 is numerically calculated, then S-matrix is representated in the form

$$\mathrm{Re}\,S(k) = I_n - 2V_0^2\frac{1}{k^2+1} + \Re(k), \tag{20}$$

$$\mathrm{Im}\,S(k) = 2V_0\frac{k}{k^2+1} + \Im(k). \tag{21}$$

The cos- and sin- transformations of the functions $2V_0^2/(k^2+1)$ and $2kV_0/(k^2+1)$ are determined exactly, and of the functions $\Re$ and $\Im$ are calculated numerically (the asymptotic behaviour of the functions $\Re$, $\Im$ provides the stable Fourier fast transformations).

II. The Wiener-Hopf matrix problem is solved. For this goal the trapesoid formula and inversion of block-Toeplitz matrix are used ([14]). Note that the matrix functions $F(t)$ and $K(t)$ exponentially decrease for $t \to \infty$.

III. Matrices $G^{-1}(k)$ and $\left[G^T(-k)\right]^{-1}$ are calculated by means of formula (10). For the stable calculation of the matrix functions $H(t)$ and $H'(t)$, by means of formulae (11) and (12) the asymptotic regularization described above (at the step I) is used.

IV. At the last step the family of the matrix second kind Fredholm integral equations (13) are solved. For this goal a special inversion of the block-Toeplitz matrices is done. It permits to extract the asymptotic behaviour of the inverse matrix with respect to the discretization step h. After such inversion the potential $V(r)$ can be stable calculated.

References

[1] Z.S.Agranovich, and V.A.Marchenko, Inverse Scattering Problem. – 1960.

[2] V.V.Babikov. Method of the Phase Functions in Quantum Mechanics. – Moscow, 1976.

[3] K.Chadan, and P.C.Sabatier. Inverse Problems in Quantum Scattering Theory. – 1989.

[4] A.M.Denisov. Introduction to Theory of Inverse Problems. – Moscow, 1994.

[5] H.V.Von Geramb. (Ed.) Quantum Inversion Theory and Applications. – Springer-Verlag, 1994.

[6] M.G.Krein, DAN SSSR **105**, 433 (1955).

[7] M.G.Krein, DAN SSSR **111**, 1167 (1956).

[8] B.M.Levitan. Sturm-Liouville Inverse Problems. – Moscow, 1984.

[9] R.G.Newton, J.Math.Ph. **2**, 188 (1961).
[10] A.G.Ramm. Multidimensional Inverse Scattering Problems. – 1992.
[11] A.A. Samarskii. Introduction to Numerical Methods. – Moscow, 1982.
[12] A.N. Tikhonov, and V.Ya.Arsenin. Methods of the Solving of Incorrect Problems. – Moscow, 1979.
[13] J.Vizner, E.P.Zhidkov, V.Lelek, R.V.Malyshev, B.N.Khoromskii, E.Kh.Khristov, and I.Ulegla, Particles and Nucleus, JINR **9**, 710 (1978).
[14] V.V.Voevodin, and E.E.Tyrtyshnikov. Calculating Processes with Toeplitz Matrices. Moscow, 1987.
[15] B.N.Zakhar'ev, V.N.Mel'nikov, B.V.Rudyak, and A.A.Suz'ko, Particles and Nucleus, JINR **8**, 290 (1977).
[16] B.N.Zakhar'ev, and A.A.Suz'ko. Potentials and Quantum Scattering: Direct and Inverse Problems. – Moscow, 1985.
[17] E.P.Zhidkov, Yu.M.Kazarinov, G.I.Makarenko, and A.V.Rakitskii. – Preprint, 1970, Dubna, JINR, R1-5306.

SOME APPLICATION OF THE CONTINUATION METHOD FOR THE NUMERICAL SOLVING OF THE THREE-DIMENSIONAL POLARON PROBLEM [1]

P.G. Akishin, I.V. Puzynin, Yu.S. Smirnov

LCTA, JINR, 141980 Dubna, Russia
e-mail: akishin@lcta28.jinr.dubna.su
puzynin@lcta15.jinr.dubna.su
smirnov@lcta6.jinr.dubna.su

Abstract. The numerical approach for study of the 3D-nonlinear problem for the spherically-nonsymmetric polaron is considered. An expansions in spherical harmonics are used for the approximation of the solution. The iterative Newton's scheme with an additional parametrization of the initial equation for the solving of the nonlinear problem is proposed. The results of numerical modelling are discussed. The comparison of the obtained eigenvalues with the known ones confirms the efficiency of the elaborated algorithms.

1 Introduction

We consider the problem of numerical solving of the spherically-nonsymmetrical model of polaron [1] in the limit of strong coupling. Such a model of polaron describes the behavior of a non-relativistic particle (electron) in the field, created by the interaction with medium.

Let us formulate a mathematical statement of the problem. The wave function ψ and potential u satisfy the following system of partial differential equations in the space R^3 :

$$\Delta\psi(\mathbf{x}) + (u(\mathbf{x}) - v(\mathbf{x}))\psi(\mathbf{x}) - \lambda\psi(\mathbf{x}) = 0, \tag{1}$$

$$\Delta u(\mathbf{x}) + 4\pi\psi^2(\mathbf{x}) = 0, \tag{2}$$

$$\Delta v(\mathbf{x}) - c^2 v(\mathbf{x}) + 4\pi\psi^2(\mathbf{x}) = 0, \tag{3}$$

where Δ is the Laplace operator and c is the constant of coupling.

The wave function $\psi(\mathbf{x})$ satisfy the normalization condition:

$$\int_{R^3} \psi^2(\mathbf{x}) dV_{\mathbf{x}} = 1. \tag{4}$$

In spherically symmetrical case (when $\psi(\mathbf{x})$ depends on $r = |\mathbf{x}|$ only), the system (1)-(3) may be reduced to a boundary problem for the system of ordinary differential equations on semi-axis [2],[3]. This problem has been studied quite well. In paper [4] the authors consider the axi-symmetrical solutions of the polaron model. The papers [5],[7] are devoted to numerical investigations of the problem for the nonsymmetrical case.

In the present paper to approximate equations (1)-(3) we use the approach proposed in [5],[6]. To solve the nonlinear discrete problem, a parametrization of initial equations is introduced by means of the additional continuous parameter t [8]. At the initial moment

[1]This work was supported in part by the Russian Foundation of Basic Researches under Grant No.94-01-01119.

$t = 0$ the problem is reduced to the sufficiently simple spectral Helmholtz problem for a ball. All eigenvalues and eigenfunctions of this problem may be easily found. After that, the solution of the original nonlinear problem (1)-(2) may be obtained if we use a movement with respect to continuous parameter t.

2 Discretization

In this section we consider the problem of discretization of the system of equations (1)-(3). For this aim we use the Bubnov - Galerkin method [9]. Let functions $\psi(\mathbf{x})$, $u(\mathbf{x})$ and $v(\mathbf{x})$ be expanded in spherical harmonic series [5]:

$$\psi(\mathbf{x}) = \sum_{k=0}^{\infty} \sum_{l=-k}^{k} \psi_{kl}(r) Y_{kl}(\theta, \varphi), \tag{5}$$

$$u(\mathbf{x}) = \sum_{k=0}^{\infty} \sum_{l=-k}^{k} u_{kl}(r) Y_{kl}(\theta, \varphi), \qquad v(\mathbf{x}) = \sum_{k=0}^{\infty} \sum_{l=-k}^{k} v_{kl}(r) Y_{kl}(\theta, \varphi), \tag{6}$$

where Y_{kl} are spherical harmonics [10].

Taking into account a finite number K of terms in the expansions (5),(6) for approximations of functions $\psi(\mathbf{x}), u(\mathbf{x}), v(\mathbf{x})$, multiplying (1)-(3) by the spherical harmonics and integrating over the angle variables, we get the following approximate system of equations for the functions $\tilde{\psi}_{kl} = \psi_{kl} r$, $\tilde{u}_{kl} = u_{kl} r$, $\tilde{v}_{kl} = v_{kl} r$:

$$\tilde{\psi}''_{kl} - \frac{k(k+1)}{r^2} \tilde{\psi}_{kl} + \frac{1}{r} \sum_{k_1, l_1} \sum_{k_2, l_2} W^{l l_1 l_2}_{k k_1 k_2} \tilde{\psi}_{k_1 l_1} \tilde{u}_{k_2 l_2} = \lambda \tilde{\psi}_{kl} \tag{7}$$

$$\tilde{u}''_{kl} - \frac{k(k+1)}{r^2} \tilde{u}_{kl} + 4\pi \frac{1}{r} \sum_{k_1, l_1} \sum_{k_2, l_2} W^{l l_1 l_2}_{k k_1 k_2} \tilde{\psi}_{k_1 l_1} \tilde{\psi}_{k_2 l_2} = 0, \tag{8}$$

$$\tilde{v}''_{kl} - (c^2 + \frac{k(k+1)}{r^2}) \tilde{v}_{kl} + 4\pi \frac{1}{r} \sum_{k_1, l_1} \sum_{k_2, l_2} W^{l l_1 l_2}_{k k_1 k_2} \tilde{\psi}_{k_1 l_1} \tilde{\psi}_{k_2 l_2} = 0, \tag{9}$$

$$k = 0, 1, 2, ..., K; l = 0, \pm 1, \pm 2, ..., \pm k.$$

Here coefficients $W^{l_1 l_2 l_3}_{k_1 k_2 k_3}$ equal to the integral over the unit sphere for the product of spherical harmonics:

$$W^{l_1 l_2 l_3}_{k_1 k_2 k_3} = \int_0^{2\pi} \left[\int_0^{\pi} Y_{k_1 l_1}(\theta, \varphi) Y_{k_2 l_2}(\theta, \varphi) Y_{k_3 l_3}(\theta, \varphi) sin\theta d\theta \right] d\varphi. \tag{10}$$

The normalization condition (4) may be written in the following form:

$$\sum_{k,l} \int_0^{\infty} \tilde{\psi}^2_{k,l} dr = 1. \tag{11}$$

Let $[0, R]$ be an interval of the variation of r. We assume that R is large enough, so functions $\tilde{\psi}_{kl}(r), \tilde{u}_{kl}(r), \tilde{v}_{kl}(r)$ have a behavior, which may be approximated analytically for $r \geq R$. We choose a uniform grid of nodes for variable r : $\{r_i; i = \overline{1, N}\} \left(r_i = \right.$

$(i-1)h, h = \frac{R}{N-1}\Big)$ to discretize the system (7)-(9). The second order accuracy finite difference scheme with respect to step h is used for the approximation of equations (7)-(9). Let the functions $\psi^i_{kl}, u^i_{kl}, v^i_{kl}$ be $\psi^i_{kl} = \tilde{\psi}_{kl}(r_i), u^i_{kl} = \tilde{u}_{kl}(r_i), v^i_{kl} = \tilde{v}_{kl}(r_i)$. Then the discrete system of equations has the following form:

$$\frac{\psi^{i+1}_{kl} - 2\psi^i_{kl} + \psi^{i-1}_{kl}}{h^2} - \frac{k(k+1)}{r_i^2}\psi^i_{kl} + \frac{1}{r_i}\sum_{k_1,l_1}\sum_{k_2,l_2} W^{ll_1l_2}_{kk_1k_2}\psi^i_{k_1l_1}u^i_{k_2l_2} = \lambda\psi^i_{kl} \qquad (12)$$

$$\frac{u^{i+1}_{kl} - 2u^i_{kl} + u^{i-1}_{kl}}{h^2} - \frac{k(k+1)}{r_i^2}u^i_{kl} + 4\pi\frac{1}{r_i}\sum_{k_1,l_1}\sum_{k_2,l_2} W^{ll_1l_2}_{kk_1k_2}\psi^i_{k_1l_1}\psi^i_{k_2l_2} = 0 \qquad (13)$$

$$\frac{v^{i+1}_{kl} - 2v^i_{kl} + v^{i-1}_{kl}}{h^2} - (c^2 + \frac{k(k+1)}{r_i^2})v^i_{kl} + 4\pi\frac{1}{r_i}\sum_{k_1,l_1}\sum_{k_2,l_2} W^{ll_1l_2}_{kk_1k_2}\psi^i_{k_1l_1}\psi^i_{k_2l_2} = 0 \qquad (14)$$

$k = 0, 1, 2, ...K; l = 0, \pm1, \pm2, ..., \pm k; i = \overline{2, N-1}$. For $r_1 = 0$ we have the following left-hand boundary conditions: $\psi^1_{kl} = 0, u^1_{kl} = 0, v_{kl} = 0$. Taking into account exponential vanishing of the function ψ_{kl} on infinity, we get the right-hand boundary condition for ψ^i_{kl}:

$$\frac{\psi^N_{kl} - \psi^{N-2}_{kl}}{2h} = (-\sqrt{\lambda} + \frac{k}{r_{N-1}})\psi^{N-1}_{kl}. \qquad (15)$$

Analogously, for u^i_{kl} and v^i_{kl} we have

$$\frac{u^N_{kl} - u^{N-2}_{kl}}{2h} = \frac{k}{r_{N-1}}u^{N-1}_{kl}, \qquad \frac{v^N_{kl} - v^{N-2}_{kl}}{2h} = (-c + \frac{k}{r_{N-1}})v^{N-1}_{kl}. \qquad (16)$$

By using some numerical approximation for (11) the normalization conditions may be written in the form:

$$\sum_{k,l}\sum_{i=1}^{N}\alpha_i(\psi^i_{kl})^2 = 1, \qquad (17)$$

where α_i are the coefficients of the quadrature formula for calculation of integrals.

To find the approximate solutions of the original problem (1)-(2), it is necessary to solve the nonlinear system of algebraic equations (12)-(17) for the unknown variables $\{\psi^i_{kl}, u^i_{kl}, v^i_{kl}, i = \overline{1, N}\}$ and the spectral parameter λ.

3 The method for solving the nonlinear discrete systems

Let $\vec{\Psi}_i$ be the vector with the components ψ^i_{kl}, $\vec{U}_i$ and $\vec{V}_i$ be the vectors with the components u^i_{kl} and v^i_{kl} correspondingly. Analogously we define the vectors corresponding to the nonlinear terms (double sums) in the equations (12)-(14), as $\vec{F}_i = \vec{F}_i(\vec{\Psi}_i, \vec{U}_i), \vec{G}_i = \vec{G}_i(\vec{\Psi}_i, \vec{\Psi}_i)$. Then the system (12)-(14) can be written in the form:

$$\frac{\vec{\Psi}_{i+1} - 2\vec{\Psi}_i + \vec{\Psi}_{i-1}}{h^2} + [D_i]\vec{\Psi}_i + \vec{F}_i(\vec{\Psi}_i, \vec{U}_i) = \lambda\vec{\Psi}_i, \qquad (18)$$

$$\frac{\vec{U}_{i+1} - 2\vec{U}_i + \vec{U}_{i-1}}{h^2} + [D_i]\vec{U}_i + \vec{G}_i(\vec{\Psi}_i, \vec{\Psi}_i) = 0, \qquad (19)$$

$$\frac{\vec{V}_{i+1} - 2\vec{V}_i + \vec{V}_{i-1}}{h^2} - c^2\vec{V}_i + [D_i]\vec{V}_i + \vec{G}_i(\vec{\Psi}_i, \vec{\Psi}_i) = 0, \tag{20}$$

where $i = \overline{2, N-1}$; the matrices $[D_i]$ are diagonal matrices with coeffi cients on the main diagonal depending on the number k of the correspondent harmonic. The normalization condition (17) may be written in new variables as

$$\sum_{i=1}^{N} \alpha_i |\vec{\Psi}_i|^2 - 1 = 0, \tag{21}$$

The left-hand boundary conditions reduce to the following equations:

$$\vec{\Psi}_1 = 0, \qquad \vec{U}_1 = 0, \qquad \vec{V}_1 = 0. \tag{22}$$

Analogously, we rewrite the right-hand boundary conditions:

$$\frac{\vec{U}_N - \vec{U}_{N-2}}{2h} = [B_1]\vec{U}_{N-1}, \qquad \frac{\vec{V}_N - \vec{V}_{N-2}}{2h} = (-c[E] + [B_2])\vec{V}_{N-1}, \tag{23}$$

$$\frac{\vec{\Psi}_N - \vec{\Psi}_{N-2}}{2h} = (-\sqrt{\lambda}[E] + [B_2])\vec{\Psi}_{N-1}, \tag{24}$$

where $[E]$ is an identity matrix, $[B_1]$, $[B_2]$ are diagonal matrices.

Let us consider an auxiliary problem. We can write the right boundary condition for the wave functions in the following form:

$$\vec{\Psi}_N = 0. \tag{25}$$

Then the system of equations (18)-(23),(25) also approximates the problem (7)-(11), but with the lesser accuracy. To solve the nonlinear problem (18)-(23),(25), we use the continuation method from [8]. In equations (18)-(20) we introduce a parameter t, ($t \in [0, 1]$) by the following way:

$$\vec{F}_i(t) = t\vec{F}_i(\vec{\Psi}_i, \vec{U}_i), \qquad \vec{G}_i(t) = t\vec{G}_i(\vec{\Psi}_i, \vec{\Psi}_i).$$

After this substitution the system (18)-(20) may be transformed to the fo lowing system of equations:

$$\frac{\vec{\Psi}_{i+1} - 2\vec{\Psi}_i + \vec{\Psi}_{i-1}}{h^2} + D_i\vec{\Psi}_i + t\vec{F}_i(\vec{\Psi}_i, \vec{U}_i) = \lambda\vec{\Psi}_i \tag{26}$$

$$\frac{\vec{U}_{i+1} - 2\vec{U}_i + \vec{U}_{i-1}}{h^2} + D_i\vec{U}_i + t\vec{G}_i(\vec{\Psi}_i, \vec{\Psi}_i) = 0, \tag{27}$$

$$\frac{\vec{V}_{i+1} - 2\vec{V}_i + \vec{V}_{i-1}}{h^2} - c^2\vec{V}_i + D_i\vec{V}_i + t\vec{G}_i(\vec{\Psi}_i, \vec{\Psi}_i) = 0, \tag{28}$$

with boundary conditions (22)-(23),(25). In this case the solutions $\vec{\Psi}_i, \vec{U}_i, \lambda$ of the problem (21)-(23),(25)-(28) will be the functions of the parameter t ($t \in [0,]$). Obviously,

for $t = 1$ we have the original system of equations (18)-(23),(25), and for $t = 0$ we obtain the sufficiently simple non-connected linear spectral problem for each harmonic separately.

To find the values of functions $\vec{U}(t), \vec{V}(t), \vec{\Psi}(t), \lambda(t)$ at the moment $t = 0$ for the fixed orbital momentum k_0, we solve the one-dimensional spectral problem for one equation only:

$$\frac{y_{i+1} - 2y_i + y_{i-1}}{h^2} - \frac{k_0(k_0+1)}{r_i^2} y_i = \lambda y_i, \; i = \overline{2, N-1} \tag{29}$$

with boundary and normalization conditions

$$y_1 = 0, \qquad y_N = 0; \qquad \sum_{i=1}^{N} \alpha_i y_i^2 = 1. \tag{30}$$

Let $\{y_i^*\}, \lambda^*$ be a solution of the problem (29)-(30). For the fixed number $l_0 : -k_0 \leq l_0 \leq k_0$ we assume

$$\psi_{kl}^i(t)|_{t=0} = \begin{cases} y_i^*/\sqrt{2}, & l = l_0, k = k_0 \\ y_i^*/\sqrt{2}, & l = -l_0, k = k_0 \\ 0, & l \neq l_0 \\ 0, & k \neq k_0. \end{cases} \tag{31}$$

This case corresponds to the real initial approximation for determing the solutions which are even functions with respect to angle φ. The case corresponding to the odd initial approximation can be written by the following way:

$$\psi_{kl}^i(t)|_{t=0} = \begin{cases} y_i^*/\sqrt{2}, & l = l_0, k = k_0 \\ -y_i^*/\sqrt{2}, & l = -l_0, k = k_0 \\ 0, & l \neq l_0 \\ 0, & k \neq k_0. \end{cases} \tag{32}$$

Accordingly for the spectral parameter we have $\lambda(t) = \lambda^*$ for $t = 0$.

So, for every $t \in [0, 1]$ we have obtained the nonlinear boundary problem. Let $\{t_j; j = \overline{0, M}\}(t_0 = 0, t_M = 1)$ be some partition of the interval [0.,1.]. To find the solutions $\psi_{kl}(t_j)$, $u_{kl}(t_j)$, $v_{kl}(t_j)$, $\lambda(t_j)$ of the problem (29)-(30) in the point t_j the Newton method is used. Supposing the difference $|t_j - t_{j+1}|$ is small enough, we have the good initial approximation from previous time step t_{j-1} for Newtonian iterative procedure.

Let $\vec{\Psi}_i^m$ be the vector of components ψ_{kl}^i on $m-$th step of Newtonian iterative process. Analogously let $\vec{U}_i^m$ and $\vec{V}_i^m$ be the vectors with components u_{kl}^i and v_{kl}^i. Then we have the following boundary problem for the vectors of corrections $\delta\vec{\Psi}_i^m = \vec{\Psi}_i^{m+1} - \vec{\Psi}_i^m, \delta\vec{U}_i^m = \vec{U}_i^{m+1} - \vec{U}_i^m, \delta\vec{V}_i^m = \vec{V}_i^{m+1} - \vec{V}_i^m$:

$$\frac{\delta\vec{\Psi}_{i+1}^m - 2\delta\vec{\Psi}_i^m + \delta\vec{\Psi}_{i-1}^m}{h^2} + [D_i]\delta\vec{\Psi}_i^m + [A_{11}](\vec{u}_i^m)\delta\vec{\Psi}_i^m + [A_{12}](\vec{\Psi}_i^m)\delta\vec{U}_i^m - \lambda_m\delta\vec{\Psi}_i^m +$$

$$+\delta\lambda_m\Psi_i^m = -\left(\frac{\vec{\Psi}_{i-1}^m - 2\vec{\Psi}_i^m + \vec{\Psi}_{i+1}^m}{h^2} + [D_i]\vec{\Psi}_i^m + \vec{f}_i(\vec{\Psi}_i^m, \vec{U}_i^m) - \lambda_m\vec{\Psi}_i^m\right),$$

$$\frac{\delta\vec{U}_{i+1}^m - 2\delta\vec{U}_i^m + \delta\vec{U}_{i-1}^m}{h^2} + [D_i]\delta\vec{U}_i^m + [A_{21}](\vec{U}_i^m)\delta\vec{U}_i^m + [A_{22}](\vec{U}_i^m)\delta\vec{U}_i^m =$$

$$= -\left(\frac{\vec{U}_{i-1}^m - 2\vec{U}_i^m + \vec{U}_{i+1}^m}{h^2} + [D_i]\vec{U}_i^m + \vec{G}_i(\vec{\Psi}_i^m, \vec{\Psi}_i^m)\right),$$

$$\frac{\delta\vec{V}_{i+1}^m - 2\delta\vec{V}_i^m + \delta\vec{V}_{i-1}^m}{h^2} + ([D_i] - c^2[E])\delta\vec{V}_i^m + [A_{31}](\vec{V}_i^m)\delta\vec{V}_i^m + [A_{32}](\vec{V}_i^m)\delta\vec{V}_i^m =$$

$$= -\left(\frac{\vec{V}_{i-1}^m - 2\vec{V}_i^m + \vec{V}_{i+1}^m}{h^2} + ([D_i] - c^2[E])\vec{V}_i^m + \vec{G}_i(\vec{\Psi}_i^m, \vec{\Psi}_i^m)\right),$$

where $[A_{11}], [A_{12}], [A_{21}], [A_{22}], [A_{31}], [A_{32}]$ are the matrices corresponding to differential operators for functions $\vec{F}_i$ and $\vec{G}_i$ from (18)-(20). After that, the boundary conditions can be written in the following form:

$$\delta\vec{\Psi}_0^m = 0, \delta\vec{\Psi}_N^m = 0; \qquad \delta\vec{U}_0^m = 0, \delta\vec{U}_N^m = 0; \qquad \delta\vec{V}_0^m = 0, \delta\vec{V}_N^m = 0.$$

To find the correction $\delta\lambda_m$ for spectral parameter, we have the equation:

$$(\delta\vec{\Psi}^m, \vec{\Psi}^m) = 0$$

To solve this boundary problem, we apply the matrix sweep method [13]. The found solutions $\{\vec{\Psi}_i^*(t_M), \vec{U}_i^*(t_M)\, \vec{V}_i^*(t_M)\}, \lambda^*(t_M)$ are used as the initial approximations for Newtonian iterative process to solve the problem (18)-(23) with the non-zero boundary conditions (24). Initial approximations $\{y_i^*\}, \lambda^*$ for $t = 0$ have been found by using the procedure from [12].

4 The results of numerical simulation

On the base of the proposed techniques the FORTRAN code to calculate the three-dimensional egenvalues and eigenfunctions for the polaron model has been created. These programs allow one to find as spherically-symmetrical and axi-symmetrica solutions as essentially three-dimensional solutions.

Now we define some classification of solutions. Let the class of solutions $\Omega_{k_0 l_0}$ correspond to the solutions obtained from initial guesses with $k = k_0, l = l_0$ in (31),(32). It should be noted that from iteration process we have functions ψ_{kl}, u_{kl} equal to 0 for $k < k_0$. So the class Ω_{00} includes the spherically-symmetrical solutions of the polaron problem. The case with $l_0 = 0$ corresponds to the axial symmetry of the problem. In Table 1 the summary of the obtained results for polaron model in the limit of strong coupling ($c = \infty$) and the comparison with results from [5] are shown. From this Table we see that the fast convergence of solutions in respect to the number of harmonics takes place near the ground state, but there is a dynamics of the numerical results for the low part of spectrum. In order to obtain the reliable results in this case it is necessary to increase the dimension of the solving discrete problem.

Table 1. The calculated eigenvalues λ in dependence on the number of harmonics $K = k_0$.

Eigenvalue	Class	$K=1$	$K=3$	$K=5$	results from [5]
λ_0	Ω_{00}	0.08139	-	-	0.0814
λ_1	Ω_{00}	0.01540	-	-	0.0154
λ_2	Ω_{00}	0.00626	-	-	0.0062
λ_3	Ω_{00}	0.00337	-	-	-
λ_0	Ω_{10}	0.02705	0.03392	0.03443	0.0343
λ_1	Ω_{10}	0.00880	0.00833	0.01327	0.0126
λ_2	Ω_{10}	0.00432	0.00407	0.00415	-
λ_3	Ω_{10}	0.00255	0.00263	0.00276	-
λ_0	Ω_{21}	-	0.01490	0.01642	0.0159
λ_1	Ω_{21}	-	0.00618	0.00573	-
λ_2	Ω_{21}	-	0.00342	-	-

Table 2. Eigenvalues λ in dependence on the value of the constant of coupling c.

Eigenvalue	Class	$C=0.1$	$C=1.0$	$C=10.0$	$C=\infty$
λ_0	Ω_{00}	0.0347194	0.0769554	0.0813419	0.0813949
λ_1	Ω_{00}	0.0127952	0.0153239	0.0153978	0.0153986
λ_2	Ω_{00}	0.0058969	0.0062565	0.0062632	0.0062638
λ_0	Ω_{10}	0.0215822	0.0333973	0.0339144	0.0339199
λ_1	Ω_{10}	0.0078265	0.0083211	0.0083275	0.0083276
λ_2	Ω_{10}	0.0039902	0.0040679	0.0040688	0.0040688
λ_3	Ω_{10}	0.0026181	0.0026338	0.0026339	0.0026339
λ_0	Ω_{21}	0.0126877	0.0148575	0.0148967	0.0148970
λ_1	Ω_{21}	0.0059044	0.0061771	0.00618088	0.00618088
λ_2	Ω_{21}	0.0033554	0.0034145	0.0034152	0.0034152

In Table 2 the numerical results in dependence on the constant of coupling c for different classes Ω_{kl} are given. From this Table one can see the fast convergence of the solutions for large value c to the solutions in limit of strong coupling.

This investigation has been supported by the Russian Foundation of Basic Researches, grant N 94 – 01 – 01119.

References

[1] Pekar S.I Issledovania po elektronnoi teorii kristallov, Moscow, Gostekhizdat, 1951 (in Russian).

[2] Amirkhanov I.V. et.al. Report of Science Investigation Biological Center of USSR Science Academy , Pushchino, 1988.

[3] Amirkhanov I.V. et.al. Report of Science Investigation Biological Center of USSR Science Academy , Pushchino, 1990.

[4] Amirkhanov I.V.,Zemlyanaya E.V.,Puzynina T.P. JINR Report, **P-11-91-139**, Dubna, 1991.
[5] Gabdullin R.R. Report of Science Investigation Biological Center of USSR Science Academy , Pushchino, 1991.
[6] Amirkhanov I.V.,Puzynin I.V.,Puzynina T.P.,Zemlyanaya E.V. Preprint of JINR, **E-11-92-205**, Dubna, 1992.
[7] Gabdullin R.R. - DAN Rossii, 1993, **333, 1**, p.23 (in Russian).
[8] Ortega J.M., Rheinboldt W.C. Iterative solution of nonlinear equations in several variables, Academic Press, New York and London, 1970.
[9] Fletcher C.A.J. Computational Galerkin Methods, Springer-Verlag, New York, Berlin, Heidelberg, Tokyo, 1984.
[10] Korn G.,Korn T.. Mathematical handbook for scientists and engineer, McGraw-Hill book comp., New York, Toronto, London, 1961.
[11] Samarsky A.A. Teorija raznostnykh skhem, Moscow, Nauka, 1977 (in Russian).
[12] Akishin P.G.,Puzynin I.V. JINR Report, **5-10992**, Dubna, 1977.

ON SOME PROBLEMS OF NUMERICAL INVESTIGATION OF EIGENVALUE PROBLEMS IN MOMENTUM SPACE [1]

I.V.Amirkhanov,E.V.Zemlyanaya,I.V.Puzynin,T.P.Puzynina,T.A.Strizh

Joint Institute for Nuclear Research
Laboratory of Computing Techiques and Automation
141980 Dubna, Moscow Reg., Russia
e-mail: puzynin@lcta48.jinr.dubna.su, puzynina@lcta15.jinr.dubna.su,
zemel@lcta7.jinr.dubna.su, str@cv.jinr.dubna.su

Abstract. Some modifications of Coulomb and increasing potentials are considered. Proposed approaches are suitable for numerical investigation of eigenvalue problems in momentum space. Their efficiency is demonstrated for the Schroedinger equation as well as relativistic generalization of the latter in the framework of a QCD potential model Some properties of considered equations solutions with the specified potentials are discussed.

1 Introduction

A numerical study of some relativistic potential QCD models reduces to a solving of the problems in a momentum space (for instance, relativistic generalizations of Schroedinger equation [1],[2], Schwinger-Dyson and Bethe-Solpeter equations [3]). One of the difficulties is that in the mentioned models a combination of the increasing V_I and Coulomb V_C potentials is usually used as an effective potential of interquark interactions. That leads to a singularity in the kernels of the integral equations. To prevent this, one can modify the initial potential at the coordinate representation by way of its approximation by some functions with "good" characteristics. Thus, the initial potential $V(r)$ is replaced in the equation by a modified potential $V_M(r)$, the properties of which can be investigated on an example of the non-relativistic Schroedinger equation [4]. Solving the equation in a coordinate representation with the potentials $V(r)$ and $V_M(r)$, we can compare the eigenvalues E_n, E_n^M obtained for n-state ($n = 0, 1, 2, ...$) and corresponding eigenfunctions ϕ_n, ϕ_n^M. The modified potential $V_M(r)$ should be selected so that the spectral characteristics agree with the given accuracy ε for the given finite number of states N, i.e.

$$|E_n - E_n^M| < \varepsilon, \qquad ||\phi_n - \phi_n^M|| < \varepsilon, \qquad 0 \le n \le N. \tag{1}$$

Only if the relationships (1) are fulfilled, the Fourier-image of the modified potential can be used as approximation of the initial potential for the solving of relativistic equations in momentum space. It should be noted that the given potential $V(r)$ can be modified in a different ways. Moreover, some modification procedures appear to be more convenient for numerical study. Therefore, the goal of this investigation is to perform a comparative numerical analysis of accuracy of the most used ways of approximation of Coulomb and linear potentials.

[1]This work was supported in part by the Russian Foundation for Basis Research under Grant No 94-01-01119

Programs developed on the basis of the continuous analog of the Newton s method [5] - SLIPH4 [6] and SYSINT (SYSINTM) [7] was used for numerical investigation.

2 General statement of the problem

The eigenvalue problem for Schroedinger equation in a coordinate representation can be written as follows [8]:

$$[d^2/dx^2 + 2M\epsilon_{nl} - l(l+1)/x^2 - 2MV(x)]\phi_{nl}(x) = 0. \tag{2}$$

Here x is a dimensionless variable, ϵ_{nl} is a system of energy levels, $n = 0, 1, 2, ..., l = 0, 1, 2, ...$, M is a dimensionless parameter, $V(x)$ is a given potential. Below we will use a Coulomb potential

$$V(x) = \alpha/x \tag{3}$$

and increasing potentials

$$V_l(x) = \sigma_\gamma x^\gamma, \qquad \gamma > 0. \tag{4}$$

Boundary conditions and normalization conditions are defined as

$$\phi_{nl}(0) = 0, \quad \phi_{nl}(\infty) = 0, \quad \int_0^\infty \phi_{nl}^2(x)dx = 1. \tag{5}$$

In momentum space eq.(2) has the form

$$[Q(p) - \epsilon_{nl}]\phi_{nl}(p) + \int_0^\infty dq V_l(p,q)\phi_{nl}(q) = 0, \qquad Q(p) = p^2/2M \tag{6}$$

with a normalization condition

$$\int_0^\infty \phi_{nl}^2(p)dp = 1. \tag{7}$$

Here

$$\phi_{nl}(p) = \sqrt{\frac{2}{\pi}} \int_0^\infty dx g_l(px)\phi_{nl}(x), \quad \phi_{nl}(x) = \sqrt{\frac{2}{\pi}} \int_0^\infty dp g_l(px)\phi_{nl}(p), \tag{8}$$

$$V_l(p,q) = \frac{2}{\pi} \int_0^\infty dx V(x) g_l(px) g_l(qx) = 0, \tag{9}$$

where g_l is a Rickatti-Bessel function.

Let us note that a great variety of relativistic generalizations of Schroedinger equation differ from the problem (6) by the form of the function $Q(p)$ only [1,2].

Let's take one feature of Schroedinger equation with potential (4). Having transformed eq.(2), one can rewrite it in the form

$$[d^2/dy^2 + \lambda_{nl} - l(l+1)/y^2 - y^\gamma]\phi_{nl}(y) = 0, \qquad \lambda_{nl} = 2M\epsilon_{nl}/b^2, \tag{10}$$

where

$$y = bx, \qquad b = [2M\sigma_\gamma]^{\frac{1}{\gamma+2}}. \tag{11}$$

The important property of spectrum of eq.(10) is that the relation

$$\lambda_{nl}/\lambda_{0l} = \epsilon_{nl}/\epsilon_{0l} \tag{12}$$

does not depend upon the parameters M and σ_γ of the equation.

As Coulomb potential (3) can be presented in form (4) with $\gamma = -1$ and $\sigma_{-1} = \alpha$ everything mentioned above is related to it, too.

This property looks especially important for a numerical study of the mesons' spectra for comparing the calculated data with the experimental results. It is of importance to know the relationships between the energy levels in various states [3]. For this it is enough to solve the equation once at each fixed value of the parameter γ. The γ-dependence of the eigenvalues λ_{nl} of the problem (10) is tabulated below.

TABLE 1.

l	n	$\gamma = 0.5$	$\gamma = 1$	$\gamma = 1.5$	$\gamma = 2$	$\gamma = 2.5$
0	0	1.833	2.338	2.708	3.000	3.242
	1	2.551	4.088	5.856	7.000	8.313
	2	3.051	5.521	8.227	11.000	13.737
	3	3.229	6.787	10.732	15.000	19.383
1	0	2.300	3.361	4.251	5.000	5.639
	1	2.854	4.884	6.966	9.000	10.938
	2	3.286	6.208	9.521	13.000	16.496
	3	3.646	7.406	11.969	17.000	22.244
2	0	2.658	4.248	5.705	7.000	8.141
	1	3.120	5.630	8.311	11.000	13.609
	2	3.502	6.869	10.796	15.000	19.286
	3	3.833	8.013	13.194	19.000	25.124

Note that the condition (12) is not held for the potential combination and for the relativistic generalization of Schroedinger equation. However, it will be useful for selection of initial approximation for the eigenvalue.

3 Modification of Coulomb potential

In momentum space the Coulomb potential (3) can be written as follows:

$$V(p,q) = \frac{\alpha}{2\pi}\int_{-1}^{+1} dt \frac{P_l(t)}{y-t}, \qquad y = \frac{p^2+q^2}{2pq}. \tag{13}$$

Here P_l is Legendre polynomial.

An analytical expression for the eigenvalues of Schroedinger equation at $M = 0.5$ has the form [4]

$$\epsilon^*_{nl} = -\alpha^2/(n+l+1)^2.$$

Let us consider some modifications of potential [3]. Note that in some cases one has been succeeded in obtaining an analytical expression for integral (9). That lessens drastically

a computing time. The computations given below, show a dependence of solutions upon the parameters of approximating potentials. This allows to estimate their applicability for the solving of the problems in momentum space taking the required accuracy of the results into account.

First modification.

$$V_C^M(x) = \alpha[\exp(-\mu_1 x) - 1]/x. \tag{14}$$

Here $\mu_1 >> 1$ is a parameter the value of which should be chosen so that for the given number of states the conditions (1) are fulfilled. The eigenvalues dependence on the parameter μ_1 for problems (2) and (6) is given in Table 2.

TABLE 2. $\alpha = -1, M = 0.5$

n	μ_1	ϵ_{n0} (2)	ϵ_{n1} (2)	ϵ_{n2} (2)	ϵ_{n0} (6)	ϵ_{n1} (6)	ϵ_{n2} (6)
0	5	-0.87529	-0.24962	-0.11111	-0.87603	-0.24997	-0.1111
0	10	-0.95198	-0.24907	-0.11111	-0.95192	-0.24910	-0.1111
1	5	-0.23400	-0.11098	-0.06247	-0.23490	-0.11100	-0.0624
1	10	-0.24813	-0.11111	-0.06249	-0.24809	-0.11111	-0.0624
2	5	-0.10631	-0.06238	-0.03741	-0.10629	-0.06240	-0.0374
2	10	-0.10931	-0.06247	-0.03747	-0.11057	-0.06243	-0.0374

Here and further the parameters of the computational scheme are selected in order to provide a coincidence of the results of the numerical solving of the problems (2) and (6) with accuracy 10^{-3}. Thus, a disagreement between the numerical results and the analytical solution is connected rather with the choice of the approximating potential than with the calculation errors.

Second modification.

$$V_C^M(x) = \alpha[\exp(-\mu_1 x) - \exp(-\mu_2 x)]/x \tag{15}$$

$\mu_1 >> 1$ и $\mu_2 << 1$ are parameters. The results of calculations are tabulated in Table 3.

TABLE 3. $\alpha = -1$, $M = 0.5$

n	μ_1	μ_2	ϵ_{n0} (2)	ϵ_{n1} (2)	ϵ_{n2} (2)	ϵ_{n0} (6)	ϵ_{n1} (6)	ϵ_{n} (6)
0	5	0.010	-0.85548	-0.23012	-0.09212	-0.85513	-0.23012	-0.09212
0	10	0.001	-0.94991	-0.24795	-0.10912	-0.94922	-0.24800	-0.10913
1	5	0.010	-0.21462	-0.09218	-0.04436	-0.21463	-0.09213	-0.04432
1	10	0.001	-0.24190	-0.10911	-0.06051	-0.24089	-0.10919	-0.06049
2	5	0.010	-0.08765	-0.06237	-0.03157	-0.08769	-0.06213	-0.03150
2	10	0.001	-0.10730	-0.06050	-0.03538	-0.10733	-0.06050	-0.03532

Third modification.

$$V_C^M(x) == -\begin{cases} \alpha/x_1, & 0 \le x < x_1 \\ \alpha/x, & x_1 \le x \le x_2 \\ 0, & x_2 < x < \infty \end{cases}, \tag{16}$$

where $x_1 < x_2$ are parameters. Table 4 gives the results for this potential.

TABLE 4. $\alpha = -1$, $M = 0.5$

n	x_1	x_2	ϵ_{n0} (2)	ϵ_{n1} (2)	ϵ_{n2} (2)	ϵ_{n0} (6)	ϵ_{n1} (6)	ϵ_{n2} (6)
0	0.10	10	-0.98889	-0.25000	-0.11111	-0.98889	-0.25000	-0.11111
0	0.01	20	-0.99999	-0.25000	-0.11111	-0.99999	-0.25000	-0.11111
1	0.10	10	-0.24865	-0.11111	-0.06248	-0.24862	-0.11110	-0.06247
1	0.01	20	-0.25000	-0.11111	-0.06249	-0.25000	-0.11111	-0.06249
2	0.10	10	-0.11073	-0.06249	-0.03745	-0.11072	-0.06248	-0.03743
2	0.01	20	-0.11111	-0.06250	-0.03749	-0.11110	-0.06250	-0.03748

4 Modification of increasing potentials

Among the increasing potentials of type (4) a linear potential ($\gamma = 1$) and an oscillator potential ($\gamma = 2$) are usually applied to hadrons spectroscopy. Note that the use of the oscillator potential unlike the linear one, involves no difficulties. That's why this work presents the results for a case of $\gamma = 1$, although the formulas of the approximating potentials are given in general form. The eigenvalues of Schroedinger equation for $\gamma = 1, \sigma_1 = 1, M = 0.5$ presented in [4], are

$$\epsilon^*_{00} = 1.338, \epsilon^*_{10} = 4.088, \epsilon^*_{20} = 5.521, \epsilon^*_{30} = 6.787, \epsilon^*_{40} = 7.944, \epsilon^*_{50} = 9.023.$$

Let us consider some modifications of approximation for potential (4).

First modification.

$$V^M_\gamma(x) = V_{1\gamma}(x) + V_{2\gamma}(x), \tag{17}$$

$$V_{1\gamma}(x) = \sigma_\gamma \begin{cases} x^\gamma - x^\gamma_M, & 0 \le x < x_M \\ 0, & x_M \le x < \infty \end{cases}, \quad V_{2\gamma}(x) = \sigma_\gamma x^\gamma_M, 0 \le x < \infty, \tag{18}$$

where x_M is a parameter. The results of the calculations are listed in Table 5.

TABLE 5. $\gamma = 1$, $\sigma_1 = 1$, $M = 0.5$

n	x_M	ϵ_{n0} (2)	ϵ_{n1} (2)	ϵ_{n2}(2)	ϵ_{n0} (6)	ϵ_{n1} (6)	ϵ_{n2}(6)
0	3.	2.31865	3.12290	3.84161	2.31862	3.12285	3.84160
0	5.	2.33808	3.36075	4.24345	2.33804	3.36072	4.24343
0	7.	2.33811	3.36125	4.24818	2.33808	3.36121	4.24815
0	10.	2.33811	3.36125	4.24818	2.33810	3.36124	4.24817
1	3.	3.15919	4.62245	3.77400	3.15913	4.62213	3.77489
1	5.	4.08088	4.82777	5.30294	4.08085	4.82776	5.30290
1	7.	4.08794	4.88435	5.62885	4.08790	4.88431	5.62883
1	10.	4.08795	4.88445	5.62971	4.08793	4.88442	5.62968
2	3.	4.81313	3.88672	4.35192	4.81309	3.88670	4.35190
2	5.	5.17286	5.37217	6.56140	5.17283	5.37215	6.56132
2	7.	5.51954	6.20067	6.83369	5.51950	6.20066	6.83366
2	10.	5.52056	6.20763	6.86896	5.52052	6.20763	6.86893
3	3.	4.23014	4.65850	5.15077	4.23009	4.65845	5.15075
3	5.	5.62870	6.06615	6.50373	5.62865	6.06613	6.50371
3	7.	6.74819	7.25408	7.64995	6.74817	7.25407	7.64994
3	10.	6.78679	8.01273	8.01276	6.78675	7.40562	8.01269

It is clear that for each set (n, l) the computed eigenvalues ϵ_{nl} coincide with those given in [4] if $x_M > \epsilon_{nl}$. Thus, correctness of the computations can be easily controlled

by means of computations on a sequence of increasing values of the parameter x_M. Note that this way of approximation of the initial potential looks the most convenient one.

Second modification.

$$V_\gamma^M(x) = \sigma_\gamma[(2/\mu_1^2)(\exp(-\mu_1 x) - 1 + \mu_1 x)/x)]^\gamma. \tag{19}$$

Here $\mu_1 << 1$ is a parameter. The results are tabulated in Table 6.

TABLE 6. $\gamma = 1$, $\sigma_1 = 1$, $M = 0.5$

n	μ_1	ϵ_{n0} (2)	ϵ_{n1} (2)	ϵ_{n2}(2)	ϵ_{n0} (6)	ϵ_{n1} (6)	ϵ_{n2}(6)
0	0.050	2.28978	3.26822	4.10422	2.28975	3.26820	4.10413
0	0.001	2.33713	3.35938	4.24528	2.33708	3.35935	4.24525
1	0.050	3.94074	4.67991	5.36406	3.94070	4.67987	5.36402
1	0.001	4.08498	3.35938	5.62434	4.08494	3.35933	5.62431
2	0.050	5.25293	5.87421	6.46686	5.25291	5.87413	6.46683
2	0.001	5.51514	4.88032	6.86073	5.51509	4.88030	6.86067
3	0.050	6.38338	6.93001	7.45939	6.38335	6.93007	7.45934
3	0.001	6.77852	6.20086	7.99851	6.77845	6.20082	7.99848

Third modification.

$$V_\gamma^M(x) = \sigma_\gamma[2(\mu_1^2 - \mu_2^2)(\exp(-\mu_1 x) - \exp(-\mu_2 x) + (\mu_1 - \mu_2)x)/a]^\gamma, \tag{20}$$

$\mu_1 << 1$ and $\mu_2 >> 1$ are parameters. The calculation results are presented in Table 7.

TABLE 7. $\gamma = 1$, $\sigma_1 = 1$, $M = 0.5$

n	μ_1	μ_2	ϵ_{n0} (2)	ϵ_{n1} (2)	ϵ_{n2}(2)	ϵ_{n0} (6)	ϵ_{n1} (6)	ϵ_{n2}(6)
0	0.050	5	2.28887	3.26685	4.10253	2.28886	3.26682	4.10250
0	0.001	10	2.33696	3.33931	4.24216	2.33694	3.33928	4.24213
1	0.050	5	3.93885	4.66230	5.36327	3.93880	4.66227	5.36323
1	0.001	10	4.08339	3.31894	5.62378	4.08335	3.31890	5.62375
2	0.050	5	5.25137	5.86448	6.46584	5.25133	5.86445	6.46580
2	0.001	10	5.51381	4.86996	6.85990	5.51378	4.86991	6.85987
3	0.050	5	6.38244	6.92142	7.45671	6.38241	6.92137	7.45668
3	0.001	10	6.77680	6.18772	7.99355	6.77678	6.18768	7.99350

5 Numerical solution of the relativistic equation

Let us consider a relativistic equation [2]:

$$[2\sqrt{p^2 + m^2} - (2m + E_{nl})]\phi_{nl}(p) = -\int_0^\infty dq V_l(p,q)\phi_{nl}(q). \tag{21}$$

Here $m = 2M$ is a parameter of the model, V_l is a linear potential for representation of which the form (17)-(18) is used. Note that the equation (21) reduces to (6) if to use $\hat{Q}(p) = 2(\sqrt{p^2 + m^2} - m)$ in place of $Q(p)$. Table 8 gives the values of $E_0^R = \epsilon_{n0}$ and E_{n0}^R $(n = 0, 1)$ for various value of the m parameter for the relativistic problem (21) and for the nonrelativistic Schroedinger equation. The latter case gives the computations both in coordinate and in momentum space. In doing so, for the coordinate representation of Schroedinger equation we used a non-modified linear potential. The parameters of the computational scheme were selected in order to provide a coincidence of three signs after the point for solutions of the problems (2) and (6).

TABLE 8.

m	E_{00}^{NR}	E_{00}^{R}	E_{10}^{NR}	E_{10}^{R}
0.01	10.888	3.134	18.983	4.687
0.10	5.037	2.971	8.807	4.522
0.50	2.946	2.513	5.150	3.986
1.00	2.338	2.143	4.088	3.554
2.00	1.856	1.782	3.245	3.033
3.00	1.621	1.580	3.828	2.718
5.00	1.367	1.350	2.391	2.337

It is evident from the Table 8 that the value of E tends to decrease with increase of m. Moreover, the solutions of the relativistic and the nonrelativistic equations converge. Note that for $n = 1$ the convergence of the eigenvalues with increase of m is more slow than for $n = 0$.

It is easy to find a relationship between solutions $\{E_{nl}^{NR}, \phi_{nl}^{NR}\}$ of the nonrelativistic equation (6) and solutions $\{E_{nl}^{R}, \phi_{nl}^{R}\}$ of the relativistic equation (21)

$$E_{nl}^{NR} - E_{nl}^{R} = [\int_0^{R_M} dp v_R(p)\phi_{nl}^{NR}(p)\phi_{nl}^{R}(p)]/[\int_0^{R_M} dp \phi_{nl}^{NR}(p)\phi_{nl}^{R}(p)], \tag{22}$$

where

$$v_R(p) = (p^2/m)[1 - 2/(\sqrt{1 + p^2/m^2} + 1)], \quad v_R(p) \geq 0, \quad 0 \leq p \leq R_M.$$

Applying the theorem about average, for $n = 0$ the relation (22) can be rewritten as

$$E_{0l}^{NR} - E_{0l}^{R} = v_R(\xi), \qquad 0 < \xi < R_M. \tag{23}$$

Note that

$$\min_{[0,p_M]} v_R = v_R(0) = 0, \qquad \max_{[0,p_M]} v_R = v_R(p_M).$$

In a similar way for $n = 0$ one can get a relationship of the eigenvalues of eq.(21) for two values of the parameter m: m_1 and m_2.

$$E_{0l}^{R}(m_1) - E_{0l}^{R}(m_2) = w_R(\xi), \tag{24}$$

$$u(\xi) = 2(m_1 - m_2)[1 - (m_1 + m_2)/(\sqrt{\xi^2 + m_1^2} + \sqrt{\xi^2 + m_2^2})].$$

As $E_{0l}^{NR}(m) \sim E_{0l}^{R}(m)$ for large m, the equality (24) can be written as follows:

$$E_{0l}^{R}(m_1) \sim E_{0l}^{NR}(m_2) + w_R(\xi). \tag{25}$$

For $n > 0$ the equality (23) and (24) are fulfilled approximately.

Relations (23)-(25) are useful for control of the numerical calculations as well as for approximate estimation of relativistic effects without solving the equation. Analogous estimates of the relativistic correction of the spectrum can be also obtained for other relativistic generalizations of Schroedinger equation.

Table 9 gives the relationships E_{nl}/E_{0l} calculated. To compare the results, we present the calculations performed in [8]. Our results prove the convergence of the solutions of the relativistic and the nonrelativistic equations with increase of the parameter m. Besides, it is seen from the tables that the relation E_{nl}/E_{0l} does not depend upon the value of m in the nonrelativistic case because of (12).

TABLE 9.

$M_{(eV)}$	n	$\frac{E_{n0NR}}{E_{00NR}}$	$\frac{E_{n0R}}{E_{00R}}$	$\frac{E_{n0R}}{E_{00R}}$ [8]	$\frac{E_{n1NR}}{E_{01NR}}$	$\frac{E_{n1R}}{E_{01R}}$	$\frac{E_{n1R}}{E_{01R}}$ [8]	$\frac{E_{n2NR}}{E_{02NR}}$	$\frac{E_{n2R}}{E_{02R}}$	$\frac{E_{n2R}}{E_{02R}}$ [8]
300	1	1.75	1.69	1.63	1.45	1.40	1.43	1.33	1.28	1.31
	2	2.36	2.22	2.11	1.85	1.74	1.80	1.62	1.53	1.58
	3	2.90	2.68	2.51	2.20	2.04	2.13	1.89	1.75	1.82
500	1	1.75	1.71	1.67	1.45	1.42	1.39	1.33	1.30	1.27
	2	2.36	2.28	2.18	1.85	1.82	1.71	1.62	1.56	1.51
	3	2.90	2.76	2.62	2.20	2.14	2.00	1.89	1.80	1.72
1500	1	1.75	1.74	1.72	1.45	1.44	1.36	1.33	1.32	1.25
	2	2.36	2.34	2.30	1.85	1.83	1.67	1.62	1.60	1.47
	3	2.90	2.87	2.80	2.20	2.17	1.93	1.89	1.86	1.67

6 Conclusion

This work represents an approach to numerical study of the effect of the potentials' approximation on the accuracy of solving the problems in momentum space. The performed numerical analysis of dependence of solving Schroedinger equation upon a sort of an approximating potential allows to estimate applicability of one or another modif cation taking the required calculation accuracy into account. It has been shown that with n the class of the potentials discussed the relations between the eigenvalues are independent of the parameters of Schroedinger equation. The efficiency of one of the proposed modifications is demonstrated on an example of solving a concrete relativistic equation. The approximated estimations of the relativistic correction for the spectrum obtained for the given equation allow to determine, without solving the equation, the contribution of relativistic effects on change in the spectrum of the equation depending on the parameters of the model.

The authors are sincerely grateful to Prof. E.P. Zhidkov for his permanent attention to the work and useful remarks. We also would like to thank Dr. A.A. Machavariani for materials presented for our tests.

References

[1] F.Gross, Phys.Rev. **B 6**, 125 (1968); F.Gross and J.Milane, Phys.Rev. **D 43**, 2401 (1991); F.Gross and J.Milane, Phys.Rev. **D 45**, 969 (1992); V.G.Kadyshevsky, Nucl.Phys. **B 6**, 125 (1968).

[2] R.H.Thompson, Phys.Rev. **D 1**, 110 (1970).

[3] A.Le Yaouanc et al., Phys.Rev. **D 29**, 1233 (1984); Phys.Rev. **D 31**, 137 (1985); S.L.Adler and A.C.Davis, Nucl.Phys. **B 244**, 469 (1984). I.V.Amirkhanov et al., Math. Model. **6** N7, 55 (1994); **7** N7, 34 (1995).

[4] G.Bethe and E.E.Salpeter. Quantum Mechanics for one or two electrons. Fizmatgiz, Moscow (1960).

[5] E.P.Zhidkov, G.I.Makarenko and I.V.Puzynin, PhEPAN **4** N1, 127 (1973); I.V.Puzynin and T.Zhanlav, Comput.Math.and Math.Phys. **32** N6, 846 (1992).

[6] I.V.Puzynin, T.P.Puzynina and T.A.Strizh, JINR Comm. P11-87-332, Dubna (1987).

[7] E.V.Zemlyanaya, JINR Comm. P11-94-120, Dubna (1994).

[8] K.H.Maung, D.E.Kahana and J.W.Norbury, Phys.Rev. **D 47** N3, 1182 (1993).

NUMERICAL INVESTIGATION OF QUANTUMFIELD MODEL OF STRONG-COUPLING BINUCLEON [1]

I.V.Amirkhanov, E.V.Zemlyanaya, I.V.Puzynin, T.P.Puzynina, T.A.Strizh

Joint Institute for Nuclear Research
Laboratory of Computing Techniques and Automation
141980 Dubna, Moscow Reg., Russia
e-mail: puzynin@lcta15.jinr.dubna.su, puzynina@lcta15.jinr.dubna.su, zemel@lcta7.jinr.dubna.su, str@cv.jinr.dubna.su

V.D.Lakhno

RAN Computer Center (IMPB), 142292 Pushchino, Moscow Reg., Russia
e-mail: LAK@impb.serpukhov.su

Abstract. The quantumfield binucleon model for the case of the nucleon spot interaction with the scalar and pseudoscalar mesons fields is considered. For the strong coupling limit the nonlinear equations describing two nucleons in the meson field are developed. Methods of numerical investigation and obtained results are discussed.

The calculation of binucleon states has been the objective of many investigations, references to which can be found in books and reviews concerned with the binucleon problem [1,2] This problem is of interest because this is its solution that gives direct information on nuclear forces.

In this paper we make an attempt to calculate a bound state of two nucleons in a meson field by using a consistent translation - invariant theory of strong interaction between a particle and a quantum field. It is well known that in the limit of strong particle - field interaction the field can be considered as classical in zero approximation [3].

The theoretical consideration of this problem was done in [4] by use of the canonical Bogolubov-Tyablikov transformation in the limit of strong coupling.

1 Differential equations for binucleon

The Schroedinger equation for binucleon with potential

$$\Pi_C(r) = -\frac{2g^2}{\pi}\int d^3r'|\psi_0(r')|^2\frac{e^{-\mu_C|r-r'|/2}}{|r-r'|}, \tag{1}$$

where g is the constant of the interaction with the scalar field, μ_C is the scalar pion mass, can be presented in the form of variational principle with the functional:

$$F = \frac{\hbar^2}{m}\int|\nabla\psi|^2dV - \frac{g^2}{\pi}\int\frac{e^{-k_C|r-r'|/2}}{|r-r'|}|\psi(r)|^2|\psi(r')|^2dVdV' - \tag{2}$$
$$-\tilde{f}^2\int(\sigma\nabla_r)(\sigma\nabla_{r'})\frac{e^{-k_{\Pi C}|r-r'|/2}}{|r-r'|}|\psi(r)|^2|\psi(r')|^2dVdV',$$

[1]This work was supported in part by the Russian Foundation for Basis Research under Grant No 94-01-01119

where

$$\tilde{f}^2 = \frac{3f^2}{4\pi\mu^2}, \tag{3}$$

$$k_C = \mu_C, \qquad k_{\Pi C} = \mu = \mu_\pi, \tag{4}$$

Here m is a nucleon mass, μ is the pion mass, f is the interaction constant and σ is spin operator. Taking into account the condition for the functional (2) minimum on the class of functions $\psi(\vec{r})$ with the normalizing conditions

$$\int d\vec{r}|\psi(\vec{r})|^2 = 1, \tag{5}$$

we get the following system of equations

$$\begin{cases} \frac{\hbar^2}{m}\Delta\psi - \varepsilon\psi + \frac{2g^2}{\pi}V_1\psi + 2\tilde{f}^2[(\sigma\nabla_r)V_2]\psi = 0, \\ \Delta V_1 - (\frac{k_C}{2})^2 V_1 = -4\pi|\psi|^2, \\ \Delta V_2 - (\frac{k_{\Pi C}}{2})^2 V_2 = -4\pi(\sigma\nabla_r)|\psi|^2. \end{cases} \tag{6}$$

Using the notation

$$V_3 = (\sigma\nabla_2)V_2$$

and assuming $\hbar^2 = 1$, we can present (6) as a following system

$$\begin{cases} \frac{1}{m}\Delta\psi - \varepsilon\psi + \frac{2g^2}{\pi}V_1\psi + 2\tilde{f}^2 V_3\psi = 0, \\ \Delta V_1 - (\frac{k_C}{2})^2 V_1 = -4\pi|\psi|^2, \\ \Delta V_3 - (\frac{k_{\Pi C}}{2})^2 V_3 = -4\pi(\sigma\nabla_r)^2|\psi|^2. \end{cases} \tag{7}$$

Energy levels ε and wave functions ψ of the binucleon can be determined by solving the system (7) on the set of functions $\psi(\vec{r})$, which satisfy the normalization condition (5) and are limited for $0 < \vec{r} < \infty$, with boundary conditions

$$\lim_{r\to\infty}\psi(\vec{r}) = 0, \qquad \lim_{r\to\infty}V_3(\vec{r}) = 0, \qquad \lim_{r\to\infty}V_1(\vec{r}) = 0. \tag{8}$$

(Here $r = \sqrt{x^2+y^2+z^2}$.) Then for the binucleon radius $\bar{R}$ and the quadrupole moment $\bar{Q}$ we have

$$\bar{R} = \int d\vec{r}r|\psi|^2, \qquad \bar{Q} = \int d\vec{r}(3z^2 - r^2)|\psi|^2. \tag{9}$$

Turning in eqs.(7) to dimensionless variables

$$\vec{r} = r_0\bar{\vec{r}}, \qquad \psi = \gamma\varphi, \tag{10}$$

we get the following system of equations (bar mark for vector $\bar{r}$ was omitted)

$$\begin{cases} \Delta\varphi - \lambda\varphi + (V_1 + BV_3)\varphi = 0, \\ \Delta V_1 - k_1^2 V_1 = -|\varphi|^2, \\ \Delta V_3 - k_2^2 V_3 = -(\sigma\nabla_r)^2|\varphi|^2, \end{cases} \tag{11}$$

where

$$\lambda = \varepsilon m r_0^2, \quad k_1^2 = (\frac{k_C}{2})^2 r_0^2, \quad k_2^2 = (\frac{k_{\Pi C}}{2})^2 r_0^2, \quad B = \frac{\tilde{f}^2}{g^2}\frac{\pi}{r_0^2}, \quad 8g^2 m r_0^4 \gamma^2 = 1. \tag{12}$$

Using the last equation we can determine γ, for example, supposing $r_0 = 10^{-13}$cm.

Taking the change of variables (10) into account, we rewrite the Eqs. (5),(9) and (2) in the form:

$$\hat{N} = \int d\vec{r}|\varphi|^2 = \frac{1}{r_0^3\gamma^2}, \qquad R = r_0^4\gamma^2\hat{R}, \quad \hat{R} = \int d\vec{r} r|\varphi|^2, \tag{13}$$

$$Q = r_0^5\gamma^2\hat{Q}, \quad \hat{Q} = \int d\vec{r}(3z^2 - r^2)|\varphi|^2, \qquad F = \frac{\varepsilon}{\lambda}\frac{(\hat{T} - \frac{1}{2}W_{13})}{\hat{N}}, \tag{14}$$

$$T = r_0\gamma^2\hat{T}, \quad \hat{T} = \int d\vec{r}|\nabla\varphi|^2, \qquad W_{13} = \int d\vec{r}(V_1 + BV_3)|\varphi|^2. \tag{15}$$

Then we will use the next relations between the variables:

$$R = r_0(\frac{\hat{R}}{\hat{N}}), \qquad Q = r_0^2(\frac{\hat{Q}}{\hat{N}}). \tag{16}$$

Here the expressions in the brackets are dimensionless variables.

The system of differential equations for bipolaron (11) is the nonlinear three-parameter eigenvalue problem. Numerical integration of (11) was perform in [5] in the limit $k_{\Pi C} \to \infty$. Thus, the solution of (11) for finite $k_{\Pi C}$ is an actual problem.

2 Statement of the boundary value problem for the system of equations

Here we discuss the statement of the boundary problem for the system (11) in the special case of axial symmetry, namely $\Phi(\rho, z, \phi) \to \Phi(\rho, z)$ and for the finite interval: $-z_M \le z \le z_M$, $0 \le \rho \le \rho_M$. We can formulate the statement of problem in terms of the ordinary differential equations. The solution of the system (11) in cylindrical coordinates we present as

$$\varphi(\rho, z) = \sum_{n=1}^{\infty} y_n(z)\Phi_n(\rho), \qquad V_1(\rho, z) = \sum_{n=1}^{\infty} V_{1n}(z)\Phi_{1n}(\rho), \qquad V_3(\rho, z) = \sum_{n=1}^{\infty} V_{3n}(z)\Phi_{3n}(\rho), \tag{17}$$

where function $\{\mu_n, \Phi_n\}$ are determined as the solution of a spectral problem

$$\begin{cases} \Phi_n'' + \dfrac{1}{\rho}\Phi_n' + \mu_n\Phi_n = 0, \quad 0 \le \rho \le \rho_M, \\ \Phi_n'(0) = 0, \quad \Phi_n'(\rho_M) - A_{\rho_M}\Phi_n(\rho_M) = 0, \\ \int d\rho\rho\Phi_n^2 = 1. \end{cases} \tag{18}$$

Here A_{ρ_M} is determined from

$$A_{\rho_M} = \left.\frac{\frac{d}{d\rho}[\Phi_A(\rho)]}{\Phi_A(\rho)}\right|_{\rho=\rho_M}, \quad \Phi_A(\rho) = \int_{-z_M}^{z_M} dz y_0(z)\varphi_A(z, \rho).$$

The function φ_A is replaced by V_{1A} or V_{3A}, and the function $y_0(z)$ is replaced by V_{10} or V_{30}, respectively, for the sets of $\{\mu_{1n}, \Phi_{1n}\}$ or $\{\mu_{3n}, \Phi_{3n}\}$. In special case we can use only one set of functions $\{\Phi_n\}$. Substituting the expansion (17) into the system (11), we have

$$\begin{cases} y_n'' - (\lambda + \mu_n) y_n + \sum\limits_{m=1}^{\infty} W_{nm} y_n = 0, \quad n = 1, 2, \cdots \\ V_{1j}'' - (k_1^2 + \mu_{1j}) V_{1j} = -J_{1j}, \quad j = 1, 2, \cdots \\ V_{3j}'' - (k_2^2 + \mu_{3j}) V_{3j} = -J_{3j}, \end{cases} \tag{19}$$

where

$$W_{nm} = \sum_{j=1}^{\infty} [Q_{1njm} V_{1j} + B Q_{3njm} V_{3j}], \qquad J_{1j} = \sum_{n,m=1}^{\infty} Q_{1njm} y_n y_m, \tag{20}$$

$$J_{3j} = \sum_{n,m=1}^{\infty} Q_{3njm} \frac{d^2}{dz^2}(y_n y_m), \qquad Q_{injm} = \int_0^{\rho_M} d\rho\rho \Phi_n(\rho) \Phi_{ij}(\rho) \Phi_{im}(\rho), \quad i = 1, 3. \tag{21}$$

Functions y_n, V_{1j}, V_{3j} satisfy the following boundary conditions

$$\begin{cases} y_n'(z_M) - A(z_M) y_n(z_M) = 0, \\ V_{1j}'(z_M) - A_1(z_M) V_{1j}(z_M) = 0, \\ V_{3j}'(z_M) - A_3(z_M) V_{3j}(z_M) = 0, \end{cases} \tag{22}$$

$$\begin{cases} y_n'(-z_M) - A(-z_M) y_n(-z_M) = 0, \\ V_{1j}'(-z_M) - A_1(-z_M) V_{1j}(-z_M) = 0, \\ V_{3j}'(-z_M) - A_3(-z_M) V_{3j}(-z_M) = 0, \end{cases} \tag{23}$$

where

$$A(z_M) = \left. \frac{\frac{d}{dz}[y_A(z)]}{y_A(z)} \right|_{z=z_M}, \quad y_A(z) = \int_0^{\rho_M} d\rho\rho \Phi_0(\rho) \varphi_A(z, \rho),$$

$$A_1(z_M) = \left. \frac{\frac{d}{dz}[V_{1A}(z)]}{V_{1A}(z)} \right|_{z=z_M}, \quad V_{1A}(z) = \int_0^{\rho_M} d\rho\rho \Phi_{10}(\rho) V_{1A}(z, \rho),$$

$$A_3(z_M) = \left. \frac{\frac{d}{dz}[V_{3A}(z)]}{V_{3A}(z)} \right|_{z=z_M}, \quad V_{3A}(z) = \int_0^{\rho_M} d\rho\rho \Phi_{30}(\rho) V_{3A}(z, \rho).$$

The coefficients $A(-z_M)$, $A_1(-z_M)$ and $A_3(-z_M)$ are determined in the similar way.

The normalization conditions for the system (19) has the form:

$$\sum_n \int_{-z_M}^{z_M} dz y_n^2(z) = \frac{1}{2\pi r_0^3 \gamma^2} = \hat{N}_c,$$

if we use the boundary conditions (22),(23).

Expressions (2) and (9) for the physical values calculations using the expansion (17) can be written as:

$$\hat{R} = 2\pi \int_{-z_M}^{+z_M} dz \int_0^{\rho_M} d\rho\rho\sqrt{(z^2+\rho^2)}\phi^2(z,\rho), \qquad \hat{Q} = 2\pi \int_{-z_M}^{+z_M} dz \int_0^{\rho_M} d\rho\rho(2z^2-\rho^2)\phi^2(z,\rho), \tag{24}$$

$$F = \frac{\varepsilon}{\lambda}[\hat{T}_c - \frac{1}{2}\hat{W}_c]/(N_c * 2\pi), \qquad \hat{T}_c = 2\pi \int_{-z_M}^{+z_M} dz \int_0^{\rho_M} d\rho\rho[|\frac{\partial\phi^2}{\partial\rho}| + |\frac{\partial\phi^2}{\partial z}|], \tag{25}$$

$$\hat{W}_c = 2\pi \sum_{jnm} \int_{-z_M}^{+z_M} dz[V_{1j}(z) + BV_{3j}(z)]y_n(z)y_m(z),$$

Using the problem symmetry the system (19) can be solved on the interval $0 \leq 0 \leq z_M$, with appropriate change of the boundary conditions and normalization.

3 Numerical methods and analysis of numerical results

Calculations of the basic functions $\{\Phi_n\} \equiv \{\Phi_{1n}\} \equiv \{\Phi_{3n}\}, n = 1, L, \quad L = 1, 2, 3, 4, 5,$ for the given ρ_M was performed. The code SLIP1 [6] was used for the numerical solving of the problem (18) for the base functions $\{\Phi_n\}$ with normalization (18). The algorithm suggested in [7] was used for solving the problem (19). This algorithm is based on the combination of continuous analogue of the Newton's method (CANM)[8] and the continuation method and allows to reduce the solving of the complicated system of nonlinear equations to the sequential solving of the relatively simple linear problems.

The next iterative scheme was used:

- The problem (18) was solved with the boundary conditions

$$\Phi_n(\rho_M) = 0 \tag{26}$$

  for the set of basic functions determination.
- For the given initial approximation for $y_n(z)$ and a set of basic functions calculated on the first step, we evaluate the right site of second and third equations of system (19) using eqs.(20-21). Then we determine functions V_{1j} and V_{3j} $(j = n)$ as the solution of linear boundary problem by using the sweep method.
- Substituting the given approximations for V_{1j} and V_{3j} into the first equations of system (19), we have an eigenvalue problem for the system of linear differential equations, which can be solved by using CANM in the modified code based on the program START [9]. The function $exp(-|z|)$ was used as the initial approximation for $y_n(z)$ together with the solutions founded for the other set of parameters. The number of Newtonian iterations does not exceed 10-13 usually.
- Solutions y_n from the previous step are used for the calculations of J_{1j} and J_{3j} on the next iteration and so on. The process of iterations stops when the solutions, obtained after two sequential iterations, would coincides with each other with a given accuracy ε. Actually the 5–8 iterations are enough to reach the accuracy $\varepsilon = 10^{-3}$.

- Parameter A_ρ is calculated with the help of the obtained solutions y_n. Then the problem (18) is solved again with respective boundary conditions, and a new set of basic functions is determined. The solution to the problem (19) are calculated again for this set accordingly to the algorithm described above. Such a process will continue until a stabilization, i.e. until the newly-obtained solutions and the value A_{ρ_M} coincide with the preceding ones with the accuracy given a priory. Actually 3-4 iteration steps are required to reach the stabilization.

The discrete approximation of the problem was made on a h-step uniform grid $\{z_i\}$ with the help of the difference schemes of the second order convergence $O(h^2)$. The calculations on a sequence of the twice compressible grids were shown in Tab.1 and confirm the 2nd order of the convergence.

Table 1. $\hat{N}_c = 10, \quad \rho_M = 5, \quad k_1^2 = k_2^2 = B = 0.1, \quad L = 4, \quad z_M = 10.$

—	grid step	number of points	λ	$y_1(0)$
h	0.100	101	0.6362802	1.624439
h/2	0.050	201	0.6360388	1.633727
h/4	0.025	401	0.6359781	1.636136
σ	$\frac{h-h/2}{h/2-h/4}$	—	4.	4.

Table 2 lists the results obtained on a sequence of expanding intervals $[-z_M, z_M]$ showing convergence on parameter z_M.

Table 2. $\hat{N}_c = 10, \quad \rho_M = 5, k_1^2 = k_2^2 = B = 0.1, \quad h = 0.05, \quad L = 4.$

z_M	λ	$y_1(0)$	$y_2(0)$	$y_3(0)$	$y_4(0)$
5.	0.538	1.641	0.428	0.033	0.003
10.	0.532	1.634	0.423	0.032	0.003
15.	0.531	1.630	0.422	0.032	0.003
20.	0.531	1.630	0.422	0.032	0.003

Table 3 lists the results for different number of basic functions L. In the top raw of each column the value λ, normalization N_n and amplitude $\phi_n(z)$ for $n = 1, L$, obtained for the boundary conditions (26) was given. The same values after the solution stabilization for parameter A_{ρ_M} was given in the bottom row.

Table 3. $\hat{N}_c = 10, k_1^2 = k_2^2 = B = 0.1, \rho_M = 5$

L	λ	$N_n, n = \overline{1, L}$	$y_n(0), n = \overline{1, L}$
2	0.538	8.88,1.12	1.78,0.62
	0.632	8.87,1.13	1.61, 0.59
3	0.616	8.86,1.10,0.04	1.78,0.62,0.16
	0.635	8.01,1.94,0.05	1.59,0.79,0.15
4	0.626	8.82,1.13,0.05,0.001	1.65,0.45,0.33,0.03
	0.636	8.86,1.10,0.04,0.001	1.59,0.42,0.32,0.02
5	0.637	$8.82,1.13,0.05,0.001,10^{-5}$	1.65,0.43,0.33,0.03,0.002
	0.636	$8.85,1.10,0.05,0.001,10^{-5}$	1.63,0.42,0.32,0.03,0.003

Now let us turn our attention to experimental data. According to [10] radius is $R = 1.963 \cdot 10^{-13}$cm, quadrupole momentum is $Q = 2.86 \cdot 10^{-27}$cm^2 and binding energy is equal to 2.246MeV.

It follows from (24)–(25) that at $r_0 = 10^{-13}$ the calculated dimensionless values correspond to the experimental ones if the following relations are fulfilled

$$\hat{Q}/(\hat{N}_c \cdot 2\pi) \simeq 0.29, \quad \hat{R}/(\hat{N}_c \cdot 2\pi) \simeq 1.96, \quad F \simeq 2.246 \quad . \tag{27}$$

According to (12), the parameters of the model must be $k_1^2 \sim 3$, $k_2^2 \sim 0.1$, and the B and $\hat{N}_c$ must be free. However, we failed to get satisfactory results for these parameters. Considering k_1 as arbitrary parameter and fitting its value from the conditions of fulfillment of the mentioned relations (27), we succeeded in obtaining the best results at $\hat{N}_c = 8.5$, $k_1^2 = 0.1$, $B = 0.001$. Hence, we have $\lambda = 0.41$, $\bar{R}/(2\pi N_c) = 2.19$, $\bar{Q}/(2\pi N_c) = 0.27$, $\varepsilon = \lambda * 40 \cong 25.6$. Turning to dimension values we find $R \simeq 2.2 \cdot 10^{-13} cm$, $Q \simeq 2.7 \cdot 10^{-27} cm^2$, $F = 2.43 eV$.

4 Conclusion

Calculating numerically the nonlinear equations for two nucleons in a scalar and pseudoscalar field, we evaluate binding energy, effective radius and deuteron quadrupole momentum, which fit the experimental data, if we assume the effective mass of scalar meson to be near the mass of pseudoscalar meson $\mu_s \approx 140 Mev$ and $k_1^2 = k_2^2 = 1$. At the same time, this fitting yields the coupling constants g and f one order less than experimental data determined by the local model of nucleon-meson interactions. But our model is nonlocal and nonlinear.

Note that we investigated only the simplest case of spot interaction. The nucleon formfactor included in the equations can change significantly the final results. The other important trend is to generalize the theory of the case of intermediate coupling constants and take into account relativistic effects.

References

[1] T.Ericson, W.Weise. Pions and nuclei. Claredon Press, Oxford, 1988.

[2] G.E.Brown, A.D.Jackson. The nucleon - nucleon interaction. North - Holland P.C., Amsterdam, 1976.

[3] E.P.Gross. Ann.Phys. **19**, 219, 1962.

[4] N.N.Bogolubov. Ukr.Mat.Zh. 2, 3, 1950.

[5] V.D.Lakhno. Teor.Mat.Fiz. v.100, N2, p.219, 1994.

[6] I.V.Puzynin, T.P.Puzynina, in coll.:"Algorithms and programs for solving some problems of physics", KFKI-74-34, Budapest (93-111), 1974.

[7] I.V.Amirkhanov et al. In "Polaron and Applications", Ed. V.D.Lakhno, John Willey & Sons Ltd., p. 445, 1994.

[8] E.P.Zhidkov, G.I.Makarenko and I.V.Puzynin, PhEPAN **4** N1, 127 (1973); I.V.Puzynin and T.Zhanlav, Comput.Math.and Math.Phys. **32** N6, 846 (1992).

[9] Yu.S.Smirnov.JINR Communications, P11-88-912, Dubna, 1988.

[10] K.N.Mukhin. Experimental nuclear physics. v.I, Energoizdat, 1983.

NONLINEAR DYNAMICS OF PARTICLE BEAMS WITH SPACE CHARGE [1]

Serge Andrianov

St.Petersburg State University
e-mail: serge@asn.apmath.spb.su

Abstract. This paper is devoted to problems of modelling of particle beams with space charge. For some models of space charge distribution functions the necessary motion equations are written. For solving of these equations the matrix formalism for Lie algebraic tools is used. This approach permits to obtain solutions of the motion equations in symbolic forms. These solutions have forms of block-matrices (solution matrices) which can be created in advance for different control beam systems and the distribution functions. All symbolic information about model distribution functions and solution matrices form is involved to some base-dates. In this paper we suggest two methods of space charge description: the nonlinear rms-envelope matrices language and the distribution functions language. The necessary symbolic computations are performed with the help of computer algebra codes *REDUCE* and *MAPLE*.

1 Introduction

Many approaches to the space charge modelling problem are known. The basic difficulty which arises for this problem is a nonlinearity of corresponding motion equations even for simplest cases. In this paper evolution of long beams with an elliptical cross-section in transverse phase space is discussed. As examples some special forms of phase space distribution function are described [1]. For these distribution functions the Ferrer's integrals technique are used. It allows to obtain motion equations in symbolic forms as expansions in series. These equations are written in a matrix form. The truncated matrix equation (up to a certain order) are solved with the use of the matrix formalism for Lie algebraic tools [2]. Nonlinear nature of these equations lead to a necessity to use the successive approximations method. Obtained convergences bounds and algorithms give opportunity to estimate a current step value in advance and to create necessary software for modelling. More over the suggested approach allows to use object–oriented ideology and to realize the dynamic modeling paradigm [3, 4].

2 A Space Charge Description

2.1 The Initial Space Charge Distributions

A researcher can use two types of space charge description. The first is based on space charge distribution in the phase space and the second on the distribution in the configuration space [1]. Note that in general a phase-space distribution function (or a particle

[1] This work was supported in part by the Russian Foundation for Fundamental Researches under Grant No. 96-02-17336

density distribu ion) $f_0(X) = f(X, s_0)$ can be written in the form $f_0(X) = \sum_{k=0}^{\infty} f_k^0 X^{[k]}$ or in the case of el iptical symmetry

$$f_0(X) = \phi_0(\kappa^2) = \sum_{k=0}^{\infty} a_k^0 \kappa^{2k} = \sum_{k=0}^{\infty} a_k^0 (X^{[k]})^* \mathbf{A}_0^{\{k\}} X^{[k]}, \qquad (1)$$

where $X^{[k]} = \underbrace{X \otimes \cdot \ldots \cdot \otimes X}_{k-times}$ is the Kronecker power of the phase vector $X = (x, p_x, y, p_y)^*$, $\dim X^{[k]} = \binom{k+3}{k}$, $\mathbf{A}_0^{\{k\}}$ is the symmetrical Kronecker power of the initial form matrix $\mathbf{A}_0$: $\kappa^2 = X^* \mathbf{A}_0 X$:

$$(\mathbf{A}^{\{k\}})_{il} = b_i(k)(\mathbf{A}_0^{[k]})_{il}, \quad i, l = \overline{1, d[k]}, \quad d[k] = \binom{n+k-1}{k}$$

and $b_i(k)$ are the polynomial coefficients of k-th order, which can be calculated from the known expressic n $k!/k_1! \ldots k_n!$ using the lexicographic order $k_1 \succ k_2 \succ \ldots \succ k_n$.

2.2 The Self–Field of Space Charge

Using the Ferrers's integrals we can calculate the desired potential for the beam in the form

$$V = -\frac{\pi ab}{\epsilon_0} \int_0^{\infty} \Phi(\kappa_1^2(u)) \frac{du}{\Delta(u)},$$

when $\Phi(\kappa^2) = \int_{\zeta}^{\kappa^2} \rho(t)dt$,

$\Delta^2(u) = (a^2 + u)(b^2 + u)$, $\kappa_1^2(u) = \frac{\xi^2}{a^2+u} + \frac{\eta^2}{b^2+u}$. After some transformations we obtain the components of the vector of self–electrical field

$$E_\xi = \frac{2\pi ab}{\epsilon_0} \rho_0 \int_0^{\infty} \Big\{ \frac{\xi}{a^2+u} \Phi(\kappa_1^2(u)) \Big\} \frac{du}{\Delta(u)}, \quad E_\eta = \frac{2\pi ab}{\epsilon_0} \rho \int_0^{\infty} \Big\{ \frac{\xi}{b^2+u} \Phi(\kappa_1^2(u)) \Big\} \frac{du}{\Delta(u)}.$$

and for our par icle density distributions we can write

$$E_{\xi,\eta} = E_{\xi,\eta}^0 + \triangle E_{\xi,\eta}, \quad E_\xi^0 = \frac{4\pi\rho_0}{\epsilon_0} \frac{ab}{a(a+b)} \xi, \quad E_\eta^0 = \frac{4\pi\rho_0}{\epsilon_0} \frac{ab}{b(a+b)} \eta.$$

The values of $\triangle E_\xi, \triangle E_\eta$ can be easily calculated with the help of *MATHEMATICA* (or *REDUCE*) codes. Note that for the K-V distribution we have $\triangle E_\xi = \triangle E_\eta = 0$. Besides, if the arbitrary distribution $\rho(x, y) = \rho_0 \Phi(\kappa_1^2)$ is a polynomial of n-th order with respect to the variables κ_1^2 the functions $\triangle E_\xi$ and $\triangle E_\eta$ are polynomials of $(2n+1)$-th order with respect to the variables ξ and η. Then we can return to the coordinates referenced to a desired system.

3 The motion equation

3.1 The motion equation for a single particle

As an example we consider the transport system with the rectilinear optical axis (there are no dipole magnets). For a nonbunched beam (the longitudinal self–e ectric field is missing) the motion equation for single particle can be written in the form

$$x'' = \frac{q}{p}T\Big\{y'B_s - (1-x'^2)B_y + x'y'B_x + T\frac{E_x}{\beta c\gamma^3}\Big\},$$

$$y'' = -\frac{q}{p}T\Big\{x'B_s - (1-y'^2)B_x + x'y'B_y - T\frac{E_y}{\beta c\gamma^3}\Big\},$$

where $T = \sqrt{1+x'^2+y'^2}$, $' = d/ds$. Note that the self–magnetic field is considered with the help of the coefficient $1/\gamma^3$.

These motion equations can be rewritten in following matrix form

$$\frac{dX}{ds} = \sum_{k=1}^{\infty}\Big\{\mathbf{P}^{1k}_{ext}(s) + \mathbf{P}^{1k}_{self}(s)\Big\}X^{[k]}.$$

The matrices $\mathbf{P}^{1k}_{ext}$ and $\mathbf{P}^{1k}_{self}$ describe the external and self–fields correspondingly. As an example the corresponding motion equations for a quadrupole lense system in the presence of space-charge up to third order on phase space variables are considered in [1].

3.2 Transfer map in the presence of space-charge

The transfer map approach has been very useful for studying beam dyn mics without space-charge. In this case we can write

$$\mathcal{M}: X_0 \to X = \mathcal{M}(s|s_0)\cdot X_0.$$

where $\mathcal{M}$ is called the transfer map (or Lie map) between moments s_0 and s. If the beam is an ensemble of interacting particles then the map $\mathcal{M}(s|s_0)$ depend only on control system parameters. In the frame of the matrix formalism [2] we can represe it this map in the form

$$X = \mathcal{M}(s|s_0)\cdot X_0 = \sum_{k=1}^{\infty}\mathbf{M}^{1k}(s|s_0)X_0^{[k]}. \qquad (2)$$

where $\mathbf{M}^{1k}$ are matrices which can be calculated with the help of the matrix formalism tools.

According to the previous sections our beam has the elliptical cross-sect on. So let M_0 be an initial phase set occupied by particles:

$$M_0 = \Big\{X_0 :\ X_0^*\mathbf{A}_0^{11}X_0 \le 1\Big\}.$$

Let $\sigma_0^{11} = (\mathbf{A}_0^{11})^{-1}$. The other definition of the σ-matrix (so-called *rms*-envelope matrix) is following

$$\tilde{\sigma}_0^{11} = \int\limits_{M_0} f_0(X)XX^*dX,$$

where f_0 is an initial phase space distribution. The similar formula can be used for the definition

$$\tilde{\sigma}_0^{ik} = \int\limits_{M_0} f_0(X) X^{[i]} (X^{[k]})^* dX.$$

These definitions allow the beam dynamics to be investigated in the terms of envelopes. Indeed, according to the matrix formalism for both cases of the definition of the σ-matrix we have

$$\sigma^{11}(s) = \sum_{i=1}^{\infty} \sum_{k=1}^{\infty} \mathbf{M}^{1k} \sigma_0^{kj} (\mathbf{M}^{1j})^*.$$

In the presence of space-charge the Lie map will depend on the parameters of the beam too. In this case the operator equation for the map $\mathcal{M}$ is a nonlinear equation. The solving this equation requires the another approach. In this paper we suggest the method of step-by-step approximations. The basic idea of this method in our case is to calculate the envelope matrix according to the following algorithm (compare with [5]).

Step 0:

1) Choose the initial model distribution function $f_0(x)$ according to the Eq.(1) and the order of our approximation equal to N.

2) Define the set of matrices σ_0^{ik} according to the corresponding equations and then form the init al *nonlinear* envelope matrix σ_0^N:

$$\sigma_0^N = \left\{ \sigma_0^{ik} \right\}_{i,k=\overline{1,N}}.$$

3) Define the sequence of subintervals $I_i = [s_i, s_{i+1})$, $\bigcup\limits_{i=0}^{m} I_i = I_{tot} = [s_0, s_*]$, $i = \overline{0,m}$, $s_m = s_*$.

Step 1:

Calculate the auxiliary envelope matrices by

$$\sigma_1^{ik}(s_0) = \alpha \sigma_0^{ik}(s_0) + (1-\alpha)\sigma_0^{ik}(s_1), \quad 0 < \alpha < 1,$$

$$\sigma_0^{ik}(s_1) = \sum_{j,l=1}^{N} \mathbf{M}_1^{ij} \sigma_0^{jl}(s_0) (\mathbf{M}_1^{lk})^*,$$

$\mathbf{M}_1^{ij} = \mathbf{M}^{ij}(s_1|s_0; \{\sigma_0^{kl}\})$ is the block matrices calculated according to the matrix formalism by constant matrices σ_0^{kj} (self-consistent motion equations or self-consistent Hamiltonian functions), see the Eq.(2).

Step 2:

If $\|\sigma_1^{11}(s_1) - \sigma_0^{11}(s_1)\| < \varepsilon$ then go to finish else change s_1 on s_2 and s_0 on s_1 and return to the **Step** ..

Note that in the frames of the matrix formalism we can also calculate the particle-density distribution function. According to the Eq.(2) and the properties of the Kronecker product we can write the transformation of $X^{[k]}$:

$$X^{[k]} = \left(\mathcal{M} \cdot X_0\right)^{[k]} = \left(\sum_{j=1}^{\infty} \mathbf{M}^{1j} X_0^{[j]}\right)^{[k]} = \sum_{j=1}^{\infty} \mathbf{M}^{kj} X_0^{[j]},$$

$$\mathbf{M}^{ml} = \sum_{\substack{k_1+\ldots+k_m=l \\ k_i \geq 1}} \bigotimes_{i=1}^{m} \mathbf{M}^{1k_i}, \; l \geq m.$$

¿From the known properties of Lie maps we can write for an arbitrary function of an initial distribution $f_0(X) = f(X, s_0)$: $f(X, s) = f_0(\mathcal{M}^{-1}(s|s_0) \cdot X)$. In our case we have $\mathcal{M}^{-1}(s|s_0) \cdot X = \sum_{k=0}^{\infty} \mathbf{T}^{1k}(s|s_0) X^{[k]}$, where $\mathbf{T}^{10} = -\mathbf{M}^{10}$, $\mathbf{T}^{11} = (\mathbf{M}^{11})^{-1}$ and other matrices $\mathbf{T}^{1k}$ for $k > 1$ can be calculated with the help of the recurrent generalized Gauss's algorithm using the matrices $\mathbf{M}^{1k}$. It is worthy to note that according to this algorithm one should inverse only the matrix $\mathbf{M}^{11}$ and then use only matrix operations (with the Kronecker operations extension) for calculation the necessary matrices $\mathbf{T}^{1k}$ up to the desired order. So after some calculations we can obtain the following equation (see the Eq.(1))

$$f(X, s) = f_0(\mathcal{M}^{-1} \cdot X) =$$

$$a_0^0 \sum_{k=1}^{\infty} a_k^0 (\mathcal{M}^{-1} \cdot X^{[k]})^* \mathbf{A}_0^{\{k\}} (\mathcal{M}^{-1} \cdot X^{[k]}) =$$

$$a_0^0 + \sum_{k=1}^{\infty} a_k^0 \sum_{l=k}^{\infty} \sum_{j=k}^{\infty} (X^{[l]})^* \mathbf{B}_k^{lj} X^{[j]},$$

$$\mathbf{B}_k^{lj} = (\mathbf{T}^{kl})^* \mathbf{A}_0^{\{k\}} \mathbf{T}^{kj}.$$

For the convergence of our approximations methods for envelope matrices the following condition has to be satisfied

$$\|\sigma_k^{11} - \sigma_{k-1}^{11}\| \leq \beta \|\sigma_{k-1}^{11} - \sigma_{k-2}^{11}\|$$

for $\beta < 1$. The constant of this method β can be calculated as a function of the initial beam characteristics and the transport system parameters. The condition $\beta < 1$ allows the limitations on the step values $|s_k - s_{k-1}| = \Delta s_k$ to be calculated which guarantee the fulfillment of the inequality $\beta < 1$. The preliminary calculations showed that it is sufficient to use two – four steps of this iteration process for some value of Δs_k. Note that the convergence condition does not depend on the parameter α.

References

[1] S.N.Andrianov, Proc. of the First Int. Workshop Beam Dynamics and Optimization, July 4–8,1994, St.Petersburg, Russia, St.Petersburg, 19, (1995).

[2] S.N.Andrianov, Proc. of IY Int. Conf. on Computer Algebra in Physical Research, Dubna, USSR,1990, World Sci., Singapore, 447, (1991).

[3] Hiroshi Nishimura, Comp. in Phys., **6**, 456, (1992).

[4] S.N.Andrianov, Absrts. of the Int. Conf. on Interval and Computer-Algebra Methods in Science and Engineering, INTERVAL'94, March 7-10, 1994, St.-Petersburg, Russia, 41, (1994).

[5] A.J.Dragt, R.D.Ryne, Part.Accel., **35**, 129, (1993).

DYNAMIC MODELING PARADIGM AND COMPUTER ALGEBRA [1]

Serge Andrianov

St.Petersburg University
e-mail: serge@asn.apmath.spb.su

Abstract. This paper is devoted to description of basic concepts for modelling process of such kind of nonlinear dynamical systems as beam line systems. There are two different types of beam physics problems: long time evolution of beams and synthesis of beam line systems with desired characteristics. Both these problems have some general lines. Indeed, a researcher can include different kind of effects step by step. In this paper the matrix formalism for Lie algebraic tools for these mentioned problems is used. This formalism allows to create databases of necessary matrices for advance. For this purpose it is necessary to use computer algebra codes. They allow not only to do symbolic calculations but to do necessary investigation on early modelling steps. In particular this helps to realized a process of synthesis of dynamical systems. This approach can be realized with the help of expert systems ideology. Some basic concepts of this process are described.

1 Introduction

At first, let us consider some problems which appear during any process of dynamical systems modelling. A researcher obviously must pass following steps:

– *to define an initial set of decisions (preliminary judgements on a design problem);*

– *to applicate one or more mathematical models of various properties of the dynamical system (in our case of the beam line system).*

– *to analyse the solution and then accept or reject corresponding model, and in the later case, define a new set of decisions and return to the second step.*

These pointed steps can be put into new paradigm for modelling process – **dynamic modeling,** if this process will be ensured by necessary attributes among the we should emphasize the following:

– *the using mathematical methods have to be adequate both to physical models and modern programming methods, to put it in another way the mathematical methods have to permit using of artificial intelligence methods and facilities;*

– *objects which are manipulated by a researcher must maintain a natural extension of the set of processes and effects being investigated;*

– *objects which are manipulated by a researcher must admit maximal simple (from the a computational point of view) operations necessary for calculations.*

These and other reasons as well as author's experience show that one of the methods satisfying these requirements is the matrix formalism of Lie algebraic methods [1]. In this work we discuss how this approach can be used for beam physics, how the modelling process can be formulated and what place in this process we must assign computer algebra methods.

[1] This work was supported in part by the Russian Foundation for Fundamental Researches under Grant No. 96-02-1733

2 The Basic Definitions and Concepts

2.1 Lie Transformations

It is known that for any dynamical system we can write

$$\mathcal{M}(t|t_0): X_0 \mapsto X(t), \tag{1}$$

where $X_0 = X(t_0)$, $X(t)$ are an initial and a current states of our dynamical system (i.e. corresponding phase vectors), $\mathcal{M}(t|t_0)$ is a dynamical mapping. Such representation we can write for both autonomous and time dependence systems. In the second case we must use either chronological ordered operators or so called Magnus's representation [2]. For both cases the map $\mathcal{M}$ can be written as an exponential operator in the form

$$\mathcal{M}(t|t_0) = \exp\left\{\mathcal{L}_{G(X,t|t_0)}\right\},$$

where $\mathcal{L}_{G(X,t|t_0)}$ is a Lie operator associated with a vector function $G(X,t|t_0)$. For example, for an autonomous system we have $G(X,t|t_0) = (t-t_0)F(X_0)$, where $F(X)$ is a right part of evolution equation for the dynamical system. The knowledge of these operators is a main goal of our modelling process.

2.2 The Matrix Representation for Lie Transformations

The polynomial representation for Lie transformation is well known [3]. But this representation does not have necessary properties for above mentioned steps of modelling process in the full scope. That is why we suggest to use the matrix formalism for Lie algebraic tools [1]. The basic concepts of this formalism is to compare a set of matrices to an operator representation for Lie transformation according the following rule:

$$\mathcal{M}(t|t_0) \circ X_0 = \sum_{k=0}^{\infty} \mathbf{M}^{1k}(t|t_0) \cdot X_0^{[k]}, \tag{2}$$

where $\mathbf{M}^{1k}$ are matrices to be found and $X_0^{[k]}$ – the Kronecker power of k-th order for the phase vector X_0. The basic theory of this comparison is published in previous works by the author (see, also [4]). Here we must mention that if there is a symbolic form for $\mathbf{M}^{11}$ (for a matricant of linearized motion equation) then we can calculate all (up to desired order) matrices $\mathbf{M}^{1k}$, $k \geq 2$ in symbolic form too. Here let us point following properties of the matrix representation for Lie transformation which are very important for modelling process:

– the simplicity of the mathematical nature of used objects – matrices;

– the possibility of unifying description of a subject of control (a beam line system) and an object of control (a particle beam) using the matrix tools including space-charge effects (see, for instance, [5]);

– the possibility of extensive application of such modern computer approaches as *object–oriented programming, computer algebra, intelligent interfaces, expert systems.*

– truncated Lie maps (Lie approximants) can be easily corrected for ensuring necessary properties (for example, symplectic properties for Hamiltonian systems);

– a possibi ity to create databases for corresponding matrices in symbolic form and include them t› a calculation module.

– the runni ıg calculation experiments allow to create knowledge bases, for example in the form of re›ommendations and exclusions which can help to select models (here the reference is to :orresponding matrices) in the future.

3 Computer Algebra Calculation Problems

3.1 The Linear Matricant Calculation

As we have mentioned above a knowledge of symbolic form for a linearized equations matricant plays a main role. It is known that symbolic forms can be obtained in the case for autonc mous linear systems. There are some other cases of time dependence when such forms exi:t. As examples we may mention so called bell-like model (in the terms of trigonometrica functions) and linear dependence (in the terms of Airy's functions). There is an interest in an approach based on series representation for elements of matricants. This is very useful for some limit cases (for example, the thin lens approximation in beam physics). For this purpose the Baker approximants representation is useful too. The truncated serie› (polynomials) allow to create necessary symbolic forms for matricant very simply. But in this case one must point limits of polynomials approximations. There is an interest ap›roach (see for eaxample, [6]). According to this approach we can create some set of linear differential equations with time dependence coefficients which permit symbolic form of solutions. In particular, this approach can be use for investigation of fringing fields in magnetic lenses.

3.2 The Block Matrices Calculation

The second ste›p of matrix representation calculation is to obtain symbolic form for block matrices $\mathbf{M}^{1k}$, $k \geq 2$. For this one can investigate systems with time dependent nonlinear part. Indeed, if we know symbolic form of the matricant then we can calculate symbolic form for $\mathbf{M}^{1k}$ using an integrator procedure. For this purpose we can use either built in integrator packa›es for computer algebra codes or generated integrator packages. In the last case one can sa›e computer resources. We should mention that necessary calculations for every block ma›trix are performed only once. The corresponding results must be included into two types of databases. The first type consists of symbolic forms of matrices $\mathbf{M}^{1k}$ for analytical investigations on early steps of the modelling process. The second type of databases is a ›et of procedures (in one of high level programming languages) for numerical modelling. De›ending on the types of matricant representations and time dependence of nonlinear terms one use corresponding integrator procedures.

3.3 The Structure of Databases and Knowledge Bases

Here we must point out the block structure of the matrix representation, see the Eq.(2). But there is a subblock structure for the matrices $\mathbf{M}^{1k}$. This structure is defined from symmetry pro›erties of the system under study. The knowledge of these symmetries allows to create a hier›rchy of calculation procedures for different kind of objects. For example the

manipulation with noncommutative variables in the frame of Lie algebra descriptions for our objects (CBH and Zassenhause formulae, their analogies and so on). This approach allows to create more economical calculation procedures and databases. For this one can create several levels of databases. Every level corresponds to special set of objects: noncommutative abstract variables, block matrices and block vectors (without knowledge of their inner structures) and at last the databases of block matrices which correspond to some specific effects (spherical or chromatic aberrations) or stirring elements (dipole magnet and so on). If these databases are supplied with corresponding choosing rules and comments (help files) then we obtain a set of knowledge bases. In particular, these correspondence can be obtained by using oriented graphs methods. This approach allows to use object–oriented projecting process on all levels of the modelling process [7].

3.4 The Basic Advantages of the Matrix Representations

The use of the matrix representation for transfer maps allows to apply the power of computer algebra methods and codes more completely. Indeed, we can point out the following moments of such approach.

1). All calculation have several levels. Each level can be used separately and independently. For example: the level of manipulating by non-commutative variables of different nature and the level of calculation of Kronecker products and sums. One can generate corresponding databases for every level and fill up them independently.

2). All objects of databases have similar nature – they are matrices.

3). All operations are algebraic, necessary integration operation are realized in some set of function and can be done without using integrator module of a computer algebra codes.

4). Block structures of matrix representations of Lie transformation allow to work with matrices of low dimensions.

5). Every elemental matrix can be compared some physical effect. For using of such comparision one can build oriented graphs. These graphs help to create knowledge bases.

6). A user of such approach has a possibility to fill up his databases step-by-step in different ways, using his own calculations and results from different sources.

6). Symbolic forms of results permit to make corresponding numerical calculation (for modelling and/or optimization of beam dynamics) more effective and flexible. Indeed manipulation by matrix objects (corresponding to any physical effect) is more simple than by another objects.

4 Conclusion

In the conclusion we want to do some remarks. At first, an application of methods and codes computer algebra allows to go to another level of modelling – dynamic modeling. Indeed, in this case symbolic form of all necessary for solving process information leads to a possibility of more flexible parameters manipulation. Every "physical" object has its mathematical image in the form of a matrix. Such matrices are *bricks* of corresponding databases and knowledge bases. Secondly, there appears an opportunity to investigate qualitative properties using well designed symmetry packages which are built into the

most of computer algebra codes. The matrix structure of the representation of dynamic systems information allows to use parallel processing both in symbolic and numerical modes.

For practical goals as a kern of such approach we use the *REDUCE* codes. These codes are compact and sufficient. For different tests we used the *MAPLE* codes. But to our opinion these codes are excess for our problems and they can be used for helping to *REDUCE*.

References

[1] S.N.Andrianov, Abstrs. of the Int. Congress on Computer Systems and Applied Mathematics, CSAM'93, St.Petersburg, July 19-23, 1993, St.Petersburg, 14, (1993).

[2] S.N.Andrianov, Mechanics and Control Problems, issue 15, The Theory of Control Systems, St.Petersburg University, St.Petersburg, 7, (1995), in Russian.

[3] A.J.Dragt Physics of High Energy Particle Accelerators, eds. by R.A.Carrigan, F.R.Husor, M.Month, AIP Conf. Proc., **87**, 147, (1982).

[4] S.N.Andrianov , Proc. of the Second Int. Workshop Beam Dynamics & Optimization BDO'95, July 4–8 1995, St.Petersburg (Russia), St.Petersburg, SPSU, 16 (1995).

[5] S.N.Andrianov , Proc. of the First Int. Workshop Beam Dynamics & Optimization BDO'94, July 4–8 1994, St.Petersburg (Russia), St.Petersburg, SPSU, 19 (1995).

[6] T.A.Antonie, A.A.Al–Maaitah, J.of Math.Phys., **33**, 331, (1992).

[7] S.N.Andrianov , Abstrs. of the Int. Conference on Interval and Computer Algebra Methods in Science and Engineering, INTERVAL'94, March 7–10, 1994, St.Petersburg, Russia, St.Petersburg, 41, (1994).

Calculations of contrast structures in singularly perturbed equations and systems [1]

Anikeeva V.A.

IHEP, Protvino, RUSSIA
e-mail: anikeeva@mx.ihep.su

Vasil'eva A.B., Plotnikov A.A.

MSU, Moscow, RUSSIA
e-mail: abvas@mathabv.phys.msu.su

Abstract. Some techniques and examples of the numerical calculation of n.s.p.e. (nonlinear singularly perturbed equations) solutions with *interior layer*, called *contrast structure*, are proposed. These techniques are based on the combination of numerical and asymptotical methods. The latter ones allow to make up the good qualitative comprehension of the solution being searched and to choose the most effective numerical method and suitable initial approximation for iterations. The results are represented by graphs.

In this paper we discuss solutions of boundary problems for singularly perturbed equations and systems of equations. Equation is called *singularly perturbed* if it contains small parameter in the highest derivative. As a rule the solution rapidly changes near the boundary in the form of *boundary layer*. Solutions containing *interior layer* can appear for nonlinear equation (so called *contrast structures* - c.s.). There are two main types of c.s.:

1). step-type c.s. (fig. 1)

2). spike-type c.s. (fig. 2)

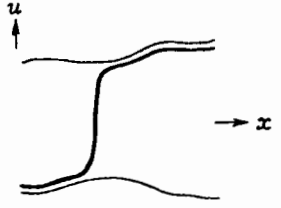


Figure 1:

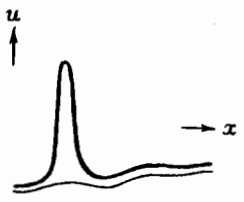


Figure 2:

Asymptotical theory of c.s., i.e. finding the expansion of the solution with respect to the small parameter is now under intensive developement ([1, 2]).

Last years numerical methods are often used, they play double role:

1) they illustrate the analytical results,

2) they describe expected phenomena which are partially or not at all investigated in analytical manner.

[1] supported in part by the RFFI (Grant No.96-01-00694)

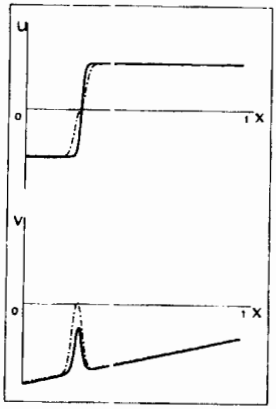


Figure 3:

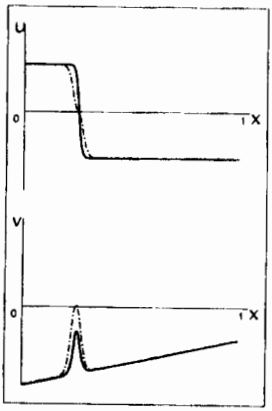


Figure 4:

Numerica integration of singularly perturbed equations, particularly for boundary problems meets specific difficulties – standard routines do not work in the proper manner because of the transition layer exists. So one has someway take into account the specifics of asymptotics.

Application of asymptotical theory for numerical solution of boundary problems for singularly perturbed equations and systems of equations was used earlier (see for instance [3, 4]).

The problem with contrast structures is more complicated because one does not know a priori the point where interior transition layer occurs. Some information can be provided only on the basis of asymptotical theory. This theory prompts the succesfull choice of initial approximation for the solution being searched for.

Following problems were considered.

I. The system

$$\begin{aligned} \varepsilon^2 u'' &= F(u, v, x, \varepsilon) \\ \varepsilon^2 v'' &= G(u, v, x, \varepsilon) \end{aligned}$$

is not still investigated in the sence of appearance of c.s. with the interior layer in both u and v components at the same point.

We proved the existence and found the asymptotic of c.s. for two specific forms of F and G.

As illustration plots are given in figs. 3,4 for the computated c.s. of the equations system

$$\begin{aligned} \varepsilon^2 u'' &= (u^2 - 1)(u - 0.5x + 0.125) - 2\varepsilon uv \\ \varepsilon^2 v'' &= v - (1.7 - x)u^2 + \varepsilon u^2 v^2 \end{aligned}$$

under boundary conditions of II kind

$$\begin{aligned} u'(0) &= u'(1) = 0 \\ v'(0) &= v'(1) = 0. \end{aligned}$$

The right side of the last system was chosen in the special form, so that u-component has a step, while v-component has a spike. In agreement with boundary conditions the boundary layer was absent for the zero approximation in points $x = 0$ and $x = 1$. The small parameter was $\varepsilon = 0.01$.

The solution of this system was calculated as stationary solution of parabolic system

$$\begin{aligned} -u_t + \varepsilon^2 u_{xx} &= (u^2 - 1)(u - 0.5x + 0.125) - 2\varepsilon uv \\ -v_t + \varepsilon^2 v_{xx} &= v - (1.7 - x)u^2 + \varepsilon u^2 v^2 \end{aligned}$$

under boundary conditions

$$u_x(0,t) = u_x(1,t) = 0, \qquad v_x(0,t) = v_x(1,t) = 0$$

and some initial conditions (one can see them on the plots as a dotted line).

The Krank-Nicolson schema, quasilinearization and "progonka" were used. Transition from layer to layer in the variable t was performed when the difference between two sequential iterations in each component was less than 10^{-5}. The maximal number of iterations was 3 and occured on the first layer. The solution was considered stationary when its variation was less then 10^{-5} at least on 20 000 layers.

As example of the system of more general form we considered the system

$$\begin{aligned} \varepsilon^2 u'' &= (u^2 - 1)(u - 0.5x + 0.125) + 0.9v \\ \varepsilon^2 v'' &= (v^2 - 1)(v - 0.5x + 0.25) + 4.5u \end{aligned}$$

under boundary conditions of I kind

$$\begin{aligned} u(0) &= u(1) = 0 \\ v(0) &= v(1) = 0. \end{aligned}$$

Computations were done in the same manner, some c.s. are given in the fig. 5,6. Here, as it must be in agreement with boundary conditions of I kind the boundary layer exists in points $x = 0, \quad x = 1$. The dependence of step-type c.s. in both components from initial conditions is evident. Worth to be noted that transition in both u- and v- components takes place in the same point in x (with precision $5 \cdot 10^{-3}$).

II. The parabolic equation is considered in the domain $D = \{(0 < x < 1) \times (t > 0)\}$

$$\varepsilon u_t - \varepsilon^2 u_{xx} = f(u, x, t)$$

under boundary $u|_{x=0} = u|_{x=1} = 0$ and initial conditions $u|_{t=0} = \sin \pi x$, f is a periodical function in t.

This problem can have the solution both step-type and spike-type. The step (to be definite we discuss step) can change the place when t changes. Situations when step-type c.s. survives as a step when t changes were asymptotically investigated. But there are situations when there appears one or more steps or step disappears at all, as well exist

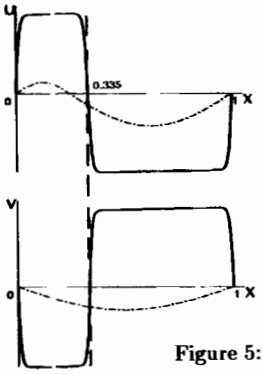


Figure 5:

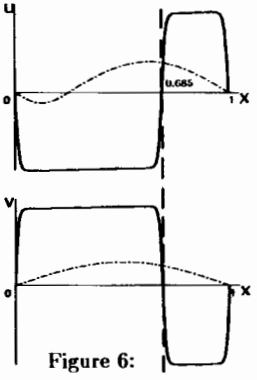


Figure 6:

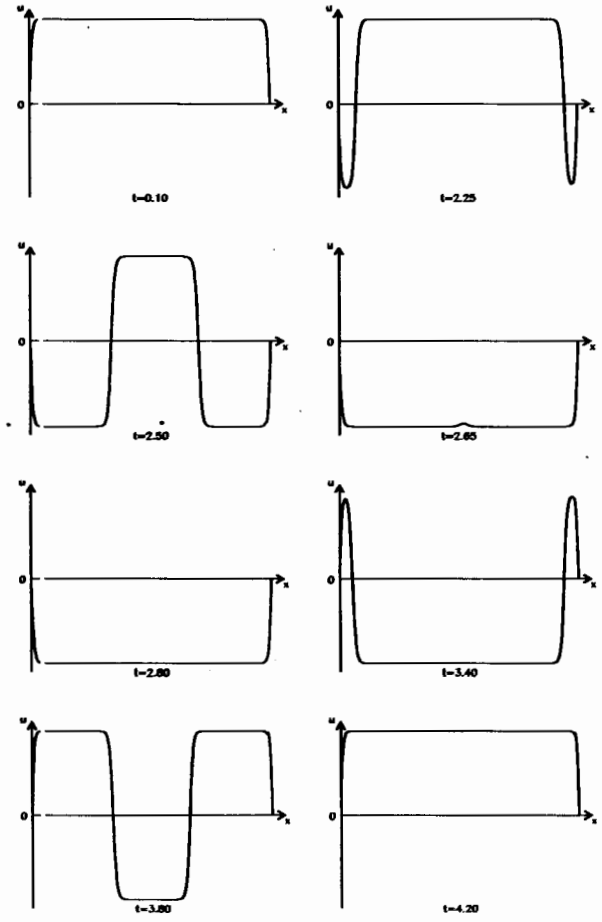


Figure 7:

situations when the spike appears. They call these c.s. *changing structures*. They are yet not studied analytically.

You can see the results of computations performed by A.A. Plotnikov for

$$f(u,x,t) = (1-u^2)(u - 0.5\sin\pi x + 0.73\sin 2t + 0.25), \quad \varepsilon = 0.01.$$

in the fig. 7. for some values of t. The quasilinearization method was used here too.

References

[1] A.B. Vasil'eva, V.F. Butuzov. Asymptotic Methods in the Theory of S ngular Perturbation. Moscow. Vysshaja Shkola. 1990.

[2] A.B. Vasil'eva, V.F. Butuzov, L.V. Kalachev. The Boundary Function Method for Singular Perturbation Problems. SIAM. Phyladelphia. 1995.

[3] A.M. Il'in. Finite differences schema for the differential equation conta ning the small parameter with the highest derivative. Matematicheskije zametki, **6**, vyp. 2. (1969). 237-248.

[4] I.V. Amirkhanov, E.P. Zhidkov, I.E. Zhidkova. The singular perturba ed differential equations system investigation with the application of numerical-analytical methods. JINR preprint R11-93-350.

Taking into account the finite statistics of Monte Carlo simulation for parameters estimation from experimental data [1]

Anykeyev V.B, Popov A.V., Zhigunov V.P.

IHEP, Protvino, RUSSIA, 142284
e-mail: anikeev@mx.ihep.su

Abstract. One of the most important problems in modern high energy physics experiments on measuring cross sections is the problem of parametric analysis, which includes testing the hypothesis about agreement between some parametric model and experimental data. In this problem we must take into account statistical fluctuations in Monte-Carlo sample, that is used for correcting experimental data by limited acceptance and finite resolution of the detector. These fluctuations lead to the distortions of the final distribution used in χ^2 criterion (because of these fluctuations, the final distribution, used in the criterion, differs from the χ^2 distribution). In this report we propose the method to take into account these fluctuations and, as a consequence, correction of the final distribution, that allows to use the χ^2 criterion correctly in the problem of the hypothesis testing. Our numerical experiments show that this method provides significant improvement of the final distribution and it's proximity to the corresponding true χ^2 distribution.

1 Model of measurements [1,2,3]

We consider $s(x) = \frac{1}{\sigma_{tot}} \frac{d\sigma}{dx}$ - p.d.f. for some differential cross section.

$L = N_{tot}/\sigma_{tot}$ - integral luminosity of experiment.

Expected number of events in histogram bins will be

$$\overline{Y}_n = L \int_{x'_n}^{x'_{n+1}} dx' \tilde{E}(x') \int dx E(x) R(x'|x) \frac{d\sigma}{dx}.$$

Introducing

$$S(x) = L d\sigma/dx, \quad A_n(x) = \int_{x'_n}^{x'_{n+1}} dx' \tilde{E}(x') R(x'|x) E(x),$$

we can rewrite $\overline{Y}_n$

$$\overline{Y}_n = \int dx A_n(x) S(x), \quad n = 1, ..., N.$$

We introduce smearing matrix

$$A_{nm} = \frac{\int_{x_m}^{x_{m+1}} dx A_n(x) S(x)}{\int_{x_m}^{x_{m+1}} dx S(x)}.$$

Model of measurements can be rewritten

$$Y_n = \overline{Y}_n + \varepsilon_n = \sum_{m=1}^{M} A_{nm} f_m + \varepsilon_n,$$

[1] supported by the RFFI grant #96-02-16495

where $f_m = \int_{x_m}^{x_{m+1}} dx S(x)$ – expected number of events in interval $[x_m, x_{m+1}]$ corresponding true kinematical variable x; ε_n – fluctuation of events number Y_n in n-th bin of experimental histogram.

Apparatus functions, i.e. rows of smearing matrix (2D-histogram $M_{true} \times N_{observed}$) can be obtained from Monte Carlo simulation of the detector.

$$\hat{A}_{nm} = S_{nm}/I_m; \quad n = 1, ..., N; \quad m = 1, ..., M;$$

where S_{nm} – contents of n, m-th bin of 2D histogram; I_m – contents of m-th bin of input vector (uniform high statistics vector preferred, but it depends on the variable and simulation program).

2 Parameter estimation

Our aim is to estimate $S(x)$. Suppose we know functional form of $S(x)$: $S(x|\vec{a})$ and want to estimate unknown parameters $\vec{a}$. To estimate unknown parameters and to test the hypothesis that chosen model corresponds to experimental data we use the functional:

$$F = \sum_{n,l=1}^{N} \left[Y_n - \hat{\overline{Y}}_n(\vec{a})\right] K_{nl}^{-1} \left[Y_l - \hat{\overline{Y}}_l(\vec{a})\right],$$

where K_{nl}^{-1} is the element of weight matrix for this functional,

$$\hat{\overline{Y}}_n(\vec{a}) = \sum_{m=1}^{M} \hat{A}_{nm} f_m(\vec{a}),$$

$$f_m(\vec{a}) = \int_{x_m}^{x_{m+1}} dx S(x|\vec{a}).$$

First we find estimate of parameters $\hat{\vec{a}}$ from minimization of the functional F, then we use obtained minimal value of the functional

$$F_{min} = \sum_{n,l=1}^{N} \left[Y_n - \hat{\overline{Y}}_n\left(\hat{\vec{a}}\right)\right] K_{nl}^{-1} \left[Y_l - \hat{\overline{Y}}_l\left(\hat{\vec{a}}\right)\right]$$

to test the hypothesis if the chosen model corresponds to experimental data.

Usually elements of covariation matrix

$$K_{nl} = \overline{\left(Y_n - \overline{Y}_n\right)\left(Y_l - \overline{Y}_l\right)},$$

supposed to be

$$K_{nl} \approx \delta_{nl} Y_n, \qquad \text{or}$$
$$K_{nl} \approx \begin{cases} Y_n(1-\hat{p}_n), & \text{where } \hat{p}_n = Y_n / \left(\sum_{i=1}^{N} Y_i\right), \quad \text{for } n = l; \\ -Y_n \hat{p}_l, & \text{for } n \neq l. \end{cases}$$

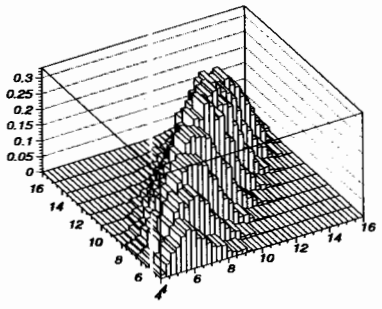


Figure 1: Smearing matrix for 40 000 MC events.

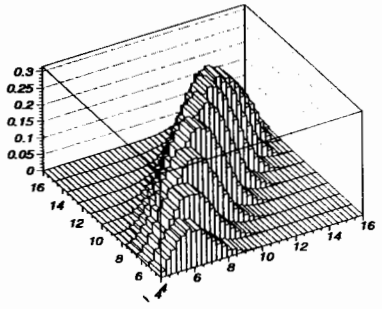


Figure 2: Smearing matrix for 400 000 MC events.

and the value F_{min} to be distributed in accordance with χ^2 distribution with $N - r - 1$ degrees of freedom, where r is the number of estimated parameters if the model adopted fits experimental data.

But the problem is that we use *estimates* of measurements instead of expectations. We also use *estimates* of elements of smearing matrix, obtained from simulation. Sometimes the statistics of Monte Carlo simulation is limited or is comparable with the statistics of experiment. You can see the role of statistical fluctuations from the figures (Monte Carlo statistics differs in one order of magnitude: 40 000 and 400 000 events). We use here the standard "camel" numerical test.

As the result the distribution of the functional differs from χ^2_{N-r-1} and we cannot use this statistics to test hypothesis if our model (chosen parameters) corresponds to experimental data. We found this problem in DELPHI experiment at CERN [6, 7].

We write the expression for the matrix inverse to weight matrix:

$$K_{nl} = \overline{\left[Y_n - \hat{\overline{Y}}_n(\vec{a})\right]\left[Y_l - \hat{\overline{Y}}_l(\vec{a})\right]},$$

and we expect that functional will follow χ^2 distribution when the model fits experimental data. The number of degrees of freedom will be discussed later.

After the simple transformation

$$\left[Y_n - \hat{\overline{Y}}_n(\vec{a})\right] = \left[Y_n - \overline{Y}_n\right] + \left[\overline{Y}_n - \hat{\overline{Y}}_n(\vec{a})\right].$$

1-st term describes experimental fluctuations, second – Monte Carlo fluctuations. As they are independent

$$K_{nl} = \overline{\left[Y_n - \hat{\overline{Y}}_n(\vec{a})\right]\left[Y_l - \hat{\overline{Y}}_l(\vec{a})\right]} = \overline{\left[Y_n - \overline{Y}_n\right]\left[Y_l - \overline{Y}_l\right]} + \overline{\left[\overline{Y}_n - \hat{\overline{Y}}_n(\vec{a})\right]\left[\overline{Y}_l - \hat{\overline{Y}}_l(\vec{a})\right]}.$$

The estimate of the second term is

$$\overline{Y}_n - \hat{\overline{Y}}_n(\vec{a}) = \sum_{m=1}^{M} A_{nm} f_m(\vec{a}) - \sum_{m=1}^{M} \hat{A}_{nm} f_m(\vec{a})$$

We divide and multiply by I_m (contents of m-th bin of input vector) in the assumption $I_m \approx \overline{I}_m$.

$$\bar{Y}_n - \hat{\overline{Y}}_n(\vec{a}) \approx \sum_{m=1}^{M} \left(A_{nm}\overline{I}_m - \hat{A}_{nm} I_m \right) \frac{f_m(\vec{a})}{\overline{I}_m}.$$

From the definition of smearing matrix

$$A_{nm}\overline{I}_m = \overline{S}_{nm},$$

$$\hat{A}_{nm} I_m = S_{nm}.$$

Finally

$$\overline{Y}_n - \hat{\overline{Y}}_n(\vec{a}) \approx \sum_{m=1}^{M} \left(\overline{S}_{nm} - S_{nm} \right) \frac{f_m(\vec{a})}{\overline{I}_m}.$$

$S_{nm}, n = 1, ..., N$ for fixed m is the 1-D histogram, the slice of 2-D histogram used to get the smearing matrix in x' axe. Suppose the number of bins is big enough, neglecting bin-bin correlations we can rewrite the second term:

$$\overline{\left[\overline{Y}_n - \hat{\overline{Y}}_n(\vec{a})\right]\left[\overline{Y}_l - \hat{\overline{Y}}_l(\vec{a})\right]} \approx \sum_{m=1}^{M} \overline{\left[S_{nm} - \overline{S}_{nm}\right]\left[S_{lm} - \overline{S}_{lm}\right]} \left[\frac{f_m(\vec{a})}{\overline{I}_m}\right]^2 \approx \sum_{m=1}^{M} \overline{\left[S_{nm} - \overline{S}_{nm}\right]\left[S_{lm} - \overline{S}_{lm}\right]} \left[\frac{f_m(\vec{a})}{I_m}\right]^2.$$

As it was postulated before S_{nm} for fixed m is 1-D slice in x' of 2-D histogram. Though we have a fixed number of events in input vector, the number of events in each slice S_{nm} for fixed m $S_m^{tot} = \sum_{n=1}^{N} S_{nm}$ will be the random value. **Two approaches** can be considered (see [4]): We can consider S_m^{tot} as fixed value, so we consider the number of events in bin S_{nm}, $n = 1, ..., N$ $m = const$ **as multinomial**. We obtain

$$\overline{\left[S_{nm} - \overline{S}_{nm}\right]\left[S_{lm} - \overline{S}_{lm}\right]} \approx \begin{cases} S_{nm}\left(1 - \hat{P}_{nm}\right), & \text{for } l = n; \\ -S_{nm}\hat{P}_{lm}, & \text{for } l \neq n \end{cases}$$

where

$$\hat{P}_{nm} = S_{nm}/S_m^{tot}.$$

We can consider S_m^{tot} as the random value and consider the distribution S_{nm} as **Poisson**:

$$\overline{\left[S_{nm} - \overline{S}_{nm}\right]\left[S_{lm} - \overline{S}_{lm}\right]} \approx \delta_{nl} S_{nm}.$$

Finally for the **first, multinomial** approach full covariation matrix (experimental and Monte Carlo errors included) will be

$$K_{nl} \approx \begin{cases} Y_n(1 - \hat{p}_n) + \sum_{m=1}^{M} S_{nm}\left(1 - \hat{P}_{nm}\right)\left[\frac{f_m(\vec{a})}{I_m}\right]^2, & \text{for } n = l; \\ -Y_n\hat{p}_l - \sum_{m=1}^{M} S_{nm}\hat{P}_{lm}\left[\frac{f_m(\vec{a})}{I_m}\right]^2, & \text{for } n \neq l. \end{cases}$$

where $\hat{p}_n = Y_n / \left(\sum_{i=1}^{N} Y_i\right)$. It is obvious that the distribution of the functional is χ^2_{N-r-1} if the model describes experimental data in the correct way (r is the number of parameters $\vec{a}$).

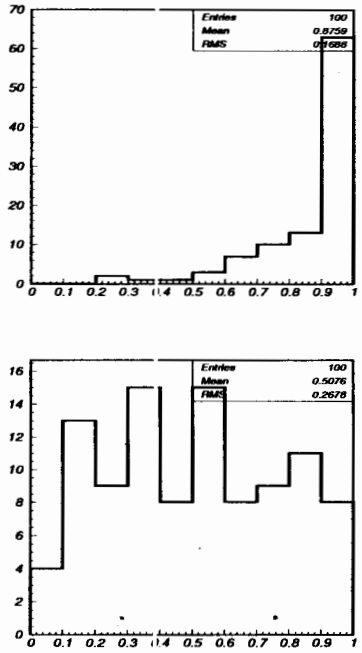


Figure 3: Distribution of Φ neglecting fluctuations in the smearing matrix (top histogram) and taking them into account (Poisson approach, bottom) for 40 000 Monte Carlo events.

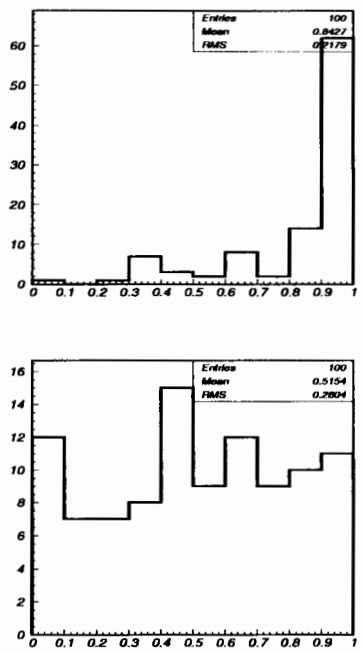
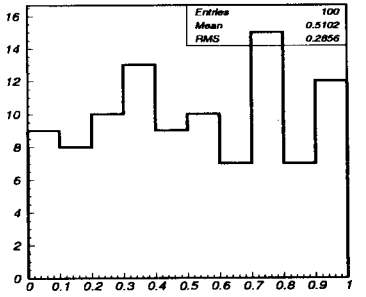


Figure 4: Distribution of Φ neglecting fluctuations in the smearing matrix (top histogram) and taking them into account (multinomial approach, bottom) for 40 000 Monte Carlo events.



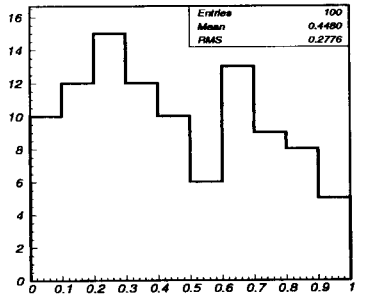
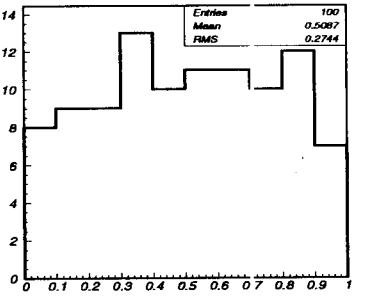


Figure 5: Distribution of Φ neglecting fluctuations in the smearing matrix (top histogram) and taking them into account (Poisson approach, bottom) for 400 000 Monte Carlo events.



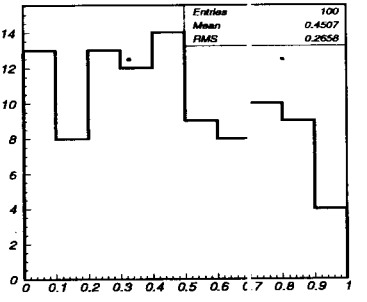


Figure 6: Distribution of Φ neglecting fluctuations in the smearing matrix (top histogram) and taking them into account (multinomial approach, bottom) for 400 000 Monte Carlo events.

For the **second, Poisson** approach

$$K_{nl} \approx \delta_{nl} \left(Y_n + \sum_{m=1}^{M} S_{nm} \left[\frac{f_m(\vec{a})}{I_m} \right]^2 \right).$$

The functional will follow χ^2_{N-r} distribution.

Numerical experiment was [1, 2, 3]:

$$S(x) = C \left[A_1 \frac{G_1^2}{(x-x_1)^2 + G_1^2} + A_2 \frac{G_2^2}{(x-x_2)^2 + G_2^2} \right],$$

$$E(x) = 1 - \left[\frac{(x-x_1)}{x_2 - x_1 + 2G_2} \right]^2, \quad \tilde{E}(x') = 1,$$

$$R(x'|x) = N(x'|x, \sigma),$$

where $x \in [4, 16]$, $A_1 = 2$, $A_2 = 1$, $G_1 = G_2 = 1$, $x_1 = 10$, $x_2 = 14$, $\sigma = 1.5$, C from the condition $\int dx S(x) = 4 \cdot 10^4$ (number of events in experiment). As unknown parameters of the model we chose amplitudes A_1, A_2 and peaks position x_1, x_2. Input vectors for Monte Carlo were uniform with $4 \cdot 10^4$ events (equal to experimental statistics) and $4 \cdot 10^5$. Minimization was done with MINUIT [5]. In figures one can see the distribution

$$\Phi = \int_0^{F_{min}} d\chi_l^2 \, \varphi(\chi_l^2),$$

where $\varphi(\chi_l^2)$ – p.d.f. of χ^2 distribution with l degrees of freedom ($l = N - r - 1$ for multinomial approach and $l = N - r$ for Poisson). It is evident that if F_{min} really follows χ_l^2, the distribution of Φ will be uniform. In our experiment $N = 10$, $r = 4$, $l = 5$ or 6. ¿From Figures it is evident that when Monte Carlo statistics volume is close to the experimental, one should take into account fluctuations in the smearing matrix; this makes possible to use χ^2 criterion for hypothesis testing.

It also must be noted that estimated parameters obtained from for all three approaches are very close.

References

[1] Zhigunov V.P. *Nucl. Instr. and Meth.*, 1983, **216**, 189

[2] Anykeyev V.B., Spiridonov A.A., Zhigunov V.P. *Nucl. Instr. and Meth.*, 1991, **A303**, 350

[3] V.B. Anikeev, V.P. Zhigunov, *Fiz. Elem. Chastits At. Yadra* 1993, **24**, 989-1055

[4] W.T.Eadie at al., *Statistical Methods in Experimental Physics*, North Holland, 1971.

[5] F.James and M.Roos. CERN Program Library D506. *MINUIT – Function Minimisation and Error Analysis*, 1988.

[6] Abreu.P at al., – CERN-PPE-94-130, 1994

[7] Abreu.P at al., – CERN-PPE-95-130, 1995

THE METHOD OF CONSIDERATION OF A STRONG SCATTERING ANISOTROPY IN TRANSPORT EQUATION [1]

E.N. Aristova, V.Ya. Gol'din.

Institute for Mathematical Modelling RAS,
Miusskay sq. 4-A, Moscow 125047, Russia
e-mail: goldin#6@imamod.msk.su, aristova#6@imamod.msk.su

Abstract. A new method is proposed to consider the strong scattering anisotropy. The method makes possible to take into account some peculiarities and ruptures in the scattering indicatrix as a result of a change of variables into the scattering integral. The presence of δ-pecularity may be naturally taken into account in the method framework. The method permits to reduce the integral to the Stieltjes form that is more convenient for real calculations. The quasi-diffusion system, which includes three first terms of the Legendre polinomials decomposition of the indicatrix and term corresponding δ-pecularity, was written. Boundary conditions, which take into account the reflecting anisotropy, were written. The convergency rate of iterations was compared with other methods.

1 Introduction

Let us consider a steady equation of radiation or particle transport in 1D plane geometry for one group with regard to scattering anisotropy [1] in usial notations:

$$\mu\frac{\partial\varphi}{\partial z}+\sigma\varphi=\frac{\sigma_s}{2\pi}\int\limits_{-1}^{1}\int\limits_{0}^{2\pi}w(\mu_0)d\chi\,\varphi(z,\mu')d\mu'+q(z,\mu). \tag{1}$$

Here $\mu=\cos\theta$, μ_0 is a cosine of an angle between the directions of falling and scattering particles $\mu_0=\mu\mu'+\sqrt{(1-\mu^2)(1-\mu'^2)}\cos\chi$, σ is a full cross-section, $\sigma_s<\sigma$ is a scattering coefficient, $w(\mu_0)$ is a scattering indicatrix fulfilling the norming condition:

$$\int\limits_{-1}^{1}w(\mu_0)d\mu_0=1. \tag{2}$$

The boundary conditions with regard to anisotropy of reflection and falling from outside radiation are [1]:

$$\mu\varphi_L(\mu)|_{\mu>0}=\mu f_L(\mu)+\frac{1}{2\pi}\int\limits_{-1}^{0}\int\limits_{0}^{2\pi}\mu' g_L(\mu,\mu',\chi)\,\varphi_L(\mu')d\chi d\mu', \tag{3}$$

$$\mu\varphi_R(\mu)|_{\mu<0}=\mu f_R(\mu)+\frac{1}{2\pi}\int\limits_{0}^{1}\int\limits_{0}^{2\pi}\mu' g_R(\mu,\mu',\chi)\,\varphi_R(\mu')d\chi d\mu'. \tag{4}$$

At strong scattering anisotropy, $w(\mu_0)$ can have a singularity near to δ-function at $\mu_0=1$ as well as to be nearly discontinuous function [1, 2] (ref. Fig.1a made according to data from [2]). These facts make it difficult to calculate an integral in the right hand side of (1), besides, the iteration process of solving integro-differential equation (1) is strongly accomplished.

[1] This work was supported in part by ISTC under Grant No. 115-95

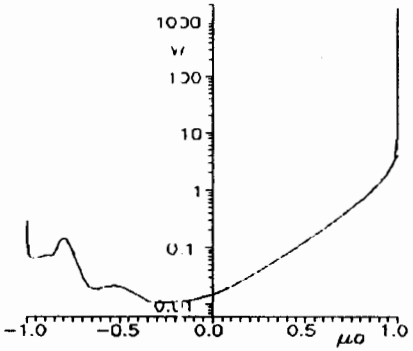


Fig.1a). Scattering indicatrix is reconstracted by data from [2].

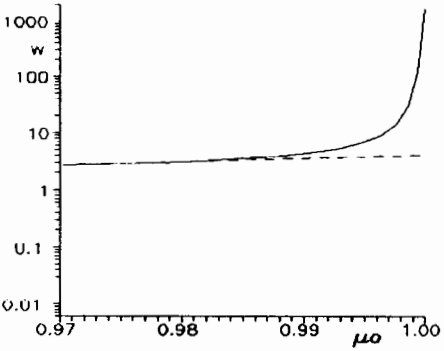


Fig.1б). Separation of δ-pecularity in scattering indicatrix. Smooth part $u(\mu_0)$ is drew by dashed line.

2 Method

It was shown in [4] that at strong scattering anisotropy $w(\mu_0) = (\frac{1+\mu_0}{2})^{10}$ a convergence of quasi-diffusion iterations (for multi-group problem) slightly depends on anisotropy. However the direct integration of remainder of expansion in terms of Legendre polynomials of scattering indicatrix is still expansive in computational respect because of discontinuities and δ-singularity. Let's make the following transformation of equation (1) for obtaining an efficient method.

1) Isolate δ-like singularity in the scattering indicatrix (ref. Fig.1b):

$$w(\mu_0) = u(\mu_0) + \delta_w(\mu_0), \tag{5}$$

where $u(\mu_0)$ is a smooth part of the scattering indicatrix, $\mu_0 \in [-1, 1]$, and $\delta_w(\mu_0)$ is the isolated singularity, $\mu_0 \in [\tilde{\mu}_0, 1]$. Partly expand in terms of Legendre polynomials the function $u(\mu_0)$:

$$u(\mu_0) = \frac{1}{2}u_0 + \frac{3}{2}u_1\mu_0 + \frac{5}{2}u_2\frac{(3\mu_0^2-1)}{2} + \frac{1}{2}\tilde{u}(\mu_0) \tag{6}$$

and transform the equation (1) to the form:

$$\begin{aligned}\mu\frac{\partial\varphi}{\partial z} + \sigma\varphi &= \frac{\sigma_s}{2}\left(u_0\psi + 3u_1 W\mu + 5u_2\frac{(3D-1)(3\mu^2-1)}{4}\psi\right) + \\ &+ \frac{\sigma_s}{2\pi}\int_{-1}^{1}\int_{0}^{2\pi}\{\frac{1}{2}\tilde{u}(\mu_0) + \delta_w(\mu_0)\}d\chi\,\varphi(z,\mu')d\mu' + q,\end{aligned} \tag{7}$$

$$\psi = \int_{-1}^{1}\varphi(z,\mu)d\mu, \qquad W = \int_{-1}^{1}\mu\varphi(z,\mu)d\mu, \qquad D = \int_{-1}^{1}\mu^2\varphi(z,\mu)d\mu \Big/ \int_{-1}^{1}\varphi(z,\mu)d\mu. \tag{8}$$

2) Replace $\delta_w(\mu_0)$ by δ-function with conservation of scattering indicatr x norming (2):

$$\tilde{w}(\mu_0) = u(\mu_0) + (1 - u_0)\,\delta(\mu_0 - 1). \tag{9}$$

$$\text{As} \quad \delta(\mu_0 - 1) = \sum_{n=0}^{\infty} \frac{2n+1}{2} P_n(\mu_0), \quad \text{then :} \tag{10}$$

$$\tilde{w}_n = u_n + 1 - u_0 \qquad n = 1,2,\cdots \tag{11}$$

A comparison of w_n and $\tilde{w}_n$ for the indicatrix shown in Fig.1, ($\tilde{\mu}_0 = 0.98$) gives:

$w_1 = 0.8732 \quad \tilde{w}_1 = 0.8822 \qquad w_2 = 0.8109 \quad \tilde{w}_2 = 0.8183.$

For an integral of δ-function we can write:

$$\frac{\sigma_s}{2\pi} \int_{-1}^{1} \int_{0}^{2\pi} \delta_w(\mu_0) d\chi\, \varphi(z,\mu') d\mu' \simeq \sigma_s (1 - u_0)\, \varphi(z,\mu) \tag{12}$$

and carry the corresponding term in the left side of the equation.

3) In the remaining integral, let us make a change of variables: from χ, μ' pass to variables χ, μ_0. This transformation is not an injection and is carried out according to standard technique for these transformations [10].

An inner integration with respect to angle variable leads to a function of two variables $\Phi(\mu, \mu_0)$, and thus makes it possible to reduce the original integral to Stieltjes integral which is more appropriate for real calculations at presente of singularities of function $\tilde{u}(\mu_0)$:

$$\frac{\sigma_s}{2\pi} \int_{-1}^{1} \varphi(\mu') d\mu' \int_{0}^{2\pi} \frac{1}{2} \tilde{u}(\mu_0) d\chi = \int_{-1}^{1} \Phi(\mu_0, \mu) \tilde{u}(\mu_0) d\mu_0 = \int_{-1}^{1} \Phi(\mu_0, \mu) d\tilde{U}(\mu_0), \tag{13}$$

$$\text{where} \qquad \tilde{U}(\mu_0) = \int_{-1}^{\mu_0} \tilde{u}(\bar{\mu}) d\bar{\mu}. \tag{14}$$

4) Transform the integral of indicatrix smooth part expansion remainder by introduction of linear fractional functional:

$$\eta = \frac{1}{2\pi} \int_{-1}^{1} \int_{-\pi/2}^{3\pi/2} \tilde{u}(\mu_0) d\chi\, \varphi(z,\mu') d\mu' \Big/ \int_{-1}^{1} \varphi(z,\mu') d\mu', \tag{15}$$

the first integral is calculated according to variable change formulae.

As a result of these transformations, the transport equation with the boundary conditions (3,4) is solved with the help of iteration process of the following form:

$$\mu \frac{\partial \varphi^k}{\partial z} + (\sigma - \sigma_s(1 - u_0))\varphi^k = \frac{\sigma_s}{2}(u_0 \psi^{k-1} + 3u_1 W^{k-1}\mu \; + \; +5u_2 \frac{(3D^{k-1} - 1)(3\mu^2 - 1)}{4} \psi^{k-1} + \eta^{k-1}\psi^{k-1}) \; + \; q. \tag{16}$$

This equation solving gives a possibility to find a quasi-diffusion coeficient D^k (8) as well as η^k (15) for the next iteration for the transport equation.

The simple iterations of the form (16) converge slowly. A transition to the quasi-diffusion, when ψ and W are determined independently from the system

$$\frac{dW^k}{dz} + (\sigma - \sigma_s)\psi^k = q_0, \tag{17}$$

$$\frac{dD^k \psi^k}{dz} + (\sigma - \sigma_s \tilde{w}_1) W^k = q_1, \tag{18}$$

leads to a fast convergent iteration process.

The boundary conditions for this system at presence of reflection anisotropy are derived in [10] and have the form (for example, for the right boundary)

$$\frac{W_R - W_{\bar{R}1}^-}{\psi_R - \psi_{\bar{R}1}^-} = \frac{c_R(1 + g_{R2}^+)}{1 + g_{R1}^+ c_R}, \tag{19}$$

where $\psi_{\bar{R}1}^-$, $W_{\bar{R}1}^-$ are defined by falling from outside particles, and c_R, g_{R1}^+, g_{R2}^+ are linear fractional functionals of distribution function of particles with some different weights.

Note that the suggested method is practically without changes applicable for solving transport problems in multi-resonance medium by Lebesgue averaging method [7].

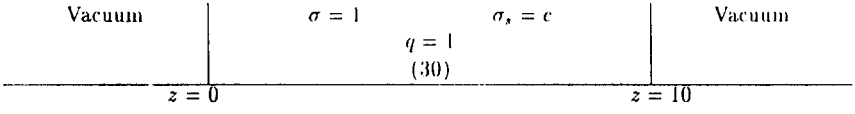
3 Calculation results

We applied a scheme with quasi-analytical solution interpolation for quasi-diffusion equations led to Rhimann invariants. Its application necessity for diffusion and quasi-diffusion equations in the optically thick media was pointed out long ago [5, 8], and a transition to Rhimann invariants was proposed by A.V.Shilkov. Provided equation coefficients constancy inside a difference cell, a solution of quasi-diffusion problem (17)-(18) can be constructed exactly and this solution is put into right hand side of equation (16).

Problem I A problem of anysotropical scattering with Henie- Greenstain indicatrix:

$$w_{H-G}(\mu) = \frac{1}{2}\frac{1-g^2}{(1+g^2-2g\mu)^{3/2}} \qquad w(1) = \frac{1}{2}\frac{1+g}{(1-g)^2}.$$

Here g is a parameter responsible for the sharpness of a pick at $\mu \to 1$. As nearer the value of g to one, then the singularity is stronger. The problem geometry is presented at the following scheme [9]:



We adduce a table of the iteration numbers for different parameters of the problem with the convergence criterion:

$$\max_i \left| \frac{\psi_i^k - \psi_i^{k-1}}{\psi_i^k} \right| < \varepsilon = 10^{-4} \tag{20}$$

		Iteration number						
					$\tilde{\mu}_0$			
g	c	P_1SA^a	N/A^b	$P_1SA-\delta^c$	0.95	0.99	0.999	1.
0.9	0.9	20 ÷ 15	66 ÷ 45		9	11		12
	0.99	66 ÷ 23	392 ÷ 89		12	17		19
0.99	0.9	28	48				7	30
	0.99	75	120	13 ÷ 21			10	101
0.999	0.99			5 ÷ 7			7	

a,b,c) are the data taken from [9]. a) corresponds to completely consistent P_1-synthetic scheme with acceleration; iteration number depends on the choice of quadratures. b)

corresponds to a scheme without iteration acceleration. c) corresponds to a scheme with separation of δ-function. The last four columns are the results of present paper. Here $\tilde{\mu}_0$ is a point of indicatrix rectification according to the derivative in this point for smooth part separation. $\tilde{\mu}_0 = 1$ corresponds to an absence of singularity separation. A case $g \to 1$, $c \to 1$ corresponds to direct particle transit forward without absorption a computation without singularity separation is senseless.

Problem II. A problem of anysotropical scattering with indicatrix shown at Fig.1. The problem geometry is presented at the following scheme:

Vacuum	$\sigma = 3$	$\sigma_s = c\sigma$	Vacuum
	$q = 0$ | $q = 5(z-1)(2-z)$ | $q = 0$		
$z = 0$	$z = 1$	$z = 2$	$z = 3$

Iteration convergence for different parameters of the problem at the convergence criteria (20) is adduced in the table:

ε	10^{-4}	10^{-5}	10^{-7}	10^{-9}
c=0.9	8	10	16	21
c=0.99	11	14	20	27

References

[1] *V.V. Sobolev.* Light Scattering in Planet Atmospheres. ; M., Nauka, 1972, 111 p.

[2] *Dairmendjian D.* Electronics Scattering on Spherical Polidispersions American Elsevier Publishing Company, Inc., New York, 1969.

[3] *V.Ya. Gol'din.* Quasi-Diffusion Method of Solving Kinetic Equation. // USSR Journ. of Comp. Math. and Math. Phys., **4** (1964), N 6, p. 1078-1087.

[4] *D.Yu. Anistratov, V.Ya. Gol'din.* Solving of Multi Group Transfer Equation by Method of Quasi Diffusion . // Preprint of Keldysh Inst. Appl. Math., 1986, N 128.

[5] *V.Ya. Gol'din, A.V. Kolpakov, A.V. Sokolov.* Methods of Solving of Multi Group Transport Equation by Method of Quasi Diffusion . // Preprint of Keldysh Inst. Appl. Math., 1978, N 80.

[6] *V.Ya. Gol'din.* On Mathematical Modeling of Problems of Continuous Medium with Non Equilibrium Transfer. // Modern Problems of Mathematical Physics and Computational Mathematics. M., Nauka, p.113-127.

[7] *A.V. Shilkov.* Generalized Multigroup Approximation and Lebesque Averaging Method in Particle Transport Problems. // J. Transp. Theory and Stat. Ph., **23**, N6, 781 (1994).

[8] *A.A.Samarskii.* Introduction to Theory of Finite Difference Schemes. M., Nauka, 1971, 552 p.

[9] *A.M. Voloschenko.* Completely Consistent P_1 Synthetic Acceleration Sheme for Charged-Particle Transport Calculations. // Proceedings 1996 Topical Meeting Radiation Protection and Shielding, No. Falmouth, Massachusets, April 21-25, 1996, **1**, p.408-415. [Oral report.]

[10] *E.N. Arisova, V.Ya. Gol'din.* Method of Taking into Account Strong Anisotropy of Scattering in Transport Equation. // Mathematical Modeling [in publishing].

Spline approximation and smoothing using the adaptive projective filters

M.V.Avdeev, N.D.Dikoussar

Abstract

A new algorithm for determining the adaptive cubic spline (ACS) which approximates or smoothes functions on a given interval, is proposed. The automatic search for the spline nodes is realized in the algorithm. To search for the spline nodes, the local one-parameter three-point cubic splines are constructed and fitted to a given function using the adaptive projective filter (APF). The boundary points of the intervals, on which the function is approximated sufficiently well by the local splines, are chosen as the nodes of the global spline, which is further determined by the two- or three-point model.

The first version of the library in C++ for working with ACS is presented.

The examples of the treatment of model data are given.

1. Introduction

In spite of a tremendous growth in the theory of spline approximation (interpolation) and function smoothing [e.g. 1,2], the fundamental problem of sampling the spline nodes [3,4] remains to be solved. Up to now, there has been no strict approach either to the choice of the number of the nodes, or their position, and this question is still within the power of the art.

We propose to use the adaptive projective filters (APF) based on the four-point transformation to choose the nodes of the spline, which approximates (interpolats) or smoothes the function. The ideas and mathematical formalism of the four-point transformation were developed in [5,6,7]. Such kind of APFs were used in [6] to find linear and quadratic tracks. Theoretic consideration of the proposed algorithm presents a problem, since every so often one has to check its working capacity for different classes of functions by experimental way.

The object of our work was to create both algorithmic and software basis for investigations of the four-point transformation, specifically as regards the above-mentioned problem.

In the present paper we shall introduce the concept of the four-point transformation in accordance with [6,7]; consider the algorithm for constructing the local adaptive projective filter and the algorithm for constructing the adaptive cubic spline based on this filter; report the first results of its application to the model data; present the library for working with adaptive cubic splines.

2. Four-point transformation

The four-point transformation [6,7] is based on the cross-ratio functions. This transformation is defined by the mark R: $\{(x_0, f_0), (x_1, f_1), (x_2, f_3)\}$, (Fig.1). Le us introduce the following designations:

$$\lambda = x_1 - x_0,\ L = x_2 - x_0,\ \tau = x - x_0,\ f_\lambda \equiv f(x_0 + \lambda),\ f_L \equiv f(x_0 + L). \quad (1)$$

Then, by definition, the direct and inverse four-point transformations are gi ven by:

$$h(\tau) \equiv f^{\triangleleft}(x, R) = (\vec{P}, \vec{F}) = \sum_{i=1}^{3} p_i(\tau, \lambda, L) f_i \quad (2)$$

and

$$f(x) \equiv h^{\triangleright}(\tau, R) = (\vec{D}, \vec{Z}) = \sum_{i=1}^{3} d_i(\tau, \lambda, L) z_i \quad (3)$$

where:

$\vec{F} = (f_\lambda, f_L, f(x))^{\mathrm{T}}$, $\vec{Z} = (f_\lambda, f_L, h(\tau))^{\mathrm{T}}$, $\vec{P} = (p_1, p_2, p_3)^{\mathrm{T}}$ and $\vec{D} = (d_1, d_2, d_3)^{\mathrm{T}}$ are the 3D-vectors of the direct and inverse cross-ratio functions:

$$p_1 = \frac{L\tau}{(\tau - \lambda)(L - \lambda)},\ p_2 = \frac{-\lambda\tau}{(\tau - L)(L - \lambda)},\ p_3 = \frac{\lambda L}{(\tau - \lambda)(\tau - L)} \quad (4)$$

and

$$d_1 = \frac{-\tau(\tau - L)}{\lambda(L - \lambda)}, \qquad d_2 = \frac{\tau(\tau - \lambda)}{L(L - \lambda)}, \qquad d_3 = \frac{(\tau - \lambda)(\tau - L)}{L\lambda} \quad (5)$$

Let us list the peculiarities of the four-point transformation, formin; the basis of the presented algorithm:

1) use of the projective transformation based on the cross-ratio of four coll near points;

2) stability to errors everywhere but for two points;

3) reduction of the polinom power by 2;

4) possibility to parametrize functions using the coordinates of the pivot (experimental) points;

5) comparatively small number of numerical operations.

Items 2) and 4) make it possible to work with experimental data. Items 1) and 3) allow us to set a problem of finding the intervals where the function is described by the cubic parabola to a specified accuracy. The endpoints of the found intervals are used as the nodes of the global cubic spline, which is determined using the standard methods.

3. Local adaptive projective filter for the cubic model

When applied to the cubic parabola:

$$Y_3(x) = \alpha x^3 + \beta x^2 + \gamma x + \delta \quad (6)$$

the four-point transformation gives us a linear function with respect to τ:

$$Y_3^{\alpha}(x;\ R) \equiv h(\tau) = \alpha L\lambda\tau + Y_3(x_0) \tag{7}$$

Let us shift the origin to the $(x_0,\ Y_3(x_0))$ point. Then the function (6) changes to:

$$h(\tau) = \alpha L\lambda\tau \tag{8}$$

and using the inverse transformation (3) Y_3 can be expressed in the new system of coordinates as:

$$Y_3(\tau) \equiv h^{\alpha}(\tau;\ R) = Y_3(\lambda)\ d_1 + Y_3(L)\ d_2 + \alpha\lambda L\tau\ d_3. \tag{9}$$

This expression is the cubic model for APF, which treats a set $\left\{(x_i, \tilde{f}_i)\right\}$ of experimental points in the case of smoothing, or a set of support points in the case of approximation. The filter looks for the interval where the function can be described by cubic parabola with the precision T_v. The functional of the form:

$$\sum [\tilde{f}_i - \hat{\theta}_1 d_1(x_i) - \hat{\theta}_2 d_2(x_i) - \hat{\theta}_3 \lambda L x_i d_3(x_i)]^2 \to \min_{\hat{\theta}} \tag{10}$$

is minimized, where $\hat{\bar{\theta}} = (\hat{\theta}_1, \hat{\theta}_2, \hat{\theta}_3)$ $(\hat{\theta}_3 \equiv \hat{\alpha})$ is the vector of the estimated parameters. When N points are treated, $\hat{\bar{\theta}}$ can be expressed as:

$$\hat{\bar{\theta}}^T = \left(A^T A\right)^{-1} B \tag{11}$$

where:

$$A^T A = \begin{pmatrix} \sum d_1^{\,2}(x_i) & \sum d_1(x_i) d_2(x_i) & \lambda L \sum x_i d_1(x_i) d_3(x_i) \\ \cdots & \sum d_2^{\,2}(x_i) & \lambda L \sum x_i d_2(x_i) d_3(x_i) \\ \cdots & \cdots & (\lambda L)^2 \sum x_i^{\,2} d_3^{\,2}(x_i) \end{pmatrix} \tag{12}$$

and

$$B = \left(\sum_i \tilde{f}_i d_1(x_i),\ \ \sum \tilde{f}_i d_2(x_i),\ \ \lambda L \sum x_i \tilde{f}_i d_3(x_i)\right)^T,\ i = 1,\ 2,\ \ldots,\ N \tag{13}$$

However, we do not treat all N points at once like in the ordinary least-squares method, but use the adaptive algorithm to determine the $\hat{\bar{\theta}}$ vector [6] by the following seqence:

step 1) the initial mark and x_0, λ, L are set;

step 2) the initial values of $\hat{\bar{\theta}}$ are estimated:

$$\hat{\theta}_1 = \tilde{f}_\lambda \quad \hat{\theta}_2 = \tilde{f}_L, \quad \hat{\theta}_3 = \left(\lambda L \sum \tau_i^{\,2}\right)^{-1} \sum \tau_i \tilde{h}_i, \tag{14}$$

where

$$\tilde{h}_i = \tilde{f}_\lambda p_1(\tau_i) + \tilde{f}_L p_2(\tau_i) + \tilde{f}_i p_3(\tau_i) \tag{15}$$

and summation is taken over the points between x_1 and x_2;

step 3) the non-mark point $(x_i, \tilde{f}_i)$ is taken and the prediction is made:

$$\tilde{f}_{pred} = \hat{\theta}_1 d_1(x_i) + \hat{\theta}_2 d_2(x_i) + \hat{\theta}_3 \lambda L x_i d_3(x_i) \qquad (16)$$

step 4) if the prediction is good, *i.e.*:

$$\left|\tilde{f}_i - \tilde{f}_{pred}\right| < T_v, \qquad (17)$$

where T_v is the given threshold, then the point is entered into $\left(A^T A\right)$ and $\hat{A}$ in the required way and the vector of the parameters is corrected according (11).

step 5) move to the next point (*step 3*).

This is continued until we have a number of unsuccessful predictions, which means that the function cannot be further described as a cubic parabola. The last endpoints x_l and x_r, where the obtained cube parabola satisfies (17), are the *boundaries of the desired interval.* An example of smoothing the sine function with an additive noise $\varepsilon = N(0, \sigma)$, $\sigma = 0.15$, is presented in Fig.2. The adaptation of the vector of the parameters is presented in Fig.3. One can see that the parameters become stabilize long before the fitting process is finished. To save the number of numerical operations we can introduce, in principle, some criteria for the stability of the parameters, and hence do not correct $\hat{\theta}$ at all points.

4. Construction of the global spline

Here we give a simplified algorithm for finding the nodes of the global spline (Fig.4):

step 1) set the first mark:

$x_1^{(1)}$ is the left boundary point,

$x_0^{(1)}$ is an arbitrary point at the right, $x_0^{(1)} > x_1^{(1)}$, which is not so far away from $x_1^{(1)}$,

$x_2^{(1)}$ is the point to the right from $x_0^{(1)}$, $x_2^{(1)} > x_0^{(1)}$;

step 2) determine the $x_l^{(1)}$ and $x_r^{(1)}$ points using the local APF;

step 3) set the next mark in the following way:

$x_1^{(2)} = x_0^{(1)}$,

$x_0^{(2)} = x_r^{(1)}$,

$x_2^{(2)}$ - the point at the right from $x_0^{(2)}$, $x_2^{(2)} > x_0^{(2)}$;

step 4) determine the $x_l^{(2)}$ and $x_r^{(2)}$ points using the local adaptive filter;

and so on until you reach the right end of the interval, where the function is defined. Then the $\left\{x_r^{(i)}\right\}$ points and also the $x_1^{(1)}$ and $x_0^{(1)}$ points are used as the nodes for the standard spline interpolation. In fact, the algorithm is more complicated, since we move the basis points so that the $x_r^{(i)}$ point moves to the right as far as possible.

5. Examples.

Let us give an example of sampling nodes when approximating the sine function. The algorithm sets the nodes *uniformly* to make a good spline interpolation. It is interesting that the number of nodes N for the sine function grows as the accuracy of approximation increases as:

$$N \sim -\ln(\varepsilon), \qquad (18).$$

where ε is the respect precision of approximation, $0 < \varepsilon < 1$. This problem as regards other functions is to be studied thoroughly.

The interpolation using the obtained nodes can be performed in the following ways:

- 2-point cube spline with continuous 1st and 2nd derivatives;
- 3-point cube spline with continuous 1st derivative [7];
- 2-point cube spline with continuous 1st derivative and local χ^2 fitting.

The choice of one of these ways depends on an experimenter, since we often deal with the situation when one method works well while the other does not.

The results of the preliminary investigations of the algorithm let us conclude that it works satisfactorily for approximating (interpolating) to any possible accuracy and for smoothing data with additive noise (Fig. 5).

The situation is more complicated for the data with multiplicative errors. We know that the four-point transformation is stable for the additive noise. The extent to which the smoothing by the four-point transformation is affected by the multiplicative errors is being studied at the moment. Preliminarily we can say that the points obtained as the result of the application of the described algorithm to the regression data, define the neighbourhoods where one should look for the nodes of the global spline.

6. Library for working with ACS

The first version of the "acs1" library for working with ACS has been designed using the C++ object-oriented programming language. The library consists of four small libraries:

"discr" - library for working with "discrete" functions, *i.e.* the functions which are defined on a discrete set of points.

"dpt4" - library for working with the four-point transformation;

"apf" - library for working with the adaptive projective filters based on the four-point transformation;

"spline"- library for working with the spline interpolation.

These are the sets of classes using which you could easily realize and modernize the above-described algorithm, as well as investigate the applications of the

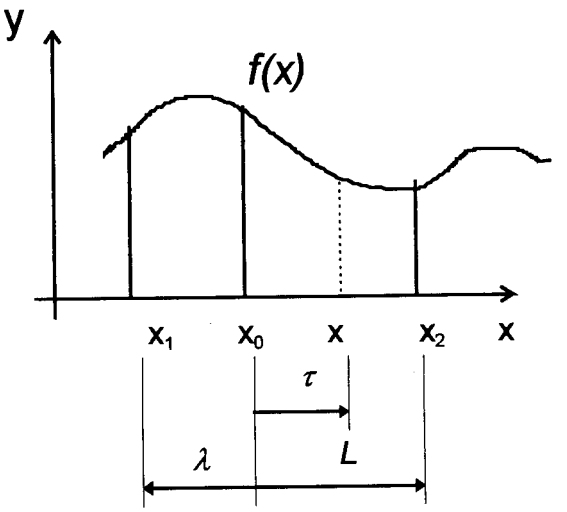


Fig. 1. Mark of the 4-point transformation.

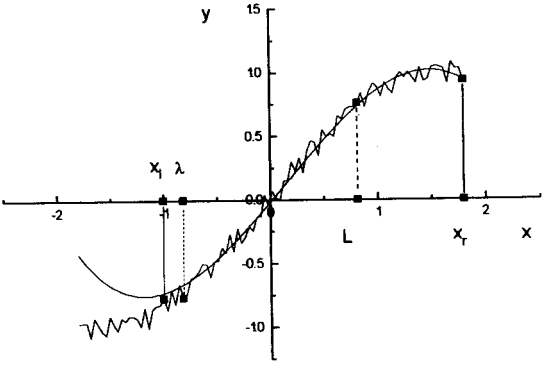


Fig. 2. Local smoothing with the adaptive projective filter for the sine function with additive normal distributed error 0.15.

The starting (λ; L) and the ended (x_l ; x_r) points of the cubic approximation are indicated.

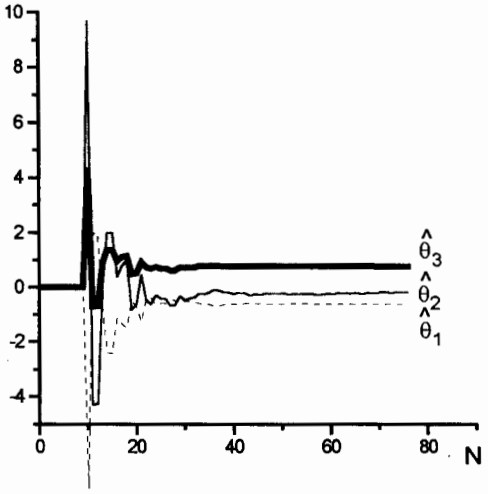


Fig. 3. Adaptation of vector of the parameters when smoothing.

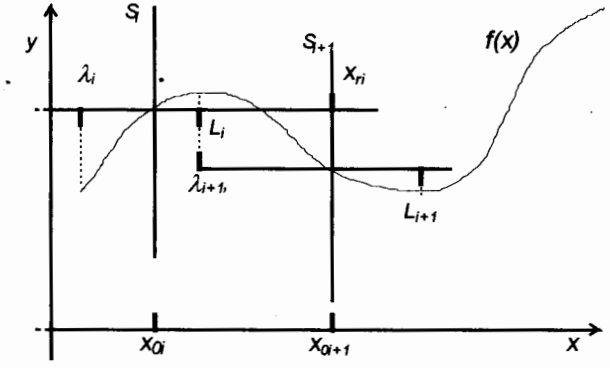


Fig. 4. Scheme of the algorithm for searching the nodes of the global spline.

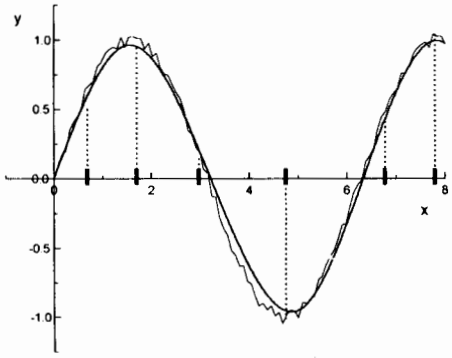


Fig. 5. 3-point ACS for the sine function with additive normal distributed error 0.05.

four-point transformation to different problems. Those who are interested in the library and algorithm, please, contact: avd@nf.jinr.dubna.su or dikoussar@main1.jinr.dubna.su

7. Conclusions

The purpose of the paper is to present the algorithmic and software bases for investigating the four-point transformation, particularly in respect of sampling the nodes for spline approximation and smoothing.

Some peculiarities of the four-point transformation makes it possible to design the adaptive projective filters. The algorithm for constructing the adaptive cube spline is presented. It determines the intervals where a function can be well approximated by the cube parabola. Then, the ends of these intervals are used as the nodes for spline interpolation.

The preliminary investigations of the algorithm let us conclude that it works satisfactorily for approximating to any possible accuracy and for smoothing data with additive normal distributed errors.

When smoothing regressions we can preliminarily say that the points which are the result of the application of the presented algorithm define the neighbourhoods where one should look for the nodes of the global spline.

The first version of the library in the object-oriented C++ programming language for working with ACS is presented. It contains the classes using which you can easily realize and modify the algorithm, as well as investigate the applications of the four-point transformation to different problems.

The approach of the four-point transformation as applied to the problems of approximation, interpolation and smoothing holds much promise.

References

[1] J.H. Ahlberg, E.N. Nilson & J.L. Walsh, The Theory of Splines and their Applications
(Academic Press, New York, 1967).

[2] Y.S. Zavjalov, B.I. Kvasov & V.L. Miroshnichenko, Methods of Spline-functions (Moscow, Nauka, 1980) (in Russian).

[3] G.A.F. Seber, Linear regression analyses (Wiley, 1977).

[4] Svante Wold, Technometrics. 16, No1, (1974) 1.

[5] N.D. Dikoussar, Math. Model. 3 (10) (1991) 50 (in Russian).

[6] N.D. Dikoussar, Comput. Phys. Commun. 79 (1994) 39.

[7] N.D. Dikoussar, JINR. Commun. E10-96-118 (Dubna, JINR, 1996) (to be published in Comput. Phys. Commun.)

SOLITON–LIKE SOLUTIONS OF LINEAR DIFFERENTIAL EQUATIONS

V. S. Barashenkov

Joint Institute for Nuclear Research. Dubna Moscow Region, 141980, Russia
e-mail: barashenkov@lcta30.jinr.dubna.su

L. G. Sapogin

Department of Physics of Technical University (MADI). Leningradsky pr. 64, A-319, Moscow, 125829. Russia
e-mail: sapogin@cnf.madi.msk.su

Abstract. Attention is focused on non-dispersive, form-steady solutions of Klein-Gordon, Schrödinger and homogeneous wave equations which are very perspective from the viewpoint of practical applications. The group velocity of packets made of superposition of such solutions can exceed the plane light wave velocity in vacuum. That seems to be strong contradictory to Einsteinean theory of relativity. The physical meaning of such superluminal solutions is discussed. Though the faster-than-light wave beams predicted by mathematic can be launched not only in gedanken experiments but also in practice, they are incapable of energy and information transfer.

Here is a widely accepted standpoint that solitons (steady wave packets) are peculiar to non-linear equations only. The reason of a spreading of wave solutions of any linear differential equations is a dependence of wave velocity on the wave frequency (wave length). For example, the group velocity of a wave $\exp(ikz - i\omega t)$ where the frequency $\omega = c\sqrt{k^2 + m^2}$

$$v_g = d\omega/dk = c^2 k/\omega \tag{1}$$

diminishes when ω is up.

Nevertheless, using such a spreading wave, one can make wave packets which are solutions to the linear equations are beeng stable non-dispersive objects. Perhaps, H.Bateman was the first who proved [1] a possibility of such unexpected solutions for the homogeneous equation.

$$\left(\nabla^2 - \frac{1}{c_*^2}\frac{\partial^2}{\partial t^2}\right)\Phi = 0, \tag{2}$$

In the well known Stratton's book [2] the cylindrically symmetrical solution to eq. (1) is adduced also developing with no change of its form:

$$\Phi(\rho, z, t) = AJ_o(\alpha\rho)e^{i(\omega t - kz)} \tag{3}$$

where A is a constant coefficient, ω and k are connected by the relation

$$\omega^2/c_*^2 - k^2 = \alpha^2 \tag{4}$$

and the separation constant α defined by boundary conditions takes the part of an effective mass of the packet. Averaging these solutions over an appropriate transmitting transfer function $F(\omega)$, one can make up the soliton-like packet in any desired form.

Practically, of course, we can not launch absolutely steady beams since the solutions of type (3) concern the infinity emitter aperture while in practice it is always finite and creates diffraction effects destroying the developing beam. However though one can say only on quasi-soliton packets, they are rather long-living.

In spite of the remarkable properties of such wave beams they remained known for a long time only for particular specialists. The intensive study began in the mid-80s in connection with expectations to use the soliton-like packets in communication networks, in optical and ultrasound instruments (see papers [3, 4, 5, 6] and especially [7] where one can find the detailed bibliography). The quasi-non-dispersive pulses can be used for essentially more accurate ultrasound location in medicine and technical, for seismic "exploratory drilling", even in quantum mechanics [8].

One should stress that up to now only a small part of pulse shapes has been investigated. It is of specials interest to investigate the particle beams which must possess also non-dispersive properties because after separation of time variable $[\Phi(x.t) = \psi(x)\exp i\omega t]$ Klein-Gordon and even Schrödinger equations have the some form

$$(\nabla^2 - U)\psi = 0. \tag{5}$$

Having in mind its practical significance, we may say that the discovery of a family of the non-dispersive beams is one of the most important achievements of mathematical physics.

One shows now a particular interest to non-dispersive beams with an imaginary separation constant α. From the physical point of view it means that we consider an object with an imaginary mass and superluminal speed $v_g > c_*$. It is especially interesting that such faster-than-light objects can exist in vacuum and have properties of independent space-localized material bodies ("wave torpedo", "purely electromagnetic particles") [5, 7, 8, 9, 10] [1] A modeling of such striking objects by means of acoustic pulses in water confirmed that under appropriate conditions quasi-undistorted wave groups traveling with a superluminal (n the case supersound!) speed $v > c_*$ can be launched indeed [7].

However, as it is known (see, e. g. a reviews [11,12]), the hypothesis of an existence of superluminal bodies (tachyons) with $V > c$ contradicts drastically the causality principle. An appropriate transition of the observer to another co-ordinate frame turns immediately the tachyon receiver into a transmitter sending messages spontaneously, by itself. This conclusion is a strong consequence of Lorentz transformations in the region of the customary subluminal observer's velocities where these transformations are precisely checked. One would have to wait for an appearance of tachyons only in very small space-time intervals where an opposition of the Past and the Future becomes relative or in some exotic events where the theory of relativity and Lorentz formulae are inapplicable. What meaning have then the solutions with $v > c$?

One may maintain that the faster-than-light pulses can be launched in a medium only, the group speed of all non-dispersive beams in vacuum is always less than the plane wave light velocity $c = 3.10^{10}$ cm/c.

Indeed, in an appropriate media, as it is, for example, for acoustic pulses in water [7] the group speed of massive packet [2] v_g can exceed the phase light velocity $c_* = c/n$ where

[1] In the paper [7] the principle possibility of superluminal speeds exceeding the vacuum light velocity ($v > c$) is emphasized even by the paper title

[2] Let us remember that the constant α takes formally the part of the packet mass.

n is the refraction coefficient. It is the well known Cherenkov's effect [3] when the pulse speed v_g remains, however, always less than c.

An energy transfer speed which in the case of electromagnetic waves is defined usually by the relation

$$v_e = |\text{Poyting vector}|/\text{energy} \tag{6}$$

and, as a rule, coincides with the group speed v_g is also less than light velocity c.

The appropriate value of the "mass" α can be obtained by means of suitable chosen boundary conditions using, for instance, a superficially charged cylinder as it is proposed in the Band's papers [9]. Formally, from the viewpoint of the equation (5) and the dispersion relation (4) such an approach is equivalent to a charge of the medium properties (i. e. U and $c_* = c/n$). In this case the pulse speed $v_g = c^2\omega/nk$ can exceed the vacuum light wave velocity c (in this case $n < 1$). Nevertheless, theory of relativity says that the energy transmission speed has again to be less than c.

By way of illustration one may point out the light field created by a string of neon advertisement where each letter flashes irrespective of the preceding one. In that way, by using the independent switches, one can attain a superluminal alteration of superficial charges in the gedanken experiment discussed in the Band's paper [9].

Here we encounter a very interesting circumstance. We do not learn yet to change properties of vacuum, however, one can simulate such alteration with the help of an appropriate choice of the constant α. That allows to launch in vacuum a soliton-like faster-than-light pulse. It would be a real, observable motion — an electromagnetic wave excited by the charged Band's cylinder or by the just mentioned string of advertisement letters, a reflection of a light beam or something analogous. However, in all these cases such a motion can not transfer any energy and information. One can not send Morse code signals or launch tachyon wave "torpedo" [4].

Of course, these conclusions are valid within the bounds of Lorentz transformations and the causality principe only, i. e. they suppose that the cause precedes always their effect. Any superluminal transmitting of energy ($v_e > c$) violates at once these conditions [11, 12]. They may be wrong in not yet investigated regions of ultrasmall space-time intervals or, for example, in extremely strong gravitation fields.

References

[1] H. Bateman. Electrical and Optical Motion. Cambridge Univ. Press. Cambridge,1915.

[2] J. A. Stratton. Electromagnetic Theory. Mc Graw-Hill Co. New York & -London, 1941.

[3] J. Durnin at al. Phys. Rev. Lett. **58** 1499 (1989).

[4] J. N. Brittingham. J. Appl. Phys. **54** (1983) 1179 ().

[3] Such an effect is impossible when $c_* > c$, i.e. if $n < 1$. For the speeds ($c_* < v < c$) instead of superluminal the introduced by W. Band term "superoptic" [9] is more suitable.

[4] In the paper [9] an astonishing result is obtained: the superluminal electromagnetic pulses transfer energy but with subluminal speeds, i. e. the vawe packet rushes forward but its energy remains behind! The paradox is avoided if we take into account that in comparison to a subluminal body in the case of tachyon (would such bodies exist really) the momentum and energy switch the role [?, 12] and in the relation (6) $v_e \to 1/v_e$

[5] R. W. Ziolkowski. *Phys. Rev.* **A39** 2005 1983.
[6] J.-Y. Lu, J.F. Greenleaf. IEEE Trans. Utras. Ferroel. Freq. Contr. **39** 19 (1992) and **39** 441 (1992).
[7] W. A. Rodrigues Jr and J.-Y. Lu. On the existence of undistorted progressive waves (UPWs) of arbitrary speeds $0 < v < \infty$ in nature. BR 12/96 IMECC — UNICAMP, Campinas, 1996.
[8] L.G.Sapogin. Ann. de la Foundation Louis de Broglie. **5** 285 (1980).
[9] W. Band. Found. Phys. **18** 549 (1988) and **18** 625 (1988).
[10] F. O. Barut, H. C. Chandola. Phys. Lett. **A180** 5 (1993).
[11] V. S. Barashenkov. Uspechi Fiz. Nauk. **114** 133 (1974).
[12] V. S. Barashenkov and M. Z. Yur'iev. Hadronic J. **18** 69 (1995).

On the numerical solution of Shrödinger equation

B.Batgerel and T.Zhanlav

Department of Mathematics, National University of Mongolia, P.O. 46/687, Ulaanbaatar 210646, Mongolia
E-mail : numelect@magicnet. mn

Abstract. A family of high order finite difference schemes is developed for the numerical integration of the one-dimensional Schrödinger equation. With the aid of proposed method his been calculated the resonance energy, phase shift and eigenvalue with higher accuracy. Numerical results also indicate that the new methods are much more accurate than the Numerov-type exponentially fitted methods and predictor-corrector methods .

1 Introduction

The numerical solution of the radial Schrödinger equation has attracted continuous interest over many years. The radial Schrödinger equation can be written

$$u'' = -f(r)u, \qquad r \in (0, \infty), \tag{1.1}$$

where $f(r) = E - W(r)$ and $W(r) = l(l+1)/r^2 + V(r)$ is an effective potential and potential $V(r)$ is sufficiently smooth function with satisfies, $V(r) \to 0$ as $r \to \infty$. E is real number denoting the *energy*, l is a nonnegative integer denoting the *orbital moment*.

The boundary conditions are:

$$u(0) = 0 \tag{1.2}$$

and a second boundary condition, for large values of r, determined by physical considerations.

The solution of the equation (1.1) for large values of r has the following asymptotic form:

(1) For positive values of the energy, E,

$$u(r) \sim C\sin(\sqrt{E}r - l\pi/2 + \delta), \qquad r \to \infty, \tag{1.3}$$

where δ, the *phase shift* is to be determined.

(2) For negative values of the energy, E,

$$u(r) \sim C\exp(-\sqrt{-E}r), \qquad r \to \infty. \tag{1.4}$$

In the traditional approach for the construction of finite difference schemes with higher accuracy, only sufficiently smoothness of the solution is assumed, but the whole information from (1.1) namely the behavior of the coefficient is discarded. On the other hand, many problems of practical interest call for very fast, accurate and reliable algorithms for solving such problems. The growing abilities of modern computer algebra systems allow to review the traditional approach in an attempt to construct algorithms which use not only the coefficient of (1.1) but also their derivatives up to the necessary order. Tracking the above idea, we constructed high order difference scheme for solving the Schrödinger equation and applied the obtained scheme to the computing of resonance energy, phase shift and eigenvalue.

2 The difference scheme

We construct an uniform mesh $\omega_h = \{\tau_i = ih, 0 \le i \le N, hN = R\}$ on the interval $[0, R]$. Because of the sufficiently smoothness of $f(r)$, all derivatives $u_i^{(n)}$ $(n \ge 2)$ of the solution can be expressed using u_i, u_i' and the derivatives of f at the point τ_i i. e.

$$u_i^{(2n)} = (-1)^n f_i^n u_i + A_{2n} u_i + B_{2n} u_i', \quad i = 0, 1, ..., N,$$

$$u_i^{(2n+1)} = (-1)^n f_i^n u_i' + A_{2n+1} u_i + B_{2n+1} u_i', \tag{2.1}$$

where some values of coefficients A_n and B_n are given in appendix.

Using the Taylor expansion and presentation (2.1), we obtain

$$u_{i+1} - 2u_i + u_{i-1} = 2p_1 u_i + 2q_1 u_i' + O(h^{2m+2}), \quad i = 1, ..., N-1, \tag{2.2}$$

$$u_{+1} - u_{i-1} = 2p_2 u_i + 2q_2 u_i' + O(h^{2m-1}), \quad i = 1, ..., N-1, \tag{2.3}$$

where

$$p_1 = \psi(f_i, h) + \sum_{n=2}^{m} A_{2n} \frac{h^{2n}}{(2n)!},$$

$$q_1 = \sum_{n=2}^{m} B_{2n} \frac{h^{2n}}{(2n)!}, \tag{2.4}$$

$$p_2 = \sum_{n=1}^{m-2} A_{2n+1} \frac{h^{2n+1}}{(2n+1)!},$$

$$q_2 = \varphi(f_i, h) + \sum_{n=1}^{m-2} B_{2n+1} \frac{h^{2n+1}}{(2n+1)!},$$

$$\psi(f, h) = \begin{cases} \cos\sqrt{f}h - 1 & f \ge 0 \\ \cosh\sqrt{|f|}h - 1 & f < 0 \end{cases} \qquad \varphi(f, h) = \begin{cases} \dfrac{\sin\sqrt{f}h}{\sqrt{f}} & f > 0 \\ \dfrac{\sinh\sqrt{|f|}h}{\sqrt{|f|}} & f < 0 \\ h & f = 0 \end{cases}$$

If eliminate u_i' from (2.2) and (2.3), then we obtain

$$(1 - \frac{q_1}{q_2}) u_{i+} \ - 2(1 + p_1 - p_2 \frac{q_1}{q_2}) u_i + (1 + \frac{q_1}{q_2}) u_{i-1} = O(h^{2m+2}), \quad i = 1, ..., N-1.$$

Thus we obtain the following three-point difference scheme

$$(1 - \frac{q_1}{q_2}) y_{+1} - 2(1 + p_1 - p_2 \frac{q_1}{q_2}) y_i + (1 + \frac{q_1}{q_2}) y_{i-1} = 0, \quad i = 1, ..., N-1. \tag{2.5}$$

which has $2m$-order approximation.

Note that the scheme constructed above gives exact solution of the equation (1.1), when $W(r) \equiv 0$

3 A numerical example

In this section we present some numerical results to illustrate the performance of our method.

1. We consider the one-dimensional oscillator problem

$$u'' + (\lambda - r^2)u = 0, \qquad u(\pm\infty) = 0.$$

Eigenvalue and eigenfunction of problem (3.1) has form

$$\lambda_n = 2n + 1, \qquad u_n(r) = H_n(r)e^{-r^2/2}, \quad n = 0, 1, \ldots$$

where $H_n(r)$ is the Hermite polynomial. For the purpose of our numerical computation we take the domain of integration to be [-9,9]. In the Tab.1 are given absolute errors of fourteenth eigenvalue and given the values

$$\sigma_h = \left| \frac{\lambda^{(h)} - \lambda^{(h/2)}}{\lambda^{(h/2)} - \lambda^{(h/4)}} \right|$$

which must be theoretically equal to $\sigma = 2^{2m}$ for the $2m$-order scheme.

Table.1

h	$m=3$	$m=4$	$m=5$	$m=6$
1/8	8.7E-6	9.9E-8	6.5E-10	2.8E-12
1/16	1.3E-7	3.7E-10	6.0E-13	6.4E-16
1/32	2.0E-9	1.4E-12	5.7E-16	1.7E-18
σ_h	66.2	267.7	1078.5	4336.5
σ	64	256	1024	4096

2. We consider Schrödinger equation with $l = 0$ and with Wood-Saxon potential

$$V(r) = \frac{u_0}{1+t} - \frac{u_0 t}{a(1+t)^2}$$

where $t = \exp((r - r_0)/a), a = 0.6, r_0 = 7.0$ and $u_0 = -50$. As in [1]-[3] the finite domain $0 \le r \le 15$ is chosen for the purpose of illustration.

For positive energies one has the so-called resonance problem. This problem consists either of finding the *phase shift* $\delta = \delta(E)$ or finding those E at which δ equals $\pi/2$.

Table.2 *Absolute errors of the resonance energy in* 10^{-7} *units*

E	h	[1]	[2]	[3]	$m=4$	$m=5$	$m=6$
53.5888719	1/8	987	-	3	49	1	0
	1/16	12	6	0	0	0	0
	1/32	4	0	0	0	0	0
341.4958743	1/8	9318	-	8	46816	4216	245
	1/16	372	977	0	24	0	0
	1/32	23	70	0	0	0	0
989.7019159	1/8	45678	-	12	1937352	65557	9918
	1/16	6813	98623	1	1583	104	4
	1/32	549	3044	0	1	0	0

Table.2 *Absolute errors of the phase shift in* 10^{-6} *units*

E	h	[4]	[5]	[6]	$m=4$	$m=5$	$m=6$
53.588872	1/8	126121	8578	74	1	0	0
	1/16	6698	12	0	0	0	0
	1/32	405	0	0	0	0	0
341.495374	1/8	-	-	864	50	4	0
	1/16	59154	1460	1	0	0	0
	1/32	3170	7	1	0	0	0
989.701916	1/8	-	-	543289	450	15	2
	1/16	400726	213334	4060	0	0	0
	1/32	15071	10	1	0	0	0

Table 4.*Absolute errors of the eigenvalue in* 10^{-9} *units*

E	h	[1]	[2]	[3]	$m=4$	$m=5$	$m=6$
-41.232607772	1/8	0	0	0	0	0	0
	1/16	0	0	0	0	0	0
-26.873448916	1/8	12	87	0	7	0	0
	1/16	0	0	0	0	0	0
-8.67608 .670	1/8	45	504	0	58	1	0
	1/16	1	5	0	0	0	0

Tab.2 shows the absolute errors of the eigenenergies in 10^{-7} units for different choices of constant stepsize , which are shown in column 2.

In Tab.3 we present the absolute errors of the phase shifts from the exact value $\pi/2$ in 10^{-6} units.

In Tab.4 we list the true eigenvalues correct to nine decimal places and the absolute errors obtained by our method.

The difference problem were solved by shooting method. We need initial value y_1. Since equation (1.1) is linear and homogeneous then y_1 can be chosen arbitrarily. For convenience we use $y_1 = h$.

4 Appendix

Here are given some coefficients in presentation (2.1).

$$
\begin{aligned}
A_3 &= -f' \\
B_3 &= 0 \\
A_4 &= -f'' \\
B_4 &= -2f' \\
A_5 &= 4ff' - f''' \\
B_5 &= -3f'' \\
A_6 &= 7ff'' + 4f'^2 - f^{IV} \\
B_6 &= 6ff' - 4f''' \\
A_7 &= -9f^2f' + 11ff''' + 15f'f'' - f^V \\
B_7 &= 13ff'' + 10f'^2 - 5f^{IV}
\end{aligned}
$$

$$A_8 = -22f^2f'' - 28ff'^2 + 16ff^{IV} + 26f'f''' + 15f''^2 - f^{VI}$$

$$B_8 = -12f^2f' + 24ff''' + 48f'f'' - 6f^V$$

$$A_9 = 16f^3f' - 46f^2f''' - 148ff'f'' + 22fff^V - 28f'^3 + 42f'f^{IV} + 56f''f''' - f^{VII}$$

$$B_9 = -34f^2f'' - 52ff'^2 + 40ff^{IV} + 98f'f''' + 63f''^2 - 7f^{VI}$$

$$A_{10} = 50f^3f'' + 100f^2f'^2 - 86f^2f^{IV} - 338ff'f''' - 211ff''^2 + 29ff^{VI} - 232f'^2f'' + 64f'f^V + 98f''f^{IV} + 56f'''^2 - f^{VIII}$$

$$B_{10} = 20f^3f' - 80f^2f''' - 320ff'f'' + 62fff^V - 80f'^3 + 180f'f^{IV} + 280f''f''' - 8f^{VII}$$

$$A_{12} = -95f^4f'' - 260f^3f'^2 + 296f^3f^{IV} + 1878f^2f'f''' + 1201f^2f''^2 - 239f^2f^{VI} + 2972ff'^2f'' - 1292ff'f^V - 2288ff''f^{IV} - 1376ff'''^2 + 46ff^{VIII} + 280f'^4 - 1260f'^2f^{IV} - 3530f'f''f''' + 130f'f^{VII} - 675f''^3 + 255f''f^{VI} + 372f'''f^V + 210f^{IV2} - f^X$$

$$B_{12} = -30f^4f' + 200f^3f''' + 1200f^2f'f'' - 314f^2f^V + 600ff'^3 - 1840ff'f^{IV} - 2920ff''f''' + 128ff^{VII} - 2180f'^2f''' - 2790f'f''^2 + 490f'f^{VI} + 1026f''f^V + 1440f'''f^{IV} - 10f^{IX}$$

References

[1] J.R.Cash, A.D.Raptis and T.E.Simos. J.Comput. Phys. **91**,413-423(1990).

[2] A.C.Allison, A.D.Raptis and T.E.Simos. J.Comput. Phys.**97**,240-248(1991).

[3] R.M.Thomas, T.E.Simos and G.V.Mitsou. J.Comput.Appl.Math.**67**,255-270(1996).

[4] P.J.Van der Houwen, B.P.Sommeijer, K.Strehmler and R.Weiner. Computing **37**,195-218(1986).

[5] P.J.Van der Houwen and B.P.Sommeijer. IMA.J.Number.Anal.**7**,407-422(1987).

[6] T.E.Simos. J.Comput.Appl.Math.**58**,337-344(1995).

ENERGETIC PARTICLE FLUXES DATA BASE OF THE SPACECRAFT.

A. Beliaev, A. Dmitriev, S. Kuznetsov, O. Nechaev, M. Panasyuk

NPI MSU, Russia
e-mail:dalex@srdlan.npi.msu.su

Abstract. The data base of primary space physical data obtained on spacecrafts (SC DB) oriented on archiving and primary analysis of large sets of data containing informat on about particle fluxes in the near Earth space. SC DB software is used to perform detail scientific analysis and interpretation. The system is integrated with UNIX based software "Model SC and Physical Fields in the NES DB". Primary data are archived to binary files in predefined SC DB format. Data files contain the time ordered information about particle fluxes and spacecraft orbital elements (geographic co-ordinates, B and L-shell) and local time (LT and MLT). Orbital elements appended to data on the stage of data conversion. The primary analysis includes: data validity cheeks, detector calibration, on line calculating of magnetospheric and heliospheric physical parameters. The physical parameters values and diagnostic messages are presented in terminal of SC DB operator. The additional SC DB software packages may be used by scientists to perform multidimensional graphic analyses of data in interactive regime and to develop various semi-empirical models of space environment. The flexible SC DB structure permits to develop the software for additional data processing. OS DOS and high level programming languages (PASCAL, FORTRAN) were used to develop the SC DB software. The SC DB is currently used to perform data analysis for data obtained on CORONAS-I satellite and on MIR station.

1 INTRODUCTION

The data bases (DB) for near Earth spacecraft experiment are an important tool in accumulation, representation and investigation of information on the structure of magnetosphere and on the radiation environment. Since the number and dimension of measured parameters is well defined the data base storage format and data accumulation operations are uniform and simple. In this case we do not need to use a complicated DB product such as 'CDF'.

The alternative project is so-called a local 'mini' DB. This DB has to have firstly a simplest data appending and effective data packing operations which provide the shortest read-out time of large data volumes. Secondly a special operations for automatic calculation of standard detector capability indexes and of some necessary physical parameters have to be used in on line regime. Thirdly a complex data representation and data operation software has to be developed for data interpretation and for scientific analysis of experimental results. The SC DB has been created in manner of the described 'mini' DB.

2 INSTRUMENTATION

The scientific instrument package RIABINA operates since 1990 on board MIR. The detectors measure a fluxes of neutrons (E<15 MeV) and high energy charged particles (relativistic electrons and protons with energy more than 50 MeV) [1]. The CORONAS-I

spacecraft was launched on March 2, 1994 into a 500 km circular orbit with 83 degree inclination. The scientific instrument package SCR on board CORONAS-I contains the detectors for measuring gamma- spectra (E=0.1-200 MeV), relativistic electron spectra (E=0.5-124 MeV) and high energy charged particle fluxes [2]. Using the data obtained by the RIABINA and SCR instrument packages a data base SC DB for the time period of 1991 - 1996 has been developed in the Institute of Nuclear Physics.

3 DATA BASE CAPABILITIES

1. The data base has sufficiently flexible software developed using high-level languages PASCAL and FORTRAN and works in the DOS operating system. Transition to other operating systems is achieved via compilation of the initial code in the required operating system.

2. The data base may be stored on any device with a memory capacity not less than 600 Mb (e.g. hard disk or CD- ROM).

3. When used in a computer network the data base may be utilised simultaneously by any number of users.

4. The data base permits easy update or modification of the stored data, including total substitution of the formal parameters and corresponding data. The data base concept may be applied with minimum alterations to many other data sets, which are organised in terms of monotonously increasing time.

5. The special SC DB software package permits on line automatic calculation of standard detector capability indexes and of necessary physical parameters (averaged particle fluxes, radiation belt boundary positions and so on) .

6. The standard SC DB Management System permits the users to: 1) browse the required data parameters and time intervals; 2) quickly view the browse parameter time profile; 3) plot intensity graphs in both L, B, MLT and geographical co-ordinates; 4) determine the range of local times (LT, MLT) and L values in which the detectors operated; 5) select, sort and average data in the graphics mode; 6) store the data for the selected parameters in an external file; 7) run software packages, developed in the SC DB standard; 8) run alternative computer programs and program packages, which were created by the user for further data processing. SC DB FLOWCHART

The SC database is a set of sub-directories, containing the data files, text files, and executable files. All the sub- directories are united into one main directory.

When the SC DB is run the main resident program code (less than 30 KB) is loaded into the resident memory. When the standard DB operations are called, the necessary codes will be loaded into the operating memory, overlapping those which were loaded earlier and are no longer in use. Hence, the DB performs dynamic memory allocation. The high memory is not used during DB operation.

4 SC DB DATA LOADING

All the data in the SC DB is stored according to telemetry-read-out sessions about 1.5-12 hours long, recorded as binary sequential access files. The data division on the small

binary files provides the fast data searching operation. The data file list is contained in the text file which can be read by the SC DB software.

The primary data is telemetry-read-out data accepted on the RIABINA and SCR instrument packages. This data consists of the satellite operation mode, session descriptor and primary count rates of detectors. The satellite operation mode contain the information about operation regimes of instrument packages and about operation modes of detectors (on/off). The session descriptor describes the recount coefficients for restoring of primary count rates of detectors which are written in telemetric code.

The software package for adding primary data into the SC DB makes the following automatic operations:

- finding and getting of a new primary data using a UNIX based software "Model SC and Physical Fields in the NES DB";
- restoring of a real count rates of detectors from the primary data;
- calculating of a satellite co-ordination (Altitude Alt, Latitude Lat and Longitude Lon), magnetic field parameters (magnetic Latitude mLat and Longitude mLon, L-shell and magnetic field strength B) and local times (geographic LT and magnetic MLT) using the standard software packages for co-ordinate and magnetic field model (IGRF) calculations;
- creating of a new binary data file in SC DB standard;
- loading of a new data into the SC DB: loading the binary data file in the data directory and modifying the data file list.

5 SC DB ON LINE DIAGNOSTIC

The new data loaded into the SC DB is cultivated automatically by On Line processing and Calculating software package. This package checks the satellite operation mode and forms the message about operation regimes of detectors. The calculation of some averaged particle fluxes are necessary to compare these fluxes with standards fluxes. If the calculated flux and standard flux are non equal with some accuracy than the diagnostic message is formed. The boundaries of radiation belts are determined as a boundary between relatively high and low charged particle fluxes. The boundary location in magnetic co-ordination system permits to estimate the geomagnetic activity and to check the accuracy of co-ordination calculations. The result of last procedure is message too. All these operations are made immediately after the loading of new data into the SC DB therefore the operator has an opportunity to recognise a failure in operation of detectors with shortest time delay.

The diagnostic messages are used by investigators in the next operations by means of DB Management System and of Complementary Software or in the next steps of data restoring by means of special Initial Data Investigation program package.

6 SC DB MANAGEMENT SYSTEM

The core of SC DB Management System is 'Main Menu'. The DB permits the user to choose from several standard operation modes: fast data browsing, output of information on spacecraft location, data processing in the graphics mode, data output into an external file, LB-diagram plotting, raining users program from users menu. In order to conduct

these operations the user should identify the input parameters (define the time interval of interest, state the time-averaging period for the data, and denote the required parameters) in the 'Parameter Menu'. Since these initial settings may be needed later, they are stored in the so-called scratch directory, for storing the files with all the DB current settings , and DB operation results.

This scratch directory is created prior to the beginning of the user work with the DB, by running the SC DB Installation code. The program creates for user all the necessary directories if they did not exist earlier, as well as all the necessary files with the initial settings. In addition a batch file for running the SC DB 'Main menu' from the scratch directory is created. The general path structure for the SC DB is described in the respective DB readable file in the scratch directory.

7 COMPLEMENTARY SOFTWARE

The record string in the SC DB data file contains the record number, time, data error code, counts in the channels of the parameters, corresponding to the various types and energies of the measured particles, and co-ordinates of the observation point. The description of the data recording structure in the binary file is given in the respective text file and is software readable. The description includes parameter names, data read out time with preliminary re-count coefficients and geometric factors for corresponding particle parameters. The time interval between each record string in the data file (DB time resolution) is about 2.5 seconds. The high time resolution permits investigation of the spatial structure of particle fluxes in detail.

The binary file is equivalent to a matrix where the columns are the independent parameter values ordered by time dependence in the rows (record string). This independent columns structure adds the capability of multidimensional analysis of any DB parameters. SC DB software operates with columns of the matrix. The matrixes and column numbers are browsed by the investigator in the SC DB 'Parameter Menu' as an operating time interval and parameter names respectively. The time averaging is set in this menu also. The initial settings are saved in the service files and are software readable for later access.

The program packages of SC DB Complementary Software are run in the 'Users program menu'.

7.1 Plotting of Intensity Diagrams

For the selected SCR and RIABINA parameter this program plots intensity diagrams in different co-ordinate systems: L-B, LAT-LON, mLAT-mLON, invLAT-mLon, invLAT-MLT (invLAT is invariant Latitude); in various intervals in LT and MLT. Both middle latitudes and polar regions are presented on this diagrams. These diagrams permit the user to trace the structure and dynamics of particles in the Earth's magnetosphere. The capability to average the particle fluxes in the bins of different co-ordination matrix (Lat-Lon, L-B, mLat-mLon, invLat-mLon) on the certain altitude region permits to use the SC DB in comparable analysis of near Earth radiation distribution observed by different satellites.

7.2 A Semi-Empirical Model of Energetic Particle Fluxes in the Inner Radiation Belt

For a given time period in a given range of MLT, averaging of proton intensities is performed in the nodes of the L-B grid. The resulting matrix is plotted in L-B co-ordinates. For the selected L-shell values additional data processing is performed: the cosmic ray background is subtracted from the proton flux values (latitude flux dependence forming) and the cases of detector saturation are treated. The resulting altitude dependencies for protons are approximated with a running exponent and for regions of equal intensities are interpolated via same interpolation. The approximated altitude dependencies on the some L-shells and L-B diagrams of equal intensity contours in the inner radiation belt region are the final results of the operation. This contours can be used for verification of existing trapped radiation models (AP-8, AE-8) and for development of new models including the dynamic effects in trapped particle distribution.

8 CONCLUSIONS

1. The SC DB permits to store a permanently increasing extensive volume of uniform experimental data.

2. The SC DB provides fast input/output data operations.

3. The SC DB generates on line detector capability and data diagnostic.

4. The SC DB permits the user 1) to view experimental data in different co-ordinate systems, 2) to extract the necessary data to external files, 3) to select and modify the selected data.

5. The SC DB Complementary Software can be used for data redistribution which permits 1) to compare the data obtained in different experiments 2) to build new empirical models and to verify existing theoretical and semi- empirical models.

References

[1] Bratolyubova-Tsulukidze, L.S., Yu. P. Gordeev, V. I. Lyagushin, O. Yu. Nechaev, M. I. Panasyuk, and P. I. Shavrin, Multiobject instrument package for space neutron detection, Kosmicheskie Issledovaniya, **29**, 3, 487 (1991).

[2] Kuznetsov S. N., Yu. V. Gotselyuk. A. N. Podorolsky, S. P. Ryumin, S. Fisher, and K. Kudela. Scientific Goals and Specific Devices of the "Solar Cosmic Rays" (SCR) Scientific Set of Instruments in the "CORONAS" Project, Coronas Information **3** (1991).

CALCULATION OF DECAY RATES OF THE $HeH\mu$ SYSTEMS

V.B.Belyaev, O.I.Kartavtsev, V.I.Kochkin, E.A.Kolganova

Joint Institute for Nuclear Research, Dubna, Russia
e-mail: belyaev@thsun1.jinr.ru, oik@thsun1.jinr.ru, ckoch@cv jinr.ru, kea@thsun1.jinr.ru

Abstract. Study of long-lived exotic molecular systems $He\,H\mu$ is important for for the understanding the muon kinetic in hydrogen media and nuclear interactior at very low energies. The method of hyperspherical "surface" functions has been appl ed to to the calculation of the nonstable three-body systems $He\,H\mu$, contained the $^{3,4}He$ and $^{1,2}H$ isotopes. Energies of these systems, radiationless decay rates and γ-ray spectra of radiative decay have been obtained and compared with the recent expe imental data.

During last decades, study of charge-nonsymmetric muonic molecules attracts a continuous interest both theoretically and experimentally. The first measurements of the yield of γ-rays due to the decay of the Hedμ molecules [1, 2] revive interest in the investigation of this system. In the recent experiments [3] γ-rays spectra of ^{3,4}Hedμ systems have been measured with very high resolution. Besides, the first measurement of γ-rays spectra from the ^{3,4}Hepμ - molecules has been done [4]. These results call for theoretical treatment of comparable accuracy.

All the above mentioned experiments give evidence of the significant role of the nonradiative decay channel. Strong isotopic dependence of this process was found both experimentally and in the different calculations [5]-[8].

The competition of two decay channels and the strong isotopic dependence of the nonradiative decay rates are of special interest. A comparison of the calculated γ-rays spectra with recent experiments allows the determination of the total angular momentum of HeHμ molecules. Besides, the investigation of the nonradiative decay rate provides information on the behavior of the wave function at small distances, which is important for description of nuclear transitions. The unified treatment of radiative and nonradiative decay rates and eigenenergies of the ^{3,4}He^{1,2}Hμ systems for all possible values of the total angular momentum are given in this paper.

The approach based on the hyperspherical "surface" functions (HSF) method [10] has been applied in this paper. This method is well known in literature [10]-[15], allows the reliable treatment of the posed problems, and is described in details in the paper [9]. For this reason the notation and units of this paper will be used below.

The following properties of the systems under consideration should be taken into account in the calculation of the resonance energy and wave function. The dominant cluster structure of the type He + Hμ is a reason to use in this calculation the second effective potential $\varepsilon_2(\rho)$. This potential has an attractive part and supports the resonant state.

Next, the widths of such resonances are rather narrow. Therefore, coupling with the open channel described by the eigenpotential $\varepsilon_1(\rho)$ can be neglected. Indeed, the calculated coupling terms $Q_{12}(\rho)$ and $P_{12}(\rho)$, as can see from Fig. 1 [9], are small in the region, where the component u_2 of the resonant wave function is localized. This consideration brings

to the well-known "uncoupled adiabatic approximation" for the calculation of energy E_R and the radial wave function $u_2(\rho)$ as an eigenvalue problem:

$$\begin{cases} \left[\frac{d^2}{d\rho^2} - \frac{15}{4\rho^2} - \varepsilon_2(\rho) - P_{22}(\rho) + E_R\right] u_2(\rho) = 0 \\ u_2(0) = u_2(\infty) = 0. \end{cases} \tag{1}$$

Possible final states in the decay of HeHμ are the scattering ones of hydrogen nucleus on the Heμ atom. Eigenpotential $\varepsilon_1(\rho)$ asymptotically approaches the Heμ atom ground state energy $\varepsilon_{1\infty}$ and describes the dynamics in the open channel.

The scattering wave function $u_1(\rho)$ can be found as a solution of the boundary problem at fixed HeHμ resonance energy $E = E_R$:

$$\begin{cases} \left[\frac{d^2}{d\rho^2} - \frac{15}{4\rho^2} - \varepsilon_1(\rho) - P_{22}(\rho) + E\right] u_{1k}(\rho) = 0, \\ u_{1k}(0) = 0, \\ u_{1k}(\rho) \underset{\rho\to\infty}{\sim} \sin(k\rho + \delta), \end{cases} \tag{2}$$

where $k = (E - \varepsilon_{1\infty})^{1/2}$.

Due to the smallness of coupling the interchannel transition can be defined by the matrix element M of the coupling operator:

$$M_k = \int_0^\infty d\rho\, u_{1k}(\rho) \left[Q_{12}(\rho)\frac{d}{d\rho} + \frac{d}{d\rho}Q_{12}(\rho) - P_{12}(\rho)\right] u_2(\rho). \tag{3}$$

The radiationless decay rate λ has the form:

$$\lambda = \frac{1}{k}|M_k|^2. \tag{4}$$

The radiative decay of the systems under the consideration proceeds by means of the dipole transition γ-ray spectra and total rates of all dipole transitions from the states with angular momenta $L = 0, 1$ have been calculated. The final state radial wave function $u_{1k}(\rho)$ describing relative motion of a hydrogen isotope nucleus and Heμ - atom is a solution of equation (2) at the energy carried by these particles $E = E_R - E_\gamma$.

The dipole transition rate per unit energy (spectrum of γ-rays) is given by

$$\frac{d\lambda_\gamma}{dE_\gamma} = \frac{4}{3}\frac{(\alpha E_\gamma)^3}{k_1} \cdot \frac{1}{2L_i + 1} \sum_{L_f} | < \Psi_f ||\mathbf{d}|| \Psi_i > |^2 \tag{5}$$

where α is the fine structure constant, $k^2 = E_R - \varepsilon_{1\infty} - E_\gamma$ and L_i, L_f are total angular momenta of the initial and final states, respectively. Dipole operator $\mathbf{d}$ is expressed in terms of the scaled Jacobi coordinates $\mathbf{x}, \mathbf{y}$

$$\mathbf{d} = \mathbf{y}\frac{-(m_2 + m_3) - 3m_1}{(m_1 + m_2 + m_3)\sqrt{m_2 + m_3}} + \mathbf{x}\frac{2m_2 - m_3}{\sqrt{m_2 m_3 (m_2 + m_3)}} \tag{6}$$

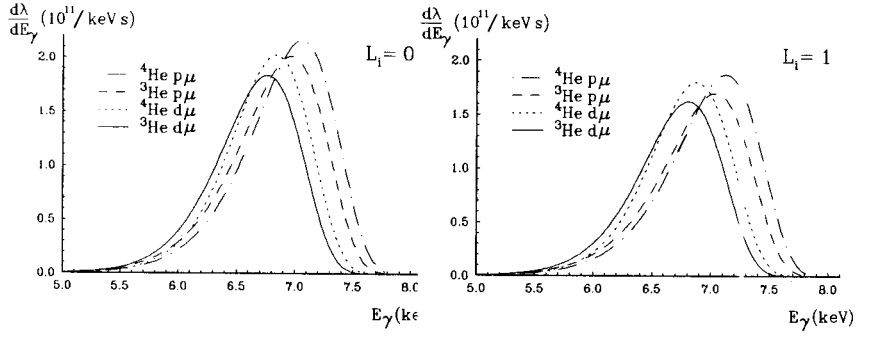


Figure 1: γ-ray spectra of the ^{3,4}HeHμ system in the initial states with the total angular momenta $L_i = 0$ and $L_i = 1$.

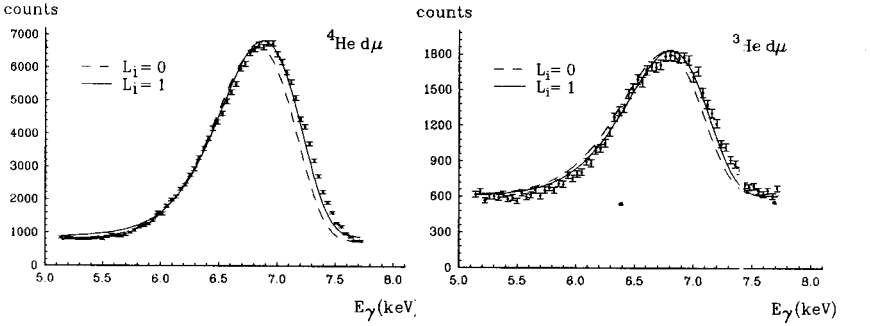


Figure 2: Comparison of the normalized theoretical γ-ray spectra with the experiment for the ^{4}Hedμ and ^{3}Hedμ systems.

The binding energies $E_B = \varepsilon_{2\infty} - E_R$, γ-ray spectra of radiative decay $\frac{d\lambda}{dE_\gamma}$, radiative decay rates $\lambda_\gamma = \int \frac{d\lambda}{dE} dE$ and nonradiative decay rates λ have been calculated for the systems ^{4}Hedμ, ^{4}Hepμ, ^{3}Hedμ, ^{3}Hepμ in the states with the total angular momenta $L = 0, 1, 2$.

Calculated values and available results of other authors are presented in Table 1. The type and number of trial functions, interval of radial integration, masses of particles and resulting accuracy of the numerical procedure are the same as in [9].

To demonstrate the isotopic dependence all theoretical γ-ray spectra for the decay from $L_i = 0, 1$ states are presented on Fig. 1.

At the moment new and accurate γ-ray spectra from the decay of the ^{4}Hedμ, ^{3}Hedμ and ^{4}Hepμ systems are measured in the experiments [3] and [4]. Since the absolute yield of γ-ray was not obtained in the experiments and due to the presence of a background, only shapes of the experimental and theoretical spectra can be compared. Taking these facts

into account, one can assume that the following two-parameter normalization procedure should be applied to fit the experimental spectrum

$$\left(\frac{d\lambda}{dE}\right)_n = A + B\left(\frac{d\lambda}{dE}\right)_{th},$$

where the parameters A and B are determined through the least squares procedure. The normalized theoretical spectra $\left(\frac{d\lambda}{dE}\right)_n$ for $L_i = 0, 1$ are presented in comparison with the experiment in Figs. 2.

Table 1. Calculated binding energies $E_B(eV)$ and decay rates λ, λ_γ $(10^{11}s^{-1})$ of the systems $^{3,4}He^{12}H\mu$, * - present results.

system	ref.	L=0			L=1			L=2	
		E_B	λ	λ_γ	E_B	λ	λ_γ	E_B	λ
$^4Hed\mu$	[5]				58.221	1.67			
	[8]				58.225303	1.80			
	[16]				57.84		1.69		
	[6]	77.96	2.3		56.10	2.4			
	[7]	78.7	1.85	1.94	57.6	1.38	1.74	20.3	0.9
	*	77.49	0.73	1.76	55.74	1.20	1.58	17.49	1.04
$^3Hed\mu$	[5]				48.419	5.06			
	[8]				48.420890	5.29			
	[16]	70.74		1.75	47.90		1.55		
	[6]	69.96	8.0		46.75	7.0			
	[7]	70.6	3.58	1.80	48.2	2.77	1.58	9.6	1.54
	*	69.37	2.87	1.64	46.31	3.22	1.44	7.11	1.74
$^4Hep\mu$	[16]				50.0		1.92		
	[6]	74.36	71		41.17	38			
	[7]	75.4	35.4	2.24	45.4	24.8	1.89		
	*	80.64	70.7	2.02	47.45	38.5	1.74		
$^3Hep\mu$	[6]	67.70	100		33.85	46			
	[7]	69.0	47.3	2.11	38.1	31.6	1.74		
	*	72.76	97.7	1.91	38.82	46.7	1.60		

All the calculations (see Table 1) support the strong isotopic dependence of the nonradiative decay rate noted in experiments [1]-[4]. This rate decreases rapidly with increasing reduced mass of two heavy particles. Qualitatively, the muon transfer in the nonradiative decay process takes place mainly at small internuclear distances. Due to the cluster $He + H\mu$ structure, two heavy particles should penetrate under the potential barrier to bring together. As a consequence, the transition probability decreases exponentially with increasing reduced mass. Besides, the nonradiative decay rate is sensitive to the total angular momentum and reduced mass due to the interference of the initial and rapidly oscillating final state wave functions.

Opposite to the nonradiative decay channel, the radiative decay rate (as it is clear from Table 1) depends slightly on isotope masses. The different isotopic dependence can

be explained by the different spatial structure of the transition operators. Considering the position of γ-ray spectra presented in Fig. 1, one can conclude that the total radiative decay rate increases with the γ-ray energy both for $L_i = 0$ and $L_i = 1$. A change of the spectrum position for different isotopes is a consequence of a change of the threshold energies for the initial and final states. As a matter of fact, the transition energy increases slightly with increasing mass of the helium isotopes due to lowering of the first eigenpotential. At the same time, increasing the mass of the hydrogen isotope gives rise to a bit more decreasing in the transition energy due to the lowering of the second eigenpotential.

The definite conclusion about the total angular momentum of the ^{3,4}Hepμ molecules can not be made due to the low γ-ray yields and, therefore, the comparatively low experimental accuracy [4]. In this case the nonradiative decay dominates over the radiative one.

The shapes of γ-ray spectra for different total angular momenta $L_i = 0, 1$ are very similar and the $L_i = 1$ spectra are slightly shifted to higher energies in comparison with the $L_i = 0$ ones. However, the accuracy of recently measured γ-ray spectra of ^{3,4}Hedμ [3] is high enough to distinguish the $L_i = 0, 1$ theoretical spectra on qualitative level. For this reason, the experimental and theoretical γ-ray spectra of these systems for the total angular momenta $L_i = 0, 1$ have been compared in Fig. 2. From the comparison, one has to conclude that radiative transitions take place from the $L_i = 1$ state of the ^{3}Hedμ and ^{4}Hedμ molecules. One should mention that this result is the first definite assignment of the total angular momentum of the ^{3,4}Hedμ - molecules observed in experiment.

The providing of the experimental data by B. Gartner and S. Tresch is gratefully acknowleged.

References

[1] T.Matsuzaki, K.Ishida, K.Nagamine, Y.Hirata and R.Kadono, Muon Cat. Fusion **2**, 217 (1988).
[2] K.Ishida, S.Sakamoto, Y.Watanabe, T.Matsuzaki and K.Nagamine, Hyperfine Interactions **82**, 111 (1993).
[3] B.Gartner et al., Hyperfine Interactions, **101/102**, 249 (1996).
[4] S.Tresch et al., Hyperfine Interactions, **101/102**, 221 (1996).
[5] Y.Kino and M.Kamimura, Hyperfine Interactions **82**, 195 (1993).
[6] S.S.Gershtein and V.V.Gusev, Hyperfine Interactions **82**, 185 (1993).
[7] A.V.Kravtsov, A.I.Mikhailov and V.I.Savichev, Hyperfine Interactions **82**, 205 (1993).
[8] V.I.Korobov, Hyperfine Interactions, **101/102**, 329 (1996).
[9] V.B.Belyaev, O.I.Kartavtsev, V.I.Kochkin and E.A.Kolganova, Phys. Rev. A**52**, 1765 (1995).
[10] J.H.Macek, J. Phys. B**1**, 831 (1968).
[11] M.Fabre de la Ripelle, M.I.Haftel and S.Y.Larsen, Phys. Rev. A**44**, 7084 (1991).
[12] B.J.Archer, G.A.Parker and R.T.Pack, Phys. Rev. A**41**, 1303 (1990).
[13] C.D.Lin, Adv. Atom. Mol. Phys. **22**, 77 (1986).
[14] V.V.Gusev et al, Few-Body Systems **9**, 137 (1990).
[15] M.Decker, W.Sandhas and V.B.Belyaev, Phys. Rev. A**53**, 726 (1996)
[16] S.Hara and T.Ishihara, Phys. Rev. A**39**, 5633 (1989).

PHYSICS, ACCELERATOR and SOFTWARE DATA BASES at the WWW/Oracle-server of JINR.

V.A.Berezhnoj,
Institute of Nuclear Physics RAS (Troizk),

B.S.Volkov,
Institute of High Energy Physics (Protvino),

Yu.N.Isaev, A.P.Sokolov,
Institute of Theoretical and Experimental Physics (Moscow),

V.F.Borisovskij, G.Musulmanbekov, V.V.Koren'kov, E.G.Nikonov, M.S.Plyashkevich, A.B.Ribalkin, V.P.Shirikov.
141980 Dubna Moscow region, Russia JINR LCTA.

1. Introduction.

Information hypertext system (IHS) for access to physical data bases and information servers in the world is designed at the WWW/ORACLE-server of Joint Institute for Nuclear Research (JINR). This work is a part of the project BAPHYS-96, granted by Russian Fund of Fundamental Research for seven physical institutes : Joint Institute for Nuclear Research (Dubna), Institute of High Energy Physics (Protvino), Institute for Nuclear Research (Moscow), Institute of Theoretical and Experimental Physics (Moscow), Institute of Nuclear Physics of Moscow State University, Institute of Nuclear Physics (Sanct-Petersburg), Institute of Nuclear Physics (Novosibirsk). The main purpose of this project is elaboration of a unique network system of information servers in the field of fundamental matter properties and applied nuclear physics.

The network of information servers, which will be created under this project, should provide the following basic possibilities:

- information support of the participation of physicists in the international experimental collaborations such as ATLAS, CMS, ALICA, EAST, ZEUS, NA48, HERMES etc.;
- Effective usage of the information and search systems as WWW, WAIS, GOPHER etc., easy access to some scientific data bases such as INIS, PPDS, RPP, PREP, INSPEC, COMPUSCIENCE, MATHDI, VINITI, STN Int. etc.;
- representation of data and knowledge available for all scientists in Russia, Europe and in other world.

The structure of IHS is given in Section 2. An Oracle Data Base (DB) for accelerators of Russia and former Soviet Union designed at WWW/Oracle-server of JINR is described in Sections 2 and 3. An access to this DB is realized via IHS using special Web/Oracle gateway.

2. A structure of IHS.

IHS designed as a system of files in HTML-format. This set of HTML-pages provides a fast and easy access to a wide variety of information, service and data on nuclear and particle physics at low, medium and high energies, which is available via World Wide Web. The main HTML-page consists of the following sections.

- **Physics Servers and Services** around the world give access to Physics Encyclopedia, Educational Resources, News in Physics, Data and Tables on Physical Constants et cetera.

- **Physics Conferences, Workshops and Summer Schools**. This WWW-server contains a list of various conference, workshops and schools. It is possible to give your own announcement about your conference etc.

- **Institutions around the world**. This page consists of the following HTTP-references:
  1. Institutions in Russia;
  2. Institutions by Country;
  3. Institutions by Field of Research;
  4. Other (non-university) Research Laboratories.

- **High energy physics (HEP)** section, which includes current news, educational materials, information about HEP centers, programs and service packages on data analysis and modeling, detector simulation, event generators etc. Data bases, supplied by friendly interfaces, give direct access to information about conferences, institutions, accelerators, experiments, publications and numerical data obtained in HEP experiments.

- **Low and medium energy physics** section, which includes servers with bibliographic and numeric data covering nuclear structure and decay, nuclear reactions, group constants, neutron capture and other information in computerized files. These services are maintained by National Nuclear Data Center (NNDC) of Brookhaven National Laboratory in collaboration with other institutions, and Nuclear Data Section (NDS) of the International Atomic Energy Agency..

- **Russian WWW-Servers**. Now there are available the following Russian WWW-servers: Astrakhan, Barnaul, Bryansk, Chernogolovka, Chelyabinsk, Dubna, Ekaterinburg, Gelendzhik, Irkutsk, Ivanovo, Izhevsk, Kaliningrad, Kazan, Khabarovsk, Kostroma, Krasnoyarsk, Lipetsk, Moscow, Murmansk, Nizhny-Novgorod, Novgorod, Novosibirsk, Obninsk, Orel, Perslavl-Zalessky, Petrozavodsk, Penza, Perm, St. Petersburg, Pskov, Rostov-on-Don, Ryazan, Samara, Saratov, Serpukhov, Severodvinsk, Simbirsk, Surgut, Taganrog, Tambov, Togliatti, Tomsk, Tula, Tver', Vladimir, Voronezh, Vyatka, Ufa, Yaroslavl.

- **Publishing offices**. This section consists of a List of **Publishing Offices** home pages which gives an access to book and journal catalogs, their tables of contents and abstracts of articles, covering wide areas of science and technique. Some offices maintain electronic versions of issued journals. This section includes **(La)TeX Navigator**, as well.

- **A list** of WWW-servers with an information concerning **accelerators** around the world. This HTML-page contains three HTTP-references.

- **A list of accelerators around the world** is a reference to WWW-server with the information on all accelerators in foreign physical centers.

- **Accelerators** Data Base placed at the JINR WWW/Oracle-server is a reference to the Oracle data bases concerning various information on the most important machines in Russia and former Soviet Union.

- **Software** for calculations and modeling of accelerator technique consists of descriptions and references to various programs for simulating and modeling experiments, magnetic and electric systems of accelerators, detectors and other experimental technique etc.

The basic idea of IHS is to order HTTP-references to various data available via INTERNET at Russian and foreign sites. IHS is available via the following HTTP-reference:

1. Directly, via the URL **http://sunct2.jinr.dubna.su/~nikonov/phbd.html**;

2. Via the INFO-section of JINR Home Page **http://www.jinr.ru**;

3. Via the section "Special home pages and Web-servers" of LCTA Home Page **http://jicom.jinr.ru/LCTA/.**

3. Accelerator and Software Data Bases.

Section **Accelerators** of main HTML-page of IHS consists of hypertext references to the following DBs. The first one is a hypertext DB of software for physics research. It is based on the collection of annotations of programs for calculations of accelerators devices and its components, for modeling a dynamics of charged particles beams in these devices. The second reference points to Accelerator DB which is placed at the JINR WWW/Oracle-server. It was created to provide easy access to the information on big accelerator complexes by a queries about types of accelerator machines and a main trends of their development. This DB also has to give the information on a beam energy, intensity, luminosity, a type of particles and other parameters, which could be interesting for both to realize physical experiments and to find directions of a progress of accelerator science and technique.

It is supposed to renew these Data Bases and, in particular, to enlarge it by including of more and more machines and selecting of perspective directions of a development of accelerators of charged particles. It is foreseen creating new sections on every machine such as measurement, management, high frequency, magnetic, vacuum systems and other elements of accelerators. It is important that a concrete list of these sections on every complex will be defined by specialists responsible for preparing corresponding Web-pages who is directly working at described machines or with projects of new accelerators.

The Data Bases will promote getting more information for specialists and help to access to new accelerator and physics technologies. This system of data bases is open for all persons who is interested in. By means of its public character and under the condition of an interest of accelerators specialists in creating DB, these DB will be able to get such properties which will provide principally new environment for Scientific and Technical Designs. This new environment

will enlarge co-operation between various organizations and application a new experience. It will improve tutoring of stuff and their professional skills.

4. A structure of Accelerator DB.

This DB consists of projects of accelerators. All projects grouped in 9 main *directions*:

- Proton and electron complexes of high energy;
- Linear colliders;
- B-factory;
- High current accelerators of protons and positrons;
- Accelerators of heavy ions;
- Synchrotron radiation sources of third generation;
- Superconductor resonator in big electron accelerator complexes;
- Accelerators of radioactive beams (Acceleration of particles);
- Accelerators of radioactive beams (Acceleration of radioactive beams);

A description of a project consists of a general information, physics and technical parameters and also can include photographs, graphical and other materials. Structure of Accelerator DB is represented by Figure 1.

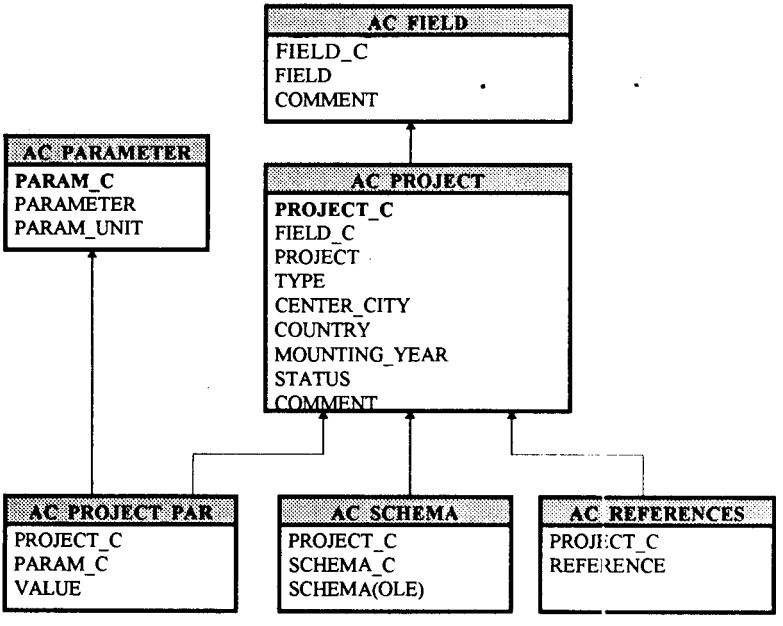


Figure 1.

An access to Accelerator DB via IHS is realized by ORALink gateway. ORALink is a multithreaded CGI (Common Gateway Interface) 1.1 application that uses HTTP POST method to get output from HTML forms. More information on ORALink gateway is available at **http://oradb1.jinr.ru/software/oralink/.** Scheme of access to Accelerator DB by ORALink is shown at the picture bellow.

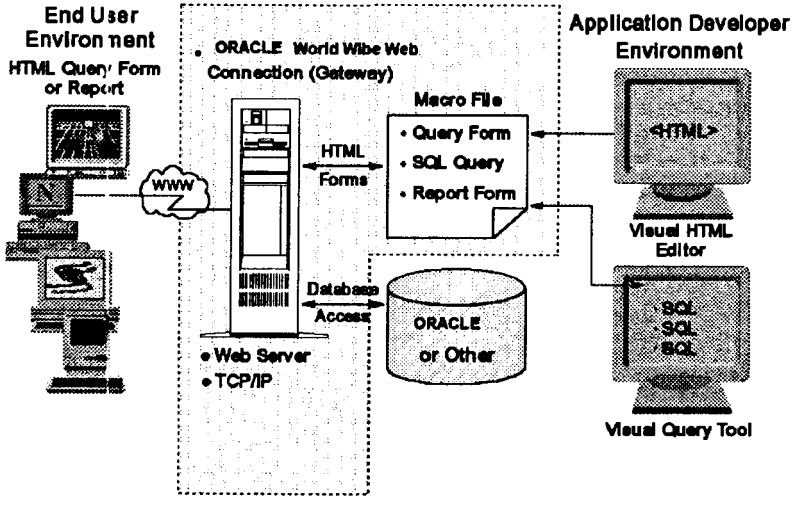


Figure 2.

THE U(1) GAUGE-INVARIANT HEISENBERG ANTIFERROMAGNET MODEL AND MULTIDIMENSIONAL TOPOLOGICAL SOLITONS

I.L.Bogolubsky and A.A.Bogolubskaya

141980, Dubna, Moscow region, JINR, LCTA, Russia
e-mail: bogoljub@main1.jinr.dubna.su

Abstract. The 3-component $U(1)$ gauged sigma model with the broken $Z(2)$ symmetry ("the A3M model") is discussed, which admits search for topological solitons in $(D+1)$-dimensional space-time ($D = 2, 3$). For $D = 2$ string-like localized solutions are studied numerically using the "hedgehog" ansatz. Three-dimensional particle-like solutions of the A3M model and its extension ("the Dirac-Heisenberg-Maxwell model") could be considered as classical prototypes of massive quantum particles.

1. Presently, one century later since the discovery of the electron, there is still no commonly accepted theoretical description of its internal structure. One of possible ways to imagine a plausible model of the electron is to conjecture it to be the bunch (i.e., a 3D soliton) of interacting fundamental fields, localized within a small space region. However till now, a satisfactory soliton model of the electron has not yet been constructed. Nevertheless, we believe that the considerable progress can be achieved here if one considers the topological solitons (see, e.g. [1,2,3]) within nonliear sigma models. Classical sigma models describe the evolution in time of N-component unit isovector field $s_a(\mathbf{x}, t)$, $s_a s_a = 1, a = 1, .., N$, in $(D+1)$-dimensional space-time; field manifolds of these models are unit spheres S^{N-1}. In sigma models with some specific pair combinations of D and N (for example, $D = 2, N = 3$; $D = 3, N = 3$; $D = 3, N = 4$) the localized distributions of unit isovector $s_a(\mathbf{x})$ are divided into disjoint classes with different integer topological "charges" Q_t; solitons with the nonzero Q_t are referred to as "topological solitons". It is worth noting that the existence of solitons is not yet guaranteed for the sigma model with nontrivial topological properties, though in the cases of nonzero Q_t additional possibilities arise for solitons to exist.

2. First let us consider the 3-component unit isovector field $s_a(\mathbf{x})$ with the "easy-axis" anisotropy defined by the Lagrangian density

$$\mathcal{L} = (\partial_\mu s_a)^2 - V(\mathbf{s}), \quad s_a s_a = 1, \quad V(\mathbf{s}) = 1 - s_3^2, \tag{1}$$

$$\mu = 0, 1, ..., D, \quad a = 1, 2, 3.$$

(it was called the A3-field in [4]; "A3" stands for "anisotropic 3-component"). It is important to note that the Lagrangian (1) possesses $U(1) \times Z(2)$ internal symmetry; its vacuum manifold comprises two points on the S^2 sphere: $s_3 = 1$ and $s_3 = -1$ and possesses discrete $Z(2)$ symmetry [4], which is broken both on vacuum solutions and topological solitons for $D = 1, 2$ (see, e.g., [5-8]).

The Lagrangian density (1) can be derived when describing in continuous approximation easy-axis Heisenberg antiferromagnets [6] and ferroelectrics [9] with the easy-axis anisotropy; thus, the pattern of the symmetry breaking under discussion is realized in condensed matter physics. It seems interesting to investigate possible implications of this

pattern (which differs from that used in the Standard Model) for high-energy physics, especially having in mind intensive use of sigma models in modern particle physics. The scalar (with respect to Lorentz transformation) A3-field used by us to this end can be regarded either as one belonging to the set of basic quantum fields or as coherent state [10] of more fundamental (e.g., spinor) fields.

3. Next consider "minimal" interaction of the A3-field with the Maxwell field $A_\mu(x)$ in $(D+1)$-dimensional space-time ($D = 1, 2, ...$), described by the $U(1)$ gauge-invariant Lagrangian ("the A3M model"):

$$\mathcal{L} = \eta^2(\partial_\mu s_a)^2 - V(s_a) - \frac{1}{4}F_{\mu\nu}^2 + 2g\eta^2 A_\mu(s_2\partial^\mu s_1 - s_1\partial^\mu s_2) + g^2\eta^2(s_1^2 + s_2^2)A_\mu A^\mu. \quad (2)$$

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu, \quad V(s_a) = \beta^2(1 - s_3^2), \quad (3)$$

where β^2, η^2, are constants, $[\eta^2] = L^{(1-D)}$, $[\beta^2] = L^{-(1+D)}$, g is a coupling constant, $[g^2] = L^{(D-3)}$, $\mu, \nu = 0, 1, ..., D$, and summation over repeated indices μ, ν is meant. The Lagrangian (2),(3) of the A3M model possesses both $U(1)$ *local* symmetry and spontaneously broken $Z(2)$ symmetry. It should be noted that the existence of 2D and 3D solitons in this *two-field* model is not forbidden by the Derrick theorem [11].

It is instructive to present the equations of the A3M model in terms of angular variables θ, ϕ on the unit sphere S^2,

$$s_1 = \sin\theta\cos\phi, \quad s_2 = \sin\theta\sin\phi, \quad s_3 = \cos\theta, \quad (4)$$

and to make rescaling $x_\mu \to g^{-1}\eta^{-1}x_\mu$, $A_\mu \to \eta^{-1}A_\mu$. As a result the Lagrangian density (3) takes the form (in rescaled x_μ, A_μ):

$$g^{-2}\eta^{-4}\mathcal{L} = \partial_\mu\theta\partial^\mu\theta + \sin^2\theta\,[\partial_\mu\phi\partial^\mu\phi - 2A_\mu\partial^\mu\phi + A_\mu A^\mu - p] - \frac{1}{4}F_{\mu\nu}^2. \quad (5)$$

and the Euler-Lagrange equations become (the Lorentz gauge, $\partial_\mu A^\mu = 0$, is chosen):

$$\partial_\mu\partial^\mu\theta + \frac{1}{2}\sin 2\theta\,[p - \partial_\mu\phi\partial^\mu\phi + 2A_\mu\partial^\mu\phi - A_\mu A^\mu] = 0, \quad (6)$$

$$\partial_\mu\left[\sin^2\theta(\partial^\mu\phi - A^\mu)\right] = 0, \quad (7)$$

$$\partial_\mu\partial^\mu A_\nu + 2j_\nu + 2A_\nu\sin^2\theta = 0, \quad j_\nu = -\sin^2\theta\partial_\nu\phi, \quad \mu, \nu = 0, 1, ..., D. \quad (8)$$

Note that Eqs. (6)-(8) can be satisfied if

$$\phi(x) = \phi(\mathbf{x}) - \omega t, \quad A_0 = \omega = const, \quad A_k(x) = A_k(\mathbf{x}), \quad \theta(x) = \theta(\mathbf{x}), \quad k = 1, ..., D, \quad (9)$$

where A_k, θ satisfy equations which do not contain ω ($k, m = 1, ..., D$, summation over repeated k):

$$\partial_k^2\theta - \frac{1}{2}\sin 2\theta\left[p + (\partial_k\phi - A_k)^2\right] = 0, \quad (10)$$

$$\partial_k\left[\sin^2\theta(\partial_k\phi - A_k)\right] = 0, \quad (11)$$

$$\partial_k^2 A_m + 2\sin^2\theta(\partial_m\phi - A_m) = 0. \quad (12)$$

4. Below we shall study localized solutions to Eqs. (10)-(12) for $L = 2$ using the "hedgehog" ansatz for the unit isovector field $s_i(\mathbf{x})$, $i = 1, 2, 3$,

$$s_1 = \cos m\chi \sin\theta(r), \quad s_2 = \sin m\chi \sin\theta(r), \quad s_3 = \cos\theta(r) \tag{13}$$

$$\sin\chi = \frac{y}{r}, \quad \cos\chi = \frac{x}{r}, \quad r^2 = x^2 + y^2,$$

where m is an integer number, and the "vortex" ansatz for the Maxwell field $A_\mu(\mathbf{x})$,

$$A_0 = 0, \quad A_1 = A_x = -m\alpha(r)\frac{y}{r^2}, \quad A_2 = A_y = m\alpha(r)\frac{x}{r^2}. \tag{14}$$

As a result we obtain equations for $\theta(r)$ and $\alpha(r)$ [12],

$$\frac{d^2\theta}{dr^2} + \frac{1}{r}\frac{d\theta}{dr} - \sin\theta\cos\theta\left[\frac{m^2(\alpha-1)^2}{r^2} + p\right] = 0, \tag{15}$$

$$\frac{d^2\alpha}{dr^2} - \frac{1}{r}\frac{d\alpha}{dr} + 2(1-\alpha)\sin^2\theta = 0. \tag{16}$$

with boundary conditions

$$\theta(0) = \pi, \quad \theta(\infty) = 0, \tag{17}$$

$$\alpha(0) = 0, \quad \frac{d\alpha}{dr}(\infty) = 0. \tag{18}$$

Note that field configurations $s_i(\mathbf{x})$ given by Eqs.(13),(17) correspond to maps from R^2_{comp} to S^2 with integer homotopic indices ("winding numbers")[2,3], and we shall refer to solutions of the problem (15)-(18) as topological solitons with the "topological charges" $Q_t = m$.

Using series expansion of $\theta(r)$ and $\alpha(r)$ at $r \to 0$, we find from Eqs. (15) and (16) for $m = 1$

$$\theta(r) = \pi - C_1 r + o(r), \tag{19}$$

$$\alpha(r) = r^2(E_1^2 - \frac{1}{4}C_1^2 r^2) + o(r^4). \tag{20}$$

and for $m = 2$

$$\theta(r) = \pi - C_2 r^2 + o(r^2), \tag{21}$$

$$\alpha(r) = r^2(E_2^2 - \frac{1}{12}C_2^2 r^4) + o(r^6). \tag{22}$$

Equations (19),(20) and (21),(22) are useful when searching for solutions of the problem (15)-(18) with $m = 1$ and $m = 2$, respectively, by the shooting method. Careful numerical investigation shows that for small positive p values there exists a unique solution to the problem (15)-(18); the functions $\theta(r)$ and $\alpha(r)$ of solitons with $m = 1$ are presented in Fig.1 for p =0.01, 0.1, 0.15. Note that the asymptotic value $\alpha_\infty = \alpha(\infty)$ decreases monotonically as p is increased, with $\alpha_\infty \to 1$ when $p \to 0$. With the function $\alpha_\infty(p)$ at hand one can find the asymptotic form of the soliton solution for $r \to \infty$:

$$\theta(r) \approx \frac{T}{\sqrt{r}}\exp(-\sqrt{p}r), \quad T = const, \tag{23}$$

$$\alpha(r) \approx \alpha_{\infty} - (1 - \alpha_{\infty})\frac{T^2}{2r}\exp(-2\sqrt{p}r).$$

Distributions of the energy density

$$\mathcal{H}(r) = \left(\frac{d\theta}{dr}\right)^2 + \sin^2\theta\left[p + \frac{m^2(\alpha-1)^2}{r^2}\right] + \frac{m^2}{2}\left(\frac{1}{r}\frac{d\alpha}{dr}\right)^2, \qquad (24)$$

and the magnetic field, $-B(r) = (d\alpha/dr)/r$, for solitons with $m = 1$ are plotted in Figs. 2 and 3, respectively. The magnetic flux of the 2D solitons in the A3M model is uniquely determined by values of p and the topological charge $Q_t = m$,

$$\Phi = \int \mathbf{B}d\mathbf{S} = \int A_k dx_k = m\int_0^{2\pi} \alpha(\infty)d\varphi = 2\pi m\alpha_{\infty}(p) = 2\pi Q_t\alpha_{\infty}(p). \qquad (25)$$

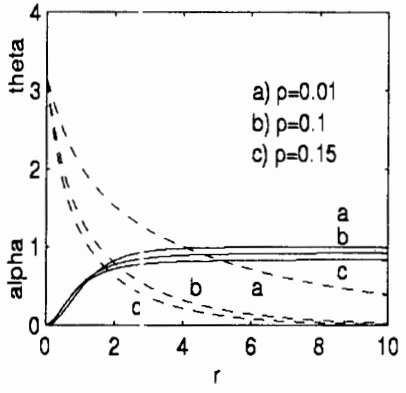


Figure 1: Functions $\alpha(r)$ and $\theta(r)$ of the 2D A3M solitons with $m = 1$ for various p.

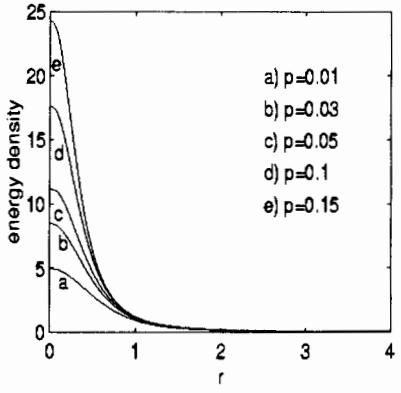


Figure 2: Energy density of the 2D A3M solitons vs r for various p; $m = 1$.

The dependence of the soliton energy $E = 2\pi\int\mathcal{H}(r)rdr$ on p is depicted in Fig.4; it is natural to conjecture that the dependence $E(\log p)$ is exactly linear. It is important to note that $E(p) < 8\pi$ for $p < p_0 \approx 0.4$ (recall that 8π is the energy value of the Belavin-Polyakov localized solutions in the 2D isotropic Heisenberg ferromagnet [13]). It means that at least for $p < p_0$ the particle-like solutions (18)-(23) of the A3M model describe spatially localized *bound* states of the A3- and the Maxwell fields, and hence it is natural to conjecture these 2D solitons to be stable for $p < p_0$ (indeed, to destroy them it is necessary to supply an amount of energy exceeding the appropriate "mass defect").

4. One can consider the 2D A3M solitons as strings with finite radia embedded into 3D space. Such strings-solitons can be used when designing cosmological models of evolution of the early Universe and its present large-scale structure (for review see, e.g., [14]). From mathematical point of view these solitons are well suited for computer simulation of interaction of extended cosmic strings, indeed, to investigate the interaction of M strings within

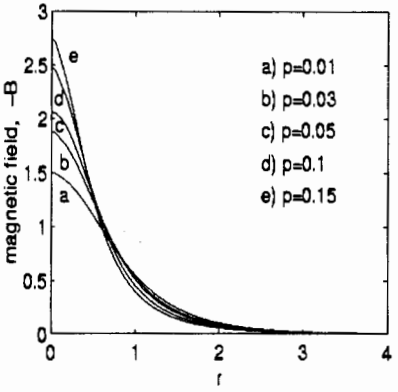


Figure 3: Magnetic field of the 2D A3M solitons with $m = 1$ for various p.

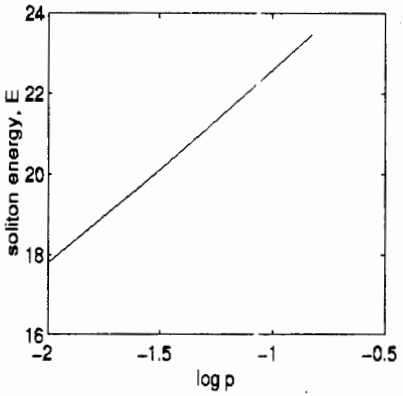


Figure 4: The dependence of the 2D A3M soliton energy E on $log(p)$; $m = 1$.

the A3M model initial data for the evolutionary PDE problem could be chosen as a set of M solitons well separated from each other. Notice that this physically natural prescription cannot be used for computer simulation of interactions of the Abrikosov-Nielsen-Olesen strings-vortices[15] within the abelian Higgs model[16], which are widely discussed (along with their non-abelian generalizations) in cosmological models[14].

Further computer studies of the 2D solitons within the A3M model are in progress, in particular, investigation of the boundary value problem (15)-(18) for various m and p. Next, we are starting computer study of dynamical processes in which the 2D A3M topological solitons take part, using evolutionary PDEs of the model presented above.

5. From the viewpoint of particle physics the most important line of further studies within the A3M model is certainly the search for stable 3D solitons. Our preliminary analysis shows that existence of stationary 3D A3M solitons with nonzero integer values of the topological charge of the $A3$-field (the Hopf index, see, e.g., [2a,17]) is quite likely. If such solitons are found in computer experiments that we are preparing, then the next natural step will be the search for 3D localized "electron-like" solutions with nonzero values of electric charge and internal angular momentum (spin) within the 3-field ("Dirac-Heisenberg-Maxwell") model, or the "extended electrodynamics" (EED), defined by the $U(1)$ gauge-invariant Lagrangian density (ϵ, κ are constants)

$$\mathcal{L} = \eta^2 \left(\bar{\mathcal{D}}(g)_\mu s_- \mathcal{D}(g)^\mu s_+ + \partial_\mu s_3 \partial^\mu s_3\right) + i\kappa\bar{\psi}\gamma_\mu \mathcal{D}(\epsilon)^\mu \psi - V(s_a) - \frac{1}{4}F_{\mu\nu}^2, \qquad (26)$$

$$\bar{\mathcal{D}}(g)_\mu = \partial_\mu + igA_\mu, \quad \mathcal{D}(g)_\mu = \partial_\mu - igA_\mu, \quad \mathcal{D}(\epsilon)_\mu = \partial_\mu - ie\mathcal{A}_\mu,$$

$$s_+ = s_1 + is_2, \quad s_- = s_1 - is_2.$$

However, serious mathematical difficulties should be overcome first in order to make appropriate computer investigations reliable and successful.

References

[1] T.H.R.Skyrme, Proc. Roy. Soc., **A260** 127 (1961); Nucl. Phys. **31**, 556 (1962).

[2] L.D.Faddeev, in: Proc. IV Intern Conf. on Nonlocal and Nonlinear field theories, Alushta, 1976. Joint Institute for Nuclear Research, Dubna, D-29788, 1976, p.207; Lett. Math Phys. 1 289 (1976).

[3] R.Rajaraman, Solitons and instantons (North-Holland, Amsterdam, 1982).

[4] I.L.Bogolubsky and A.A.Bogolubskaya, Soliton analogs of Abrikosov-Nielsen-Olesen vortices, in: Proc. of the 1st Workshop "Nonlinear Physics. Theory and Applications", Gallipoli (Lecce), 1995, Eds. E.Alfinito, M.Boiti, L.Martina and F.Pempinelli (World Scientific, Singapore, 1996).

[5] I.V.Baryakhtar and B.A.Ivanov, Zh.E.T.F., 85 (1983) 328.

[6] A.M.Kosevich, B.A.Ivanov and A.S.Kovalev, "Nonlinear magnetization waves. Dynamical and topological solitons" (Naukova Dumka, Kiev, 1983) [in Russian]; Phys. Reports, 194 (1990) 1.

[7] I.L.Bogolubsky and A.A.Bogolubskaya, JINR, E5-95-273, Dubna, 1995 (submitted to Physica D).

[8] V.P.Voronov and A.M.Kosevich, Zh.E.T.F., 90 (1986) 2145.

[9] J.Pouget and G.A.Maugin, Phys. Rev., B30 (1984) 5306 .

[10] J.R.Klauder and B.S.Skagerstam, Coherent States (World Scientific, Singapore, 1985); A.M.Perelomov, Generalized Coherent States and Their Applications (Springer, Berlin, 1986).

[11] G.H.Derrick, J. Math. Phys. **5** 1252 (1964).

[12] I.L.Bogolubsky and A.A.Bogolubskaya, JINR, E5-96-73, Dubna, 1996 (to be published in Annales de la Fondation Louis de Broglie).

[13] A.A.Belavin and A.M.Polyakov, Pis'ma v Zh.E.T.F., 22 (1975) 503 [JETP Lett., 22 (1975) 245].

[14] A.Vilenkin and E.P.S.Shellard, Cosmic strings and other topological defects (Cambridge University Press, Cambridge, 1994).

[15] A.A.Abrikosov, Zh.E.T.F., 32 (1957) 1442; H.B.Nielsen and P.Olesen. Nucl. Phys., B61 (1973) 45 .

[16] P.Higgs, Phys. Lett., 12 (1964) 132.

[17] D.A.Nicole, J.Phys., G4 (1978) 1363; H.J.de Vega, Phys.Rev., D18 (1978) 2945; I.L.Bogolubsky, Phys. Lett., A126 (1988) 511.

THE HALL EFFECT IN TWO-DIMENSIONAL MHD PROBLEMS WITH TRANSVERSE MAGNETIC FIELD [1]

K.V.Brushlinsky, T.A.Ratnikova

Keldysh Institute for Applied Mathematics, Russia Acad.Sci, Moscow
e-mail: oyufed@spp.keldysh.ru

Mathematical models, based on the MHD-approach, are successfully used in solving pure and applied problems in the hot and dense plasma physics, including nuclear fusion and astrophysics. But investigation of some phenomena, e.g. interaction of plasma flows with the electrodes, need more complicated models, at least because the MHD-equations cannot recognize the electrode polarity: they remain unchanged, when the electric current and the magnetic field invert simultaneously their directions. We consider below a modified MHD-model taking into account the Hall effect, that is a first step at the way from the usual ("one-fluid") MHD to the strickly two-fluid one. It is derived from the transport equations of ions and electrons [1] jointly with the Maxwell ones, if we neglect there the electron mass m_e and the displacement current $\partial E/c\partial t$ [2]. This model describes a continuous medium with the density $\rho = m_i n$, equal (for simplicity!) ten peratures and pressures of the components ($T_i \equiv T_e \equiv T$; $p_i \equiv p_e$; $p \equiv p_i + p_e$), different velocities $v_i \neq v_e$ and in general with the finite plasma conductivity. It is based on the set of equation:

$$\frac{\partial \rho}{\partial t} + \nabla(\rho \mathbf{v}) = 0 \ ;$$

$$\rho \frac{d\mathbf{v}}{dt} + \nabla p = \mathbf{j} \times \mathbf{H} \ ;$$

$$\rho T \frac{ds}{dt} = \nu j^2 + \frac{\xi}{2} T \, \mathbf{j} \nabla s \ ; \qquad (1)$$

$$\frac{\partial \mathbf{H}}{\partial t} = \nabla \times (\mathbf{v} \times \mathbf{H}) - \nabla \times (\nu \mathbf{j}) - \xi \nabla \times \frac{\mathbf{j} \times \mathbf{H} - \nabla p_e}{\rho} \ ;$$

$$p = \frac{\beta}{2} \rho T = \frac{\beta}{2} \rho^\gamma e^s \ ; \quad \mathbf{j} = \nabla \times \mathbf{H} \ ; \quad \frac{d}{dt} = \frac{\partial}{\partial t} + \nabla \mathbf{v},$$

that are written in a dimensionless form: units of measurement of all varial les are formed from their characteristic values in specific problems. There are here three dimensionless parameters:

$$\beta = \frac{8\pi p_0}{H_0^2} \ , \quad \nu = \frac{1}{Re_m} = \frac{c^2}{4\pi \sigma r_0 v_0}, \quad \xi = \frac{c}{e r_0} \frac{m_i}{\sqrt{4\pi \rho_0}}.$$

The exchange parameter ξ [2-4] is related to the Hall effect, that describes interaction

[1] This work was supported in part by International Science Foundation jointly with the Russia Federation Goverment under Grants No M3X000 and No M3X300 and by the Russia Fundamental Research Foundation under Grant No 94-01-01083.

between the current and the magnetic field in the Ohm's law and supposes the ion and electron velocities to be different:

$$\mathbf{v} \equiv \mathbf{v}_i; \qquad \xi \mathbf{j} = \rho(\mathbf{v} - \mathbf{v}_e) \ . \tag{2}$$

The generalized Ohm's law

$$\nu \mathbf{j} = \mathbf{E} + \mathbf{v} \times \mathbf{H} + \frac{\xi}{\rho} (\nabla p_e - \mathbf{j} \times \mathbf{H}) \ . \tag{3}$$

is the electron dynamic equation, where the electron inertia $m_e d\mathbf{v}_e/dt$ is neglected. It follow from (3) in particular, that the current and electric field are not collinear in general, and the plasma conductivity is nonisotropic, tensor-type.

The MHD-model, taking into account the Hall effect, is being used for example in investigation of plasma flows in channels of plasma accelerators, magnetoplasma compressors and MHD-generators [2-6], when the electrode polarity and near-electrode processes are essential. It is realized by the authors [7-10] in flow computations, connected with the elaboration of recent high-current quasi-steady-state plasma accelerator (QSPA) [11-15]. The accelerator channel (fig.1) is formed by two coaxial electrodes.

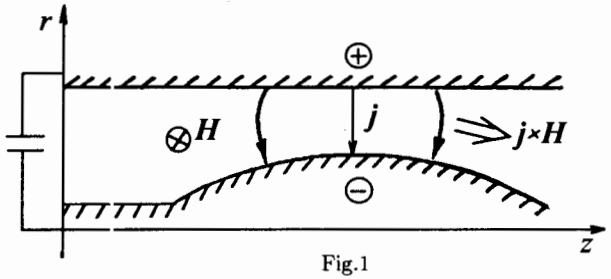


Fig.1

The electric current between them generates azimuthal magnetic field and, interacting with it, induces plasma to move in the z-direction. The nozzle-type channel shape insures the transonic plasma acceleration: the velocity gets over the fast magnetic sound speed in the minimal cross-section. Basic flow features are investigated in the MHD-approach [2,12,16], but some specific phenomena near the electrodes need to be simulated using the Hall effect.

The flows of interest are steady-state as a rule, since the discharge duration is much longer than the flow time. They are being calculated by means of the time-relaxation method. Plasma flows in coaxial channels (fig.1) may be considered as axisymmetric:

$$\frac{\partial}{\partial \varphi} \equiv 0 \ ; \quad v_\varphi \equiv 0 \ ; \quad H_z \equiv H_r \equiv 0 \ ; \ ,$$

i.e. they belong to the class of two-dimensional flows across the magnetic field. The equations (1) in general, second order, quasilinear, of parabolic type, degererate, being applied to this class and, when $\nu \equiv 0$, make up a first order nonlinear system: the Hall term in the fourth equation (1) consists only of first derivative products. By this reason

problems with finite plasma conductivity ($\nu > 0$) and infinite one ($\nu \equiv 0$) are different in setting boundary conditions at the electrodes. When $0 < \nu \ll 1$, solutions include Hall boundary layers, that are in direct relation with flow anomalies, detected n experiments: current sliding and potential jumps near the electrodes.

The paper [9] presents an analysis of the nonlinear system (1) ($\nu \equiv 0$) in terms of attendant linearized equations and gives conditions of its being hyperbolic (all characteristic velocities in any direction are real). These condition are usually fulfilled for steady-state plasma flows in QSPA, when $\xi \ll 1$. The characteristic velocities are found, and electrode boundary conditions in flow computation are agreed with them. The number and the content of the boundary conditions are different for the anode and the cathode. In particular, it is shown, that the equipotential electrode condition contradicts with the regular acceleration of the infinite conductive plasma.

In investigation of plasma flows in a channel including near-electrode zones, we take into account the finite high conductivity ($0 < \nu \ll 1$) and consider three variants of boundary conditions, that correspond to three principal types of flow-electrode nteraction.

1) Electron current transfer (ECT) regime

$$v_n = 0; \qquad E_\tau = 0, \tag{4}$$

where n and τ are the normal and tangent directions relative to the boundary. The electrodes are equipotential and unpenetrable for ions, the current is carried by electrons.In this case steady-state flow regimes may be obtained only if $\xi < \xi^*$, where the critical value ξ^* depends on ν and β. If $\xi > \xi^*$, the computation loses stability and leads to an explosion type phenomenon at the anode. Before the explosion we can see thin layers, where the current slides along the electrodes (fig.2) and the plasma density falls off near the anode.

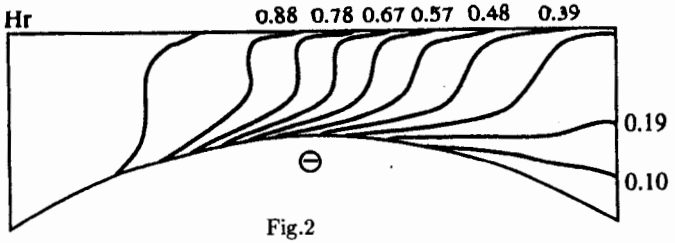


Fig.2

2) Ion current transfer (ICT) regime

$$v_{en} = 0; \qquad E_\tau = 0, \tag{5}$$

The plasma (ions) is allowed to penetrate through the electrode surfaces in amount $\rho v_n = \xi j_n$ (see(2),(5)), that is proportional to the current and the exchange parameter. This regime may be realized in cases of rod electrodes or more complicated devices – transformers [11]. In such a model steady-state flows are obtained under the condition $\xi < \xi^{**}$, but the restriction is less strong here: $\xi^* < \xi^{**}$. In the opposite case computation leads again to an explosion type instability and electrode boundary layer, as a rule, at the

cathode. A series of computations established the dependence of the exchange parameter critical values ξ^* and ξ^{**} on the magnetic viscosity ν in the both electrode polarity cases and for different values of parameter β [7]. They confirm the advantage of ICT ($\xi^* < \xi^{**}$) and showed, that the central position of the cathode admits steady-state flows with more strong Hall effect then the inverse polarity.

3) The most efficient plasma acceleration regime is connected with an "ideal" current distribution, ortl ogonal to unpenetrable streem boundaries

$$v_n = 0; \qquad j_\tau = 0. \tag{6}$$

As a result of co nputation with boundary conditions (6) there are obtained steady-state flow regimes witl finite values of ν, including $\nu \equiv 0$ [8]. No boundary layers appear in this case. A distinctive feature of such flows is that the stream boundaries are not equipotential: the electric field has a positive component, hence the potential decreases down the boundaries. This result may be useful for plasma accelerators with electrodes, consisting of separate sections: the voltage at the sections is to be agreed with computations. As applied to the QSPA, whose electrodes have longitudinal metal elements and therfore are equipotential, the computation result may be related to the "stream kernel", that is the most part of the channel. To determine its position in the channel space and calculate complements to the remaining part – thin (of order $\xi \ll 1$) zones near electrodes – the "smooth channel" approximation for two-dimensional steady-state flows [3,4] may be used. As a result we obtained a complete flow picture and electric current distribution in the channel in the optimal acceleration regime (fig.3).

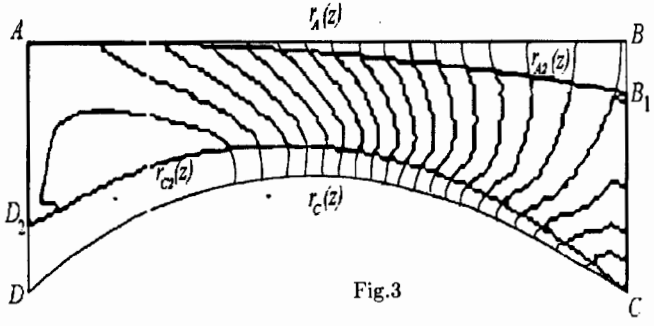


Fig.3

In particular, the flux densities of plasma passing through the anode and cathode, appropriate for this regime, are found [10].

References

[1] Braginsky S.I. Transport Phenomena in the Plasma, Reviews of Plasma Physics, ed. by M.A.Leontovich, **1**, 205, Consultant Bureau, N.Y.,(1965).

[2] Brushlinsky K.V., Morozov A.I., Reviews on plasma physics, ed. by M.A.Leontovich, **8**, 105, Consultant Bureau, N.Y.-London, (1980).

[3] Morozov A.I , Solovyev L.S., Ibid., 1.

[4] Morozov A.I. Physical basis of cosmic electric propulsion engines. M Atomizdat (in Rassian), (1978).
[5] Brushlinsky K.V., Morozov A.I., Savelyev V.V., Two-dimensional numerical plasma simulation, ed. by K.V.Brushlinsky, Moscow, Keldysh Inst.Appl.Math., USSR Ac.Sci, 7, (1979).
[6] Vatazhin A.B., Lubimov G.A., Regirer S.A., Magnetohydynamic flows in channels, M. Nauka, (1970).
[7] Brushlinsky K.V., Ratnikova T.A., Plasma Phys. Reports, **21(9)**, 743, (1995).
[8] Brushlinsky K.V., Ratnikova T.A., Fluid dynamics, **30(5)**, 690, (Consultant Bureau, N.Y.), (1995).
[9] Brushlinsky K.V., Ratnikova T.A., Matem. Modelirovanie, **8(2)**, 75, (1996).
[10] Brushlinsky K.V., Ratnikova T.A., Fizika plasmy, (in print), (1997).
[11] Morozov A.I., Sov. Y. Plasma Phys., **16(2)**, 55, (1990).
[12] Brushlinsky K.V., Zaborov A.M., Kozlov A.N. et al, Ibid., 73.
[13] Belan V.G., Levashov V.F., Mainashev V.S. et al, Fiz. plasmy, **17(12)**, 1511, (1991).
[14] Astashinsky V.M., Mankovsky A.A., Minko L.Ya., Morozov A.I. et al., Fiz. plasmy, **18(1)**, 90, (1992).
[15] Volkov Ya.F., Kulik N.V., Marinin V.V. et al., Fiz. plasmy, **18(11)**, 718, (1992).
[16] Brushlinsky K.V., Gorshenin K.P., Fiz. plasmy, **19(5)**, 682, (1993).

JAVA AND WWW FOR NEMO EXPERIMENT

A.D. Budnik and T.V. Filipova

Laboratory of Computing Techniques and Automation,
Joint Institute for Nuclear Research, 141980, Dubna, Russia
e-mail: budnik@linserv.jinr.ru
and
Laboratory of Nuclear Problems
Joint Institute for Nuclear Research, 141980, Dubna, Russia,
e-mail: filipova@nu.jinr.ru

Abstract To use the Java programming in WWW technology for NEMO experiment is proposed. We use Java code for presentation of NEMO software. Evolution of tracking algorithms is demonstrated: elastic net method and cellular automaton. Event visualization code was designed in Java to provide a monitoring system. We propose to use CGI program as a Log Book for NEMO. These applications are available through the Internet and allow all the collaboration members to be well informed about the present state of the experiment.

1 Introduction

Today large physics collaborations need a uniform and simple system to share and deliver information. WWW was invented at CERN and is progressively becoming the standard way to distribute information in High Energy Physics. The use of WWW technology in NEMO collaboration is proposed. The World Wide Web is rooted in the Internet technology. It uses the HyperText Transfer Protocol (HTTP) which allows clients to get documents from WWW servers. HTTP supports documents of different types thus permitting format negotiation between clients and servers. Documents are referred to by clients with Uniform Resource Locators (URL) based upon a naming scheme that guarantees unicity. Hypertext documents are formatted with the HTML language. The usual way to access HTML documents is using browsers: Mosaic, Netscape and others.

Interfacing to the World Wide Web becomes necessary when we want to do more then distributing static documents. Interfacing to the Web may be split up in two parts: server interfacing and client interfacing.

2 Log Book via WWW

As an example of the first way we propose practical use of the server interface program (CGI programming) as a Log Book (a Journal) of the experiment to collect all the information and the current status of the experiment[2]. We have a test program called GuestBook. You can look through the messages that somebody leaves in your database. You can leave your own message. This program runs on server site. How does it work? You fill a form via HTML document and send it to server. The server runs this GuestBook program and recognizes your information as an input for the database. It is possible to establish restricted access for collaboration members only. This is a simple example of such

program. It gives you only the list of all messages. It can be developed by introducing different search or sorting: for example, you can retrieve messages specifying the date, the author, the keyword.

3 Java programming language

3.1 What's Java?

Next step of our problem is if we want to have graphical presentation of experimental data. In this case we need a fast graphical interface with user. But the server interfacing is too slow to react on our action such as mouse move, click, key press. There is a possibility to have a client interfacing. That means the program for graphics is running on your local computer. Such a possibility is via using Java programming. Java is a new programming language. It was developed by SUN Microsystems[1].

Java is a new, object-oriented, general-purpose programming language, with elements from C, C++. It is platform independent as it is interpreted by virtual machine. Java is a system that has been designed to make it easy for users to master. It contains libraries that give simple convenient use of Internet features. Java specification includes standard on new tools of modern computing: graphics, audio, networking support. No need to have any separate graphics package like HIGZ which we use in our software. Java allows anyone to create a WWW page with code in it, that can be executed by anyone else accessing that page regardless of hardware. A program in a web page is called an "applet".

3.2 Life way of Java program

The Java source is compiled to *class* file, which contains the bytecode representation of the program. Bytecodes are very similar to the machine instructions, so Java programs can be very efficient. However, bytecodes are not specific to a particular machine, so Java programs can be executed on lots of different computers without recompiling the programs. Then this *class* file can be transported via the Internet as other documents and executed by any computer that supports Java. Java bytecodes are interpreted by the virtual machine. For example, such browser as Netscape of latest versions has a virtual machine inside itself. So the program of graphics user interface can run on your local computer and allows you to work interactively.

3.3 Security of Java

About security in Java. The reason that Java applets are secure is that Java programs are compiled into bytecode instructions that can be verified. Bytecode contains additional information that is used to verify that the program is legal. A verified program is guaranteed not to break any of the Java language constraints and can be safely executed. Verification is used by browsers to make sure that applets don't contain viruses, cannot crash or damage the computer. Java applets loaded via net have additional restrictions: they don't have an access to local file system.

4 Application of Java to NEMO problems

4.1 Events visualization

We created the Java program for visualization of experimental data for NEMO-34.1. User can download the Java code with a browser and run it. This program reads experimental data via Internet, draws setup and waits for the user action. The user can then interact directly with the event and is able to zoom, scale, move the event. So you can look at the events from any point of world and use this program for example as a remote monitoring. This program can be modified for the reading data that are taking at the moment and for the analysis of the apparatus status.

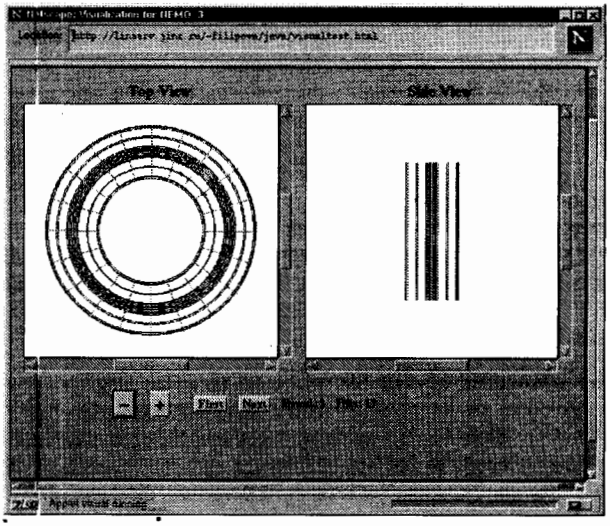


Figure 1: Visualization of NEMO 3 event using Java program.

4.2 Software presentation

One more field for application of Java to NEMO is a presentation of the NEMO software. Usually in a physics collaboration each member has his task and makes only his part of work. It is useful for everyone to have a simple way to be knowing about the programs and algorithms which are used to interpret experimental data. In NEMO experiment Elastic Net Method is used for reconstruction of charged particles tracks. We present you the applets that explain this method and demonstrate it in action[4, 3]. Any user can easily learn the idea of the algorithm and conclude that it really finds a good solution.



Figure 2: Demonstration of NEMO 3 tracking using Java program.

5 Conclusion

The analogous approach is already used at CERN. There is 'Online' Data Acquisition System of the DELPHI Experiment[5]. This is server interface system which is built as a CGI program so it allows restricted interface: only text information is available. If we would like to have something more the main problem of systems of that type arises: a slow performance.

To conclude, our suggestions are to :

- use CGI program as a Log Book of the NEMO experiment;
- use Java for visualization of NEMO 3 data;
- use Java for the NEMO software presentation.

References

[1] http://java.sun.com

[2] http://nuweb.jinr.dubna.su/LNP/NEMO/GuestBook/GuestBook.cgi?read

[3] http://nuweb.jinr.dubna.su/LNP/NEMO/nemoiiiN.html

[4] http://nuweb.jinr.dubna.su/LNP/NEMO/tspN.html

[5] http://www.ncsa.uiuc.edu/SDG/IT94/Proceedings/DDay/donszelmann/paper.html

Using cellular automata to model cave stuctures

M. Casolino*, P. Picozza
Dept. of Physics, II Univ. di Roma "Tor Vergata", Via della Ricerca Scientifica 1, I-00133, Rome, Italy
INFN, National Institute of Nuclear Physics, Sec. Roma II, Italy
e-mail: casolino@roma2.infn.it, picozza@roma2.infn.it
* Corresponding author

Abstract. A Cellular Automaton method is used to model underground caves where the structures are predominantly formed by rock collapse. The vertical cave profile is generated on a square lattice with random configurations where 70% of the sites represents rock elements and the remaining 30% represents air. The system is iterated according to rules which take into account gravity and mechanical interactions. After convergence, the scaling properties and corresponding fractal dimensions are estimated and compared with real cave profiles. In spite of the approximations used (all water erosion processes are neglected) there is a satisfactory agreement between simulated and real data.

PACS: 91.35.cb, 07.05.Tp, 93.85+q, 91.60.-x

Underground cave formation is due to several concurring processes, usually taking place over several length and time scales. These processes can be roughly divided in: 1, Limestone erosion by water; 2, Mechanical erosion by underground rivers; 3, Rock fall due to gravity. In a complex, structured cave it is possible to distinguish regions that are a, completely devoid of water; b, partially filled with water; c, totally filled with water. Processes 1 and 2 are associated with regions b and c, usually younger and below regions of type a where rock collapse is dominant and often drastically changes the shape and ramifications of the underground cavities determined by previous water erosion phenomena[1]. According to its history, a cave may have greatly varying shapes and size, ranging from the relatively simple vertical pits or underground rivers to majestically large environments of several kilometers, developing on several levels at different dephts.

Over the years several simulations and studies on the fractal nature of rivers and rocks have been carried out [2], here we present a model simulating the formation of caves: for their simplicity and speed Cellular Automata (CA)[4] have been used. Their versatility has made CA a widely employed tool in several fields, ranging from the modeling of physical systems to noise removal and signal processing in particle detectors[3]. To speed up calculations and to employ a binary state (rock/air) CA, all water erosion processes in limestone have not been considered in the present application. These simulations apply therefore only to caves where the erosion role of water has taken place so far in the past that its structural characteristics have been reshaped by rock fall.

A dynamical study of the development and formation of the cave would have to take into account a three state system (air, rock and water) where the partial pressure of CO_2 takes a critical role in determining the amount of diluted calcium carbonate[1]. In this case we are concerned with the static configurations which are produced by gravitational and mechanical forces: as initial conditions a percentage $P = 30\%$ of empty (air) sites has been assumed. In our case the two states that each site (or cell) of the binary automaton can assume represent, respectively, the presence (*State 1*) or absence (*State 0*) of rock.

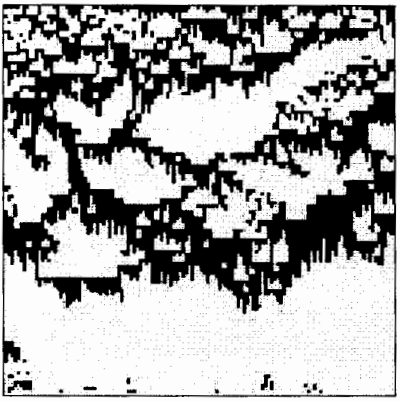


Figure 1: An example of a 128×128 structure obtained after convergence of the CA (each pixel pictures a lattice site). The vertical axis represents the depth below ground. For clarity, rock sites have light colour, while air sites are black.

Vertical cave profiles were simulated in order to compare them with those of real caves: the evolutionary rules have been devised to take into account gravity and mechanical interaction among nearby rocks. The gravitational forces move rock sites vertically toward the lower row if the rock is not stopped by its directly lower element or by at least two upper and one lateral element (on the same side).
The CA is updated until convergence in a stable state (no more rock movements). Several underground structures resembling a cave profile and showing underground chambers and passages have taken place, as shown in one case in Figure 1; the fractal nature of the resulting system is evident[1]. It is to be noted how vast air structures tend to be formed at larger depths, while smaller formations appear closer to the surface. As with real cave profiles, the presence of apparently suspended rock is to be attributed to a orthogonal projection effect, to passageways that are not necessarily one above the other but in different positions along the z axis (orthogonal to the plane of the paper). It is also clear from the resulting pictures how the system is invariant under X (horizontal) axis translations, in accordance with the properties of the update rules. This invariance also produces a great number of "cave entrances", holes leading from the surface lead to larger underground structures. As most speleologists sadly know, this is very far from real conditions, where entrances leading to huge underground complexes are extremely rare and

[1]The structures formed present some similarity with the 'class 3' (chaotic) 1D+1D automata present in [4]. The patterns that appear in that case are due to the presence of some particular combinations of cells that produce the observed 'empty' structures. The analogy with our case is due to the fact that only particular structures are stable from the mechanical point of view and do not collapse under gravity. The rock fall could be therefore be seen as a propagation of that initial structure forward in time (that is the Y axis in the cave model)

X size	Y size	D_a
64	64	1.71 ± 0.05
128	128	1.71 ± 0.03
256	256	1.76 ± 0.02
256	512	1.78 ± 0.01
512	256	1.78 ± 0.02
512	512	1.78 ± 0.01
1024	1024	1.83 ± 0.01

Table 1: Fractal dimension of the air (D_a) zones for grids of different size.

hard to find. This depends on the boundary conditions chosen; infact a layer of completely filled space close to the surface would have produced results with fewer entrances. The quantitative study of this model has been performed analyzing the scaling properties and fractal dimensions of the air sites using the box counting method. Several different grids, from 64×64 up to 1024×1024 elements, have been considered to study the dependence of the fractal dimension from the size of the system. In each case a set of $N = 100$ stuctures have been generated.

The conversion of each grid element to meters, that is the lenght represented by each lattice site, is not immediate: the same update rules considered are indeed valid for the structural and mechanical properties of rocks of different sizes. Experimental evidence for rock collapse with boulder size from $\simeq 50\ cm$ (i.e. Pozzo Arcaro, It) to $\geq 5m$ (Grotta dell'Ovito, It) can be easily found. Moreover, the overall size of the underground structure simulated must be finite: for instance, a lattice of 1024 sites - each of about $1\ m \times 1\ m$ - implies a cave depht of $1\ km$. These kind of caves, albeit found in nature, are not very common and have been formed by several different concurring processes. Therefore simulations on a larger grid would not find their geological counterpart. These limitations do not pose a real problem since the fractal dimension of the simulated caves will be shown to be constant over several scales, so that a comparison with experimental values will be possible without having to fix a precise value for size of the lattice elements. Let us come back to the scaling properties of the systems considering the empty elements (air). The scaling properties for grids of different sizes are shown in Figure 2 where the number of elements $N(\lambda l)$ present at each scale λl is shown (l is the size of a lattice site and λ takes the values of integer power of 2: 1,2,4...). $N(\lambda l)$ is the average taken over all the $N = 100$ simulations performed for each grid. Errors (usally smaller than the symbols shown) represent 1 $r.m.s.$ from mean values. Since the relation between these values and the fractal dimension D of a system is[5] $N(\lambda l)/N(l) = \lambda^{-D}$, a linear fit in double logarithmic scale yields the estimated value of D. Note that the fit is in good agreement with data points over all statistically significant lenght scales. The resulting fractal dimensions for the empty area zone D_a are shown in Table 1. It is possibile to see how the value of D_a slowly grows as the grid size increases: this is due to the formation of larger structures. The dependence of the fractal dimension from the grid size is linked to the finiteness of the systems considered and finds its counterpart in the variety of size,

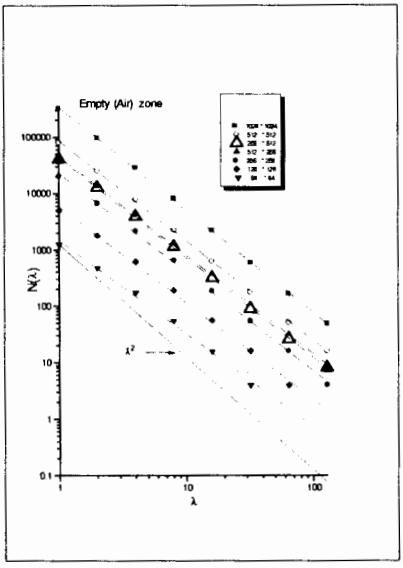


Figure 2: Scaling properties for different size grids of the simulated caves. The number of air sites present at scale λ is plotted as a function of the scale λ. The absolute value of angular coefficent of the linear fit in double logarithmic scale represents the fractal dimension D_a. Note how the fit is consistent with data points almost over all scales.

structure and value of D_a of real caves.

A detailed study of the dependence of the value of D_a from the percentage of air sites P was carried out in order to determine the values which - under these update rules - were more consistent with real caves. It is to be expected that in the limit of $P \to 0\%$ the air sites are so far away one another that there is little rock collapse and clustering of air sites with and no subsequent cave formation, whereas in the limit of $P \to 100\%$ the amount of empty space is such that all the few rocks present fall to the "bottom" of the grid. Simulations performed on a 64×64 grid comfirm this hypotesis, as shown in Figure 3 where the value of D_a tends to 2 and to 1 for $P \to 0\%$ and $P \to 100\%$, respectively. Values of $P < 25\%$ have very big statistical errors due to a very large spread in the behavior depending on the initial conditions: indeed only values of $25\% < P < 35\%$ seem to produce meaningful results. If $P > 35\%$ the empty structures formed are too big and different with real caves. There is a connection between this system and percolation models. It is indeed possible to note how, in order to obtain a model with a certain degree of complexity and physical meaning we need to retain some kind of connectivity which is related both to the past erosion role of water and the 'explorability' of real caves (that

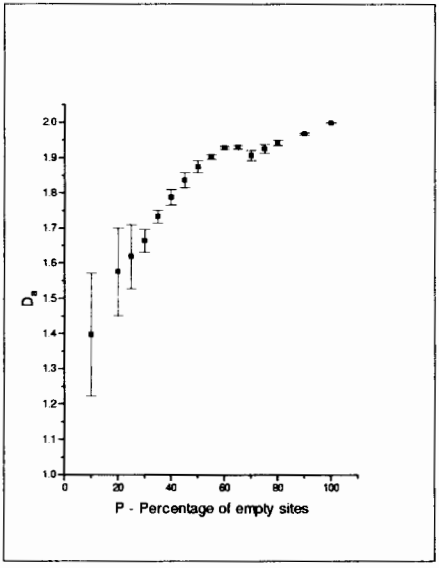


Figure 3: Values of the fractal dimension D_a for different air size percentage P in a 64*64 grid.

is, only those parts of underground complexes which are connected can be explored and studied). However, unconnected systems with $P < 25\%$ (that is below the percolation critical exponent), are not devoid of physical meaning. They represents those formations where the excavation of water has taken place over a too small time scale (or the erosion has not taken place for some reason). On the other hand systems with high P can be viewed as fully excavated karst landscapes.

The same scaling studies have been performed on real cave profiles as published in the bulletins of speleological societies. The chosen caves present absence of water and dominance of gravity effects in their structure; they have been scanned with a $300dpi$ resolution and thresholded in order to remove shades of grey. The problem in these profiles is that they arise from measurements often taken by speleologists in extremely difficult conditions and subject to errors. This is mostly true for the rock - air boundary drawn by hand and therefore not very reliable, whereas, the lenght and area size measurements are usually far more accurate. For this reason the properties for the boundary lines of the simulations have not been presented.

The values of D_a have been estimated using the same method described above for five caves of different size and origin. The overall characteristics and resulting values of D_a have been resumed in Table 2; the scaling properties are also shown in Figure 4 for

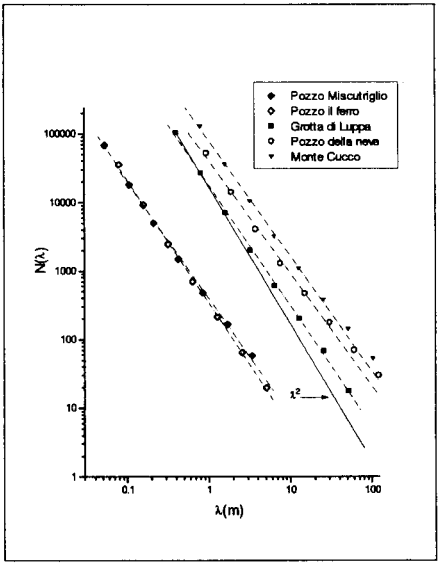


Figure 4: Scaling properties evaluated from five real cave profiles. The number of off (air) sites present at scale λ is plotted as a function of the scale λ. The absolute value of angular coefficent of the linear fit in double logarithmic scale represents the fractal dimension D_a. Also in this case the fit is representative of data over all scales.

comparison with the simulations. The errors are those calculated from the fit. There is a satisfactory agreement with the simulated data. Some trends can be seen: the cave with the most complex system of galleries (Monte Cucco) also exibits the lowest value of D_a, whereas very small and simple have higher fractal dimensions.

A more systematic work will be carried out in the near future with the analysis of the different problems arising from the determination of D_a in real cave profiles and with more complex models which take into account water erosion processes. In this paper we have shown how a simple model of cave formation may be used in the prediction and understanding of the fractal nature of underground structure formation.

Name	Ref.	h_{size}(m)	v_{size}(m)	D_a	Notes
Pozzo Miscutriglio	[6]	35	- 55	1.70 ± 0.03	simple str. due to rock collapse
Pozzo il Ferro	[6]	20	- 50	1.79± 0.02	simple str. due to rock collapse
Grotta di Luppa	[7]	1024	-174	1.76 ± 0.03	underground river str. mostly due to water eros.
Grotta di Monte Cucco	[8]	1400	-900	1.68 ± 0.03	very complex multilayered str. due to fault and rock collapse
Pozzo della Neve 0.6	[8]	1900	-800	1.64 ± 0.5	long winding canyon due to water erosion

Table 2: Fractal dimension of the air (D_a) zones for caves of different size and structure.

References

[1] M. J. Selby, *Earth's Changing surface*, Cap. XI (Clarendon Press, Oxford, 1989).
A. G. Segre, *I fenomeni carsici e la speleologia del Lazio*, (Pubbl. Ist. Geogr. Univ. Roma, 1948)
U. Sauro, in G. Castiglioni *Geomorfologia*, Cap. IX, (UTET 1987).

[2] S. Kramer, M. Marder, Phys. Rev. Lett. **68**, 205 (1992).
A. Rinaldo *et al.* , Phys. Rev. Lett. **70**, 822 (1993).
P. Wong, J. Howard, J. Lin, Phys. Rev. Lett. **57**, 637 (1986).
J.P. Hansen, A. T. Skjeltorp, Phys. Rev. B **38**, 2635 (1988).
H. F. Meng, E.G. D. Cohen, Phys. Rev. E, **51**, 3417 (1995).

[3] M. Casolino, P. Picozza, NIM **A** 364 (1995) 516-523.

[4] S. Wolfram, *Theory and Applications of Cellular Automata*, World Scientific, Singapore, (1986), p. 91.

[5] D. Veneziano, G. E. Moglen, R. L. Bras, Phys. Rev. E, **52**, 1387 (1995).
K. Falconer, *Fractal Geometry*, John Wiley & Sons, Chichester, (1990).

[6] G. Trovato, Notiziario del Circolo Speleologico Romano, **XIII**, n. 15, 5 (1968).

[7] M. Dolci, Marzolla, Notiziario del Circolo Speleologico Romano, **IX**, n. 10, 2 (1963).

[8] A. Bonucci, *Guida alla Speleologia*, Editori Riuniti, Roma (1983).

CALCULATION OF CANONICAL NORMAL DISTRIBUTIONS ON THE GROUP SO(3) AND THEIR APPLICATION

E.A. Davidzhan , T.I. Savyolova , T.M. Ivanova

MEPHI, Moscow, Russia

The orientation distribution function (ODF) of a polycrystalline sample is of special interest in texture analysis, but ODF is seldom if ever determined in a direct manner. The mathematical problem of quantitative texture analysis is calculation the ODF from experimentally obtained pole figures (PF). With any sets of PFs the solution of this problem is non-unique. Therefore we need an information about ODF structure. We consider ODF as superposition of the canonical normal distribution (CND) on the rotation group SO(3) [1].

Definition of the canonical normal distributions on SO (3) and S^2

Let us consider the normal distributions on the rotation group SO(3) as defined in [2,3].

Definition 1. A distribution $d\mu(g)$, $g \in SO(3)$, is said to be normal (Gaussian) if it admits representation of the type

$$\int_{SO(3)} T(g)d\mu(g) = \exp\Big(\sum_{i,j=1}^{3}\alpha_{ij}A_iA_j + \sum_{i=1}^{3}\alpha_i A_i\Big), \qquad (1)$$

where T (g) are the irreducible representations of the group SO(3); A_i, i=1,2,3, are the infinitesimal operators of representations T (g); (α_{ij}), i, j=1,2,3 is a real positive definite or semidefinite matrix and α_i, i=1,2,3, are real constants.

Definition 2 The normal distribution $d\mu(g)$ on SO (3) is said to be the canonical normal distribution (CND), if we have α_{ij}=0, $i \neq j$, α_i=0, i,j=1,2,3 in the expression (1).

The normal distribution $d\mu(g)$ $f(g)$ dg of general type can be transformed to CND by two rotations. Here $f(g)$ is the density of normal distribution, dg is the invariant measure on SO(3).

Let $g\{\alpha, \beta, \gamma\}$, $-\pi \le \alpha,\gamma < \pi$, $0 \le \beta \le \pi$ be three Eulerian angles that describe three rotations of Cartesian coordinates (X,Y,Z) (right-handed) by rotation through angle α, $-\pi \le \alpha < \pi$, around axis OZ, then by rotation through angle β, $0 \le \beta \le \pi$ around axis OX and by third rotation through angle γ, $-\pi \le \gamma < \pi$, about axis OZ.

We have $dg = \frac{1}{8\pi^2} \sin\beta \, d\alpha d\beta d\gamma$.

Normal distribution $d\mu(g)$ may be represented by series

$$f(g) \sum_{l=0}^{\infty} \sum_{m,n=-l}^{l} (2l+1) C_{mn}^{l} T_{mn}^{l}(g) , \qquad (2)$$

where $T_{mn}^{l}(g)$ are generalised spherical functions, C_{mn}^{l} are corresponding Fourier-type coefficients that can be calculated from expression (1) :

$$C^{l} \equiv (C_{mn}^{l}) = \exp(B^{l}) , \qquad l=0,1,2,3\ldots .$$

For CND we have

$$B^{l} = \sum_{i=1}^{3} \alpha_{ii} (A_i^{l})^2 = (b_{mn}^{l}) , \quad l=0,1,2,3\ldots , \qquad (3)$$

where $b_{mm}^{l} = -\frac{1}{2}\left[l(l+1) - m^2\right](\alpha_{11} + \alpha_{22}) - m^2 \alpha_{33}$,

$$b_{m,m-2}^{l} = b_{m-2,m}^{l} = b_{-m,-m+2}^{l} = b_{-m+2,-m}^{l} = \frac{1}{4}\left\{(l-m+1)(l+m-1)(l-m+2)(l+m)\right\}^{1/2}(\alpha_{22} - \alpha_{11}) ,$$

$m = -l, -l+1, \ldots, l-1, l$, and the rest $b_{ij}^{l} = 0$.

If $\alpha_{ij}=0$, $i \ne j$, $\alpha_i = 0$, $\alpha_{ii} = \varepsilon^2$, $i,j=1,2,3$, we obtain the central normal distribution on SO(3) from formulas (1), (2)

$$f(g) = \sum_{l=0}^{\infty} (2l+1) \exp\{-l(l+1)\varepsilon^2\} \frac{\sin[(l+1/2)t]}{\sin(t/2)} , \qquad (4)$$

where $\cos(t/2) = \cos(\beta/2)\cos[(\alpha+\gamma)/2]$.

The normal distributions on the sphere in R^3 can be obtained as "projections" from normal distributions on the group SO(3) [1] .We can get the functions on S^2 if we consider the functions on SO(3) that are constant on right-hand contiquity classes for the subgroup of rotation with respect to a certain axis.

Definition 3. We get a normal distribution $f(\xi)$, $\xi=\{\alpha,\beta\}\in S^2$, $-\pi \le \alpha < \pi$, $0 \le \beta \le \pi$, on sphere S^2 in R^3 from normal distribution $d\mu(g)$ on the group SO(3) , where $\alpha_i=0$, $\alpha_{ij}=0$, if $i=0$ or $j=0$, $i, j=1,2,3$,

$$f(\xi)d\xi = \frac{1}{8\pi^2}\int_0^{2\pi}\left[f(g)\sin\beta d\alpha d\beta\right]d\gamma = \frac{1}{4\pi}\sum_{l=0}^{\infty}\sum_{m=-l}^{l} C_{m0}^{l} T_{m0}^{l}(g)\sin\beta d\alpha d\beta . \qquad (5)$$

If $\alpha_{ij}=0$, $i\neq j$, $\alpha_i=0$, $\alpha_{ii}=\varepsilon^2$, $i, j=1,2$, in expression (5),
we have the central normal distribution on S^2

$$f(\beta)=\sum_{l=0}^{\infty}(2l+1)\exp\{-l(l+1)\varepsilon^2\}P^l(\cos\beta) \quad , \qquad (6)$$

where $P^l(\cos\beta)$ are Legendre polynomials.

The properties of normal distribution on SO(3)

Using the properties (3) of B^l we can get next note for car onical normal distributions

$$\begin{aligned} f(\alpha,\beta,\gamma) = \sum_{l=0}^{\infty}(2l+1)\{ & C_{00}^{l} P_l(\cos\beta) + 2\sum_{n=1}^{l} C_{nn}^{l} P_{nn}^{l}(\cos\beta)\cos[n(\alpha+\gamma)] + \\ & + 2\sum_{n=1}^{l} C_{n,-n}^{l} P_{n,-n}^{l}(\cos\beta)\cos[n(\alpha-\gamma)] + \\ & + 4\sum_{n=2}^{l}\sum_{k=1}^{n-1} C_{n,-n+2k}^{l} P_{n,-n+2k}^{l}(\cos\beta)\cos[k(\alpha+\gamma)]\cos[(n-k)(\alpha-\gamma)]\ \}, \end{aligned} \qquad (7)$$

where $-\pi \le \alpha,\gamma < \pi$, $-\pi \le \beta \le \pi$, $P_{mn}^{l}(\beta)$ are Jacoby polynomials.

From formula (7) we get the properties of CND

1) symmetry with respect to angles α and γ , $-\pi \le \alpha,\gamma < \pi$,

$f(\alpha,\beta,\gamma)=f(\gamma,\beta,\alpha)$;

2) property of being even with respect to angles α and γ , $-\pi \le \alpha,\gamma < \pi$,

$f(\alpha,\beta,\gamma)=f(-\alpha,\beta,-\gamma)$;

3) periodicity with respect to angles α and γ, $-\pi \le \alpha, \gamma < \pi$,

$$f(\alpha,\beta,\gamma)=f(\alpha+\pi,\beta,\gamma+\pi)=f(\alpha+\pi,\beta,\gamma-\pi).$$

From properties 1), 2), 3) follows that it is possible to reduce the volume of calculations in 8 times -instead of calculation $f(g)$ in the domain $-\pi \le \alpha, \gamma < \pi$ we can consider the domain

$$G=\left\{(\alpha,\gamma)\;\middle|\;\begin{array}{ll} -\gamma \le \alpha \le \gamma, & 0 \le \alpha \le \pi/2 \\ -\pi+\gamma \le \alpha \le \pi-\gamma, & \pi/2 \le \alpha < \pi \end{array}\right\}.$$

4) The peak of CND $f(\alpha,\beta,\gamma)$ achieves at $\beta=0, \alpha+\gamma=0$ and it is equal to

$$\max f(\alpha,\beta,\gamma)=f(0,0,0)=\sum_{l=0}^{\infty}(2l+1)\sum_{m=-l}^{l}C_{nn}^{l}=\sum_{l=0}^{\infty}(2l+1)Sp(\exp(\Lambda^{l})).$$

5) At $\beta=\pi$ we have

$$f(\alpha,\pi,\gamma)=\sum_{l=0}^{\infty}(2l+1)(-1)^{l}\{C_{00}^{l}+2\sum_{n=1}^{l}C_{n,-n}^{l}\cos[n(\alpha-\gamma)]\}.$$

Let's consider dependence of CND from parameters $(\alpha_{11},\alpha_{22},\alpha_{33})$.

We shall designate $f(g)=f(g,\alpha_{11},\alpha_{22},\alpha_{33})$.

6) Let the rotation $g_0 \in SO(3)$ leaves the invariant square-law form

$\alpha_{11}X^2+\alpha_{22}Y^2+\alpha_{33}Z^2$, where X, Y, Z are cartesian coordinates.

Then we have the equality

$$f(g)=f(g_0 g g_0^{-1}).$$

Such type of rotations have the kind $g_{01}=(\alpha,\pi,\gamma)$ when $\alpha-\gamma=\pi$, $g_{02}=(\alpha,\pi,\gamma)$ when $\alpha-\gamma=0$ and $g_{03}=(\alpha,0,\gamma)$ when $\alpha+\gamma=\pi$.

8) If we have the transformation $g \to g_1 g g_1^{-1}$, $g_1=(0,\pi/2,0)$ then the parameters α_{22} and α_{33} are replaced one by another

$$f(g,\alpha_{11},\alpha_{22},\alpha_{33})=f(g_1 g g_1^{-1},\alpha_{11},\alpha_{33},\alpha_{22}).$$

If the transformation $g \to g_2 g g_2^{-1}$, $g_2=(\pi/2,\pi/2,-\pi/2)$, is used then the parameters α_{11} и α_{33} must be replaced one by another

$$f(g,\alpha_{11},\alpha_{22},\alpha_{33})=f(g_2 g g_2^{-1},\alpha_{33},\alpha_{22},\alpha_{11}).$$

Respectively, we have

$$f(g,\alpha_{11},\alpha_{22},\alpha_{33}) = f(g_3 g g_3^{-1},\alpha_{22},\alpha_{11},\alpha_{33}),$$

where $g_3=(\alpha,0,\gamma)$ with $\alpha+\gamma=\pi/2$.

8) From the properties 4) and 7) we get $\max f(g) \; f(e)$, where e $(0,0,0)$ and does not depend from the order of the parameters $(\alpha_{11},\alpha_{22},\alpha_{33})$.

The conclusion

Owing to simplicity of approximation ODF by normal distributions the central normal distributions on the group SO (3) are used (4).

Development of computing opportunities permits to use distribution of a general kind for approximation ODF ,such as except for rare cases the experimental pole figures have distribution differs from circular character. The problem of the determination of parameters of distributions in case of approximation by the central normal distributions and in case of approximation by the CND of a general kind is nonlinear and at a initial stage is decided by a trial error method .

The present work gives the simple formula for calculation CND (7) and number of their properties 1) - 8), that are useful for approximation CDF.

The given work is executed pursuant to the program GRUNT from State Committee Institution of Higher Education in Russia on fundamental researches in the field of a geology.

REFERENCES

[1] Savyolova T.I. Textures and Microstructures, 1993, 22, 17-27.
[2] Parthasarathy K.R. Theory of probabilities and its Application. 1964, 9, 273-282.
[3] Savyolova T.I. Zavodskaya Laboratoriya , 1984,50,48-52.

USING "TG-GSS" DIALOGUE-GRAPHICS PLATFORM FOR PROCESSING THE COSMIC GAMMA-RAY SPECTRA IN CORONAS-I SPACE EXPERIMENT.

A.V. Dmitriev, S.N. Kuznetsov, S.P. Ryumin

NPI, MSU, Russia
e-mail:dalex@srdlan.npi.msu.su

P.N. Zaikin, T.B. Podosenova, I.I. Petsilyak

CMC, MSU, Russia
e-mail:pnz@cs.msu.su

Yu.D. Kotov, A.S. Glyanenko

MePhI, Russia
e-mail:kotov@mx.iki.rssi.ru

Abstract. The results of cosmic gamma-ray spectra processing by means of CASE (computer applied software environment) technology are presented. The gamma-ray spectra are detected by the SONG (SOlar Neutrons and Gamma-rays) and ATS (Amplitude and Time Spectrometer) instruments onboard CORONAS-I satellite. This computer technology is an example of screen editing the subject domain and software and allows to form the intellectual work station (IWS) for experimental data processing and operating under their graphical imaging. The particle acceleration, propagation and absorption on the Sun and other astrophysical objects are investigated by the SONG- and ATS instruments. One of the investigation methods is analysis of spectra of gamma-lines formed in nuclear interactions between the accelerated particles and environment matter. These lines are detected simultaneously with a large gamma-ray background. It arise in the interactions of protons of galactic cosmic rays and of the Earth radiation belt with Earth atmosphere, as well as with the material of the spacecraft and detector itself. The IWS for physicist-investigator is designed by using of TG-GSS platform. The main aim of the IWS is the restoration of gamma-ray spectra from gamma spectrometer data. This restoring procedure includes: gamma-ray background analysis, time dependent instrument calibration, computer searching for solar and celestial gamma-ray lines, studying their spectral and intensity dynamics. The linear-exponential dependence is chosen as the mathematical model of instrument calibration function. To calculate the calibration parameters the original version of algorithm of searching for resonance inhomogeneities (lines) which is based on the inter-covariance function analysis is used. The data processing is realised in the form of prepared by experts TG-GSS software metastructures. The last ones include, in particular, the data editing and running utilities. The operator has a possibility to check and control the main analysis stages and to get a computer system dialogue. The computer system allows in many cases to avoid the spectrum distortions caused both by external apparatuses and by own instrument electronic circuits. The IWS possibilities were presented and discussed last year in following centres: Navy Research Laboratory, Maryland, Washington DC, USA; Marshall Space Fly Centre, Space Sciences Laboratory, Huntsville, Alabama, USA.

1 INTRODUCTION

To investigate cosmic gamma-ray spectra scintillation crystal CsI(Tl) 20cm diameter by 10 cm thick was used. This crystal was installed in SONG-instrument. The large crystal

volume causes a high intensity of gamma-ray background. Intensities of gamma-ray fluxes from the Sun and other astrophysical objects are as a rule very small. Therefore a good enough data processing method is necessary to interpret the experimental data.

The space detector tuning can be changed during a long time flight. Because of that regular detector calibration based on the measuring of background gamma-line positions is necessary for restoring and studying of gamma-ray spectra.

The gamma-line spectra measurements can be accompanied by distortions connected with pollution from other devices and by operation distortion of instrument itself. The useful signal extracting is a very important stage in gamma-ray restoring process. Dialogue-Graphic system "TG-GSS" (DG "TG-GSS") [1] is applied to form the intellectual work station (IWS) for solution all this problems in the gamma-ray spectrometer data processing.

2 METHOD

The IWS is based on the sequence of successive stages of calculations. The setting up parameters of data processing and initial conditions is done by operator investigator in dialogue regime (DG TG-GSS provides this regime).

The aim of IWS first stage of calculations is selection and summarising of measured gamma-ray spectra to get the highest statistical accuracy of experimental data. Investigator has a possibility of visual selection of necessary (suitable) gamma spectra. The steps of this stage of calculations are realised by means of special algebra of operations under digital functions like time series.

The summarised averaged spectrum is used at the second IWS stage of calculations. The last one is aimed to gamma-lines searching for. This stage is based on the original version of algorithm of resonance ingomogeneities searching for and of estimating their location areas; this algorithm is dealing with the inter-covariance function analysis. The IWS possibilities allow to investigator to change the sensitivity of gamma-lines searching method and therefore permits to omit the instrument distortion effects in the gamma-ray spectrum.

At the third ISW stage one calculates a detector energy calibration function by using of gamma- lines picked out at the previous stage; the energy values corresponding to selected gamma-lines are giving by investigator. The operator has a possibility to choose the way of approximation of calibration dependence (i.e. energy as a function of channel). The calibration accuracy is calculated by using of deviations between simulated gamma-line energy values and initially given ones. The restoring procedure of energy dependent gamma- spectrum uses the obtained calibration function.

3 RESULTS

The first IWS stage is used to form the summarised averaged initial gamma spectrum. Investigator selects and summarises the low intensity gamma-ray spectra detected by SONG-instrument in the equator region in 1994.

Gamma-lines searching for is doing at the IWS second stage. The low level of method sensitivity permits to omit the instrument distortion effects and to pick out the four most

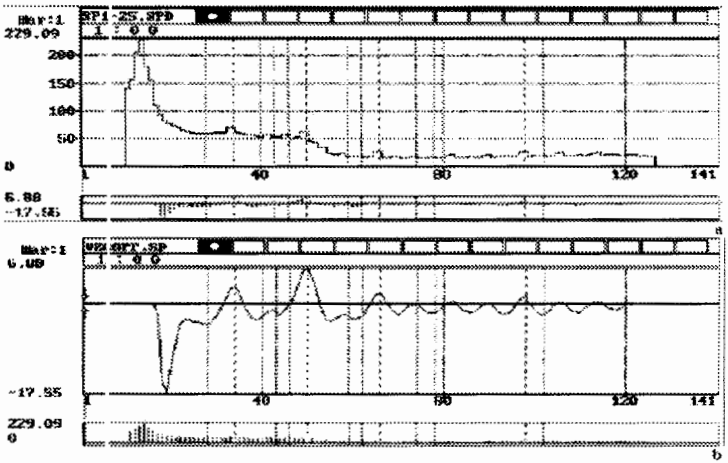


Fig. 1. The result of gamma-ray peaks searching operation in initial spectrum (histogram): a) centres and localization area bounds of picked out gamma-ray peaks are marked by dashed lines; b) statistical weight of picked out gamma-ray peaks.

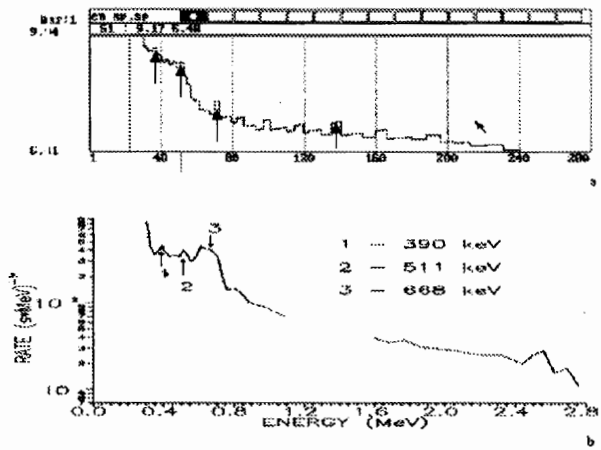


Fig. 2. Restored gamma-ray energy spectrum observed by CORONAS-I in January 1995 (a) and gamma-ray energy spectrum detected by GRS SMM instrument in 1982.

Table 1.

Line Number	Line Centre (Channel)	Line Energy (MeV)	Calculated Line Energy (MeV)	Deviation (%)
1	33.65	0.390	0.377	3.3
2	49.72	0.511	0.521	1.9
3	65.59	0.668	0.718	7.5
4	97.71	1.4	1.374	1.9

intensive gamma-ray lines. These gamma-lines correspond to galactic cosmic ray particles and trapped radiation interactions with detector crystal matter (Iodine) and with satellite matter (Aluminium) [2,3]. The results of this stage of calculation are presented at Fig. 1. The initial averaged spectrum (histogram), picked out gamma-ray lines (marked by dashed lines) and their areas bounds (marked by dotted lines) are shown at Fig. 1a. The corresponding inter-covariance function which is used to estimate the statistical weights of gamma-lines is shown at the Fig. 1b.

The gamma-spectrum calibration stage uses the values (in channels) of centres of picked out gamma-line peaks (see second column of Tab. 1) and corresponding to the peaks energy values (see the third column of Tab. 1) [2,3]. The calculated (simulated) energy values of gamma-lines and the deviation from initially given energy values are presented in the forth and fifth columns of Tab. 1 respectively. As we can conclude from Tab. 1 the calibration relative deviation is less than 10energy region is about 10-16 percentages so the accuracy of calibration stage is satisfactory.

Instrument calibration is a dependence of the gamma-quantum energy E on the order number of channel N in which the signal from gamma-quantum is detected. In the case of logarithmic impulse analisator, calibration stage presents the calibration dependence as exponent:

$$E(MeV)=0.19*\exp(0.02*N) \quad (1)$$

The result of calibration stage is a normalised energy gamma-spectrum restored from initial spectrum (Fig. 2a). The arrows at Fig. 2a mark the centres of reference gammalines. Fig. 2b shows the gamma-spectrum measured by the analogous detector GRS onboard satellite SMM in 1982 [3] for comparison with our results. As it can be seen from Fig. 2a and Fig. 2b the experimental data obtained by two detectors in the different time periods well coincide.

4 CONCLUSIONS

1. DG TG-GSS allows quickly and comfortably to select and to sum up the gamma spectra and also to identify the gamma-lines by using of dialogue mode.

2. The gamma-spectrum calibration is carrying out by TG-GSS the most effectively with accuracy that satisfies to the experiment conditions on the SONG instrument.

3. TG-GSS realises the possibility of gamma-spectra restoration by using of calculated calibrations, that allows to use them to analyse the gamma-spectra of solar and astrophysical origin.

References

[1] P. N. Zaikin, T. B. Podosenova. "TG-GSS" Manager to Generate Applied Software for Analysis of Data Like Set of Time Series and Ill-Posed Inverse Problems. (this issue).

[2] G. Share, R. Kinzer and M. Strickman. Instrumental and Atmospheric Background Lines Observed by the SMM Gamma-Ray Spectrometer, AIP Conference Proc. **186**, High-Energy Radiation Background in Space, New York (1989).

[3] D. Gruber, G. Jung and J. Matteson. Radioactivity Observed in Scintillation Counters During the HEAO-1 Mission, AIP Conference Proc. **186**, High-Energy Radiation Background in Space, New York (1989).

VISUALIZATION OF RELATIVISTIC STRING DYNAMICS [1]

Dyachin V.V., Klimenko S.V., Nikitin I.N.

Institute for High Energy Physics, Protvino, Russia
e-mail: dyachin|klimenko|nikitin_i@mx.ihep.su

Abstract. Visualization of relativistic string dynamics is performed. A set of basic tools for string world sheets visualization is developed, which enables construction of static images of the world sheets and representation of string dynamics as a film.

Relativistic string is a curve moving in the space which spans in its motion some surface in Minkowski space-time. This surface is called the ***world sheet***. The area of the world sheet should be extreme at fixed initial and final positions of the string. This requirement completely defines the motion of the string.

Physical destination of the string theory is a description of strongly interacting elementary particles, called hadrons. String models describe them as the systems of quarks, connected by string-like tubes filled by gluonic field. Break of the string looks like a decay of the hadron. Hadronic strings have a size 10^{-13} cm, a tension about 10 tons and typical energies 1 GeV

String theory considers the world sheets of various topological types. The world sheet of *open* string is obtained by a mapping of an infinite band into Minkowski space, the world sheet of *closed* string is a mapping of a cylinder, the world sheets of *Y-shaped* string is a mapping of 3 bands, sticked together along one edge. Solving the equations of motion in each topological class one obtains a simple geometrical receipt for world sheet construction. Particularly, the shape of the world sheet of open string can be reconstructed if the shape of its edge is given. The method is the following [1].

Let in Minkowski space a periodical light-like curve $Q_\mu(\sigma)$ be given:

$$Q_\mu(\sigma + 2\pi) = Q_\mu(\sigma) + 2P_\mu, \quad Q'^2(\sigma) \equiv Q_0'^2(\sigma) - \vec{Q}'^2(\sigma) = 0,$$

$2P_\mu$ is a period of the curve. This curve is shown on fig.1. We will call this curve a ***supporting*** curve.

World sheet of open string is a geometrical locus of middles of segments, connecting all possible pairs of points on supporting curve:

$$x_\mu(\sigma_1, \sigma_2) = \tfrac{1}{2}(Q_\mu(\sigma_1) + Q_\mu(\sigma_2)).$$

Surface constructed in such a way has two edges, one of which coincides with the supporting curve, another one is the supporting curve shifted onto a semi-period P_μ.

Edges of world sheet are world lines of the ends of the string. The ends of the string move at light velocity. World sheet is the periodical surface with the period P_μ. In string theory P_μ is equal to the vector of total energy-momentum of the string. The space, orthogonal to P_μ, is the rest frame of the string. One can project the world sheet into

[1] This work is fulfilled as a part of International project of Russian Ministry of Science "Visualization of Complex Physical Phenomena in Virtual Environment".

the rest frame and consider the evolution here. In the rest frame the supporting curve is closed. The string can be reconstructed from the supporting curve as follows ([2], fig.2). Let's (1) mark some initial point A on the supporting curve; (2) draw from the point A two arcs of equal lengths l; (3) mark the middle of segment, connecting the ends of arcs. When l changes, the middle of segment draws some curve from point A to point B, shifted from A by a half of total supporting curve's length. The curve constructed is the string[2]. In evolution the initial point moves at light velocity along the supporting curve and entails the described construction.

Geometrical algorithms given were realized in the complex of programs for visualization of world sheets, comprehensively described in [3]. Now we will present new features, included in this complex.

Animation. To show the string dynamics as a movie it is sufficient to draw sequentially the time slices of the world sheet (fig.1) or (more convenient in practice) use the description of string dynamics in the rest frame (fig.2). All calculations are being done between the frames. The parameters, which determine the form of the supporting curve, can be changed by user during the movie, it allows one to study the string dynamics interactively. For a DEC Alpha station (DEC AXP 3500) we are able to generate 20 frames/sec. For a more slow computer the same speed can be achieved by precalculation and bufferization of frames. The movie can be stored in a file and then played with the help of a special program on computers like VAX or IBM PC.

Stereoscopic images. We have also experimented with stereoscopic images (fig.3) of the world sheets. To obtain a stereoimage one need to create two images of the 3D object from two different points of view, related to left and right eye, respectively. They can be seen as one stereoimage by means of a stereoscope or, after some training, without it[3].

Examples of visualization

Singular points on strings. At definite conditions [1] singular points appear on strings. The most interesting case is when the supporting curve lies in a plane and has loops (winding number of tangent vector in passage along the supporting curve is not unit) or non-convex parts, bounded by inflection points. In this case *cusps* appear on strings, see fig.4. Cusps behave like point particles, they move at the velocity of light, scatter or annihilate in collisions, form bound states. If the supporting curve does not belong to a plane, the cusp on the string appears for one instant, periodically in the same point. The world sheet has in this point the singularity, called *pinch point*, see fig.5.

General properties: (1) singular points are topologically stable, they cannot be eliminated by small deformations of the supporting curve; (2) the linear density of energy-momentum of the string infinitely increases in the vicinity of singular points.

Disconnected strings. Supporting curves can be non-monotonous along the time axis

[2]Example: if the supporting curve is the circle, then the string is its diameter.

[3]If one tries to see a stereoimage without a stereoscope, the biggest problem is to focus eyes on the paper (display) and at the same time to cross them at the point *before* the paper, where the image should appear. The idea is in short as follows [4]. Look at the point somewhere in the middle between the paper and the eyes. The images, as a result, will split in two and cease to be sharp. The relative position of the paper must be chosen in such a way that two central images merged in one. (It is unnecessary to put any attention on two other images.) The central image is actually 3-dimensional. At first it won't be sharp. If one manages to hold this image for some time, it becomes sharp.

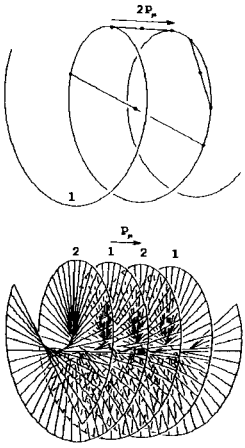


Fig.1. World sheet of open string is constructed as geometrical locus of middles of segments, connecting all possible pairs of points on the supporting curve. The world sheet has two edges: supporting curve (1) and its image in translation onto semi-period (2). Section of the world sheet by plane of constant time gives the string in this instant of time. Time axis is directed from left to right.

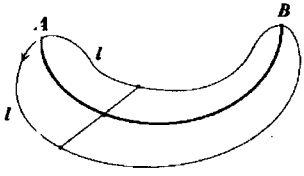


Fig.2. The shape of the string in the rest frame can be determined as follows. Let's draw two arcs of equal lengths on the supporting curve (projected to the rest frame), lying in the opposite sides from initial point A. Connect the ends of arcs by straight segment and mark its middle. The set of all such middles give the string at time instant, when its end is placed in the point A.

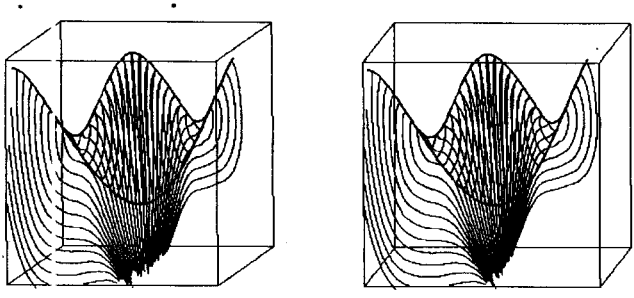


Fig.3. Stereopair.

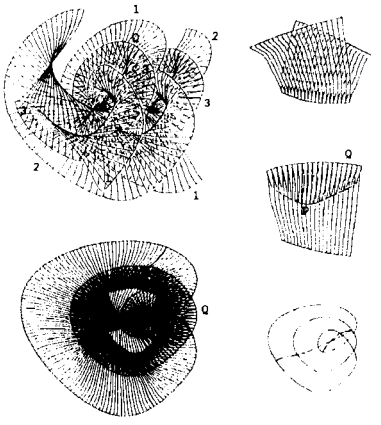


Fig.4.
For flat supporting curves the singularities have a form of cusp lines (folds) on the world sheet. Left upper part shows the world sheet. Left bottom part shows its projection into the rest frame. Numbers denote: 1,2 – the edges of the sheet, 3 – cusp line. Near the point Q the world sheet has self-intersection. Right part displays (from top to bottom): the cusp, the line of self-intersection PQ, the self-intersection on the string.

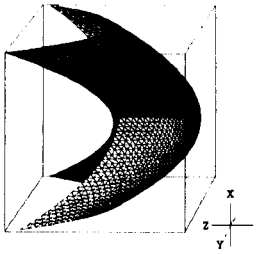


Fig.5. For non-flat supporting curves the singularities are localized in isolated pinch points. In the vicinity of pinch point the world sheet has the following canonical form: $(X, Y, Z) = (\xi, \xi\eta, \xi^2 + \eta^2)$. The string has an instantaneous cusp in passage through the pinch point.

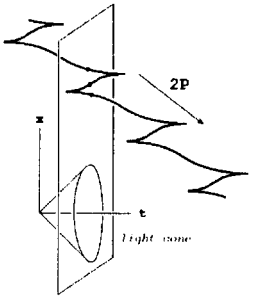


Fig.6. Supporting curve, non-monotonous along time axis. Due to light-likeness property such supporting curves necessarily have singularities (cusps), correspondent singularities are also induced on the world sheet (not shown). The supporting curve is the world line of one end of the string, we see in the figure, that the end of the string appears more than once in one time slice.

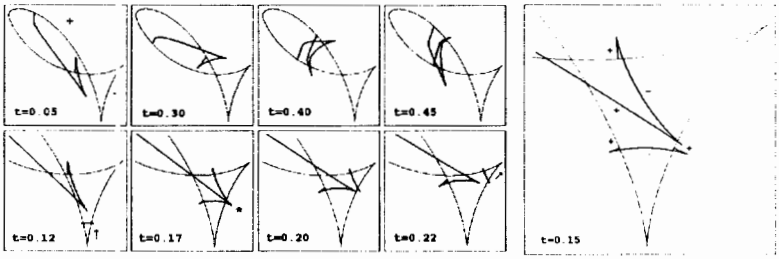


Fig.7. Motion of disconnected strings. Bold lines are the strings, thin line is the supporting curve (projected to the rest frame). The supporting curve has two segments, marked by $+$ and $-$ on the first frame, directed in Minkowski space forward and backward in time respectively. There is a long string, which constantly presents in the system. The common features of its evolution are shown on the upper-left frame sequence. Time is measured in periods. Lower-left sequence shows in details the time interval, when the end of the string passes the part $(-)$. Additional short string appears near $t = 0.10$. At $t = 0.17$ the recombination $(*)$ occurs: short string is attached to the long string and a part of long string is detached. New short string disappears at $t = 0.24$. Some parts of the string have negative energy. Density of the energy changes the sign in passing through the cusps. The right frame shows the distribution of energy on the strings near the recombination point.

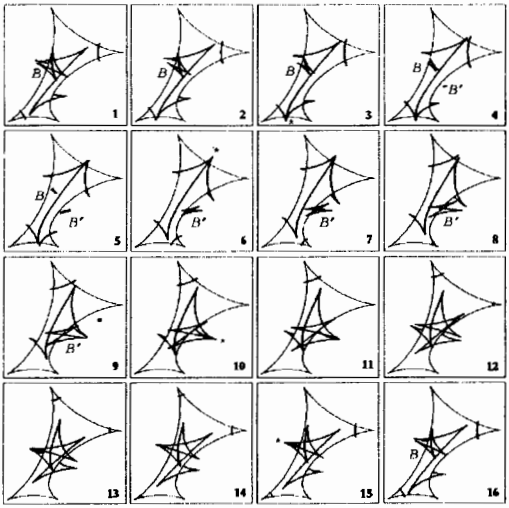


Fig.8. More complex processes in the formation of disconnected strings. On the first frame the following 4 disconnected elements can be found: Z-shaped long open string in the center, two short open strings in the corners of supporting curve and closed string B, looking like a butterfly. In the process of further evolution the long string recombinates $(*)$ with short strings, adopting the shape of reverted Z (frame 8). During this time the butterfly B is decreased and disappears on the frame 5. On the frame 4 new butterfly B' appears. On frame 10 it enters into recombination with the long string and is included in it. There are no closed strings on frames 11-14, here the long string in the center is simply connected. On the frame 15 the butterfly B is detached from the long string. The frame 16 coincides with the frame 1, and the evolution is repeated again.

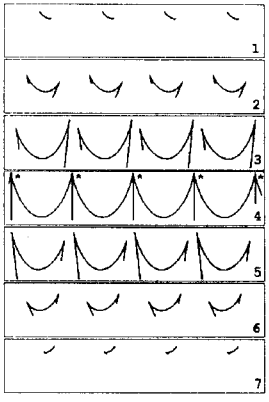


Fig.9. Disconnected strings with spacelike period in simultaneous reference frame. An infinite number of strings, periodically located in one space direction, simultaneously appear on the frame 1. On the frame 4 the strings recombine. On the frame 7 they simultaneously disappear. Further observation will show the empty space.

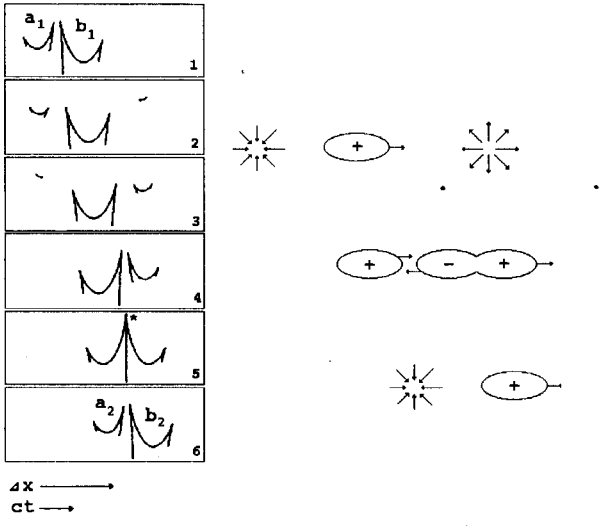


Fig.10. The same evolution observing in other reference frame. Here we see the moving string formation, which periodically repeats its form. On the frame 1 strings a_1 and b_1 are presented. The string a_1 disappears on the frame 3. The string b_1 deforms to the string a_2. The string b_2 is created from vacuum on the frame 2. The recombination (*) occurs on the frame 5. In these processes all part of the system do not exceed the velocity of light. But the mean velocity of the system is greater than the velocity of light. The simplified model of energy transfer in this motion is shown on right figure. At first, there is moving subsystem with positive energy, one zero energy subsystem is borning ahead it, and another one is disappearing behind. Then the front subsystem polarizes, and the subsystems exchange by a portion of negative energy (by string segment at the moment of recombination). After this exchange the (+) acquires zero energy and disappears. The remaining part of the front subsystem, carrying the initial positive energy, continues to move.

(fig.6). As earlier, the problem consists in reconstruction of the world sheet from the supporting curve by the method, described above, and determination of its equal-time slices. The world sheets generated by non-monotonous supporting curves are again topologically equivalent to the band, i.e. correspond to open strings. However, the mapping of the band in Minkowski space is such that its equal-time slices contain disconnected parts. In time ordered slicing such world sheet looks like a formation of the strings, which periodically appear, disappear, split, fuse and rearrange parts.

Now we will describe the common properties of evolution of this system. For non-monotonous supporting curve the vector P_μ might be either time-like ($P^2 > 0$) or space-like ($P^2 < 0$). For time-like period we can choose the rest frame for the system, directing time axis along the period. The examples of evolution are shown on fig.7,8. The typical processes here are creation of strings from vacuum, their recombination and annihilation. A study of energy flows on the world sheets shows, that some parts of strings have negative energy (energy flow through the slice of the world sheet is directed backward in time). Those components of the system, which born and disappear, have zero total energy, momentum and angular momentum. Therefore these processes do not violate the conservation laws.

The picture changes dramatically, when the period is space-like (mass squared is negative). In particle physics such objects are called the tachyons. Correspondent world sheets are stretched not in time direction, but in the space one. There is no rest frame for this system. One can use *simultaneous* reference frame, where the period has pure space direction. In this the world sheet has finite size in time direction, the evolution lasts a finite time, see fig.9. In other reference frames the world sheet is sloped to time axis, the string formation has finite sizes in space. It moves with the velocity, equal to the slope $|\vec{P}|/P_0$, which is greater than 1, the light velocity in our units. At the same time, all parts of strings in formation do not exceed the light velocity. This motion is shown on fig.10. It becomes clear from this figure, that there is no *direct* hyper-light transfer of energy or information in the system, this transfer occurs through a sequence of creation, annihilation and recombination processes. Not initial string but its exact copy arrives in the finite point.

In classical string theory such solutions are usually rejected due to their exotic properties. However, their rejection lead to extra difficulties in quantum theory [5].

Conclusion

The visualization of the relativistic string dynamics have been accomplished. To fulfil the research we created a set of programs for the visualization of the world sheets of open strings. As a next stage we plan to study singular points on strings of other topological types and the processes of string breaking.

References

[1] Klimenko S.V., Nikitin I.N. and Talanov V.V. Preprint IHEP 95-7, Protvino 1995.

[2] Nikitin I.N. Preprint IHEP 96-3, Protvino 1996.

[3] Klimenko S.V., Nikitin I.N. and Talanov V.V., Programming 4, 47 (1994); in Proc. of 2nd Dagstuhl Seminar on Scientific Visualization, Dagstuhl, Germany, May 1994.

[4] Francis G K. A Topological Picturebook – Springer-Verlag 1987, 1988.

[5] Nikitin I.N. Theor.Math.Phys. **107**, 213, (1996).

ON A VARIANT OF DOMAIN DECOMPOSITION METHOD FOR SOLUTION OF MATHEMATICAL PHYSICS PROBLEMS

Egorov A. A.

Byelorussian State University
Skoryna av., 4, Minsk 220050, Belarus
e-mail: DEP02@FPM.BSU.Minsk.By

Abstract. The new method of net domain decomposition for nonlinear parabolic problems is suggested. It involves the construction of the explicit iterative schemes in each subdomain that provides high degree of parallelizing. The convergence theorem of iterative procedure is given.

The domain decomposition methods are widely used in the solution of partial differential equations in order to reduce the problem to some separate problems in more simple subdomains [1, 2]. Such methods are especially effective when they are realized on multiprocessor parallel computing systems. Two main approaches are considered in this method. The first one is connected with using implicit difference schemes which are solved by iterative processes like component-wize splitting-up method, alternating direction method and others [3 — 5]. The second approach is characterized by constructing noniterative methods directed to nonstationary problems [6, 7]. In both cases it seems most perspective to use multicomponent alternating direction schemes [8, 9], which have several advantages in comparison with classical schemes.

In the work [10] the new decomposition method with arbitrary subdomain set based on multicomponent alternating direction schemes for two-dimensional linear elliptic equations was suggested. In the present report this result is extended to nonlinear parabolic equations in arbitrary dimensions.

In the domain $\overline{G}$ with the boundary Γ let us consider the boundary value problem for quasilinear parabolic equation

$$\frac{\partial u}{\partial t} = Lu + f(x,t), \quad (x,t) \in G \times (0,T], \tag{1}$$

$$u(x,0) = u_0(x), \quad x \in \overline{G}, \quad u(x,t) = \mu(x,t), \quad (x,t) \in \Gamma \times (0,T], \tag{2}$$

where $Lu = \sum\limits_{\alpha=1}^{p} \dfrac{\partial}{\partial x_\alpha}\Big(k_\alpha(x,t,u)\dfrac{\partial u}{\partial x_\alpha}\Big)$, $c_1 < k_\alpha(x,t,u) \le c_2$, $\alpha = 1,\ldots,p$, $c_i > 0$, $i = 1,2$. It is assumed that the functions $k_\alpha(x,t,u)$, $\alpha = 1,\ldots,p$, are continuous with respect to u.

We specify in $\overline{G}$ the net $\overline{\omega}_{h\tau} = \overline{\omega}_h \times \overline{\omega}_\tau$, $\overline{\omega}_h = \{x_i^{(\alpha)} = i_\alpha h_\alpha,\ i_\alpha = 0,\ldots,N_\alpha,\ N_\alpha h_\alpha = l_\alpha,\ \alpha = 1,\ldots,p\}$, $\overline{\omega}_\tau = \{t_n = n\tau,\ n = 0,1,\ldots\}$, and consider difference scheme approximating (1), (2):

$$y_t = \Lambda\hat{y} + \varphi(x,\hat{t}), \quad (x,t) \in \omega_h \times \omega_\tau, \tag{3}$$

$$y(x,0) = u_0(x), \quad x \in \overline{\omega}_h, \quad y(x,t) = \mu(x,t), \quad (x,t) \in \gamma_h \times \omega_\tau, \quad \gamma_h = \overline{\omega}_h \setminus \omega_h, \tag{4}$$

where a_α are functionals defined by the coefficients of the equation (1) [11]. In order to solve (3), (4) we use the iterative process

$$\frac{\overset{s+1}{y} - y}{\tau} = \tilde{\Lambda} \overset{s+1}{y} + \varphi, \quad s = 0, 1, \ldots, \quad \overset{0}{y} = y, \tag{5}$$

$\tilde{\Lambda} v = \sum\limits_{\alpha=1}^{p} \tilde{\Lambda}_\alpha v$, $\hat{\Lambda}_\alpha v = (a_\alpha(x, \hat{t}, \overset{s}{y}) v_{\bar{x}_\alpha})_{x_\alpha}$. We suppose that the operator $\tilde{\Lambda}$ involves the boundary conditions approximation. For solving of linearized equation (5) the next two multicomponent alternating direction schemes may be used [12, 13]:

$$(E + \tau A_\alpha) \overset{s+1}{y}{}^{(k+1)}_\alpha + \tau \Big\{ \sum_{\beta=1}^{\alpha-1} A_\beta \overset{s+1}{y}{}^{(k+1)}_\beta + \sum_{\beta=\alpha+1}^{p} A_\beta \overset{s+1}{y}{}^{(k)}_\beta \Big\} = y + \tau\varphi,$$

$$(E + \tau A_\alpha) \overset{s+1}{y}{}^{(k+1)}_\alpha + \tau\sigma (A_\alpha \overset{s+1}{y}{}^{(k+1)}_\alpha - A_\alpha \overset{s+1}{y}{}^{(k)}_\alpha) + \tau \sum_{\alpha=1}^{p} A_\alpha \overset{s+1}{y}{}^{(k)}_\alpha = y + \tau\varphi,$$

$$k = 0, 1, \ldots, \quad A_\alpha v = -\tilde{\Lambda}_\alpha v, \quad \overset{s+1}{y}{}^{(0)}_\alpha = \overset{s}{y}_\alpha, \quad \alpha = 1, \ldots, p.$$

These methods give successive and parallel realization of algorithm. We construct analogous methods on the basis of net domain decomposition.

We shall partition the net $\overline{\omega}_h$ into elementary subdomains $\omega_{i_1,\ldots,i_p} = \{(x_{i_1+j_1}, x_{i_2+j_2}, \ldots, x_{i_p+j_p}), j_\alpha \in \{0, 1\}, \alpha = 1, \ldots, p\}$. In each point of subdomain (there are 2^p such points) define 2^p components of solution. Each component in $\omega_{i_1,\ldots,i_p}$ is accosiated with index $j = [j_1, \ldots j_p]$. Note that anumeration of components is unessential.

Let us define in subdomain $\omega_{i_1,\ldots,i_p}$ network operator

$$A_{i_1,\ldots,i_p} y_{i_1,\ldots,i_p} = (A^{(1)}_{i_1,\ldots,i_p} y_{(1)i_1,\ldots,i_p}, A^{(2)}_{i_1,\ldots,i_p} y_{(2)i_1,\ldots,i_p}, \ldots, A^{(2^p)}_{i_1,\ldots,i_p} y_{(2^p)i_1,\ldots,i_p}),$$

$$A^{(j)}_{i_1,\ldots,i_p} y_{(j)i_1,\ldots,i_p} = -\sum_{\alpha=1}^{p} \tilde{\Lambda}^{(j_\alpha)}_\alpha y_{(j)i_1,\ldots,i_p}, \quad j_\alpha = 0, 1, \quad j = 1, \ldots, 2^p,$$

$$\tilde{\Lambda}^{(0)}_\alpha v = h_\alpha^{-1} a_\alpha v_{x_\alpha}, \quad \tilde{\Lambda}^{(1)}_\alpha v = -h_\alpha^{-1} a^{(-1_\alpha)}_\alpha v_{\bar{x}_\alpha}, \quad \alpha = 1, \ldots, p.$$

Then we can write "intrinsic" iterative processes of the decomposition method for the problem (5):

$$(E + \tau A^{(j)}_{i_1,\ldots,i_p}) \overset{s+1}{y}{}^{(k+1)}_{(j)i_1,\ldots,i_p} + \tau B^{(j)}_{i_1,\ldots,i_p} (\overset{s+1}{y}{}^{(k+\bar{j}_p)}) = y + \tau\varphi, \quad j = 1, \ldots, 2^p, \tag{6}$$

$$(E + \sigma\tau A^{(j)}_{i_1,\ldots,i_p}) \overset{s+1}{y}{}^{(k+1)}_{(j)i_1,\ldots,i_p} + (1-\sigma)\tau A^{(j)}_{i_1,\ldots,i_p} \overset{s+1}{y}{}^{(k)}_{(j)i_1,\ldots,i_p} +$$

$$+ \tau B^{(j)}_{i_1,\ldots,i_p} (\overset{s+1}{y}{}^{(k)}) = y + \tau\varphi, \quad j = 1, \ldots, 2^p, \tag{7}$$

$$B^{(j)}_{i_1,\ldots,i_p} (\overset{s+1}{y}{}^{(m)}) = A^{(\bar{j})}_{i_1+(-1)^{j_1+1},\ldots,i_p+(-1)^{j_p+1}} \overset{s+1}{y}{}^{(m)}_{(j)i_1+(-1)^{j_1+1},\ldots,i_p+(-1)^{j_p+1}}.$$

By $\bar{j} = [\bar{j}_1, \ldots, \bar{j}_s]$ we denote inversion of index, for example, if $j = (1, 0, \ldots, 0, 1)$ then $\bar{j} = (0, 1, \ldots, 1, 0)$. The method (6) is realized successively in subdomains $\omega_{i_1,i_2,\ldots,i_p-1,i_p-1}$,

$i_\alpha = 0, \ldots, N_\alpha - 1$, then in subdomains $\omega_{i_1,i_2,\ldots,i_p-1,i_p}$ and so on. The algorithm (7) admits parallel computation of the solution in all cells of the net $\overline{\omega}_h$.

It is not difficult to show that the operator $A_{i_1,\ldots,i_p}$ is positive definite. Really, by definition we have

$$(A_{i_1,\ldots,i_p} y_{i_1,\ldots,i_p}) = \sum_{\omega_{i_1,\ldots,i_p}} (Ay)_{i_1,\ldots,i_p} y_{i_1,\ldots,i_p} = \sum_{\alpha=1}^{p} a_\alpha (y_{x_\alpha})^2 \Big|_{x=(x_{i_1},\ldots x_{i_p})} +$$

$$+ \sum_{\alpha=1}^{p} a_\alpha^{(-1_\alpha)} (y_{\bar{x}_\alpha})^2 \Big|_{x=(x_{i_1+1},\ldots,x_{i_p+1})} > 0$$

for $y(x) \not\equiv 0$. Thus the system (7) into 2^p unknowns is explicit solvable under s fixed. So one can consider constructed iterative process like explicit.

For each subdomains let us introduce the error vector

$$\overset{k+1}{\Delta} = \overset{s+1}{y} - \overset{s+1}{y}(k+1), \quad \overset{k+1}{\Delta} = (\overset{k+1}{\Delta}_{(1)}, \ldots, \overset{s+1}{\Delta}_{(2^p)}), \quad \overset{s+1}{y}(k+1) = (\overset{s+1}{y}_{(1)}(k+1), \ldots, \overset{s+1}{y}_{(2^p)}(k+1)).$$

The following theorem is valid.

Theorem. *The iterative process (6) converges to approximation* $\overset{s+1}{y}$ *of the solution* $\hat{y}$ *of difference scheme (3), (4) ((7) converges to* $\overset{s+1}{y}$ *under condition* $\sigma \geq 1$*). The recurrence inequality estimating degree of convergence takes place:*

$$Q^2(\overset{k+1}{\Delta}) \leq q_k^{-1} Q^2(\overset{k}{\Delta}), \quad Q^2(v) = \sum [(Av, v) + (v, v)] + \frac{1}{2}\tau \Big(\sum Av\Big)^2,$$

$$q_k = 1 + \nu(1 - \tau\nu\mu_k), \quad \nu > 0, \quad \mu_k = \sum (A \overset{k}{\Delta})^2 / Q^2(\overset{k}{\Delta}),$$

where the summation is over all cells of the net domain.

References

[1] V. I. Lebedev, Composition Method, M., Nauka, 1986 (in russian).
[2] G. I. Marchuk, Methods of Computational Mathematics, M., Nauka, 1989 (in russian).
[3] M. Dryia, Substructuring methods for parabolic problems, N.-Y. Univ. Comp. Sci. Dept, Techn, Rept. 529, 1990.
[4] Yu. M. Layevsky, Zhurn. Vychisl. Mat. i Mat. Fiz. **32**, 11, 1744 (1992 (in russian).
[5] P. N. Vabischevich, Zhurn. Vychisl. Mat. i Mat. Fiz. **29**, 12, 1822 (1989) (in russian).
[6] A. A. Samarsky, and P. N. Vabischevich, Differ. uravn. **31**, 9, 1563 (1995) (in russian).
[7] V. N. Abrashin, Differ. uravn. **31**, 9, 1525 (1995) (in russian).
[8] V. N. Abrashin, Differ. uravn. **26**, 2, 313 (1990) (in russian).
[9] V. N. Abrashin, Differ. uravn. **31**, 5, 840 (1995) (in russian).
[10] V. N. Abrashin, Differ. uravn. **32**, 5, 652 (1996) (in russian).
[11] A. A. Samarsky, Theory of difference schemes, M., Nauka, 1986 (in russian).
[12] A. A. Egorov, Differ. uravn. **30**, 7, 1202 (1994) (in russian).
[13] A. A. Egorov, Differ. uravn. **31**, 7, 1211 (1995) (in russian).

APPROXIMATE FORMULAE FOR FUNCTIONAL INTEGRALS BASED ON INTERPOLATION OF MEASURES

Egorov A.D. and Zherelo A.V.

Institute of Mathematics
Academy of Sciences of Belarus
220072 Minsk, Surganov str., 11, Belarus
e-mail: imanb@imanb.belpak.minsk.by

Abstract. Approximate evaluation of functional integral based on measure interpolation formula is considered.

In this report we give a short description of the method of approximate evaluation of functional integrals which is based on the formula of the interpolation of measure. The method was in detail developed for integral with respect to Gaussian measure (see [1, 2]). Now we consider integrals on linear topological space $X = X([0,T])$ of functions $x \equiv x(t)$ defined on the segment $[0,T]$ of the real line w.r.t. quasimeasure μ with characteristic functional of the form

$$\chi(\xi) = \exp\left\{\frac{(-1)^p}{(2p)!}K(\xi,\ldots,\xi)\right\},$$

where $K(\xi_1,\ldots,\xi_{2p})$ is a real symmetric continuous n-linear form on $X' \times \overset{(2p)}{\cdots} \times X'$.

We suppose that

$$K(\xi,\ldots,\xi) = \int_0^T \overset{(2p)}{\cdots} \int_0^T B(t_1,\ldots,t_{2p})\,\xi(t_1)\ldots\xi(t_{2p})dt_1\ldots dt_{2p},$$

where $B(t_1,\ldots,t_{2p})$ is the $2p$-moment function of quasimeasure μ. Our approach is based on employment the measure interpolation formula (see [3, 2]):

$$\int_X f(x(\cdot))d\mu(x) = \int_X f(x(\cdot))d\mu_0(x) +$$

$$+[(2p)!]^{-1}\int_0^1\int_X\int_0^T \overset{(2p)}{\cdots} \int_0^T (B-B_0)(t_1,\ldots,t_{2p})\times$$

$$\times f^{(2p)}(x;t_1,\ldots,t_{2p})\,dt_1\ldots dt_{2p}\,d\mu_u du, \qquad (1)$$

where μ_u is the quasimeasure with the $2p$-moment function

$$B_u(t_1,\ldots,t_{2p}) = uB(t_1,\ldots,t_{2p}) + (1-u)B_0(t_1,\ldots,t_{2p}),$$

μ_0 be some fixed quasimeasure defined by $2p$-moment function $B_0(t_1,\ldots,t_{2p})$; $f^{(2p)}(x;t_1,\ldots,t_{2p})$ is functional derivative of the order $2p$ for the functional $f(x)$ at points $t_1,\ldots,t_{2p}$.

Iterating equality (1) we obtain the expansion

$$\int_X f(x)\,d\mu(x) = \sum_{k=0}^{N} \frac{1}{[(2p)!]^k k!} \int_0^{T^{2p}} \overset{(k)}{\cdots} \int_0^{T^{2p}} \prod_{j=1}^{k} (B - B_0)(t_1^{(j)}, \ldots, t_{2p}^{(j)}) \times$$

$$\times \int_X f^{(2pk)}(x; t_1^{(1)}, \ldots, t_{2p}^{(1)}, \ldots, t_1^{(k)}, \ldots, t_{2p}^{(k)})\,d\mu_0(x)\,d^{2pk}t + r_N(f) \equiv I, \qquad (2)$$

where $r_N(f)$ is the residual of the formula, $d^{2pk}t = dt_1^{(1)} \ldots dt_{2p}^{(1)} \ldots dt_1^{(k)} \ldots dt_{2p}^{(k)}$, $T^{2p} = [0,T] \times \overset{(2p)}{\cdots} \times [0,T]$.

The next important point of our approach is partition of the domain T^{2p} of moment function into subsets of the form

$$[0,T/2] \times \overset{(j)}{\cdots} \times [0,T/2] \times [T/2,T] \times \overset{(2p-j)}{\cdots} \times [T/2,T], \quad j = 0,1,2,\ldots,2p,$$

and estimate of the integrals

$$B'(T) \equiv \sum_{j=1}^{2p-1} C_{2p}^{j} \int_0^{T/2} \overset{(j)}{\cdots} \int_0^{T/2} \int_{T/2}^{T} \overset{(2p-j)}{\cdots} \int_{T/2}^{T} \prod_{l=1}^{2p} (B - B_0)(t_1, \ldots, t_{2p})\,d^{2p}t,$$

$$j = 1,2,\ldots,2p-1; \quad d^{2p}t = dt_1 \ldots dt_{2p},$$

where we choise $B_0(t) = B(t)$ for t belonging diagonal subsets and $B_0(t) = 0$ otherwise. In the case of Gaussian measures ($p = 1$), $B'(T)$ is bounded for all $T \in \mathbb{R}$ when the correlation function is $\frac{1}{2}\exp\{-|t-s|\}$, and $B'(T) \to \infty$ as $T \to \infty$ for Wiener and some other measures. In our case appropriate quasimeasures with $B'(T) \to \infty$ are the quasimeasures with $2p$-moment function

$$B(t_1, \ldots, t_p) = \exp\{-\max(t_1, \ldots, t_{2p}) + \min(t_1, \ldots, t_{2p})\}. \qquad (3)$$

For these quasimeasure $B'(T)$ is bounded for all T and $j = 1,2,\ldots,2p-1$. So, if the derivatives $f^{(2pk)}(x;t)$ increase slovely the series (2) will be convergent. By the analogy of Gaussian case we consider the integrated functionals of the form

$$f(x(\cdot)) = \exp\Big\{\lambda \int_0^T V(x(\tau))d\tau\Big\}$$

and we use the fact that under above chosen $B_0(t_1, \ldots, t_{2p})$ the functional integral in (2) may be represented in this case as a finite sum of products of integrals over space $X([0,T/2])$.

Then we can apply the partition to integrals over $X([0,T/2])$ and $X([T/2,T])$ and so on while after l steps we come to approximate expression containing the integrals over spaces $X\left(\frac{T\cdot k}{2^l}, \frac{T(k-1)}{2^l}\right)$, $k = 1,\ldots,2^l$.

As an example we consider evaluation by the proposed method the integral

$$I = \int\limits_{X([0,T])} \exp\Big\{i\lambda \int_0^T x(\tau)d\tau\Big\}d\mu.$$

This integral may be calculated exactly:

$$I = \exp\left\{ \frac{(-1)^p}{(2p)!} \int_0^T \overset{(2p)}{\cdots} \int_0^T B(t_1, \ldots, t_{2p})\, dt_1 \ldots dt_{2p} \right\}.$$

We take $p = 2$ and in this case

$$I = \exp\{(3 + 2T + T^2/2)\exp(-T) + T - 3\}.$$

The derivative in the formula (2) is evaluated exactly and it is equal λ^{4k},

$$B^{(}T) = 4! \left[3 + (3 + 2T + T^2/2)e^{-T} + (-6 - 2T - T^2/4)e^{-T/2} \right].$$

Then the integral from $B(t_1, \ldots, t_{2p})$ over $[a, b] \times \ldots \times [a, b]$ is equal to integral from $B(t_1, \ldots, t_{2p})$ over $[0, b-a] \times \ldots \times [0, b-a]$, so we only have to evaluate the functional integral over $X([0, T/2^l])$. For calculation of this functional integral we use approximate formula of given algebraic accuracy degree $2p(n+1)-1$ (see also in [4] the another formula for $B(t_1, \ldots, t_{2p}) = \min(t_1, \ldots, t_{2p})$):

$$\int_X f(x)d\mu(x) \approx L_{x_n} \ldots L_{x_1} f\left(\sum_{l=1}^{n} (c_l^{(n)})^{1/(2p)} x_l \right) \equiv L_n f, \tag{4}$$

where $c_l^{(n)}$, $l = 1, 2, \ldots, n$, are the roots of the polynomial $\sum\limits_{k=0}^{n} (-1)^k c^{n-k}/k!$; L_{x_i} is the functional L applied to f considered as a functional of the single argument $x_i \in X$; the functional $Lf(\cdot)$ is defined by the equality

$$Lf(\cdot) = \int_0^T \int_0^\infty S_4 f(\rho_1(u, \cdot)\rho_2(v, \cdot)) e^{u-v} du dv +$$

$$+ \int_0^\infty S_4 f(\rho_2(v, \cdot)) e^{-v} dv + f(0)(1 - e^T),$$

where $S_4 f(x) = \frac{1}{4} \sum\limits_{k=1}^{4} f(z_k x)$; z_k, $k = 1, \ldots, 4$, are the roots of the equation $z^4 = 1$; $\rho_1(u, t) = 1_{[0,t]}(u)$ is the characteristic function of the segment $[0, t]$; $\rho_2(v, t) = 1$ for $v > t$ and it is equal 0 for $v \le t$.

References

[1] *Egorov A.E.* and *Malyutin V.B.* On formulae for interpolation of Gaussian measures and their application to approximate functional integration. Dokl. Akad. Nauk BSSR, 1987, **31**, N 9, 788–791 (in Russian).

[2] *Egorov A.E.*, *Sobolevsky P.I.* and *Yanovich L.A.* Functional Integrals: Approximate Evaluation and Applications. Kluwer Academic Publishers, 1993, 418 p.

[3] *Egorov A.E.* Wick monomials, diagrams, formulae of infinitesimal change of moment function for quasimeasures. Dokl. Akad. Nauk Belarusi, 1992, **36**, N 2, 101–106 (in Russian).

[4] *Egorov A.D.* On approximate evaluation of integrals w.r.t. quasimeasures. Dokl. Akad. Nauk BSSR, 1976, **20**, N 8, 677–679 (in Russian).

TWO-DIMENSIONAL SOLIDIFYING AND MELTING OF NEEDLE-LIKE CRYSTALS IN THE PRESENCE OF HYDRODYNAMIC FLOW

Yuri E. Hohlov

42, Vavilova st., Steklov Mathematical Institute RAS, 117966, Moscow, Russia

Sam Howison

24-29 St.Giles, Mathematical Institute, Oxford University, Oxford, OX1 3LB, UK

Konstantin Kornev

101, prosp. Vernadskogo, Institute for Problems in Mechanics RAS, 117526, Moscow, Russia

Abstract. The problem of solidifying (melting) under the action of a forced hydrodynamic flow is considered. The problem admits a formulation in terms of a non-local Löwner-Kufarev type PDE for analytic functions. This gives a useful tool for study the behaviour of solutions. It is shown that a crystal with parabolic tip propagates without change of its shape at a steady velocity.

1 Introduction

For the last decade much attention has been paid to the problems of pattern formation in two-dimensional hydrodynamic systems [1]–[3]. One of them is concerned with the problem of dendritic crystallization in the so-called diffusion-limited regime when the growth velocity is determined by the rate of diffusion of latent heat away from the front [2]–[11]. As was first shown by Ivantsov [4], the dendrite of parabolic shape propagates at a constant velocity; however, the tip radius l and velocity v cannot be selected uniquely within the framework of Ivantsov's approach. One possible mechanism of dendritic shape selection is influence of crystal anisotropy [2]–[11]. It has been shown that the anisotropic surface tension breaks the Ivantsov family into a discrete set of possible shapes [11]. But it is not completely clear that anisotropy is the sole mechanism involved in the selection process. In real experiments the existence of hydrodynamic flow can play significant role in selection process and very little is known about the action of imposed flow on crystal growth [12]. Boisson and coauthors [12] drew attention to analysis of influence of hydrodynamic forced flow on the growth process because the results could help in better understanding of the selection problem. By the way, the problem of solidifying of moving liquids is far itself to be studied in detail. Recent theoretical studies [13]–[17] have been performed to extend the Ivantsov parabolic solution over the potential and the viscous flows. In particular, it has been found that the Navier-Stokes model will not allow an exact similarity solution any more [15]–[17] while the potential flow admits the Ivantsov-like solution [13]–[14]. The last problems appear, for example, in context of the method of artificial freezing and thawing of flow in porous media [6] –[8] and it is applicable to study the process of crystal solidifying and melting in the Hele-Shaw cell.

This work is aimed to formulate the problem of two-dimensional solidifying and melting in potential flows in terms of analytical functions. Such a formulation enables us to give

an effective too for analysis of more complicated thermo-hydrodynamic regimes and to find the common mathematical nature of patterning in different, at a glance, physical problems [1]–[3]. The semi–infinite frozen bodies will be considered. As an example the growth of parabolic dendrite will be discussed.

2 Mathematical model

We consider two-dimensional steady–state flow that satisfies Darcy's law. We assume that the liquid is incompressible and the thermophysical parameters of the matter are constant. The temperature and the velocity at infinity are denoted by T_∞ and V_∞. In a plane orthogonal to the wide side of dendrite we introduce the coordinate system $z = x+iy$ and the complex flow potential $W = \varphi + i\psi$, where φ is the velocity potential and ψ is the stream function. The velocity potential is related to the pressure p by $\varphi = -kp/V_\infty \mu l$ where k- is the permeability of the medium, μ is the viscosity of the fluid and l is the characteristic size of dendrite which must be found. The variables x, y and function W are dimensionless so that $x = x'/l, \quad y = y'/l$, where symbols x' and y' denote dimension variables. The mathematical model of the phenomenon in dimensionless variables has the form

$$\nabla \cdot \mathbf{V} = 0, \qquad \mathbf{V} = \nabla\varphi, \qquad Pe\mathbf{V} \cdot \nabla\theta = \Delta\theta, \quad z \in D(t);$$

$$\theta = 0, \quad -\frac{\partial\theta}{\partial n} = v_n, \quad \frac{\partial\varphi}{\partial n} = 0, \quad z \in \partial D(t); \tag{1}$$

$$\mathbf{V} = (1,0), \quad |z| \to \infty \quad ;$$

$$\lim_{x\to-\infty} \theta = 1, \quad \lim_{y\to\pm\infty} \frac{\partial\theta}{\partial y} = 0;$$

$$D(t)|_{t=0} = D_0.$$

Here v_n is the normal (outward with respect to liquid region) velocity of dendrite growth associated with the dimensionless time $t = t'\lambda(T_\infty - T_p)/l^2 H$, t' is the dimensional time. Note that such a normalization describes both the melting $t > 0$ and the solidifying $t < 0$. H is the latent heat, λ is the heat conductivity of liquid, and phase transition temperature is denoted by T_p. We assume it to be a physical constant of the solid-liquid pair. Function $\theta = (T - T_p)/(T_\infty - T_p)$ describes the temperature field within liquid. The Peclet number $Pe = V_\infty l/a$ is a measure of intensity of the convection heat transfer with respect to conduction and a is the thermal diffusivity of liquid. The region D is occupied by liquid, ∂D is its boundary and D_0 is the corresponding initial domain.

We use the time independent equations, because in most applications the effect of volume heat capacity is negligibly small compared with the surface heat release effect [5],[17]. We also assume that the imposed hydrodynamic velocity is much greater than the velocity caused by the effect of the solid-liquid density difference and, more strongly, the velocity of forced flow is much greater then the velocity of crystallization. Then the frozen body at any instant of time serves as an obstacle for flow and does not change the velocity distribution due to interface motion. When hydrodynamic and crystallization velocities

are comparable the problem becomes more complicated and we cannot directly apply the technique proposed below. At the same time the experimental situation mentioned above is wide spread [12, 5] so that the model is reasonable.

3 Splitting of the problem

It is remarkable that the moving boundary problem (1) may be split into two "independent" subtasks [14]–[17], where the first one is the problem of the heat exchange between the semi-infinite isothermic knife and the homogeneous flow. The second subtask is itself a free boundary problem. To demonstrate such a reduction we apply the Boussinesq transformation [19] to the convective heat transfer equation. This is equivalent to a conformal mapping of the liquid region from the physical plane onto a domain in the plane of complex potential W or use φ, ψ as independent variables. The respective balance equation for the temperature field is written in the form

$$Pe\,\frac{\partial\theta}{\partial\varphi} = \frac{\partial^2\theta}{\partial\varphi^2} + \frac{\partial^2\theta}{\partial\psi^2}. \tag{2}$$

The body cross section must map into the cut directed along the real axis $\psi = 0$ in the complex potential W-plane. In accordance with the above mentioned we have

$$\lim_{\varphi\to-\infty}\theta = 1, \quad \lim_{\psi\to\pm\infty}\frac{\partial\theta}{\partial\psi} = 0; \qquad \theta = 0, \quad \varphi\in[0,\infty), \quad \psi = \pm 0. \tag{3}$$

The problem (2)-(3) may be treated as a problem of heat exchange between the knife and homogeneous flow of velocity Pe. The problem (2)-(3) is solved separately [9] and for the solution one can find the relation

$$\left|\frac{\partial\theta}{\partial\psi}\right| = \sqrt{\frac{2Pe}{\pi\varphi}}, \qquad \psi = 0. \tag{4}$$

The heat flux at the unknown boundary will be transformed to

$$\left|\frac{\partial\theta}{\partial n}\right| = \left|\frac{\partial\theta}{\partial\psi}\right|\left|\frac{\partial\psi}{\partial n}\right| = \left|\frac{\partial\theta}{\partial\psi}\right|\left|\frac{\partial W}{\partial z}\right| = \sqrt{\frac{2Pe}{\pi\varphi}}\left|\frac{\partial W}{\partial z}\right|. \tag{5}$$

Condition (5) enables us to reformulate the initial problem in terms of analytic functions by the same way as for the Hele-Shaw problem [1, 10, 3].

4 Polubarinova-Galin equation

We introduce the auxiliary plane ω (the plane of canonical variables) so that the upper half of the plane corresponds to the region of flow and note that $W = \omega^2$. We attach the origin of coordinates to the stagnation point and the corresponding point in the auxiliary plane is also placed at the origin of coordinates $\omega = \xi + i\eta$. Let function f map the upper half of the ω–plane onto the region $D(t)$: $z = f(\omega, t)$. The unit normal vector to the moving interface $\partial D(t) = x_0(t) + iy_0(t)$ can be expressed as

$$\mathbf{n} = n_x + in_y = i\frac{\partial f}{\partial \xi}\left|\frac{\partial f}{\partial \xi}\right|^{-1}.$$

Then the normal velocity of the interface has the form

$$v_n = (\dot{x}_0 n_x + \dot{y}_0 n_y) = \Re\Big\{(\dot{x}_0 + i\dot{y}_0)(n_x - in_y)\Big\} =$$

$$\Re\Big\{\frac{\partial f}{\partial t}\overline{i\frac{\partial f}{\partial \omega}}\left|\frac{\partial f}{\partial \omega}\right|^{-1}\Big\}, \quad \xi \in (-\infty, \infty), \quad \eta \to +0. \tag{6}$$

On the other hand, the boundary of the frozen region is a stream line. Therefore, at the interface of dendrite we have

$$\pm i\left|\frac{\partial \theta}{\partial \psi}\right|\overline{(V_x - iV_y)} = \pm i\left|\frac{\partial \theta}{\partial \psi}\right|\overline{\frac{\partial W}{\partial z}} = -v_n \mathbf{n}, \quad \psi = \pm 0. \tag{7}$$

Combining Eqs.(4)-(7) we get the condition

$$\Re\Big\{\frac{\partial f}{\partial t}\overline{i\frac{\partial f}{\partial \omega}}\Big\} = -\sqrt{\frac{2Pe}{\pi W}}\,\overline{\frac{\partial W}{\partial \omega}}, \quad \xi \in (-\infty, \infty), \quad \eta \to +0. \tag{8}$$

In Eq.(8) we take into account the multivalence of the square root function so that the heat fluxes at the different sides of the cut have the different signs. Finally using

$$W = \omega^2, \tag{9}$$

we find the Polubarinova–Galin equation [3]

$$\Re\Big\{\frac{\partial f}{\partial t}\overline{i\frac{\partial f}{\partial \omega}}\Big\} = -2\sqrt{\frac{2Pe}{\pi}}. \quad \xi \in (-\infty, \infty), \quad \eta = 0 \tag{10}$$

In addition to the boundary condition (10) we must clarify the asymptotic behavior of function f at infinity and specify the initial dendrite shape. To do the former, one can use the boundary condition for velocity at infinity. There are many lot of possible hydrodynamic situations which are distinguished by boundary conditions at infinity. One of them is specified in problem statement (1). However, this is not the only possibility. We specify the initial dendrite shape as

$$f(\omega, 0) = f_0(\omega), \tag{11}$$

where f_0 is the mapping of the upper half plane onto the initial region D_0. We thus arrive at the non-linear boundary value problem (10)-(11) for analytic function f. Thus the boundary value problem (1) is reduced to the system of Eqs.(10),(11) and a boundary condition at infinity.

5 Ivantsov's dendrite.

We reproduce here the Ivantsov solution, i.e. we show that the dendrite of parabolic shape grows at constant velocity. We shall deal with a dendrite whose transverse sizes increase at infinity not faster than $\sqrt{x}$. The solution was obtained by another way in [13, 14].

To obtain such a solution we use the boundary condition for flow velocity at infinity mentioned above in (1)

$$\frac{\partial W}{\partial f} = \frac{\partial W}{\partial \omega}\frac{\partial \omega}{\partial f} = 1 \quad \omega \to \infty. \tag{12}$$

Equation (12) is written in such a form which presumes that the symmetry axis of dendrite to be aligned in x direction. Substituting Eq.(9) into Eq.(12) we arrive at asymptotes

$$f \propto \omega^2, \quad \omega \to \infty. \tag{13}$$

On the basis of problem (10),(11),(13) we seek the steady shape of dendrite growing in supercooled liquid. Consider the solution of this problem in the parabolic form

$$f = \omega^2 + iA(t)\omega + B(t). \tag{14}$$

Then the boundary condition (13) is satisfied automatically and Eq.(10) leads to formulae

$$\dot{A} = 0; \quad B = 2\sqrt{\frac{2Pe}{\pi}}\frac{t}{A}. \tag{15}$$

In the third formula of Eqs.(15) without loss of generality we can set $A = 1$. It means that all spatial variables are normalized by the tip radius l. In dimensional form we find the velocity of propagation of the parabolic dendrite as

$$v = \frac{2\lambda(T_\infty - T_p)}{l^{1/2}H}\sqrt{\frac{V_\infty}{\pi a}}. \tag{16}$$

This formula differs from the Ivantsov formula and we cannot take the limit $V_\infty \to 0$ because the time dependent solution destroys. Remind that within the framework of our statement the hydrodynamic velocity is much greater than the crystallization velocity v. Namely such a regime enables us to separate the Boussinesq problem from the free-boundary problem itself. In opposite limit the explicit solution is more complicated and has been presented in [13, 14].

6 Acknowledgements

We would like to thank Dr. J.R.Ockendon for helpful discussions. Yu.H and K.K. acknowledge support by the RFBR, Grant No 96-01-01221.

References

[1] D. Bensimon, L.P. Kadanoff, S. Liang, B.I. Shraiman, and C. Tang, Rev. Mod. Phys. **58**, 977 (1986).

[2] J.S. Langer, In *Direction in Condensed Matter Physics*, Memorial volume in honor of S.K.Ma, ed ted by G. Grinstein and G. Masenko (World Scientific, Singapore, 1986), 165 (1986).

[3] S.D. Howison, Euro. J. Appl. Math. **3**, 209 (1992).

[4] G.P. Ivantsov, Dokl. Akad. Nauk. SSSR, **58**, 567 (1947).

[5] M.E. Goldstein, and R.I. Reid, Proc. R. Soc. Lond. A, **364**, 45 (1978).

[6] V.A. Maksimov, Izvestiya Academii Nauk SSSR, Mekhanica, **4**, 41 (1965).

[7] V.A. Maksimov, Prikl. Matem. Mech., **40**, 264(1976).

[8] K. Kornev, and G. Mukhamadullina, Proc. R. Soc. Lond. A **447**, 281 (1994).

[9] L. Van Wijngaarden, Proc. Koninkl. Nederl. Akad. Wet. **69**, 263 (1966).

[10] D.Kessler, J.Koplik and H.Levine, Adv.Phys., **37**, 255(1988).

[11] E.A.Brener and V.I.Melnikov, Adv.Phys., **40**, 53(1991).

[12] P.Bouissou, B.Perrin and P.Tabeling, Phys.Rev.A, **40**, 509(1989).

[13] M.Benamar, P.Bouissou and P.Pelce, J.Crystal Growth,**92**, 97(1988).

[14] S.K.Dash and W.N.Gill, J.Heat Mass Transfer, **27**, 1345(1984).

[15] R.Ananth, and W.N.Gill, J.Fluid Mech.,**208**, 575(1989).

[16] R.Ananth, and W.N.Gill, J.Crystal Growth,**108**, 173(1991).

[17] Jian-Jun Xu, J.Crystal Growth,**128**, 219(1993).

[18] M.E.Glicksman, S.R.Coriell and G.B.McFadden, Annu.Rev.Fluid.Mech.307(1986).

[19] M.J.Boussinesq, J.Math.Pure.Appl., **1**, 285(1905).

Calculations of Entropy and Critical Phenomena by Monte Carlo Methods [1]

Chin-Kun Hu

Institute of Physics,Academia Sinica, Taipei 11529, Taiwan
e-mail: huck@phys.sinica.edu.tw

Abstract. Calculations of entropy and critical phenomena are challenging problems in computational physics, which have attracted the attention of many researchers. We have proposed histogram Monte Carlo simulation methods (HMCSM) to solve these problems. There are two versions of the HMCSM: simple sampling (SS) and important sampling (IS). Our SS-HMCSM is simpler than our IS-HMCSM; it may be applied to random percolation models on very large lattices and correlated percolations corresponding to spin models or hard-core particle models on medium-sized lattices. Our IS-HMCSM may be applied to correlated percolation on large lattices and it is still simpler than the method of Ferrenberg and Swendsen. Our HMCSM may be applied to calculate very accurate entropy of phase transition models with relatively little effort. The HMCSM may also be combined with a percolation renormalization group method to obtain very accurate critical point, critical exponents, and thermodynamic order parameter of phase transition models. Recently, we applied our HMCSM to calculate the existence probability E_p, the percolation probability P, and the probability W_n for the appearance of n, $n = 1, 2, \ldots$, top to bottom percolating clusters for bond and site percolation on finite square (sq), planar triangular (tr), and honeycomb (hc) lattices. Here $E_p(G,p)$ is the probability that the system percolates and $P(G,p)$ is the probability that a given lattice site belongs to a percolating cluster. We found that, by choosing appropriate aspect ratios for each lattice and nonuniversal metric factors for each model, we can obtain universal scaling functions for E_p, P, and W_n for six different percolation models on planar lattices.

1 Introduction

It is my pleasure and honor to come from Academia Sinica [1], Taipei [2], Taiwan [3] to give a talk at CMCP'96 at Joint Institute for Nuclear Research [4], Dubna, Russia. Taiwan is located between Korea and the Philippines and to the east of Mainland China. Although the total population in Taiwan is only about 21 millions, Taiwan ranks third in the world in manufacturing computer products, after only USA and Japan. Therefore, it seems appropriate that I come from Taiwan to give a talk at CMCP'96.

I begin by outlining some of the important and interesting current problems in statistical and computational physics.

1. Calculation of entropy by a Monte Carlo method

It follows from the second law of thermodynamics that for every thermodynamic system we may define a state function, called entropy and denoted by S, which is a fuction of the thermodynamic variables of that state and does not depend on the history of the system [5]. The second law of thermodynamics implies that the entropy is a monotonic increasing function of time. It is easy to calculate the internal energy U, the specific heat C_h, the

[1] Key lecture presented at International Conference: Computational Modelling and Computing in Physics, 16-21 September 1996 (CMCP'96), Dubna, Russia.

magnetization M, etc of a thermodynamic system by Monte Carlo methods; however, it is not so easy to calculate entropy. In the traditional Monte Carlo calculations of entropy for a thermodynamic systems at a temperature T, one should calculate the specific heat of the system from zero temperature to temperature T, then use numerical integration of the specific heat from 0 to T to obtain the entropy at T [6]. This procedure is very time consuming and tedious. In the last few years in his life, Shang-Keng Ma tried various methods to calculate entropy by Monte Carlo simulations and proposed a coincidence method [7], which may give entropy of a thermodynamic system at T from Monte Carlo data at T only. However, the method is very time consuming. The time required to obtain the entropy of a system of size N is propotional to $exp(cN)$, where c is a constant. We need more efficient ways to calculate entropy of a thermodynamic system.

2. Calculation of order parameter for phase transition models

A very important quantity in phase transition problems is the order parameter, e.g. the spontaneous magnetization in the paramagnet to ferromagnet transition, the density difference of two kinds of atoms in a sublattice of a binary alloy, etc. Famous models for phase transitions include the Ising model [8, 9, 10], the q-state Potts model [11], and the percolation model [12]. In 1952, C.-N. Yang spent much effort to obtain an exact solution for spontaneous magnetization of the Ising model on a square lattice [9]. In 1972, Yang said:"It seems to me that the prospects of finding an exactly soluble two-dimensional quantum mechanical problem or a three-dimensional classical problem of a nontrivial type are rather remote. Perhaps the combinatorial problem of paths in three dimensions is just too complicated [13]." More than 20 years have passed. The progress in the exact solutions of order parameters for phase transition models is very slow. What Yang said in 1972 is still true today. We must try to find other methods to calculate accurate order parameter for phase transition models.

3. Calculation of universal scaling functions for critical phenomena

Universality and scaling are important concepts developed in statistical physics, which have applications in many branches of sciences. In 1971, M. E. Fisher proposed the idea of finite-size scaling in static critical phenomena [14] and in 1977 M. Suzuki proposed the idea of finite-size scaling in dynamic critical phenomena [15]. Before 1984, it was widely believed that universality holds for critical exponents and not for finite-size scaling functions, e.g., all two dimensional Ising models have the same set of critical exponents, but different lattices have different finite-size scaling functions. In 1984, Privman and Fisher proposed the idea of universal finite-size scaling functions and nonuniversal metric factors in static critical phenomena [16]. For about 10 years after 1984, there have been no important progress in the calculation of the universal finite-size scaling functions. New approaches to attack this problem is needed.

We have proposed histogram Monte Carlo simulation methods (HMCSM) [17, 18, 19, 20], which are useful for solving these problems. There are two versions of the HMCSM: simple sampling (SS) [18, 21] and important sampling (IS) [19, 20]. Our SS-HMCSM is simpler than our IS-HMCSM; it may be applied to random percolation models on very large lattices and correlated percolations corresponding to spin models or hard-core

particle models [22, 23] on medium-sized lattices. Our IS-HMCSM may be applied to correlated percolation on large lattices and it is still simpler than the method of Ferrenberg and Swendsen [24]. Our HMCSM may be applied to calculate very accurate entropy of phase transition models with relatively little effort. The HMCSM may also be combined with a percolation renormalization group method [25] to obtain very accurate critical point, critical exponents, and thermodynamic order parameter of phase transition models [17, 21].

We have also applied our HMCSM to calculate the existence probability E_p, the percolation probability P, and the probability W_n for the appearance of n, $n = 1, 2, \ldots$, top to bottom percolating clusters on finite square (sq), planar triangular (pt), and honeycomb (hc) lattices. Here $E_p(G,p)$ is the probability that the system percolates and $P(G,p)$ is the probability that a given lattice site belongs to a percolating cluster. We found that, by choosing appropriate aspect ratios for each lattice and nonuniversal metric factors for each model, we can obtain universal scaling functions for E_p, P, and W_n for six different percolation models on planar lattices [26, 27, 28, 29].

2 Calculation of entropy

The free energy f of a thermodynamic system is related to the the internal energy U and the entropy S of that system by $f = U - TS$. Since it is easy to calculate U, the difficult problem of calculating S is eqivalent to the problem of calculating f. In this section, we present HMCSM to calculate f. We first present SS-HMCSM [18] for the q-state Potts model (QPM), then outline main ideas for IS-HMCSM [19, 20].

Based on the subgraph expansion of Ising-type models in external fields, Hu has shown that phase transitions of many Ising-type models can be described as geometric percolation transitions [22, 30, 31, 32, 33]. In particular Hu has shown that phase transitions of the QPM are percolation transitions of a q-state bond-correlated percolation model (QBCPM) [22, 31], in which a subgraphs G' of $b(G')$ occupied bonds and $n(G')$ clusters will appear with the probability weight:

$$\pi(G', p, q) = p^{b(G')}(1-p)^{E-b(G')}q^{n(G')}, \tag{1}$$

where $p = 1 - exp(-K) = 1 - exp(-\beta J)$ with J being the nearest neighbo ferromagnetic coupling constant and $\beta = (k_B T)^{-1}$. The spontaneous magnetization M and the magnetic susceptibility χ of the QPM are related to the percolation probability P and the mean cluster size S of the QBCPM, respectively [22, 31]. In the following, we present our method for the QPM and the QBCPM [22, 31]. The extension to other systems and the potential applications of the method are discussed at the end of [18].

In the SS-HMCSM [18] for the QBCPM on a lattice G of N sites and E nearest neighbor bonds, the bond random percolation process is used to generate subgraphs G' of G using a sequence of bond probabilities of increasing magnitudes: $0 \le p_1 < p_2 < p_3 \ldots < p_w \le 1$. For each p_j, $1 \le j \le w$, we generate N_R different subgraphs G' of G. The total number of occupied bonds in G', $b(G')$, may be calculated easily. The multiple labeling technique [34] is applied to G' to calculate the total number of clusters in G', $n(G')$, and the total number of sites in percolating clusters, $N^*(G')$, which is 0 for nonpercolating subgraphs. The data obtained from wN_R different G' are then used to construct three matries with

elements $N_p(b,n)$, $N_f(b,n)$, and $N_{pp}(b,n)$, with $0 \leq b \leq E$ and $1 \leq n \leq N$. Here $N_p(b,n)$ is the total number of generated percolating subgraphs with b occupied bond and n clusters, $N_f(b,n)$ is the total number of generated nonpercolating subgraphs with b occupied bonds and n clusters, and $N_{pp}(b,n)$ is the sum of $N^*(G')$ over percolating subgraphs with b occupied bonds and n clusters. In the large number of simulations, we expect that the total number of percolating subgraphs with b occupied bonds and n clusters, $N_{tp}(b,n)$, and the total number of nonpercolating subgraphs with b occupied bonds and n clusters, $N_{tf}(b,n)$, should be proportional to $N_p(b,n)$ and $N_f(b,n)$ with the same proportional constant $C(b)$, which may be determined from the following equation:

$$C(b)\sum_{n=1}^{N}[N_p(b,n)+N_f(b,n)]=\sum_{n=1}^{N}[N_{tp}(b,n)+N_{tf}(b,n)]=C_b^E. \tag{2}$$

Now we define $I(p,q,b,n)$ by the equation: $I(p,q,b,n)=(e^K-1)^b q^n = e^{KE}p^b(1-p)^{E-b}q^n$. The sum of $I(p,q,b,n)$ over all possible percolating subgraphs G'_p gives $R(G,p,q)$; the sum of $I(p\ q,b,n)$ over all possible nonpercolating subgraphs gives $Q(G,p,q)$. In the HMCSM, $R(G,p,q)$, $Q(G,p,q)$, and the percolation probability $P(G,p,q)$ may be calculated from following equations

$$R(G,p,q)=\sum_{b=0}^{E}\sum_{n=1}^{N}I(p,q,b,n)C_b^E\frac{N_p(b,n)q^n}{\sum_{n=1}^{N}[N_p(b,n)+N_f(b,n)]}, \tag{3}$$

$$Q(G,p,q)=\sum_{b=0}^{E}\sum_{n=1}^{N}I(p,q,b,n)C_b^E\frac{N_f(b,n)q^n}{\sum_{n=1}^{N}[N_p(b,n)+N_f(b,n)]}, \tag{4}$$

$$P(G,p,q)=\frac{1}{Z_N}\sum_{b=0}^{E}\sum_{n=1}^{N}I(p,q,b,n)C_b^E\frac{N_{pp}(b,n)q^n/N}{\sum_{n=1}^{N}[N_p(b,n)+N_f(b,n)]}. \tag{5}$$

The sum of $R(G,p,q)$ and $Q(G,p,q)$ gives the partition function $Z_N(G,p,q)$, from which we may calculate the free energy $f(G,p,q)$

$$f(G,p,q)=lnZ_N(G,p,q)/N=F_e(G,p,q)+Kz/2, \tag{6}$$

where $F_e(G,p,q)$ is the difference between $f(G,p,q)$ and $Kz/2$ with z being the coordination number of the lattice. It approaches lnq as $p \to 0$ and approaches 0 as $p \to 1$ and $N \to \infty$. The internal energy U and the specific heat C_h may be obtained from the derivatives of the free energy

$$U(G,p,q)=-\frac{\partial}{\partial\beta}f(G,p,q),\ C_h=\frac{\partial}{\partial T}U(G,p,q). \tag{7}$$

The existence probability $E_p(G,p,q)$ is defined to be the ratio of $R(G,p,q)$ and $Z(G,p,q)$. It should be noted that F_e, U, C_h, E_p, and P presented above are continuous function of p. This is quite different from the results obtained by traditional MC simulation methods [6]. We have calculated Fe, U, C_h, E_p, and P for the QPM on lattices of different linear dimensions L. Typical calculated results for the square (sq) lattice are shown in Figs. 1(a)-1(b) of [18] where the exact results of Ferdinand and Fisher [10] are shown by dotted

lines for comparison. Fig. 1(a) of [18] shows that our Fe, U, and C_h are consistent with the exact results [10] to high degree of accuracy.

For large lattices, the SS-HMCSM gives low statistics to elements of $N_p(b,n)$, $N_f(b,n)$, and $N_{pp}(b,n)$ with large n and results obtained from Eqs.(3)-(5) for large q are not accurate. To overcome this problem, we may use an importance-sampling method, e.g. the cluster algorithm [35], to calculate histograms, then use a sum rule similar to that of Eq.(2) to analyze the data [17, 19, 20]. Using such IS-HMCSM, we have obtained accurate F_e, U, and C_h for the Ising model on a large lattice [20].

3 Calculation of order parameter

Suppose we already carry out histogram MC simulations on lattices G_1 and G_2 of linear dimensions L_1 and L_2, respectively, where $L_1 > L_2$. The percolation RG transformation from lattice G_1 to lattice G_2 is given by [25]:

$$E_p(G_2, p', q) = E_p(G_1, p, q), \tag{8}$$

which gives the renormalized bond probability p' as a function of p. The fixed point of Eq.(8) gives the critical point p_c. The thermal scaling power y_t and the field scaling power y_h, which is equal to the fractal dimension D of the percolating cluster at p_c [36, 37], may be obtained from the equation:

$$\frac{1}{\nu} = y_t = \frac{(\ln \frac{\partial p'}{\partial p})_{p_c}}{\ln \frac{L_1}{L_2}}, \quad y_h = D = \frac{\ln \frac{P(G_1', p_c, q) L_1^d}{P(G_2', p_c, q) L_2^d}}{\ln \frac{L_1}{L_2}}. \tag{9}$$

Using Eqs. (8) and (9) with $L_1 = 16$ and $L_2 = 8$, we calculated p_c, y_t, and y_h for the QPM on the sq lattice. The results were shown in Fig. 2(a) of [18], where the exact solutions [11] were also shown by solid lines for comparison. Figure 2(a) shows that our method gives quite accurate results. Using the method of Ref.[25], we iterated Eq.(8) to obtain the percolation probability, i.e. the spontaneous magnetization, for the QPM on the sq lattice in the thermodynamic limit. The results were shown by solid lines in Fig. 2(b), where the exact result of Yang [9] for the Ising model was shown by a dotted line for comparison. The agreement of our result with Yang's exact solution [9] is excellent. We also used the same method to obtain very accurate order parameter for site percolation on sq and simple cubic lattices [21], which is certainly a *nontrivial type* [13] phase transition model.

4 Universal scaling functions

In 1995 $\sim$ 1996, we applied the HMCSM to calculate the existence probability E_p, the percolation probability P, and the probability W_n for the appearance of n, $n = 1, 2, \ldots$, top to bottom percolating clusters on finite square (sq), planar triangular (pt), and honeycomb (hc) lattices. Here $E_p(G,p)$ is the probability that the system percolates and $P(G,p)$ is the probability that a given lattice site belongs to a percolating cluster. We found that, by choosing appropriate aspect ratios for each lattice and nonuniversal metric

factors for each model, we could obtain universal scaling functions for E_p, P, and W_n for six different percolation models on planar lattices [26, 27, 28, 29]. We also found that the nonuniversal metric factors are independent of the boundary conditions and extra factors multiplied on aspect ratios of each lattice [27, 29]. The values of W_n at the critical point are useful for understanding σ_{xx}^{max} in the quantum Hall effects [38, 39, 40]. Very recently, Wang and I used a Monte Carlo method to find that the continuum percolation of soft disks and hard disks have the same universal scaling functions as the percolation on planar lattices [41]. Wang and I also used heat bath dynamics and Glauber dynamics to evaluate the dynamic critical exponent z and the dynamic finite-size scaling function of an Ising model on sq, pt, and hc lattices. We found convincing evidence that z is universal and, by choosing appropriate nonuniversal metric factor for each lattice, the scaling function is also universal [42]. Based on the connection between the q-state Potts model (QPM) and a q-state bond-correlated percolation model (QBCPM) [22], Chen and I used a cluster Monte Carlo simulation method [35] to calculate the universal scaling functions for the QPM on sq, hc, and pt lattices [43]. The nonuniversal metric factors for the Ising model , i.e. QPM with $q = 2$, are consistent with nonuniversal metric factors in dynamic critical phenomena obtained by Wang and Hu [42].

5 Final remarks

In summary, we have developed histogram Monte Carlo simulation methods, which may be implemented easily, yet they give fairly accurate critical points, critical exponents, and order parameters. We have used such methods to find universal scaling functions for bond and site percolation on lattices.

Many scientists and engineers are working very hard to improve computer hardware, software, and networks. Improved technology has helped computational physics to make progress very quickly. This trend will continue in the future and computers will play an increasingly important role in our research.

6 Acknowledgement

This work was supported by the National Science Council [44] of the Republic of China (Taiwan) under grant numbers NSC 86-2112-M-001-001 and NCHC-86-02-008. The authors thank the Computing Center of Academia Sinica [45] (Taipei) and National Center for High-Performance Computing (Taiwan) [46] for providing the computing and research facilities.

References

[1] For further information, contact: http://www.sinica.edu.tw/.
[2] WWW site: http://www.tcg.gov.tw/.
[3] WWW site: http://www.gio.gov.tw/.
[4] WWW site: http://www.jinr.dubna.su/.
[5] K. Huang, *Statistical Mechanics* 2nd. ed. (Wiley, New York) 1987.

[6] *Monte Carlo Methods in Statistical Physics*, 2nd ed. edited by K. Binder, Topics in Current Physics, Vol.7 (Springer-Verlag, Heidelberg, 1986).
[7] S.-K. Ma, J. Stat. Phys. **26**, 221 (1981).
[8] L. Onsager, Phys. Rev. **65**, 117 (1944).
[9] C. N. Yang, Phys. Rev. **85**, 808 (1952).
[10] A. F. Ferdinand and M. E. Fisher, Phys. Rev. **185**, 832 (1969).
[11] F. Y. Wu, Rev. Mod. Phys. **54**, 235 (1982).
[12] D. Stauffer and A.Aharony, *Introduction to Percolation Theory* 2nd. ed. (Taylor and Francis, London, 1992).
[13] C. N. Yang, in *Phase Transition and Critical Phenomena vol.1*, edited by C. Domb and M. S. Green (Academic, New York, 1972), pp. 1-5.
[14] M. E. Fisher, in *Proc. 1970 E. Fermi Int. School of Physics*, M. S. Green ed. (Academic, NY, 1971) Vol. 51, p. 1.
[15] M. Suzuki, Prog. Theor. Phys. **58** , 1142 (1977).
[16] V. Privman and M. E. Fisher, Phys. Rev. **B30**, 322 (1984).
[17] C.-K. Hu, Phys. Rev. B **46**, 6592 (1992).
[18] C.-K. Hu, Phys. Rev. Lett. **69**, 2739 (1992).
[19] C.-K. Hu, Phys. Rev. Lett. **70**, 2045 (1993).
[20] J.-A. Chen and C.-K. Hu, Phys. Rev. B **50**, 6260 (1994).
[21] C.-K. Hu, Phys. Rev. B **51**, 3922 (1995).
[22] C.-K. Hu, Phys. Rev. B **29**, 5103 and 5109 (1984).
[23] C.-K. Hu and S.-K. Mak, Phys. Rev. B **39**, 2948 (1989); **42**, 965 (1990).
[24] A. M. Ferrenberg and R. H. Swendsen, Phys. Rev. Lett. **63**, 1195 (1989).
[25] C.-K. Hu and C.-N. Chen, Phys. Rev. **B38**, 2765 (1988); Phys. Lett. A **130**, 436 (1988); Phys. Rev. **B39**, 4449 (1989) and **B43**, 6184 (1991).
[26] C.-K. Hu, C.-Y. Lin, and J.-A. Chen, Phys. Rev. Lett. **75**, 193 and 2786E (1995).
[27] C.-K. Hu, C.-Y. Lin, and J.-A. Chen, Physica A **221**, 80 (1995).
[28] C.-K. Hu, J. Korean Physical Soc. (Proc. Suppl.), **29**, S97-101 (1996).
[29] C.-K. Hu and C.-Y. Lin, Phys. Rev. Lett. **77**, 8 (1996).
[30] For a recent review, see C.-K. Hu, in *Computer-Aided Statistical Physics*, edited by C.-K. Hu (AIP, New York, 1992), pp. 79-101.
[31] C.-K. Hu, Physica **116A**, 265 (1982) and **119A**, 609 (1983); Chin. J. Phys. (Taipei), **22**, no. 1, 1 (1984).
[32] C.-K. Hu, Chin. J. Phys. (Taipei) **23**, 47 (1985).
[33] C.-K. Hu, Phys. Rev. **B44**, 170 (1991).
[34] J. Hoshen and R. Kopleman, Phys. Rev. **B14**, 3438 (1976).
[35] R. H. Swendsen and J. S. Wang, Phys. Rev. Lett, **58** 86 (1987).
[36] H. E. Stanley, J. Phys. **A10**, L211 (1977).
[37] T. A. Larsson, J. Phys. A **19**, 2383 (1986).
[38] I. M. Ruzin, N. R. Cooper, and B. I. Halperin, Phys. Rev. B. **53**, 1558 (1996).
[39] C.-K. Hu and B. I. Halperin, Phys. Rev. B **55**, 2705 (1997).
[40] N. R. Cooper, B. I. Halperin, C.-K. Hu, I. M. Ruzin, Phys. Rev. B **55**, 4551 (1997).
[41] C.-K. Hu and F.-G. Wang, preprint.
[42] F.-G. Wang and C.-K. Hu, preprint.
[43] J.-A. Chen and C.-K. Hu, unpublished.
[44] WWW site: http:/www.nsc.gov.tw/.
[45] WWW site: http:/www.sinica.edu.tw/cc/.
[46] WWW site: http:/www.nchc.gov.tw/.

FAST MULTILEVEL ITERATIVE SOLVER FOR HARMONIC BOUNDARY VALUE PROBLEM ON POLYGONAL DOMAIN.

B.N.Khoromskij and E.G.Nikonov.

141980 Dubna, Moscow region, Russia JINR LCTA.
e-mail: nikonov@cv.jinr.ru

Abstract. Fast iterative solver for harmonic boundary value problem (BVP) on polygonal domain is presented in this paper. Harmonic BVP was solved by conjugate gradient (CG) method with multilevel preconditioner[4] based on optimal algorithm of a matrix-vector multiplication. This matrix corresponds to a discrete Poincaré-Steklov operator. Results of comparative analysis with cascadic CG[3] and computers tests are presented.

Consider an application of cascadic conjugate-gradient algorithm (CCG-algorithm) to solving BVP for Laplace equation on right triangular domain. Let Ω be a triangular domain composed of M matching rectangles R_i and K regular right triangles T_i, such that $\bar{\Omega} = (\bigcup_{i=1}^{M} \bar{R}_i) \bigcup (\bigcup_{i=M+1}^{M+K} \bar{T}_i)$ (see Fig. 1).

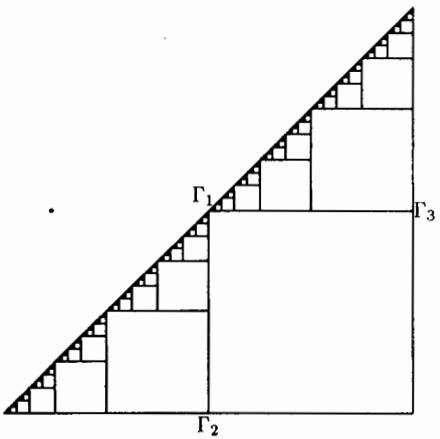


Figure 1:

Let Υ_h be the piecewise uniform triangulation of Ω aligned with the skeleton $\Gamma_{sk} = \bigcup_i \Gamma_i$, $\Gamma_i = \partial R_i$ or $\Gamma_i = \partial T_i$ including $\Gamma = \partial\Omega$. Assume the triangulation of any subdomain R_i or T_i to be uniform. Starting with the above coarse decomposition of Ω we introduce a more refined one assuming that any triangle T_i is composed of $n_i - 1$ rectangles and n_i triangles of size $O(h)$, as proposed in [1],[2]. This produces the so-called refined skeleton $\Gamma'_{sk} \supset \Gamma_{sk}$, which defines the resultant decomposition of Ω into M' rectangles R_i and

K' triangles T_i (we use just the same notations for corresponding subdomains as in the case of coarse decomposition). The fast algorithm of computations with Poincaré-Steklov operator $S_{\Gamma,h}$, defined on the refined skeleton Γ_{sk} is based on the fast computations of the harmonic Poincaré-Steklov operator S_h, defined on each rectangles R_i producing Γ_{sk} with the complexity [1]:

$$\log \varepsilon^{-1} \sum_{i=1}^{M} O(n_i \log^2 n_i),$$

where ε is the error reduction factor.

Apply CCG-algorithm to the following mixed BVP.

$$\begin{aligned} \Delta \bar{u} &= 0 \quad \text{in} \quad \Omega; \\ \bar{u}\,|_{\Gamma_1} &= 0; \\ \frac{\partial \bar{u}}{\partial \vec{n}}\,|_{\Gamma_2} &= \varphi_2; \\ \frac{\partial \bar{u}}{\partial \vec{n}}\,|_{\Gamma_3} &= \varphi_3. \end{aligned} \tag{1}$$

This BVP is equivalent to the operator equation for Poincaré-Steklov operator

$$Su = \varphi. \tag{2}$$

Here u is a trace of $\bar{u}$ on Γ_{sk}, φ is equal $\frac{\partial \bar{u}}{\partial \vec{n}}$ on exterior boundary and equal zero on interior lines of Γ_{sk}. Then consider a sequence of finite-dimensional vector spaces $M_0, M_1, ..., M_l$ corresponding a sequence of skeletons $(\Gamma_{sk})_0, (\Gamma_{sk})_1, ..., (\Gamma_{sk})_l$. Assume also that linear "interpolation" operators

$$I_i : M_i \to M_{i+1}, \quad i = 0, 1, ..., l-1, \tag{3}$$

are given. Besides we have linear invertible operators

$$S_i : M_i \to M_i, \quad i = 0, 1, ..., l. \tag{4}$$

Then the main objective of the CCG-algorithm consists in solving the following problem in a vector space M_l: *for a given* $\varphi_l \in M_l$, *find* $u_l \in M_l$ *such that*

$$S_l u_l = \varphi_l. \tag{5}$$

The main feature of the cascadic algorithms consists in successively solving the problems: *for a given* $\varphi_i \in M_i$, *find* $u_l \in M_l$ *such that*

$$S_i u_i = \varphi_i, \quad i = 0, 1, ..., l. \tag{6}$$

Now we can formulate the CCG-algorithm in more detail [3].

CCG-algorithm.

{ 1. Set $u_0 = S_0^{-1}\varphi_0$.
2. For $i = 1, 2, ..., l$ do:
{ 2.1. Set $w_i = I_{i-1}u_{i-1}$;
2.2. Do m_i iterations of the conjugate-gradient method:
$y_0 = w_i$;
$p_0 = r_0 = \varphi_i - S_i y_0$;
$\sigma_0 = (r_0, r_0)_i$;
for $i = 1, 2, ..., m_i$ do:
{ $\alpha_{k-1} = (p_{k-1}, S_i p_{k-1})/\sigma_{k-1}$;
$y_k = y_{k-1} + \frac{1}{\alpha_{k-1}} p_{k-1}$;
$r_k = r_{k-1} - \frac{1}{\alpha_{k-1}} S_i p_{k-1}$;
$\sigma_k = (r_k, r_k)_i$;
$\beta_k = \sigma_k/\sigma_{k-1}$;
$p_k = r_k + \beta_k p_{k-1}$;
} the end of the iteration;
2.3. Set $u_i = y_{m_i}$;
} the end of the level i ;
} the end of the algorithm.

Computer tests were made for the problem (1,2) with the model function:

$$u^*(x,y) = \frac{sinh(\alpha_k(y+x))}{sinh(\alpha_k a x)} sin(\alpha(y-x)), \quad \alpha_k = \frac{\frac{\sqrt{2}}{2}\pi k}{B}. \tag{7}$$

Results of computations without applaying CCG-algorithm are presented in table 1. After applying CCG-algorithm results shown in table 2 can be obtained.

i	N_{scal}	iter	Init.approx.	$\|u_i - u^*\|_C$	$\|u_i - u^*\|_{L_2}$	Time
2	9	10	0.00E+00	3.94E-02	3.15E-02	2.28E-01
3	25	20	0.00E+00	9.91E-03	8.15E-03	4.38E-01
4	65	34	0.00E+00	2.60E-03	2.11E-03	1.31E+00
5	161	51	0.00E+00	6.54E-04	5.44E-04	4.59E+00
6	385	75	0.00E+00	1.64E-04	1.39E-04	1.78E+01
7	897	108	0.00E+00	4.10E-05	3.49E-05	6.26E+01
8	2049	154	0.00E+00	1.02E-05	8.77E-06	2.29E+02
9	4609	219	0.00E+00	2.57E-06	2.20E-06	8.39E+02
10	10241	311	0.00E+00	6.67E-07	5.62E-07	3.12E+03

Table 1: CG.$\|S_i^{-1}u_i - \varphi_i\|_{L_2}/\|S_i^{-1}u_i^0 - \varphi_i\|_{L_2} \le \varepsilon_{10}$, $\varepsilon_{10} = 9.54E-07$

Using multilevel Schvarz method with BPX preconditioner [4] we obtain the following algorithm with CG as a smoother: Find the solution of $S_l u = \varphi_l$ by solving the preconditioned system

$$B_l^{-1} S_l u = B_l^{-1} \varphi_l, \tag{8}$$

i	N_{scel}	iter	Init.approx.	$\|u_i - u^*\|_C$	$\|u_i - u^*\|_{L_2}$	Time
2	9	10	6.74E-01	3.94E-02	3.15E-02	1.8)E-01
3	25	22	2.65E-01	9.91E-03	8.15E-03	4.6 3E-01
4	65	35	9.16E-02	2.60E-03	2.11E-03	1.5 E+00
5	161	48	2.85E-02	6.54E-04	5.44E-04	4.6 E+00
6	385	63	8.30E-03	1.63E-04	1.39E-04	1.5 E+01
7	897	80	2.33E-03	4.09E-05	3.49E-05	4.8 E+01
8	2049	90	6.39E-04	1.02E-05	8.78E-06	1.4 E+02
9	4609	107	1.73E-04	2.55E-06	2.23E-06	4.4 E+02
10	10241	65	4.62E-05	9.67E-07	1.23E-06	9.2 E+02

Table 2: CCG .$\|S_i^{-1}u_i - \varphi_i\|_{L_2} \leq \varepsilon_i$, $\varepsilon_{i-1} = \varepsilon_i/3$, $\varepsilon_{10} = 9.54E-09$

where

$$B_l^{-1} = \Pi_0 S_0^{-1}\Pi_0^t + \Pi_1 D_1^{-1}\Pi_1^t + ... + \Pi_{l-1} D_{l-1}^{-1}\Pi_{l-1}^t + D_l^{-1}. \quad (9)$$

Here $\Pi_i : M_i \to M_l$ is the standard nodal value interpolation operator and $\Pi_i^t : M_l \to M_i$ is the adjoint operator of Π_i. The diagonal elements of D_i contain information on the shapes of the triangles and the coefficients of the problem. The results of computations are presented in the Table 3.

m	N_{scel}	iter	Init.approx.	$\|u - u^*\|_C$	$\|u - u^*\|_{L_2}$	Time
0	1	2	5.31E-01	8.45E-01	8.45E-01	0.00 E+00
1	3	8	1.88E+00	1.28E-01	1.31E-01	0.00 E+00
2	9	13	2.66E+00	3.94E-02	3.15E-02	1.00 E+00
3	25	19	3.39E+00	9.91E-03	8.15E-03	1.00 E+00
4	65	26	3.63E+00	2.60E-03	2.11E-03	1.00 E+00
5	161	19	3.71E+00	6.54E-04	5.44E-04	6.00 E+00
6	385	16	3.82E+00	1.63E-04	1.39E-04	1.00 E+01
7	897	16	3.82E+00	4.09E-05	3.49E-05	2.20 E+01
8	2049	17	3.91E+00	1.02E-05	8.77E-06	6.30 E+01
9	4609	20	3.94E+00	2.63E-06	2.19E-06	2.32 E+02
10	10241	20	3.97E+00	6.95E-07	5.97E-07	5.44 E+02

Table 3: PCG with BPX.$\|S^{-1}u_n - f\|_{L_2} \leq 10^{-6}$

Number of iterations (IT) for obtaining the same accuracy as a function of a number of grids (N) is presented on the Fig. 2 for CG, CCG and CG with BPX preconditioner.

These computations were made on CONVEX 220.

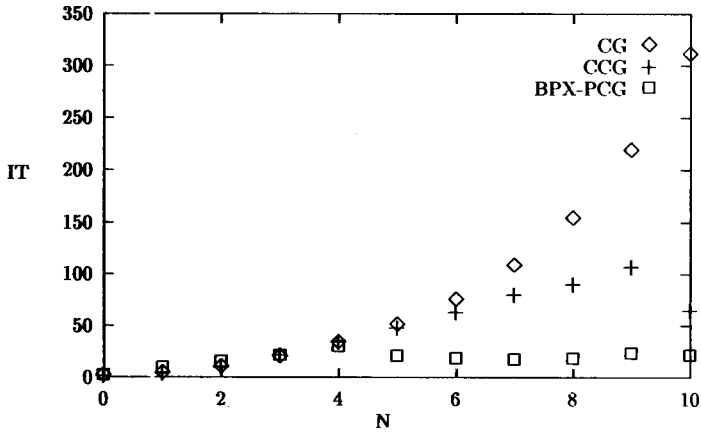


Figure 2:

References

[1] B. N. Khoromskij and G. Schmidt. Preprint No. 162, WIAS, Berlin (1995); Numer. Math. (submitted).

[2] B.N.Khoromskij. Preprint No. 233/6.Jg./TU Chemnitz (1992); Numer. Lin. Alg. with Appl. (1994) (to appear).

[3] V.Shaidurov. Problems and Methods in Mathematical Physics. B.G.Teubner Verlagsgesellschaft Stuttgart - Leipzig 185 (1994).

[4] X.Zhang. Numer. Math. **63**, 521 (1992).

THREE–BODY RESONANCES IN FRAMEWORK OF THE FADDEEV CONFIGURATION SPACE APPROACH

E. A. Kolganova, A. K. Motovilov

Joint Institute for Nuclear Research, Dubna, Russia
e-mail: kea@thsun1.jinr.ru, motovilv@thsun1.jinr.ru

Abstract. Algorithm, based on explicit representations for analytic continuation of the T-matrix Faddeev components on unphysical sheets, is worked out for calculations of resonances in the three-body quantum problem. According to the representations, poles of the T matrix, scattering matrix and resolvent on unphysical sheets, interpreted as resonances, coincide with those complex energy values where appropriate truncations of the scattering matrix have zero as eigenvalue. Scattering amplitudes on the physical sheet, necessary to construct scattering matrix, are calculated on the basis of the Faddeev differential equations. Effectiveness of the algorithm developed is demonstrated for example of searching for resonances in the system nnp and in a model three-boson system.

We make a numerical test of the approach proposed in Ref. [1] to treat the three body resonances in the case of pairwise interactions falling off in coordinate space not slower than exponentially. This approach is based on a construction of explicit representations for the Faddeev components $M_{\alpha\beta}(z)$, $\alpha, \beta = 1, 2, 3$, of the three body T matrix $T(z)$ as well as for the scattering matrices $S(z)$ and resolvent $R(z)$ continued on unphysical sheets of the energy z plane. For the sheets, the notation Π_l is used with l, an enumerating multi index (see [1]). The representations constructed demonstrate a structure of kernels of the above operators after continuation and give new capacities for analytical and numerical studies of three body resonances. In particular the representations for analytic continuation $M(z)|_{\Pi_l}$ of the matrix $M(z) = \{M_{\alpha\beta}(z)\}$ on the sheet Π_l with details omitted read

$$M|_{\Pi_l} = M|_{\Pi_0} - Q_l^\dagger \mathbf{J}^\dagger A S_l^{-1} \mathbf{J} Q_l$$

The operator $Q_l(z)$ and the "transposed" one, $Q_l^\dagger(z)$ are obviously constructed of the matrix $M(z)$ taken on the physical sheet Π_0. $A(z)$ is a number matrix, an entire function of $z \in \mathbb{C}$. By $S_l(z)$ we understand a truncation (depending essentially on l) of the total three body scattering matrix $S(z)$. Operators $\mathbf{J}(z)$ and $\mathbf{J}^\dagger(z)$ realize a restriction of kernels of the operators $Q_l(z)$ and $Q_l^\dagger(z)$ on the energy shells, respectively, in first and last momentum variables. So that the products $Q_l^\dagger \mathbf{J}^\dagger$ and $\mathbf{J} Q_l$ have half on shell kernels. Representations [1] for analytical continuation of the three body scattering matrices and resolvent follow immediately from the representations above for $M(z)|_{\Pi_l}$.

As follows from the representations constructed, the nontrivial (i.e. differing from the poles at the discrete spectrum eigenvalues of the three body Hamiltonian) singularities of the T matrix, scattering matrices and resolvent situated on an unphysical sheet Π_l, are singularities of the inverse truncated scattering matrix $S_l^{-1}(z)$. Therefore, the resonances on the sheet Π_l considered as poles of the T matrix, scattering matrix and resolvent continued on Π_l are those values of the energy z for which the matrix $S_l(z)$ has zero as eigenvalue. Thereby, to search for the resonances situated on a certain unphysical

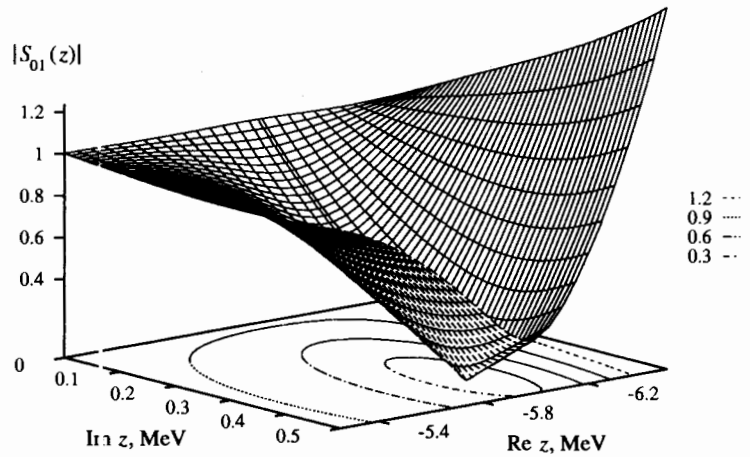


Figure 1: Surface of the function $|S_{01}(z)|$ in the model system of three bosons with the nucleon masses. The potential $V^G(r)$ is used with the barrier $V_b = 1.5$ MeV. Position of the resonance $z_{res}(3B)$ corresponds to the minimal (zero) value of $|S_{01}(z)|$.

sheet II_l, one can apply any method allowing to compute analytical continuation on the physical sheet of the elastic scattering, rearrangement or breakup amplitudes necessary for construction of respective $S_l(z)$. In particular such is the algorithm developed for $(2 \to 2,3)$ processes in Ref. [2] on the base of Faddeev differential formulation of the scattering problem in configuration space (see also [3, 4] and Refs. therein). It is only necessary to go out in this formulation on the complex plane of z including the asymptotical boundary conditions.

In the present work we utilize a code based on the algorithm [2]–[4], for computations of s-state *nnp* resonances situated on the unphysical sheet $II_{(0,1)}$ connected with physical one by crossing the spectral interval $(E_d, 0)$ between the deuteron energy $z = E_d$ and breakup threshold $z=0$. We solve the two–dimensional Faddeev integro–differential equations [4] with the $(2 \to 2,3)$ asymptotical boundary conditions [2]–[4] at complex energies z and extract the truncated s-state scattering matrix $S_{01}(z) = 1 + 2ia_0(z)$ with $a_0(z)$, the amplitude of elastic *nd* scattering continued on the physical sheet. When making a finite–difference approximation of the equations above in polar coordinates we take up to 180 points of grids in both radial and angular variables, the cut–off radius being up to 39 fm. As a *NN*–interaction, the Malfliet–Tjon potential MT I–III [5] is chosen.

Firstly, we have checked a validity of the code finding the ^3H bound–state energy E_t as a pole of the function $S_{01}(z)$. More precisely, the location of the $S_{01}(z)$ pole was found

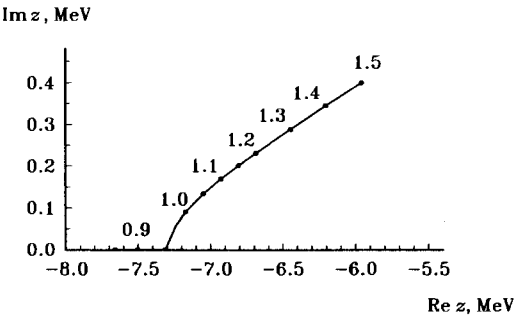


Figure 2: Trajectory of the resonance $z_{res}(3B)$ on the sheet $II_{(0,1)}$ in the model system of three bosons with the nucleon masses. The potential $V^G(r)$ is used. Values of the barrier V_b in MeV are given near the points marked on the curve.

as a root of the inverse amplitude $1/a_0(z)$. Beginning from the grid dimension 80×80, we have obtained $E_t = -8.55$ MeV. Hereafter all the energies are given with respect to the breakup threshold. Note that the value stated is in a good agreement with known results on E_t in the MT I–III model (see Ref. [6]).

Concerning the *nnp* resonances on the sheet $II_{(0,1)}$, we have inspected a domain of a range about 10 MeV in vicinity of segment $[E_d, 0]$ in the complex z plane. Especially carefully we studied a vicinity of the points $z = -1.5 \pm 0.3 + i(0.6 \pm 0.3)$ MeV interpreted in Refs. [7, 8] as a location of an exited state energy of ^3H. Unfortunately we have succeeded to find only one root z_{res} of the function $S_{01}(z)$, corresponding to the known virtual state of the *nnp* system at total spin S=1/2. On a 180×180 grid, we have found $z_{res} = -2.728$ MeV i.e. it is situated 0.504 MeV to the left from the *nd* threshold $E_d = -2.224$ MeV (in the MT I–III model). Note that the shift $E_d - z_{res}$ found from experimental data on *nd* scattering, is 0.515 MeV (see Ref. [9]). Its value [9] computed in a separabilized MT I–III model on the base of the momentum space Faddeev equations, is equal to 0.502 MeV. As one could expect (see the data on three-nucleon resonances in [9]), we have failed to find any resonances in the quartet state at $L = 0$ as well as at $L = 1$.

Also, we have studied a behavior of a resonance situated on the unphysical sheet $II_{(0,1)}$ in a model three-body system including bosons with masses of the nucleon. As a pairwise interaction between the bosons we have used the Gauss-type potential supplied with an additional Gauss repulsive barrier term,

$$V^G(r) = V_0 \exp[-\mu_0 r^2] + V_b \exp[-\mu_b (r - r_b)^2]$$

where the values $V_0 = -55$ MeV, $\mu_0 = 0.2$ fm^{-2}, $\mu_b = 0.01$ fm^{-2}, $r_b = 5$ fm have been fixed while the barrier amplitude V_b varied. A resonance (with non-zero imaginary part) on the sheet $II_{(0,1)}$ arises in the system concerned just due to the presence of the barrier term. Example of a surface of the $|S_{01}(z)|$ function for the barrier amplitude $V_b = 1.5$ MeV is shown in Fig. 1 (for a 80×80 grid). A trajectory of the resonance $z_{res}(3B)$ (a zero of

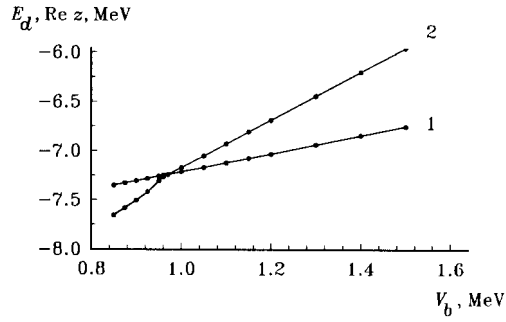


Figure 3: Dependence of the "deuteron" energy E_d (curve 1) and real part of the resonance $z_{\rm res}(3B)$ (curve 2) on the barrier value V_b.

the function $S_0(z)$) is shown for the changing barrier V_b in Fig. 2. This trajectory was watched for the barrier V_b decreasing in the interval between 1.5 MeV and 0.85 MeV. When drawing the trajectory, we have used a 160×160 grid. It can be seen from Fig. 2 that the behavior of the resonance $z_{\rm res}(3B)$ is rather expected: with monotonously decreasing real part, the imaginary part of the resonance changes also monotonously. For $V_b \geq 0.95$ MeV the energy $z_{\rm res}(3B)$ becomes real, turning into a discrete spectrum eigenvalue. In Fig. 3 we plot both the trajectories of the resonance real part $\mathrm{Re}\, z_{\rm res}(3B)$ and two-boson binding energy E_d. As one can see from this figure, the value of $|\,\mathrm{Re}\, z_{\rm res}(3B)|$ increases more quickly than $|E_d|$, coinciding with $|E_d|$ at $V_b \cong 0.97$ MeV.

Trajectory of the resonance concerned in the lower complex half-plane is symmetric to the curve shown in Fig. 2 with respect to the real axis. Respective points, symmetric to those marked in Fig. 2, correspond to the same values of V_b.

References

[1] A. K. Motovilov, Teor. Mat. Fiz. **107**, 450 (1996) (LANL E-print `nucl-th/9505028`); *Ibid*, 478 (LANL E-print `nucl-th/9505029`).

[2] S.P.Merkuriev, C.Gignoux and A.Laverne, Ann. Phys. (N.Y.) **99**, 30 (1976).

[3] A. A. Kvitsinsky, Yu. A. Kuperin, S. P. Merkuriev, A. K. Motovilov, and S. L. Yakovlev, Sov. J. Part. Nucl. **17**, 113 (1986).

[4] L. D. Faddeev and S. P. Merkuriev, *Quantum scattering theory for several particle systems.* Kluwer Academic Publishers, 1993.

[5] R. A. Malfliet and J. A. Tjon, Ann. Phys. (N.Y.) **61**, 425 (1970).

[6] J. Carbonell, C. Gignoux, and S. P. Mekuriev, Few–Body Systems **15**, 15 (1993).

[7] D. V. Aleksandrov, E. Yu. Nikolsky, B. G. Novatsky, and D. N. Stepanov, Pis'ma v ZhETF **59**, 301 (1994).

[8] A. L. Baranov, Pis'ma v ZhETF **61**, 9 (1995).

[9] K. Möller and Yu. V. Orlov, Fiz. Elem. Chast. At. Yadra **20**, 1341 (1989).

TRANSITIONS FROM PERIODIC TO GLASS-LIKE STRUCTURES IN FOAM SATURATED CAPILLARIES

Konstantin Kornev and Valerii Mourzenko

101, prosp. Vernadskogo, Institute for Problems in Mechanics RAS, 117526, Moscow, Russia
e-mail: kidin@ipm.msk.su

Raphail Dautov

18, Lenina st., Kazan University, Dep. of Numerical Math. and Cyb., 420008, Kazan, Russia
e-mail:rafail.dautov@ksu.ru

Abstract. We consider a model of 1D foam confined in a wavy capillary. The discrete model is a kind of Ulam map which admits many metastable states as governing parameters overcome certain barriers. Randomization of lamella positions speaks in favour of the glass-like patterning. The start-up pressure drop is studied in a wide range of the input physical parameters.

1 Introduction

Recent studies [1, 2] have shown that foam, because of its unique structure, reduces gas flow in porous media. This blocking action makes foam a promising blocking fluid for underground gas storage and for other engineering applications [1]. It has been shown [1, 2], that at any given instant of time, foam flows in a small fraction of the pores. The rest of the pores contain "trapped foam" which effectively blocks the flow. As a result, the permeability to gas is reduced by several orders of magnitude over that which would be assumed with a gas-liquid system without a foaming agent.

In addition to reduction of permeability, foam drastically changes the rheological behavior of the gas phase [1, 2]. In particular, in the presence of a foam, gas flows as if it were a homogeneous fluid with a start-up yield pressure drop. Such a start-up pressure drop must be applied before foam will move through porous media.

This paper concerns the physical nature of the start-up yield pressure drop.

2 Formulation

We shall discuss the nature of foam patterning under the following assumptions.

1.We shall model the porous medium as a rigid capillary with a radius (Fig. 1)

$$r = r_0 + \delta \cos\Big(\frac{2\pi x}{\lambda}\Big). \tag{1}$$

Here r_0 , λ and δ are some characteristic scales of porous medium and the x-- axis coincides with the axis of the symmetry of a capillary.

2. The pore aspect ratio is a low $\delta << r_0 << \lambda$.

3. In an initial undeformed state, foam is perfectly ordered. Under a load, all the bubbles keep the same mass, i.e., we prohibit diffusion to occur.

4. The gas n the bubbles is ideal.

5. One can gnore the Plateau borders and assume that the lamellae intersect the pore walls at 90° anjgles [2].

Under these assumptions one can write the model of 1D foam in dimensionless variables as the map [4] ı Fig. 1)

$$p_{i+1} - p_i = -2\pi\mu \sin\rho_i, \quad \mu = \frac{4\sigma\delta}{P_g r_0 \lambda}, \tag{2}$$

$$\rho_{i+1} - \rho_i = \frac{2\pi K}{p_{i+1}} - 2\pi K, \tag{3}$$

$$p_0 = P_{ext}, \qquad p_N = P, \tag{4}$$

Here p_i is tl e gas pressure within the i-th bubble, normalized by P_g. The parameter P_g is the initial gas pressure in an individual bubble. ρ_i is the displacement of the i-th lamella, normal zed by $\lambda/2\pi$. The displacement is measured from its equilibrium position at the Ki–th throat of the channel. $K\lambda$ is the "wagon length", i.e., the distance between adjacent lamellae in an initial unperturbed state. K is an integer number, in Fig. 1 $K = 1$. σ is the surface tension. P and P_{ext} are the dimensionless applied pressures. Hereafter, we shall consider the case in which $P = 1$. The parameter μ serves as a measure of the intensity of capillary forces with respect to the elastic forces.

The dimensi onless energy of the system can be expressed as

$$\frac{2F}{P_g \lambda r_0^2} = \frac{2\pi K}{N}\left\{\sum_{i=1}^{N-1}\left(\ln p_i + \frac{P}{p_i}\right) - \frac{\mu}{K}\sum_{i=0}^{N-1}\cos\rho_i\right\} + \frac{(P - P_{ext})\rho_0}{N}. \tag{5}$$

where N is the number of lamellae in the train.

We turn now to the physical picture of the displacement of the bubble train. It should be remembered that whenever the last lamella is free, i.e. $P = 1$, there is a single pinning force - capillary force, which tends to keep the last $N - 1$-th lamella at an equilibrium position. Moreover, this force resists forward displacement only within the period $\rho_{N-1} \in [0, \lambda/2]$, where the lamella bulges forward. Within this period, the profile of this force has a single extreme point – the maximum. Therefore, it might be expected that the bubble train begins motion immediately after the last lamella has reached the extreme point. The pressure drop required to overcome this barrier will be called the critical pressure drop, $G = P_{ext} - 1$. Precisely speaking, if the dependence $G(\rho_{N-1})$ is nonmonotonic, the maximum of possible $G(\rho_{N-1})$ within the period $0 < \rho_{N-1} < \lambda/2$ should be called the critical pressure drop. By this reason, the solution of the boundary value problem is replaced into the solution of the Cauchy problem: the knowledge of $P = 1$ and ρ_{N-1} determ nes recursively p_i and ρ_i for all i. Then Eqs.(2) and (3) define such a map $\mathbf{T}$ that a point $\mathbf{U}_{+1} = (p_{i+1}, \rho_{i+1})$ is transformed into point $\mathbf{U}_i = \mathbf{T}(\mathbf{U}_{i+1}) = (p_i, \rho_i)$. This map is usually called the Ulam map [3]. Within the framework of the problem of foam patterning, we seek the relation of $P_{ext} = P_{ext}(\rho_{N-1})$, where the parameter ρ_{N-1} is used to distinguish different solutions to Eqs. (2) and (3). Then we determine the maximum pressure drop among those calculated. The goal of this study consists in solving the reverse problem, i.e., the prediction of critical pressure drop as a function of the input parameters.

3 Results

The well–known Ulam model [3] contains some characteristic features of the model under consideration. The main result lies in the fact that the Ulam map has stochastic solutions which lie within stochastic domains, separated by domains of regular solutions [3]. All the trajectories, starting from the vicinity of hyperbolic point $\rho = 0$, $p = 1$, belong to stochastic layer. The stochastic layer is separated by regular trajectories from the upper and lower regions of regularity. Near the elliptic singular point $\rho = \pi$, $p = 1$ there is a "stability island" separated from the stochastic layer by a cyclic regular trajectory.

Fig. 2 shows the typical curve $P_{ext}(\rho_{N-1})$ for long trains, $N = 101$, $K = 1$, $\mu = 0.008$. In the description of the steady equilibrium state of the bubble train we start with pair $(\rho_{N-1}, 1)$, $\rho_{N-1} \to 0$ and use map **T**. The expected result, based on the analysis of the continuous model [4], is that the trajectory is unable to overcome the barrier – separatrix of the respective continuous dynamic system. But the discrete picture is more complex. Namely, when displacement ρ_{N-1} becomes nearer to zero beyond a certain critical value, chaotic behavior is observed, Fig. 2. This means that the corresponding trajectories of solutions to Eqs. (2), (3) pass through the stochastic layer on the phase plane. The randomization of the lamella distribution speaks in favour of the glass-like patterning of foam.

The energy F, as a function of ρ_{N-1}, is presented in Fig. 3 for the same parameters. From this Figure, we see that the solutions for trains with $N = 101$ and minima energies correspond to small displacements of the last lamella, ρ_{N-1}. On the other hand, the randomization of the bubble train is observed (Fig. 3) within the same range of the displacements of the last lamella as that in Fig. 2. A dashed line in Fig. 3 corresponds to the respective energy in continuous approximation.

The complexity of the foam texture can be revealed by analyzing the number of domains in the train. We use here the term "domain", in order to distinguish the region within which the displacements of the lamellae are almost constant, Fig. 4. In our numerical experiments ($N = 101$, $K = 1$, $\mu = 0.008$), we found that the states with more complex structure, containing several domains of extended or compressed bubbles, may have a smaller energy than those which have a lesser number of domains. Different labels mark the distinguishing initial data, ρ_{N-1}. This is contrary to the tendency which might be expected in the region of the regular solutions of Eqs.(2), (3), where the more domain walls are associated with the larger energy.

In this work we choose the ground state of a charged bubble chain by variation of the initial lamella positions under slow continuous variation of the pressure P_{ext}. This configuration has the smallest ρ_{N-1}, among other states with the same P_{ext}, and the simplest domain structure, $N_D = 1$.

The maximum pressure drop G is estimated by using the relation between P_{ext} and ρ_{N-1}, near the first maximum. This is the smallest estimate because we expect the onset of motion as soon as the applied pressure drop overcomes this barrier.

Figure 5 demonstrates the first maximum pressure $P_{ext}(\rho_{N-1})$ which we observed in the range $\rho_{N-1} \to 0$ for various μ and K. One can see that, for a small μ and K, the maximum admissible pressure is well predicted by the continuous model [4]. When μ or K increase, the prediction of the continuous model underestimates the admissible pressure. This can be explained by the fact that a solution of Eqs. (2), (3) follows an integral curve

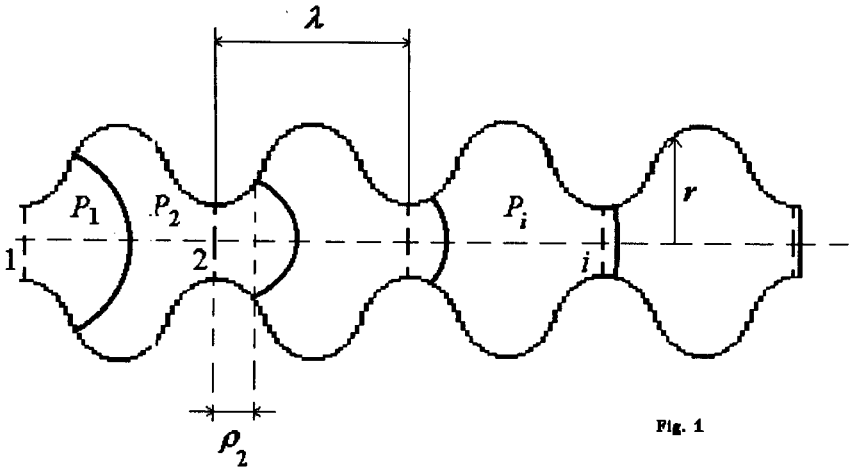


Fig. 1

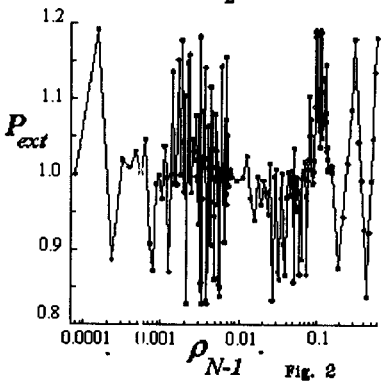


Fig. 2

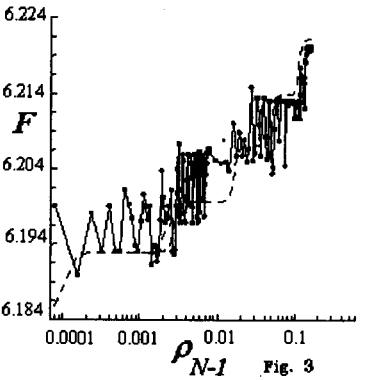


Fig. 3

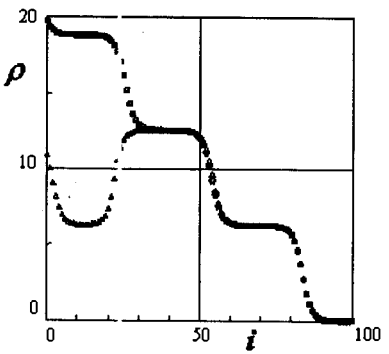


Fig. 4

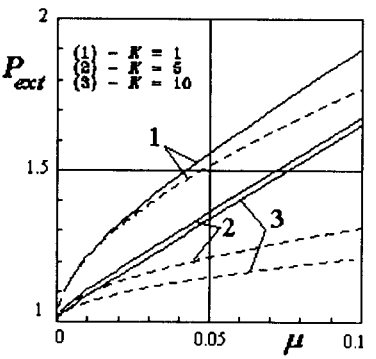


Fig. 5

of the respective continuous dynamic system only approximately. The original validity of the continuous model can be quantitatively estimated with the help of the corresponding linearized model. Moreover, to elucidate the assertion we consider the semi-infinite chain $N \to \infty$. The solution to the corresponding linearized problem can be written in the form $\rho_i = ae^{-\omega i}$, where a is a constant defined by the boundary conditions and ω satisfies the following transcendental equation

$$\cosh\omega = 1 + 4\pi^2 K\mu. \tag{6}$$

Parameter $N_{cor} \sim 1/\omega$ can be called the correlation length. But it is better to call this parameter the screening length, because it defines the characteristic distance over which the pressure in the bubbles decreases from P_{ext} to 1.

In regimes where colligative properties of foam play the crucial role, the parameter $K\mu$ is less than unit and the screening length of a caravan is estimated as

$$N_{cor} \sim \frac{1}{2\pi\sqrt{K\mu}}, \quad K\mu << 1. \tag{7}$$

In the opposite case

$$N_{cor} \sim \frac{1}{\ln(2 + 4\pi^2 K\mu)}, \quad K\mu >> \frac{1}{4\pi^2}. \tag{8}$$

Thus, in the former case, the load is distributed over a large number of lamellae. In the last case, the small number of first lamellae bear the main load. Therefore, one expects that a continuous approximation of model (2)–(3) appropriately describes the characteristic feature of the foam patterning, whenever $K\mu << 1$ and contrarily, the discrete effects should be taken into account for cases where $K\mu \geq 1/4\pi^2$.

References

[1] *Foams: Physics, Chemistry and Structure*, edited by A.J.Wilson (Springer-Verlag, Berlin Heidelburg, 1989).

[2] A.R. Kovscek and C.J. Radke in *Foams:Fundamentals and Applications in the Petroleum Industry*, edited by L.L. Schramm L.L., (Advances in Chemistry Series 242, 1994), pp.115-164.

[3] A.J. Lichtenberg and M.A. Lieberman, *Regular and Chaotic Dynamics*, 2end ed., (Springer-Verlag, New York, 1992).

[4] K.G. Kornev JETP, **80**, 1049 (1995).

MODELLING OF TOROIDAL PLASMA EQUILIBRIUM EVOLUTION WITH ACCOUNT OF NON-OHMIC CURRENTS

D.P. Kostomarov, F.S. Zaitsev, V.V. Nefedov, A.G. Shishkin, D.C. Robinson*, M. Cox*, M.R. O'Brien*

Moscow State University,
Faculty of Computational Mathematics and Cybernetics,
**UKAEA, Fusion, Culham, Abingdon, United Kingdom*
(Euratom/UKAEA Fusion Association)
e-mail: zaitsev@cs.msu.su, martin.obrien@ukaea.org.uk

Abstract. The problem of toroidal plasma equilibrium evolution with account of non-ohmic currents is considered. The mathematical formulation of the problem is given with discussion of its unusual features. A computationally tractable numerical algorithm has been developed, and the code SCoPE written to solve the problem. Results of numerical modelling for tokamaks plasma of different geometries are presented. It is shown that the intrinsic pressure-gradient-driven currents can play an important role in the evolution process. In particular, these currents can lead to the current density being greatest away from the plasma centre, which may indicate how improved plasma confinement regimes could be obtained. Results from studies of how to access to hollow current density profile and advanced confinement regimes are discussed. Formulation and solution of the current profile control problem using external current-drive systems are presented. Examples of maintaining a given radial profile in the case of strong time-variation of plasma parameters are given.

1 Introduction

The work is devoted to mathematical modelling of phenomena in toroidal systems with magnetic plasma confinement, such as tokamaks.

In "quiet" regimes the tokamak plasma is in a near-equilibrium state. This state, however, does not remain unchanged. On the one hand different plasma processes lead to a slow rearrangement of magnetic field, by changing the pressure and current, and on the other hand these plasma processes are themselves influenced by the magnetic field. Thus there is self-consistent evolution of the plasma column and its gradual transition from one quasi-equilibrium state to another.

Most existing models either give too general a formulation, which is very difficult to implement, or concentrate on one or other particular aspect of the problem, ignoring the others. In particular they do not allow accurate self-consistent description of some important phenomena, such as externally and internally driven non-ohmic currents.

This paper is devoted to modelling of toroidal plasma equilibrium evolution with account of IPGD (Intrinsic Pressure-Gradient-Driven) currents - so called bootstrap, Pfirsch-Schluter, diamagnetic currents [1-5] - and external currents, driven by high frequency electromagnetic waves. Accurate description of non-ohmic currents is very important for tokamaks since, for example, bootstrap current can be a large fraction of total plasma current. In this paper we concentrate on perspective tight aspect ratio tokamaks

(which have small ratio of major and minor radii of the torus), such as START or MAST at Culham Laboratory, UK.

2 Mathematical Formulation of the problem

With the help of a special choice of phase space variables and the introduction of special unknown functions, the problem of toroidal plasma equilibrium evolution with account of non-ohmic currents can be reduced to the coupled solution of (i) the system of two, strongly non-linear equations involving the partial derivatives for the magnetic field and (ii) the system of four non-linear transport equations that determine the evolution of quantities like the spatial distribution of plasma pressure [1,3, 6-8].

A computationally tractable numerical algorithm was developed, and the code SCoPE (**Self-Consistent Plasma Evolution**) was written to solve the problem. However, in this paper, for the purpose of separating different physical effects, we consider equations (i) only, replacing the solution of system (ii) with some model functions of time and space. Equations (i) are

$$\frac{\partial}{\partial\rho}\left(\frac{g_{22}}{r}\hat{I}\right)-\frac{\partial}{\partial\theta}\left(\frac{g_{12}}{r}\hat{I}\right)=-\frac{\mu_0\mu}{\hat{I}}\left(r\frac{\partial P}{\partial\rho}+\frac{1}{2\mu_0\mu r}\frac{\partial F^2}{\partial\rho}\right), \tag{1}$$

$$\frac{\partial\Psi}{\partial t}-\frac{\rho}{\Phi_{\max}}\frac{d\Phi_{\max}}{dt}\frac{\partial\Psi}{\partial\rho}=\frac{1}{\mu_0\mu\sigma_{\|}c_{33}^2}\frac{\partial}{\partial\rho}\left(c_{22}c_{33}\frac{\partial\Psi}{\partial\rho}\right)+\frac{1}{\sigma_{\|}\Phi_{\max}}\int_0^{2\pi}\vec{j}_{ext}\cdot\vec{B}\sqrt{g}d\theta, \tag{2}$$

Eq. (1) is equation for the instantaneous equilibrium ("the Grad-Shafranov"equation) and Eq. (2) describes the evolution. System (1), (2) is a complicated strongly non-linear system of integro-differential equations. It is an unusual system since time derivative enters Eq. (2) only. Here $\hat{I}(t,\rho,\theta)$ and $\Psi(t,\rho)$ are unknown functions of time, $\rho\equiv\Phi/\Phi_{\max}(t)$ and generalised poloidal angle θ, $(\rho,\theta)\in[0,1]\times[0,2\pi]$, $t\in[0,T_{\max}]$, $\Psi(t,\rho)$ - is poloidal magnetic flux, $\hat{I}(t,\rho,\theta)$ - is the Jacobian of transformation of (Ψ,θ) to (r,z) coordinates taken with sign minus, Φ - is toroidal magnetic flux and $\Phi_{\max}$ - its maximum value ($\Phi_{\max}(t)$ is considered as a given function). Other quantities are defined by

$$F(t,\rho)=-\frac{\Phi_{\max}}{2\pi}/\left(\frac{1}{2\pi}\int_0^{2\pi}\frac{d\theta}{r\hat{I}}\frac{\partial\Psi}{\partial\rho}\right),\quad P(t,\rho)=\sum_\alpha kn_\alpha T_\alpha,$$

$$g_{11}=\left(\frac{\partial r}{\partial\rho}\right)^2+\left(\frac{\partial z}{\partial\rho}\right)^2,\quad g_{22}=\left(\frac{\partial r}{\partial\theta}\right)^2+\left(\frac{\partial z}{\partial\theta}\right)^2,\quad g_{33}=r^2,$$

$$g_{12}=g_{21}=\frac{\partial r}{\partial\rho}\frac{\partial r}{\partial\theta}+\frac{\partial z}{\partial\rho}\frac{\partial z}{\partial\theta},\quad g_{13}=g_{23}=g_{31}=g_{32}=0,$$

$$c_{22}(t,\rho)=\frac{1}{2\pi}\int_0^{2\pi}\frac{g_{22}}{\sqrt{g}}d\theta,\quad c_{33}(t,\rho)=\frac{1}{2\pi}\int_0^{2\pi}\frac{\sqrt{g}}{g_{33}}d\theta,\quad \sqrt{g}=-\frac{r}{\hat{I}}\frac{d\Psi}{d\rho},$$

$$r(t,\rho,\theta)=r_m(t)+\delta(t,\rho,\theta)(r_p(t,\theta)-r_m(t)),$$

$$z(t,\rho,\theta)=z_m(t)+\delta(t,\rho,\theta)(z_p(t,\theta)-z_m(t)),$$

$$\delta(t,\rho,\theta)=\sqrt{\int_0^\rho\frac{\partial\Psi(t,x)}{\partial x}/\hat{I}(t,x,\theta)dx}/\int_0^1\frac{\partial\Psi(t,x)}{\partial x}/\hat{I}(t,x,\theta)dx,$$

$$r_m(t)=\lim_{\rho\to0}\frac{\partial}{\partial S}\int_S r(t,\rho,\theta)dS,\quad z_m(t)=\lim_{\rho\to0}\frac{\partial}{\partial S}\int_S z(t,\rho,\theta)dS.$$

with $(r_m(t), z_r(t))$ - coordinates of the magnetic axis, $(r_p(t,\theta), z_p(t,\theta))$ - coordinates of plasma boundary, $S = S(t,\rho)$ - area of cross-section within a magnetic surface ρ =const, P - plasma pressure, $\vec{j}_{ext}$ - external and IPGD currents. Expressions for neoclassical conductivity $\sigma_{||}$ and IPGD currents $\vec{j}_{ext} \cdot \vec{B}$ are clearly written in, e.g. Ref. [5]. These expressions involve dependence on the magnetic field geometry, plasma temperature T_α and density n_α, etc.

System (1), (2) is completed by the following boundary conditions $\lim_{\rho\to 0} g_{22}\hat{I}/r = 0$, $\hat{I}(t,\rho,0) = \hat{I}(t,\rho,2\pi)$, $\lim_{\rho\to 0}|\Psi(t,\rho,\theta)| < \infty$, $\Psi(t,1) = \Psi_p(t)$ - given poloidal flux at the edge or $\partial\Psi(t,1)/\partial\rho = -\mu_0\mu I_p(t)\ /(2\pi c_{22}(t,1))$ - given total current $I_p(t)$, and initial conditions $\partial\Psi(0,\rho)/\partial\rho$, $\partial P(0,\rho)/\partial\rho$, $\partial F^2(0,\rho)/\partial\rho$.

The code SCoPE can solve this specified problem numerically as one of its options.

3 Numerical algorithm

The problem of numerical approximation and solution of equilibrium equation (1) was studied thoroughly in Ref. [8]. So in this section we concentrate on the numerical algorithm as a whole and on the evolution equation (2) in particular.

An important feature of approximation of system of equations (1) and (2) is discretisation over time. We assign equilibrium equation (1) to the whole time points t^n and evolution equation (2) to half points $t^{n+1/2}$, calculating metric coefficients in eq. (2) at points t^n from solution of eq. (1) and taking $(\Psi^{n+1/2} + \Psi^{n-1/2})/2$ instead of Ψ in ρ derivatives. This procedure gives the second order of approximation over time at points t^n.

Eq. (2) involves ρ derivative of the first order. To obtain a monotonic numerical scheme we rewrite it as [9]

$$v^\rho\frac{\partial\Psi}{\partial\rho} = \frac{1}{\alpha}\frac{\partial}{\partial\rho}\left(\alpha v^\rho\frac{\partial\Psi}{\partial\rho}\right) - \frac{\partial}{\partial\rho}\left(v^\rho\frac{\partial\Psi}{\partial\rho}\right), \quad v^\rho \equiv \frac{\rho}{\Phi_{\max}}\frac{d\Phi_{\max}}{dt}, \quad \alpha \equiv \exp\rho.$$

At internal points of the ρ grid, the ρ derivative operator in Eq. (2) is approximated with 3-point difference operator of the second order of approximation using technique of [9].

At boundaries $\rho = 0$ and $\rho = 1$ the difference operator is calculated from the following conditions with the second order of approximation. At $\rho = \rho_{i=1} = 0$ spatial derivative operator in Eq. (2) is approximated with

$$A_i\frac{-4B_{i+1} - B_{i+2}}{4h^2}\bar{\Psi}_i + A_i\frac{B_{i+2}}{h^2}\bar{\Psi}_{i+1} + A_i\frac{4B_{i+1} - 3B_{i+2}}{4h^2}\bar{\Psi}_{i+2}\Big|_{i=1},$$

where

$$\bar{\Psi} \equiv (\Psi^{n+1/2} + \Psi^{n-1/2})/2, \quad A \equiv 1/\left(\mu_0\mu\sigma_{||}c_{33}^2\right), \quad B \equiv c_{22}c_{33}.$$

Here we used $B_1 = v_1^\rho = 0$. At $\rho = \rho_{i=N} = 1$ the boundary condition is for given poloidal flux

$$\bar{\Psi}_{i=N} = \Psi_p(t_n)$$

or for given total current

$$\frac{\bar{\Psi}_{i-2} - 4\bar{\Psi}_{i-1} + 3\bar{\Psi}_i}{2h}\Big|_{i=N} = -\frac{\mu_0\mu I_p(t^n)}{2\pi c_{22,N}(t^n,1)}.$$

The scheme can be resolved, e.g., using one of the fast algorithms for 5-diagonal matrix.

4 Access to advanced regimes

A fixed plasma boundary problem for START-like plasma was considered with account of three IPGD currents: bootstrap, Pfirsch-Schluter and diamagnetic. Calculations start from some reasonable initial equilibrium and run until a quasi-stationary regime is obtained. A large number of runs have been performed with various plasma parameters. Different plasma characterisitics were analysed, including how the safety factor $q \equiv -d\Phi/d\Psi/(2\pi)$ varies across the plasma as this variation determines the plasma stability. We discovered that in some cases monotonic initial safety factor q-profile eventually becomes non-monotonic at the magnetic axis which is appropriate to the so called advanced confinement regime. Special numerical studies allowed to draw a conclusion that the main reason for this behaviour of q is presence of relatively high bootstrap current. This conclusion can be explained by the specific radial dependence of the bootstrap current density. It has maximum off the magnetic axis where the pressure gradient is large and if its value is high enough it can make total current density to be hollow at the magnetic axis which may result in non-monotonic q-profile.

Firstly, it is important to note that in most of the considered variants the total current density was eventually forced to the outer part of the torus. The total current density has its highest value in this region. This can be explained by some compression of the flux surfaces due to the Shafranov shift on the outside of the magnetic axis resulting in higher pressure gradient which is balance by current according to Grad-Shafranov equation (equation (1)). The bootstrap current in the equatorial plane is equal to zero at the magnetic axis $r = R_o$, and has two maxima located either side of the magnetic axis ($r < R_o$ and $r > R_o$). The maximum at $r < R_o$ can form the second inner maximum of total current, and therefore a hollow total current profile within this maximum and the larger maximum at $r > R_o$. The Pfirsch-Schluter current has the opposite sign to the bootstrap current to the inside of the magnetic axis ($r < R_o$) and therefore reduces the role of the pressure current here: conversely for $r > R_o$ the two pressure-driven currents are in the same direction and the effect is greater. So the existence of a hollow total current density profile and the appearance of a non-monotonic q-profile depend on the balance between the bootstrap and Pfirsch-Schluter currents to the low major radius side of the magnetic axis.

After the effect of bootstrap current on non-monotonic q-profile was discovered we started to search for general plasma parameters responsible for access to this regime. We found that the regime is reached with increase of β_{pol}

$$\beta_{pol} \equiv \frac{2\mu_0\mu \int_S P dS}{\int_S B_{pol}^2 dS}. \qquad (3)$$

This conclusion was determined first on the basis of numerical experiment and then explained analytically using formula (3). Increase in β_{pol} with fixed plasma pressure P is due to decay of the poloidal magnetic field B_{pol}. Bootstrap current is inversely proportional to B_{pol} [1] and total toroidal current is proportional to B_{pol}. This means that the role of bootstrap current can increase in respect to other currents.

Here are examples of calculations which illustrate correlation between β_{pol} and quasi-steady-state equilibrium with non-monotonic/monotonic q-profile. The following plasma

parameters have been considered for the base variant: $R_0 = 0.3$ m, $a = 0.21$ m, $n_e(0) = n_i(0) = 2.5 \times 10^{20}$ m^{-3}, $B_0 = 0.4$ T, $T_e(0) = 0.18$, $T_i(0) = 0.15$ keV with time independent plasma pressure profiles $P(\rho) = kn_e(\rho)T_e(\rho) + kn_i(\rho)T_i(\rho))$, $n_{e,i}(\rho) = n_{e,i}(0)(0.9(1 - \rho)^{1.05} + 0.1)$, $T_{e,i}(\rho) = T_{e,i}(0)(0.9(1 - \rho)^{1.05} + 0.1)$, fixed total current I_p=121 kA and $\partial F^2(0,\rho)/\partial\rho = 0.4(1 - \rho)$. In this variant plasma evolves to a quasi-equilibrium state with non-monotonic q-profile and $\beta_{pol} = 1.21$. (One can note that finding initial plasma parameters for such a case is a separate rather complicated problem which involves many SCoPE runs). For other variants we changed different plasma parameters of the base advanced regime variant and studied their influence on quasi-equilibrium q-profile and the value of β_{pol}. These variants were: (a) lower plasma pressure due to two times drop in plasma density; (b) lower initial β_{pol} given by a different initial $\partial F^2(0,\rho)/\partial\rho = 0.6(1 - \rho)$; (c) growth of total current over time from 121 to 150 kA; (d) increased magnetic field $B_0 = 0.6$ T; (e) increased major radius $R_0 = 0.45$ m. In cases (a)-(c) quasi-equilibrium q-profile became monotonic with β_{pol} equal to 0.91, 0.86 and 0.57 appropriately. In cases (d), (e) the effect of non-monotonic q-profile became a bit stronger and β_{pol} increased to the values of 1.24, 1.33.

Investigation of how access to reversed shear regimes varies with aspect ratio has been performed. Runs with larger aspect ratios have been done, including the values of 3 and 4. In these runs we were changing the position of magnetic axis (as in case (e)) and, besides this, the value of the magnetic filed at the magnetic axis and coefficient Δ in $\partial F^2(0,\rho)/\partial\rho = \Delta(1 - \rho)$ in order to get q at the centre and minimal q in the range of 1.7-2.3 (as specified in the contract): $R_0 = 0.65$, B_0=1.3 T, Δ=0.6 and $R_0 = 0.85$ m, B_0=1.0 T, Δ=0 6. It was found that with increase of the aspect ratio the effect of reversed shear did not disappear. For aspect ratios 3 and 4 the value of β_{pol} became 0.88 and 1.51, bootstrap current was 48.3 and 56.5 kA, total current - 135 and 111 kA appropriately. Current densities and q are plotted in Figs. (e1) and (e2).

Critical values of β_{pol}, required for access to the reversed shear, verses the aspect ratio were found: aspect ratio 1.3 - $\beta_{pol} \sim 1$, 2.1 - $\beta_{pol} \sim 0.8$, 3 - $\beta_{pol} \sim 0.6$ (here we set B_0=1.5 T and Δ=0.6 in order to get initial and quasi-steady-state q near the axis in the range 1.7-2.3), 4 - $\beta_{pol} \sim 0.7$ (here we set B_0=1.0 T and Δ=0.6 in order to get initial and quasi-steady-state q near the axis approximately in the range 1.7-2.3). Critical β_{pol} were calculated with the help of several runs for each aspect ratio. In each run the value of initial total current was increased over the first 7 ms and then kept constant. Runs continued until the increment to initial current was so high that the effect of reversed shear disappeared.

However one should note that comparison of tokamaks with different aspect ratios is a complicated problem. In presented calculations we were changing several SCoPE input parameters, keeping other initial parameters the same, and it is not clear whether obtained quasi-steady-state START-like plasma with aspect ratios of 3 and 4 is realistic or not. The value of β_{pol} appears to be more important for access to the reversed shear than aspect ratio and tokamaks with different aspect ratios should be compared in respect to what values of β_{pol} can be achieved in real experiment.

The observed phenomena may give theoretical indication on how to access the advanced regimes in tight aspect ratio and other tokamaks.

5 Control of the q-profile

In this section we describe formulation of the problem for the q-profile control and present the results of calculations.

The q-profile in START-like plasmas, calculated from solution of Eqs. (1), (2) has to be maintained with 10 % accuracy at every moment of time in respect to a given profile $q_{given}(\rho)$ by varying the magnitude $u_1(t)$ and $u_2(t)$ of two non-inductive current drive systems, one centered at $\rho_{c,1}/a = 0.2$, one at $\rho_{c,2}/a = 0.75$. Both driven current density profiles were assumed to have the following form

$$j_i(t,\rho) = u_i(t)\exp(-(\rho - \rho_{c,i})^2/(\Delta\rho_i)^2), \quad i = 1,2,$$

$\Delta\rho_1/a = 0.3$ and $\Delta_2\rho/a = 0.2$.

Two different norms for measurement of difference in the q-profiles were considered

$$D_{L_1}(t) = \int_{[0,0.2]\cup[0.45,0.55]} \left| \frac{q(t,\rho) - q_{given}(\rho)}{q_{given}(\rho)} \right| d\rho / \int_{[0,0.2]\cup[0.45,0.55} d\rho,$$

$$D_C(t) = \max_{\rho\in[0,0.55]} \left| \frac{q(t,\rho) - q_{given}(\rho)}{q_{given}(\rho)} \right|.$$

Here ranges of integration over ρ are taken over the most interesting regions of plasma column.

The coordinate descent method was used to search for the minima of the functional $D_{L_1}(t)$ or $D_C(t)$. One of the difficulties in the q control problem is that for some extreme external currents equilibrium does not exist. Coordinate descent method allows to treat this by setting the range of amplitudes of externally driven currents.

In numerical implementation the amplitudes of external currents were adjusted on every time step to get the solution of Eqs. (1), (2) which gives the requested q-profile within the required accuracy. The code SCoPE was optimised in respect to the required computer time. In particular amplitudes of external currents were searched on a fixed equilibrium. The increase in CPU time is normally only ~30% compared with a run without q control.

Here we present two examples, which demonstrates the possibility of maintaining a given q-profile within 10% in time varying START-like plasma. In these examples the pressure was increased by changing linearly $T_e(0)$ from 0.1 to 0.18 keV for $t \in [0,7]$ ms, from 0.18 to 0.36 for $t \in [7,20]$ ms and stays constant for $t \in [20,30]$ ms. A fixed boundary problem (1), (2) with IPGD currents present was considered: $R = 0.3$ m, $a = 0.21$ m, $n_e(0) = 2.5 \times 10^{20}$ m^{-3}, $B_0 = 0.4$ T, $T_i(0) = 0.15$ keV. The q-profile was maintained by varying the magnitude of two non-inductive current drive systems. It was found that it is possible to achieve ~10% error with only two current drive systems. Controls obtained with error calculated in L_1 space appeared to be smoother than with error in C space, but the characteristic magnitudes of control currents and qualitative behaviour in both cases were similar.

With implemented control option SCoPE can be used to solve a wide variety of different control problems and to control different plasma parameters. In particular one can consider radial positions of external currents as additional controls. Control modules are

structured and allow easy change of the minimised functional, the number and meaning of the controls.

6 Conclusion

The mathematical formulation of the problem, numerical algorithm and results of modelling for tokamaks plasma of different geometries are presented. It is shown that the IPGD currents can play an important role in the evolution process. In particular, these currents can lead to the current density being greatest away from the plasma centre, which may indicate how improved plasma confinement regimes could be obtained. Results of studying of access to advanced regimes by varying basic plasma parameters are presented. It is found that growth of β_{pol} can lead to advanced regime with non-monotonic q-profile.

Formulation and solution of the current profile control problem using external current-drive systems is presented. The possibility of maintaining a given radial q-profile in case of strong time-variation of plasma parameters is demonstrated.

Acknowledgement

This work was performed as part of a collaboration between UKAEA, Fusion and Moscow State University. The UKAEA work was jointly funded by the UK Department of Trade and Industry and by Euratom. The MSU work was partly funded by the Russian Foundation for Basic Research, grant No. 96-01-01171a.

References

[1] F.L. Hinton, R.D. Hazeltine, Reviews of Modern Phys. **48**, No. 2, Part 1, 239 (1976).
[2] S.P. Hirshman, Phys. Fluids. **31**, 3150 (1988).
[3] Yu.N. Dnestrovskij, D.P. Kostomarov. Numerical Simulations of Plasmas. Springer-Verlag. New-York. 1986.
[4] Y. Wu, R.B. White, Phys. Fluids B. **5**, No. 9, 3291 (1993).
[5] H.R. Wilson, Report UKAEA FUS No. 271. 1994.
[6] L.E. Zakharov, V.D. Shafranov. Preprint No. 3075. I.V. Kurchatov Institute of Atomic Energy. 1978. (In Russian.)
[7] L.E. Zakharov, V.D. Shafranov, Reviews of Plasma Physics. **11**, 118 (1987).
[8] L.M. Degtyarev, V.V. Drozdov, S.Yu. Medvedev. Numerical Modelling of Equilibrium and Stability of Toroidal Plasma. Institute of Applied Mathematics. Moscow. 1989. (In Russian.)
[9] A.A. Samarskii. Theory of Numerical Schemes. Moscow, Nauka. 1983. (In Russian.)

ON NUMERICAL SOLVING OF THE PROBLEM OF PREASSIGNED MOTIONS FORMATION OF CHARGED PARTICLES IN MAGNETIC FIELD [1]

E.D.Kotina

St.Petersburg State University, RUSSIA
198904, St- Petersburg, Bibliotechnay pl. 2

Abstract. The charged particles motions formation in axially simmetric magnetic field in accordance with the preassigned velocities field is considered. The problem of magnetic field finding, initiating the preassigned charged particles dynamics can be reduced to the solving of Goursat problem for the system of quasi-linear differential equations of the first order [1, 2]. We consider the way of solving this system on the base of characteristic method. For the construction of numerical solution of Goursat problem we substitute the differential equations of characteristics for corresponding finite difference equations. The problems of characteristics network construction and the choice of initial conditions are discussed. The algorithm and the formulas for calculation of magnetic field components, providing needed particles dynamics are proposed.

1 Introduction

This paper deals with investigation of the control fields formation in electrodynamics problem. This problem attracted attention of many researchers. In particular, the solving of inverse electrodynamics problem is considered in works of G.A.Grinberg [3], V.T.Ovcharov [4, 5], V.I.Zubov [6]. In the paper the approach, suggested by V.I.Zubov is developed.

We consider the following formulation of the magnetic field determination. Let us have the velocities field

$$\begin{cases} \dfrac{dr}{dt} = \eta_r(r,\theta,z), \\ \dfrac{d\theta}{dt} = \dfrac{1}{r}\eta_\theta(r,\theta,z), \\ \dfrac{dz}{dt} = \eta_z(r,\theta,z). \end{cases} \tag{1}$$

Here r, θ, z are cylindrical coordinates, $\eta_r(r,\theta,z), \eta_\theta(r,\theta,z), \eta_z(r,\theta,z)$ are components of the given vector function η in cylindrical system of coordinates, t is time.

Let us describe charged particles motion in stationary magnetic field by the system of differential equations

[1] This work was supported in part by the Russian Foundation for Fundamental Researches under Grant No. 95-01-00926

$$\begin{cases} \dfrac{dr}{dt} = Y_r, \\ \dfrac{d\theta}{dt} = \dfrac{1}{r}Y_\theta, \\ \dfrac{dz}{dt} = Y_z, \\ m\dfrac{dY}{dt} = eY \times B + mQ, \end{cases} \tag{2}$$

where we have $Y = (Y_r, Y_\theta, Y_z)$ as particle velocity vector in cylindrical system of coordinates, m is particle mass, e is particle charge, $B = (B_r, B_\theta, B_z)$ is vector function, determining magnetic induction, $Q = Q(Y) = (\frac{1}{r}Y_\theta^2, -\frac{1}{r}Y_rY_\theta, 0)^*$ is vector column.

It is necessary to find stationary magnetic field (i.e. vector function B which satisfies Maxwell equation $divB = 0$) in which charged particles motion accords to velocities field (1). This means that in configuration space charged particles trajectories determined by system (2) under the same initial conditions will coincide whith trajectories of the system (1).

We can formulate the following theorem [1, 2].

Theorem 1 *Let η be an arbitrary twice continuously differentiable vector function. Then such magnetic field exists in which charged particle motion defined by the system (2) coincides with motion of the system (1), when initial conditions of these motion are the same , and $\eta^2 = const$. This magnetic field B can be represented in the form*

$$B = -\frac{m}{e}rot\eta + h\eta,$$

where $h = h(r, \theta, z)$ is function satisfying the following condition

$$div(h\eta) = 0.$$

In this paper we consider in details the problem of charged particles dynamics formation in the case of stationary axially symmetric magnetic field, i.e we suppose that $B_\theta = 0, B_r = B_r(r, z), B_z = B_z(r, z)$.

Charged particles motion in stationary axially symmetric magnetic field we describe

by the system of differential equations:

$$\begin{cases} \frac{dr}{dt} = Y_r, \\ \frac{d\theta}{dt} = \frac{1}{r}Y_\theta, \\ \frac{dz}{dt} = Y_z, \\ \frac{dY_r}{dt} = \frac{e}{m}Y_\theta B_z + \frac{Y_\theta^2}{2}, \\ \frac{dY_\theta}{dt} = \frac{e}{m}(Y_z B_r - Y_r B_z) - \frac{Y_r Y_\theta}{2}, \\ \frac{dY_z}{dt} = -\frac{e}{m}Y_\theta B_z. \end{cases} \tag{3}$$

In many problems of focusing and transportation of charged particles we are interested in particles behavior in the plane (r, z). Therefore, we consider the following formulation of the problem.

Let suppose, velocities field is given, as

$$\frac{dr}{dz} = f(r, z), \tag{4}$$

i.e. $f(r, z)$ is a given function, which provides the needed motion of charged particles in the plane (r, z).

It is needed to find stationary axially symmetric magnetic field in which the charged particles motion described by the system (3) coincides with the motions determined by the equations (4) in the plane (r, z).

Let us consider the following system of differential equations [1, 2] :

$$\begin{cases} -fl\frac{\partial L}{\partial z} + l\frac{\partial L}{\partial r} = d_1, \\ -q\frac{\partial h}{\partial z} + g\frac{\partial L}{\partial z} - fq\frac{\partial h}{\partial r} + fg\frac{\partial L}{\partial r} = d_2. \end{cases} \tag{5}$$

Here

$q = c_1^2 - \frac{L^2}{r^2}, g = \frac{hL}{r^2}, l = \frac{m}{e}\frac{L}{r^2}, d_1 = \frac{m}{e}\frac{L^2}{r^3} - \frac{m}{e}Fq,$

$F = \frac{\frac{\partial f}{\partial z} + \frac{\partial f}{\partial r}f}{1+f^2}, d_2 = h(c_1^2\frac{f}{r} + q(\frac{\partial f}{\partial r} - fF)), c_1$ is a total velocity of particle.

We assume, the solution of the system (5) will be twice continuously differentiable vector function $L = L(r, z)$ and continuously differentiable vector function $h = h(r, z)$ making this equations in identities.

We can formulate the following theorem [1, 2].

Theorem 2 *For arbitrary continuously differentiable function f , for which the solution of the system* (5) *exists, there is such axially symmetric magnetic field B, where charged particle motion determined by the system* (3) *coincides in the plane* (r, z) *with the motions*

of equations (4) *The considered motions had of course the same initial conditions. The components of magnetic field B are determined by the formulas*

$$B_r = \frac{n}{e}\frac{\partial L}{\partial z} + fh(\frac{c_1^2 - L^2/r^2}{1+f^2})^{1/2}, B_z = -\frac{m}{er}\frac{\partial L}{\partial r} + h(\frac{c_1^2 - L^2/r^2}{1+f^2})^{1/2}. \quad (6)$$

Further we will consider the algorithm of numerical finding of these components.

2 The characterictics equations

Thus, in the case of axially symmetric stationary magnetic field the problem of charged particles dynamics formation is reduced to solving of the system (5). It is the system of quasi-linear partial differential equations of the first order. The system (5) depends on some arbitrary function $f(r,z)$ determining the velocities field in the plane (r,z). Under any choice of this function the system (5) will be the hyperbolic system.

We consider the way of solving this system on the base of characteristics method.

Let us write out for it the equations of characteristics directions and differential correlations for them. The first family:

$$\begin{aligned} dr &= fdz, \\ qdh + gdL - d_2 dz &= 0. \end{aligned} \quad (7)$$

The second family:

$$\begin{aligned} dz &= -fdr, \\ ldL - d_1 dr &= 0. \end{aligned} \quad (8)$$

Characteristics directions equations are determined just by the function $f(r,z)$ and they do not depend on unknown functions.

3 Goursat problem

Let us consider the problem of finding solution for the system (5). For numerical solving of this problem it is necessary to assign additional conditions. Using differential correlations (7), (8) for the system (5) we will consider the Goursat problem, i.e. we will search for the solutions $L(r,z), h(r,z)$ of the system (5) under the condition that on two characteristics, outgoing from one point, values of L and h are given.

We will assume that function $f(r,z)$, defining velocities field is such that $f(0,z) = f(r,0) = 0$.It is not difficult to notice that in such case straight line $r = 0$ is characteristic for the first family and the straight line $z = 0$ is that for the second one.

Let us formulate Goursat problem in the following way. To find the solution of the system (6) assuming that function L and h are given on characteristics $r = 0$ and $z = 0$

$L(0,z) = L_{ro}(z), \quad h(0,z) = h_{r0}(z), \quad 0 \le z \le z_n$,

$L(r,0) = L_{z0}(r), \quad h(r,0) = h_{z0}(r), \quad 0 \le r \le r_m$, here $L_{ro}(z)$, $L_{z0}(r)$, $h_{r0}(z)$, $h_{z0}(r)$ are given functions, z_n, r_m are given values. In addition , values of corresponding functions, given on characteristics, coincide in the generic point

$L_{r0}(0) = L_{z0}(0), \quad h_{r0}(0) = h_{z0}(0)$.

Naturally, it is supposed, that given functions $L_{r0}(z), L_{z0}(r), h_{r0}(z), h_{z0}(r)$ on every characteristic are such that differential characteristics equations are satisfied.

Let $L_{r0} = 0$, $L_{z0}(r) = -kr^2$, here k is some constant. We determine the values of function h on mentioned above characteristics, i.e. functions $h_{r0}(z)$, $h_{z0}(r)$, using corresponding differential equations on considered characteristics.

As result we get

$h_{z0}(r) = -\frac{m}{e}k(c_1^2 - k^2r^2)^{-1/2} - \frac{m}{ekr}\frac{\partial f}{\partial z}(r,0)(c_1^2 - k^2r^2)^{1/2}$,

and by solving differential equation

$dh_{r0} + h_{r0}(\frac{f}{r}(0,z) + \frac{\partial f}{\partial r}(0,z))dz = 0$

with initial conditions $h_{r0}(0) = -\frac{mk}{ec_1} - \frac{mc_1}{ek}(\frac{\partial f}{\partial z}/r)(0,0)$ we will find the function $h_{r0}(z)$.

For the construction of numerical solution of Goursat problem we used Masso method, based on the substitution of differential equations of characteristics for corresponding finite difference equations.

As differential equations of characteristics direction not depend on functions L, h , so at first we can construct network, formed by intersections of characteristics families. Afterwards we will calculate values of functions L and h in nodes of obtained network. Then by values of L and h in nodes of network of characteristics , using formula (6) , we will obtain magnetic field components providing needed particles dynamics.

4 CONCLUSION

The proposed method allows to find the magnetic field, initiating the needed motions of charged particles. Notice, that the functions $L_{r0}(z), L_{z0}(r)$ assignment contains some difficulty. Although obtained magnetic field will satisfy the main task of providing given dynamics of charged particles, it may, however, be that this magnetic field does not satisfy one or another practical requirement connected with realization of this field. Bearing this in mind , these functions can be considered as control functions. So, we can consider not only the problem of magnetic field finding , providing given dynamics of charged particles, but needed characteristics for magnetic field.

References

[1] E.D.Kotina, Ph.D.Thesis, St.Petersburg State University, St.Petersburg, 1994.

[2] E.D.Kotina, Proc. of the Second Intern. Workshop: Beam Dynamic and Optimization, St.Petersburg, 103, (1995).

[3] G.A.Grinberg, Some Problems of Mathematical Theory of Magnetic Phenomena M., 507, (1948).

[4] V.T.Ovcharov, Dokl. of Sci.Ac. USSR, **107**, 1, 47, (1956). in Russian.

[5] V.T.Ovcharov, Radiotechnika i Electronika, **2**, 6, 696, (1957). in Russian.

[6] V.I.Zubov, Dynamics of Control Systems, M., 1982.

Cellular automaton and elastic net for event reconstruction in the NEMO-2 experiment

V. Kovalenko for NEMO collaboration

Joint Institute for Nuclear Research, Dubna, Moscow reg., 141980 Russia,
e-mail:koval@nu.jinr.dubna.su

Abstract. A cellular automaton for track searching and an elastic net for charged particle trajectory fitting are presented. The advantages of the methods are: simplicity of the algorithms, fast and stable convergence to real tracks, and a reconstruction efficiency close to 100%. Demonstration programs are available at `http://nu-web.jinr.dubna.su/LNP/NEMO` using a Java enabled browser.

1 Introduction

The goal of the NEMO collaboration[1] is to study $\beta\beta$ decays of ^{100}Mo and other nuclei to probe the effective Majorana neutrino mass down to 0.1 eV. A prototype detector NEMO-2 designed for $\beta\beta$ studies has already provided some measurements and is presently running in the Fréjus Underground Laboratory. The detector is a 1 m^3 volume made of tracking chambers composed of drift cells operating in the Geiger mode and two plastic scintillators arrays for energy and time-of-flight measurements. In the following the Geiger drift cells will be called drift tubes to avoid confusion with the cellular automaton cells.

A typical event in this experiment has a few number of tracks usually well separated in space. But this situation is complicated by essential effect of multiple scattering and even hard scattering on wires. Reconstruction of such events in the presence of the left-right ambiguity of drift tubes becomes a task lying out of typical problems of event reconstruction in high energy physics. Therefore cellular automaton and elasic net methods were choosen as flexible ones and suited to work in nonstandard situations[2].

Another reason for using algorithms different from the traditional least squares method in high energy physics has statistical ground. Usually the least squares method is used to fit a particle trajectory to position measurements in such detectors as bubble chambers where errors of measurements are independent and Gaussian distributed. But this is no longer valid for electronic experiments with discrete detectors like multiwire proportional chambers and other devices, which are essentially digital in character. Here errors of measurements are in fact correlated and their distribution is not Gaussian. In such cases it is better to use methods having statistical basis corresponding to an experimental setup[3].

The tracking procedure based on a cellular automaton and an elastic neural net was developed for the NEMO experiment during the last year. It was applied to real data of the NEMO-2 experiment and compared with the previous tracking procedure based on the Kalman filter. This comparison promoted strong confidence in the new tracking methods.

2 Searching for Tracks

Specific features of the experiment make preferable the segment model of cellular automaton[4]

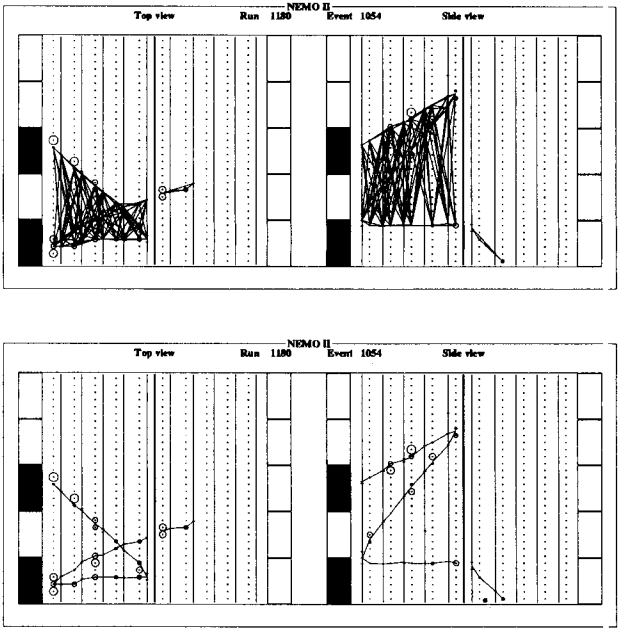


Figure 1: Initial and final configurations of the cellular automaton for a complicated NEMO-2 event. For the drift tubes perpendicular to the figure plane, the drift distance is represented by the radius of a circle. For the drift tubes parallel to the figure plane, the small segments indicate the trajectory position along the drift tube.

where an elementary unit(cell) is the segment connecting two fired wires in neighboring layers.

To construct a cellular automaton for track searching in NEMO-2 data, one follows the logic of cellular automata defining cell, neighbors, rules of evolution and time flow.

Cell. The cellular automaton is three-dimensional. A cell is identified with a straight-line segment connecting two fired wires in neighboring layers, making the cellular automaton essentially local. To take into account inefficiencies of the experimental device one must also include segments connecting wires over one layer. This initial status is depicted in the top half of Fig. 1. During evolution each cell is characterized by its position on track.

Neighbors. In establishing the criterion for assigning segments to a track it is obvious that only segments with a common extremity can be considered as neighbors. Then owing to the coordinate detectors having a discrete structure and to multiple scattering in the

material of the experimental apparatus, the angles between track segments in the real experiment are not zero, but an upper limit ϕ_{max} can be imposed.
Rules. All free cells are initialized with position value 1 and at each step of evolution they look on neighbors and increase the position value if there is a neighbor at the left (suppose) side with the same position value. This is similar to the rule SAFE-PASS[5].
Time. Time evolves discretly. All cells change their states simultaneously.

Knowing positions on track for all segments it is easy to collect all track-candidates. Upon completion of the work of the cellular automaton additional testing of the quality of tracks is carried out. This permits rejecting "phantom" tracks, which could be constructed from hits belonging to different tracks.

In lower half of Fig. 1 the final configuration of the cellular automaton is shown for a complicated NEMO-2 event. Two tracks which originate in the central source foil are shown. One is back-scattered on the front face of a scintillator and leaves the detector after crossing the central foil.

3 Fitting of Tracks

Actual fitting of charged particle trajectories with multiple scattering and hard scattering on wires is made by the elastic net method[6].

Let us

> *Define a track with multiple scattering as the "smoothest" line joining experimental points defined by hit drift tubes.*

Then try to construct the elastic net as a line that is deformed under the influence of **two types of force:**

1. the first pulls it to the edges of drift cell circles for which the radius is directly proportional to the proximity of particle path with respect to the anode wire, (depicted in Fig. 1).
2. the second smooths out the line defining the particle track.

To avoid the problem of local minima because of left-right ambiguity of drift tubes, we propose to start from two tracks which place bounds on the geometrical area where there is the real track. Effectively one of them touches the drift circles on the "upper side" while the other one on the "lower side". Next a **third type of force** is introduced:

3. an attraction between these tracks squeezes the geometrical area to get the real track.

This method allows one to find an optimal trajectory which corresponds to the track.

The elastic net can be simply modified to be able to reconstruct hard scattering tracks. Here one has only to switch off track smoothing at hard scattering points which are found during the preliminary analysis. Fig. 2 presents initial and final configurations of the elastic net for the same NEMO-2 event shown in Fig. 1.

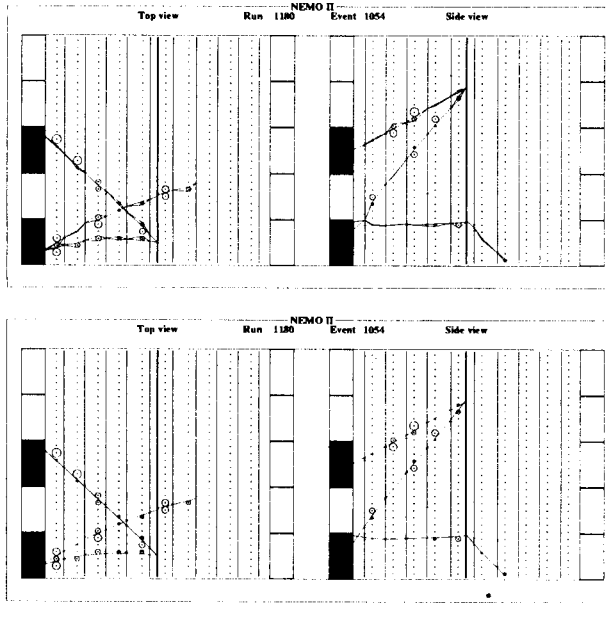


Figure 2: Initial and final configurations of the elastic net for the NEMO-2 event already shown.

4 Results of an Application

The complete data set (6588 h of running time) was previously processed using a tracking procedure based on the Kalman filter [7]. The new tracking method presented here was applied to re-process the data.

An electron is identified by a track linking the source foil and one scintillator. The maximum scattering angle along the track is required to be less than 20° to reject hard scattering. A photon is recognized as one or two adjacent fired scintillators without any associated particle track. For photons and electrons an energy deposit greater than 200 keV is required for obtaining a sufficiently good time resolution. In the analysis of two-electron events a cut on the angle between the two electron tracks is applied ($\cos\alpha < 0.6$). In the case of two-electron events, the difference between the vertices of the two tracks is required to be less than 5 cm. This condition is imposed because of multiple scattering which leads to a rms of 1 cm for a 200 keV electron.

After applying the cuts to extract the $\beta\beta2\nu$ candidates with the new tracking method, 232 two-electron events were selected in the enriched foil, and 61 in the natural one includ-

ing 14.8 calculated $\beta\beta 2\nu$ events. The Monte-Carlo simulation data of the $\beta\beta 2\nu$ decay in the NEMO-2 detector was also analysed using the new tracking program. An overall efficiency of 1.88% for extracting two-electron events is obtained. The number of events entering the half-life calculations is 185.8. A half-life of $T_{1/2} = [3.75 \pm 0.35(stat) \pm 0.21(syst)] \cdot 10^{19} y$ has been derived for both tracking programs. This supports strong confidence in the new tracking methods.

5 Conclusion

In summary the main features of the tracking program based on the cellular automaton and elastic net methods are given below:

- increase in the tracking efficiency of 9% (in the $\beta\beta 2\nu$ analysis);
- increase by a factor 35 in the processing speed;
- working in 3D space to separate close tracks in a projection;
- good reconstruction of tracks with hard scattering on wires;
- reconstruction of short tracks (2 planes involved);
- simple to modify by introducing additional criteria in the analysis.

The results of the application to real NEMO-2 events demonstrate reliable operation of the cellular automaton and elastic net methods in the case of multiple scattering in a gas and hard interactions on wires in the tracking detector.

In the next stage of this investigation a detector with a magnetic field will be taken into account to develop algorithms for the NEMO-3 experiment.

References

[1] R. Arnold et al. (NEMO Collaboration), *Nucl. Instr. and Meth.* **A354**, 338 (1995).

[2] I. Kisel, V. Neskoromnyi and G. Ososkov, *Phys. Part. Nucl.* **24** (6), November–December 1993, p. 657.

[3] N. Chernov, A. Glazov, I. Kisel, E. Konotopskaya, S. Korenchenko and G. Ososkov, *Comp. Phys. Commun.* **74**, 217 (1993).

[4] M.P. Bussa, L. Fava, L. Ferrero, A. Grasso, V. Ivanov, I. Kisel, E. Konotopskaya and G. Pontecorvo, *Nuovo Cimento A* **109**, 327 (1996).

[5] T. Toffoli and N. Margolus, Cellular Automata Machines: A New Environment for Modelling. *MIT Press, Cambridge, Mass.*, 1987.

[6] I. Kisel, E. Konotopskaya and V. Kovalenko, in proceedings of AIHENP-96, Lousanna.

[7] R. Arnold et al. (NEMO Collaboration), *Z. Phys.* **C72**, 239 (1996).

THE NEW ANALYTIC TOOLS OF THE SPACE R^n

Yu. I. Kuznetsov

Home Address: Akademicheskaya str.29, ap.44 Novosibirsk, 630090, Russia Tel.: (007 383 2) 358437 Work address: Akad. Lavrentieva ave.,6 Novosibirsk, 630090, Russia e-mail:kuzin@ommfao.comcen-1.nsk.su

Abstract. The conception of R^n - structure, combiened a strongly nonsingular symmetric matrix A, an indecomposable Hessenberg matrix T and polynomial systems, is given. In the case of Hankel matrix A (then the matrix T is tridiagonal one) the following new results are presented: 1) the Hamilton form of a positive definite (semidefinite) matrix T; 2) the complementary property of a symmetric matrix T; 3) the statistical properties of orthogonal polynomials; 4) the noneuclidean Sturm properties of the matrix T.

1. It is well known that the space R^n has good approximating properties. It has also good interpretative and representative properties. Let the matrix A be symmetric and strongly nonsingular: $|A_k| \neq 0$, $k = 1(1)n$.

Theorem 1 *Let the triangular decomposition*

$$A = R^T \mathcal{E} R, \quad A^{-1} = P \mathcal{E} P^T \tag{1}$$

be valid, where $R = (r_{ij})$, $P = R^{-1} = (p_{i-1}^{j-1})$, $i, j = 1(1)n$, $\mathcal{E} = diag(\epsilon_1, \ldots, \epsilon_n)$, $\epsilon_i = sign(|A_i|/|A_{i-1}|)$. *Then the polynomial*

$$p_j(x) = \frac{1}{\sqrt{\epsilon_k |A_j||A_{j+1}|}} \begin{vmatrix} A_j & \bar{A}_{\bullet j+1} \\ 1 \ldots x^{j-1} & x^j \end{vmatrix}, \quad j = 0(1)n-1. \tag{2}$$

has the representation

$$p_j(x) = \sum_{i=1}^{j} p_i^j x^i,$$

where p_i^j *are elements of the (j+1)-th column of the matrix* P.

The polynomials $p_j(x)$, $j = 0(1)n-1$, form a basis in the polynomials space. Therefore [1]

$$x p_{j-1}(x) = \sum_{i=1}^{j+1} t_{ij} p_{i-1}(x), \quad j = 1(1)n, t_{j+1,j} \neq 0. \tag{3}$$

The matrix $T = (t_{ij})$ is the upper Hessenberg one and

$$V_i P T = x_i V_i P, \tag{4}$$

where $V_i = (1, x_i, \ldots, x_i^{n-1})$ is the i-th row of the Vandermonde matrix V [2-3].

If the nodes x_i, $i = 1(1)n$, are distinct, then

$$R F^T = T R \tag{5}$$

where

$$F = \begin{bmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ p_n & \cdots & & p_2 & p_1 \end{bmatrix}$$

is the Frobenius matrix, i.e.,

$$R_{\bullet k+1} = TR_{\bullet k} = r_{11}T^k e_1, \quad k = 0(1)n-1. \tag{6}$$

¿From this the following assertions are valid:

Theorem 2 *The k-th column of the matrix R is proportional to the first column of the matrix T^{k-1}. The k-th column of the matrix P is the coefficient vector of the polynomial $x\mathrm{p}_{k-1}(x)$. The system of polynomials $x\mathrm{p}_j(x)$, $j = 0(1)n$, and the Hessenberg matrix T are connected by the relations*

$$(\mathrm{p}_1(x_i), \ldots, \mathrm{p}_{n-1}(x_i))T = x_i(\mathrm{p}_1(x_i), \ldots, \mathrm{p}_{n-1}(x_i)), i = 1(1)n.$$

Theorem 3 *The Hessenberg matrix T (except the last column $T_{\bullet n}$) is determined by the nonsingular matrix R. The first $n-1$ columns of T are linearly independent and*

$$T_{\bullet n} = \frac{1}{r_{nn}}(R\mathcal{P} - \sum_{j=1}^{n-1} r_{jn}T_{\bullet j}), \tag{7}$$

where $\mathcal{P} = (p_n \ \ldots, p_1)^T$ is the vector of coefficients of the polynomial $\mathrm{p}_n(x)$. The spectrum of T (at R is fixed) may be preassigned. If

$$A = \tilde{R}^T D\tilde{R},$$

where $\tilde{R} = WR$, $W = diag(w_1, \ldots, w_n)$, $W^2D = \mathcal{E}$, then

$$\tilde{T} = WTW^{-1}. \tag{8}$$

The aggregate, involving the matrix A, the matrix T and the polynomial system $\mathrm{p}_k(x)$, $k = 0(1)n$, are called the R^n-*structure*; the independent parameters determining the R^n-structure are called the degrees of freedom.

Theorem 4 *The degrees of freedom number of R^n-structure are equal to $n(n+3)/2$. For setting R^n-structure it is sufficient to set one from the following: 1) the triangular part af A and $\mathcal{P}$; 2) T and P; 3) $\mathcal{P}$ and P; 4) T; 5) $\mathrm{p}_k(x_i)$, $k = 0(1)n-1$, $i = 1(1)k+1$ and x_i, $i = 1(1)n$.*

Theorem 5 *If A is a Hankel matrix, $A = (a_{i+j-2})$, then T is symmetric in modulus one, i.e., T is tridiagonal matrix and $t_{j+1,j} = t_{j,j+1}\ \varepsilon_j\varepsilon_{j+1} = a_j$, $t_{jj} = b_j$.*

In this case (the Hankel R^n-structure) the polynomials $\mathrm{p}_k(x)$, $k = 0(1)n-1$, are the quasiorthogonal ones:

$$\sum_{i=1}^{n} c_i \mathrm{p}_{l-1}(x_i)\mathrm{p}_{k-1}(x_i) = \varepsilon_k\delta_{kl}, \quad k, l = 1(1)n, \tag{9}$$

The values x_i, c_i may be complex ones.

2. A special case of the Hankel R^n-structure is *the classical structure* with conditions $|A_k| > 0$, $k = 1(1)n$. Its elements: the orthogonal polynomials system, the positive definite Hankel matrix and the symmetric Jacobi matrix are well known.

The classical structure is the basis of the analysis on the discrete set of points: the integration operator is replaced by the Hankel matrix; the differential operator is replaced by the Jacobi matrix and its triangular decomposition factors; the function classes are replaced by the polynomial system. Let us consider the third assertion in more detail [4-5].

The system of polynomials $p_{k-1}(x)$, $k = 1(1)n$, at the fixed value x is a function of the integer variable k (for example, the Chebyshev polynomials). If $x = x_j$, $j = 1(1)n$, $x_j > x_{j+1}$, then the sequence $p_0(x_j), \ldots, p_{n-1}(x_j)$ has $(j-1)$ the changes of sign. Let this sequence (normed in the suitable way) define a polynomial of order $j-1$. If

$$\frac{1}{\gamma_j} q_{j-1}(y_k) = \frac{\kappa_j p_{k-1}(x_j)}{\gamma_1 \kappa_1 p_{k-1}(x_1)}, \quad j,k = 1(1)n, \tag{10}$$

where γ_j, κ_j are real positive numbers, then $q_0(y_k) = 1$, $k = 1(1)n$, at some unknown points y_k. Let $Q = (q_i^j)$ be the upper triangular matrix and

$$q_j(x) = \sum_{i=0}^{j} q_i^j x^i, \quad j = 0(1)n-1.$$

As $q_1(x)$ is a polynomial of order 1, then the equation

$$y_j = \frac{\gamma_2 \kappa_2 p_{j-1}(x_2)}{\gamma_1 \kappa_1 p_{j-1}(x_1)} q_1^1 - q_0^1, \quad j = 1(1)n, \tag{11}$$

follows from the relation (10).

Theorem 6 *Let the eigenvalue of the Jacobi matrix T be decreasingly ordered. Then the values y_j defined from (11) are decreasingly ordered, also.*

It is possible to point out the compatibility criteria of equation (10) with the triangular form of the matrix Q.

The polynomial system $q_j(x)$, $j = 0(1)n$, induces the symmetric Jacobi matrix T^* with the elements a_k^*, $k = 1(1)n-1$, b_k^*, $k = 1(1)n$,

$$\begin{aligned} a_k^* &= \frac{q_{k+1}^{k+1}}{q_k^k}, \quad k = 1(1)n-1, \\ b_{k+1}^* &= \frac{q_{k-1}^k}{q_k^k} - \frac{q_k^{k+1}}{q_{k+1}^{k+1}}, \quad k = 0(1)n-1, \end{aligned}$$

$q_{-1}^0 = 0$. The polynomial system $q_j(x)$, connected with $p_k(x)$ by the relations (10), may be called the *conjugate* Sturm system and the matrix T^* is a *complementary* one. The values y_j are eigenvalue of matrix T^*, i.e., the equation $V^T(y)QT^* = YV^T(y)Q$ is valid ($Y = diag(y_1, \ldots, y_n)$; $V(y)$ is the corresponding Vandermonde matrix).

Theorem 7 *The matrices T and T^* are mutial complementary, i.e., the equation $(T^*)^* = T$ is valid.*

The orthogonal matrices $O = D^{-1/2}P^TVC^{1/2}$, $O^* = D^{-1/2}(y)Q^TV(y)C^{1/2}(y)$ are connected:

Theorem 8 *The relations $O^* = O^T$ and*

$$TO = OX, \quad OT^* = YO, \tag{12}$$

are valid.

The matrix T connects three of neighbouring polynomials at one point. The complementary matrix T^* connects the values of one polynomial at three neighbouring points. In this sense the second equality (12) is an analog of the hypergeometric type equation. There is exists an analog of the Pearson equation for the orthogonality weights.

Theorem 9 *The orthogonality weights of the polynomials $p_k(x)$, $q_k(x)$ satisfy the equations*

$$K(y)\rho = D(y)d(y)\rho, \quad K(x)\rho(y) = D(x)\rho(y),$$

where $K(x)$, $K(y)$ are twodiagonal skew-symmetric matrices and $D(x)$, $D(y)$ are the diagonal matrices, $\rho = (\rho_1, \ldots, \rho_n)^T$, $\rho(y) = (\rho_1^{(y)}, \ldots, \rho_n^{(y)})^T$.

The given analysis shows that the classical structure has the approximating and interpretability properties, which is important both the direct and the inverse problems. The following example confirm this idea.

3. The Hamilton bundle of the quadratic forms is determined by the matrix $S+xD$, where x is real, $D = \mathrm{diag}(m_1, \ldots, m_n)$, $m_i > 0$, and $S = (\alpha_{i-1}, \beta_i, \alpha_i)$, $i = 1(1)n$, is the positive definite (may be semidefinite) symmetric Jacobi matrix with conditions $\alpha_{i-1}+\beta_i+\alpha_i = 0$, $i = 1(1)n$. Let $\chi_i = -1/l_i$, $i = 1(1)n-1$, and $\alpha_0 = -\delta_1/l_0$, $\alpha_n = -\delta_{n+1}/l_n$, where δ_1, δ_{n+1} equal 0 or 1; $l_i > 0$, $i = 0(1)n$. The Hamilton bundle of the quadratic forms is interpreted as an oscillation system with n degrees of freedom. The matrix $S+xD$ is congruent to the matrix $T+xI$, where T is defined by Theorem 3 at $\varepsilon_i = 1$, $i = 1(1)n$, with the elements:

$$a_i = \alpha_i(m_im_{i+1})^{-1/2}, \quad i = 1(1)n-1, \tag{13}$$

$$b_i = -(\alpha_{i-1}+\alpha_i)\frac{1}{m_i}, \quad i = 1(1)n, \tag{14}$$

Let us show [4], that elements of the Hamilton form (i.e., the values σ/l_i, m_i) may be found for any positive definite (may be semidefinite) symmetric Jacobi matrix T with the definition (13).

Let us denote

$$\theta_i = -a_i\sqrt{\frac{m_{i+1}}{m_i}} > 0, \quad i = 1(1)n-1. \tag{15}$$

The comparison of the matrices T and S gives the n equations

$$\begin{aligned} \theta_1 &= b_1 - \epsilon, \\ \theta_i &= b_i - \frac{a_{i-1}^2}{\theta_{i-1}}, \quad i = 2(1)n-1, \\ 0 &= b_n - \mu - \frac{a_{n-1}^2}{\theta_{n-1}}. \end{aligned} \tag{16}$$

¿From the first equation (16) we found $\epsilon = b_1 - \theta_1$. Using another equatio is, we found a continued fraction, from whichthe equality followes

$$\mu = \frac{\epsilon h_n - g_n}{\epsilon h_{n-1} - g_{n-1}},$$

where g_k, $k = 1(1)n$, is the matrix T major successive minor of the order k and h_k, $k = 2(1)n$, is the major successive minor of the order $k-1$ of the matrix, obtained from T by rejecting the first row and the first column.

Theorem 10 *The Hamilton form may be obtained from the Jacobi matrix T according to the definition (13) and the condition $x_1 \geq 0$ if and only if $\epsilon \in [0, \epsilon_1]$, $\epsilon_1 = \frac{g_n}{h_n}$. At this condition $\mu \in [0, \mu_1]$, $\mu_1 = \frac{g_n}{g_{n-1}}$, and*

$$m_{i+1} = m_i \prod_{j=1}^{i} \theta_j^2 / a_j^2, \quad j = 1(1)n-1, \tag{17}$$

$$\begin{aligned} 1/l_0 &= \epsilon m_1, \\ 1/l_i &= -a_i \sqrt{m_i m_{i+1}}, \quad i = 1(1)n-1, \\ 1/l_n &= \mu m_n. \end{aligned} \tag{18}$$

If $\epsilon = 0$ (or $\mu = 0$), then $\delta_1 = 0$ (or $\delta_{n+1} = 0$). In this case the value l_i (or l_n) is no difference.

4. The new view [4] on interconnection the orthogonal polynomial $p_k(x)$ and the polynomials with roots $\{x_{l_1}, \ldots, x_{l_k}\} \subset X = \{x_1, \ldots, x_n\}$ (the node polynomials). If subset $\{i_1, \ldots, i_{n-k}\}$ is complement of the subset $\{l_1, \ldots, l_k\}$ with respect to $N = \{1, \ldots, n\}$, then the polynomial $\varphi_{i_1,\ldots,i_{n-k}}(x)$ is determined by the relations $\varphi_{i_1,\ldots,i_{n-k}}(x) = \prod_{j=1}^{k}(x - x_{l_j})$. The role of node polynomials in the analysis on discret points set we note only by the following relation [4]

$$\Delta[x_1^l, \ldots, x_n^l] = \sum_{i=1}^{n} \frac{x_i^l}{\varphi_i(x_i)} = q_l,$$

where q_l is the divided difference of the order $n-1$ of the function x^l.

Theorem 11 *The relation*

$$p_k(x) = \sum_{i_1=1}^{n} \cdots \sum_{i_{n-k}=1}^{n} \tilde{c}_{i_1,\ldots,i_{n-k}} \varphi_{i_1,\ldots,i_{n-k}}(x), \tag{19}$$

where

$$\sum_{i_1=1}^{n} \cdots \sum_{i_{n-k}=1}^{n} \tilde{c}_{i_1,\ldots,i_{n-k}} = 1, \tag{20}$$

and

$$\tilde{c}_{i_1,\ldots,i_{n-k}} = r_k \prod_{j=1}^{n-k} \left(\frac{1}{c_{i_j} \varphi_{i_j}^2(x_{i_j})} \prod_{m=1}^{j-1} (x_{i_j} - x_{i_m})^2 \right) > 0, \tag{21}$$

is valid.

Let us introc uce the space Ω_{n-k} having the dimension $n-k$, and let the j-th coordinate is characterised by the index $i_j \in \mathrm{N}$. Then the orthogonal polynomial $p_k(x)$ may be interpreted as the mean value of the polynomials $\varphi_{i_1,\ldots,i_{n-k}}(x)$ having the same order, which is interpreted as the $(n-k)$-dimentional random variable from Ω_{n-k} with the distribution (21).

5. The orthogonal polynomials have the Euclidean signature

$$c_i > 0, \quad i = 1(1)n,$$

and the distinct roots: if $\mathrm{p}_k(x_i^{(k)}) = 0$ then $x_{i+1}^{(k)} < x_i^{(k-1)} < x_i^{(k)}$. Let us consider [6] the signature of Minkowski:

$$c_i > 0, \quad i \le l, \quad c_i < 0, \quad i > l, \quad l = 1(1)n-1, \tag{22}$$

at condition $x_i > x_{i+1}$, $i = 1(1)n-1$. It means that the sequence c_i, $i = 1(1)n$, has only one sign's change. Corresponding this signature, the polynomials will call as the pseudoorthogonal ones. The roots of the pseudoorthogonal polynomials have distinct real roots as well.

To show it et us introduce the condition $x_i > 0$, $i = 1(1)n$, and denote for c_i the Eucludean case and for $\tilde{c}_i$ – the Minkowski case. For any sequence $\tilde{c}_i$, $i = 1(1)n$, may be found the sequence c_i, $i = 1(1)n$, and the such b that $\tilde{c}_i = (x_i - b)c_i$, $i = 1(1)n$. On x_i, c_i, $i = 1(1)n$, and c_i, $\tilde{c}_i$, $i = 1(1)n$, as well, we construct the systems of polynomials $\mathrm{p}_k(x)$, $k = 0(1)n-1$, and $\tilde{\mathrm{p}}_k(x)$, $k = 0(1)n-1$, respectivly. Let $\tilde{\mathrm{p}}_k(\tilde{x}_i^{(k)}) = 0$, $\mathrm{p}_k(x_i^{(k)}) = 0$, $i = 1(1)k$, $k = 0(1)n$, and $I_j^{(k)} = (x_{j+1}^{(k)}, x_j^{(k)})$, $j = 0(1)k$, $\tilde{I}_j^{(k)} = (\tilde{x}_{j+1}^{(k)}, \tilde{x}_j^{(k)})$, $j = 0(1)k$, where $(x_{k+1}^{(k)}, x_0^{(l)})$ is some interval, containing all roots of the polynomial $\mathrm{p}_k(x)$ and point b. The interval $(\tilde{x}_{k+1}^{(k)}, \tilde{x}_0^{(k)})$ has the analogous sence. By definition $\tilde{x}_i^{(n)} = x_i^{(n)} = x_i$, $I_j^{(n)} = I_j$. The integer value $m = 0(1)n$ is such that $b \in I_m$.

Theorem 12 *Let $b \in I_m^{(k)}$, $m = 0(1)k$. If $\mathrm{p}_k(b) \neq 0$, $k = 1(1)n$, then the following inequlities take place:*

$$\tilde{x}_{j+1}^{(k)} < \tilde{x}_j^{(k-1)} < \tilde{x}_j^{(k)}, \quad j = 1(1)m-1, \tag{23}$$

$$\tilde{x}_{j+2}^{(k)} < \tilde{x}_j^{(k-1)} < \tilde{x}_{j+1}^{(k)}, \quad j = m(1)k-1, \tag{24}$$

if $\mathrm{p}_{k-1}(b)\mathrm{p}_k(b) < 0$;

or

$$\tilde{x}_j^{(k)} < \tilde{x}_j^{(k-1)} < \tilde{x}_{j-1}^{(k)}, \quad j = 1(1)m, \tag{25}$$

$$\tilde{x}_{j+1}^{(k)} < \tilde{x}_j^{(k-1)} < \tilde{x}_j^{(k)}, \quad j = m+1(1)k-1, \tag{26}$$

if $p_{k-1}(b)p_k(b) > 0$.

Let us remark that it is no analogy to pseudoorthogonal polynomials on the continuum.

References

[1] Ilin V.P., Kuznetsov Yu.I. Algebraic Foundation of Numerical Analysis. - Novosibirsk,: Nauka, Sibirian Branch, 1986. (in Russian)

[2] Kuznetsov Yu.I. The Relation of Positive Definite Matrices and the Hessenberg matrix. // Sib. Math. Journal, Nauka, Sib. Br. - 1986. - N 2.-p. 94-100. (in Russian)

[3] Kuznetsov Yu.I. About the Properties of Hessenberg Matrix. // Sib. Math. Journal, Nauka, Sib. Br. - 1987. - N 6. (in Russian)

[4] Kuznetsov Yu.I. The Elements of Analysis on the Discret Set of Points. - Novosibirsk,: Pub. CC SB RAS, 1994 (in Russian)

[5] Kuznetsov Yu.I. Complementary Jacobi matrix. Proceedings of the International Conference "AMCA - 95", Novosibirsk,Russia, 1995. p. 422-427

[6] Kuznetsov Yu.I. The Pseudoorthogonal Polynomials Systems. // Novos birsk, preprint N 1068 CC SB RAS, 1996 (in Russian)

CODES PACKAGE FOR THE RFQ DYNAMICS & OPTIMIZATION [1]

Anna Loukianova

St. Petersburg State University, Russia
198904, St.-Peterburg. Bibliotechnay pl. 2
e-mail:anna@louk.apmath.spb.su

Abstract. The proposed paper develops the mathematical model and optimization methods for the dynamic systems of the special form. The computer codes has been developed for the RFQ ion linac design and an analysis. The package consists of the three levels codes as well as the linac designing process has three main stages: the channel choose, the exact calculation, the simulation and the optimization. The optimization methods are developed for the presented functionals and mathematical models. The computer program realizing this mathematical simulation of the cotrol make used of the optimization in particular heavy-ion RFQ linacs and also project of compact 8 MeV deutron linac for PET isotops production.

1 Introduction

Mathematical methods of a control and an optimization are extensively used in many fields of the science and engineering. The development of the specialized software for various applications becomes of the increasing importance.

The special class of problems attracting attention of the numerous researcher is the control and optimization [1]. The proposed paper develops the optimization methods in this direction for the dynamic systems of the special form [2]. The problem of the control is the set as a problem of a trajectory ansemble control.

The computer codes has been developed for the RFQ ion linac design and an analysis. The package consists of the three levels codes as well as the linac designing process has three main stages: the channel choose, the exact calculation, the simulation and the optimization. The first level can be used for the accelerating/focusing RFQ channel calculation and help user in choosing the best channel version. The second is the program for the simulation of the charged particles beam dynamics. By using this mathematical model it is possible the problem of the optimization for the RFQ parameters. New classes of functionals are introduced for an estimation of beam dynamics. The optimization methods are developed for the presented functionals and mathematical models [3,4].

2 The model of optimization

Let us assume that beam dynamics of the charged particles is given by the system of equations [2]

$$\frac{dx}{dt} = f(t, x, u), \tag{1}$$

[1] This work was supported in part by the Russian Foundation for Fundamental Researches under Grant No. 95-01-00926

$$\frac{\partial \rho}{\partial t} + \frac{\partial \rho}{\partial x} f(t,x,u) + \rho div f(t,x,u) = 0, \tag{2}$$

$$\frac{dz}{dt} = f_3(t,x,z,u), \tag{3}$$

where

$$f(t,x,u) = f_1(t,x,u) + \int_{M_{t,u}} f_2(t,x,y_t)\rho(t,y_t)dy_t$$

with initial conditions

$$x(0) = x_0 \in \overline{M}_0^n, \rho(0,x) = \rho_0(x),$$

$$z(0,x_0,z_0) = z_0 \in \overline{N}_0^m, u \in K \subset R^n.$$

At this point M_0, N_0 are compact in R^n and $\rho_0(x)$ is continuosly differentiated function. It is conceived that $f_1, f_2, f_3, div f_1, div f_2, div f_3$ are continuous for all their arguments and have a continuous particular derivative for x, y, z. In this case the problem Cauchy (1)-(3) possess a single solution. Control $u = u(t, x_1, \ldots, x_n) \in K$ will be considered to depend on the t (time) and phase coordinates $x_1, \ldots, x_n$ (K is a convex set).

We examine the problem of the minimization for the functional $I(u)$ [1,2]

$$I(u) = \int_0^T \int_{M_{t,u}} \varphi(t,x_t,z(t,x_t),\rho(t,x_t))dx_t dt + \tag{4}$$

$$\int_{M_{T,u}} \psi(x_T, z(T,x_T), \rho(T,x_T))dx_T$$

on the solution of the system (1)-(3) by parameters $u(t,x) \in K$ ($\varphi(t,x(t),z(t,x))$, $\psi(x_T, z(T,x)) \in C^2([0,T] \times M \times N)$).

For such problem the representation for the variation of the functional and the necessary conditions of the optimality in the form of the maximum principle are received.

This mathematical model enables us to solve the several problems of practical impotance namely the problem of the optimization of beam particals in the heavy- ion RFQ linacs.

3 The mathematical model of dynamic of the beam particals in heavy-ion RFQ linacs

We'll assume that beam is an elliptical cylinder with charge distribution depended only from the longitudial variable z. Then radial particle dynamics can be described by equations [2,3]

$$\begin{aligned} \frac{ds_{11}^{x,y}}{dt} &= 2s_{12}^{x,y}, \\ \frac{ds_{12}^{x,y}}{dt} &= Q_{x,y}s_{11}^{x,y} + s_{22}^{x,y}, \\ \frac{ds_{22}^{x,y}}{dt} &= 2Q_{x,y}s_{12}^{x,y}, \end{aligned} \tag{5}$$

where

$$Q_{x,y} = \frac{e}{m_0}((\frac{Uk^2\theta}{2\pi}\sin kz \pm \frac{1}{R_0^2})\cos\omega t + \frac{I}{4\pi\varepsilon v_z\sqrt{s_{11}^x}\sqrt{s_{11}^y}}(1 \pm \frac{\sqrt{s_{11}^x} - \sqrt{s_{11}^y}}{\sqrt{s_{11}^x} + \sqrt{s_{11}^y}})). \quad (6)$$

Here U, θ, R_0, k ω are given parameters of the electric field; e, m_0 are the charge and the mass of the par icle, I is the beam current. Values $\sqrt{s_{11}^x} = r_x, \sqrt{s_{11}^y} = r_y$ are an envelopes of a beam by variables x and y correspondingly. The elements $s_{11}^{x,y}, s_{12}^{x,y}, s_{22}^{x,y}$ completely define the beam pfase condition ellipses in $x\dot{x}$ and $y\dot{y}$ planes. The quantities s_{12}^x, s_{12}^y characterize the angle between main axes of ones ellips and ox, oy axes correspondingly.

The parameters of the system (5)-(6) depend from z thus the system (5)-(6) should be coupled with ec uation

$$\frac{d^2z}{dt^2} = \frac{2eU\theta k}{m_0\pi\omega^2}\cos kz\cos\omega t + F_z, \quad (7)$$

where F_z is the beams particle interaction force. In this work one considered the Coulomb forces action in longitudinal and radial movement separately. Besides that ones assume that the beam is axisymmetric and moves in the cylindrical tube of the radius r. The disk model is used for the beam longitudinal movement modelling.

4 The numerical optimization

By using this mathematical model of the charged particl dynamics (5)-(7) it is possible to solve the prc blem of optimization for the accelerating structure [3, 4]. Lengths of the semiperiods of a modulation, the effetiveness of the acceleration θ and the "median" radius of the structure R_0 must be selected as the cotrol parameters. The functions φ and ψ in the functional (1) will be represented in the form

$$\varphi = a_1\varphi_\tau + a_2\varphi_x + a_3\varphi_y; \quad (8)$$

$$\psi = a_4\psi_\beta + a_5\psi_L. \quad (9)$$

Here $a_i, i = \overline{1,5}$ are weight coefficients,

$$\varphi_\tau = \begin{cases} z_\tau - \beta_\tau\lambda - z_t, & \text{if } z_\tau - \beta_\tau\lambda - z_t > 0 \\ 0 & \text{otherwise} \end{cases}$$

$$\varphi_{x,y} = \begin{cases} s_{11}^{x,y} - r^2, & \text{if } s_{11}^{x,y} > r^2 \\ 0 & \text{otherwise} \end{cases}$$

$$\psi_\beta = \begin{cases} (\beta_T - \beta_L)^2, & \text{if } \beta_T < \beta_L \\ 0 & \text{otherwise} \end{cases}$$

$$\psi_L = \begin{cases} (z_T - z_L)^2, & \text{if } z_T < z_L \\ 0 & \text{otherwise} \end{cases}$$

The functions $\varphi_x, \varphi_y, \varphi_\tau$ describe the capture of particls in the radial and longitudinal dynamics correspondigly (β_τ, z_τ are a relative velocity and a longitudinal coordinate for the singled out partical); functions ψ_β, ψ_L define the receiving of the preassigned energy and the limitation on the length (L) of the structure.

The gradient method of the descent is determined for the problem of this optimization (5)-(7),(8)-(9).

5 The application

The computer program realizing this mathematical simulation of the cotrol was created and make used of the optimization in particular heavy-ion RFQ linacs [3] and also, for example [4], project of compact 8 MeV deutron linac for PET isotops production is proposed. It includes surface plasma source of D^- ions, preaccelerating system of ion beam up to 60 keV, 2.5 meter 433 MHz RFQ cavity and 1.5 m 433 MHz APF (alternate phase focusing) H-type resonator with drift tubes. Output deutron enegy of RFQ is 2 MeV. APF resonator accelerates deutrons from 2 up to 8 MeV. Original software environment has been developed for the computing of an electrodes geometry of the RFQ and gaps geometry of the APF structure with using of mathematical methods of the beam dynamics optimization. RF system is based on endotron type devices. Total length of accelerator is near 6 meters.

References

[1] D. A. Ovsyannikov, Modeling and Optimization of Charged Particle Beam Dynamics. L.: Leningrad University, 1990 (in Russian).

[2] A. Lukjanova, International Symposium on computer arithmetic,scientific computation and mathematical modelling, Bulgaria, September 23-28.1990.

[3] Anna Loukianova, European Partical Accelerater Conference (EPAC), London, June 27-30, 1994.

[4] O. I. Drivotin, A. E. Loukianova, D.A. Ovsyannikov, Yu. N. Gavrish, Yu. A. Svistunov, M. F. Vorogushin, European Partical Accelerater Conference (EPAC), Barselona, 1996.

Deformable Templates for Circle Recognition

L. Mureşan*, R. Mureşan, G. Ososkov, Yu. Panebratsev**

JINR, Dubna, Moscow region, 141980 Russia
e-mail:lucian@sunhe.jinr.dubna.su, raluca@sunhe.jinr.dubna.su,
osg@lcta39.jinr.dubna.su, panebrat@sunhe.jinr.dubna.su

Abstract. An algorithm for circles recognition using deformable templates was carried out and its performances were studied. The displacement of the points from circles and the presence of noise that appear in real situations were taken into account. The deformable templates are initialized by Hough transforms. This technique was explored in a number of different situations ranging from sets of simulated circles to real data obtained in lead-gold interactions from the RICH detectors used in CERES / NA45 experiment. Results show the robustness of our algorithm to background contaminations. It can be used in many data handling problems that appear in high energy physics like Cherenkov rings recognition and tracks reconstruction.

1 Introduction

An important problem in the area of pattern recognition is curve recognition. In the context of high energy physics, Cherenkov rings finding represents such a problem. This is getting more important with inventing such modern detectors as for example RICH (Ring Imaging CHerenkov) requiring in each event the determination of parameters for tens or hundreds rings formed by Cherenkov photons. The presence of the background noise and the appearance of several overlapping rings, along with the displacement[1] of the points from the circle make inapplicable classical methods for circles fitting due to their noise sensitivity. In this work a new deformable templates algorithm for circle reconstruction is presented and its performances are evaluated. These deformable templates are initialized by Hough transform. The algorithm effectively represents a merger between neuronic decision and parameter fitting that are made simultaneously by solving the equation of the motion for the analogous dynamic system.

2 Hough Transform

The idea of the elastic arms algorithm is to match the observed events into a known parameterized model. To ensure a fast convergence towards a high quality solutions avoiding the local minima, the elastic arms algorithm must be initialized with approximate values for the positions of the centers and the radii of the circles. Hough transform is appropriate for this task.

An event is defined as a set $\vec{x} = \vec{x}_1, \vec{x}_2, ..., \vec{x}_{N_p}$ of N_p signal points. Each event also corresponds to a set of circles (Cherenkov rings) and the aim of the algorithm is to find

[1] For real data the displacement is defined as the error in determining the coordinates of the signal points

*Home Institute: Institute of Physical Chemistry "I.G. Murgulescu", Spl. Independentei No. 202, RO-77208, Bucharest, Romania

**Home Institute: Institute of Atomic Physics (IFIN) RO-76 900 P.O. Box MG 6, Bucharest, Romania

these circles knowing $\vec{x}$. If we denote the coordinates of three arbitrary non linear points by (x_i, y_i) $i = 1, 2, 3$ the circle which passes through these points has for center the (x_0, y_0) coordinates [3]:

$$x_0 = \frac{1}{2}\frac{(x_2^2 - x_3^2 + y_2^2 - y_3^2)(y_1 - y_2) - (x_1^2 - x_2^2 + y_1^2 - y_2^2)(y_2 - y_3)}{(x_2 - x_3)(y_1 - y_2) - (x_1 - x_2)(y_2 - y_3)} \quad (1)$$

$$y_0 = \frac{1}{2}\frac{(x_1^2 - x_2^2 + y_1^2 - y_2^2)(x_2 - x_3) - (x_2^2 - x_3^2 + y_2^2 - y_3^2)(x_1 - x_2)}{(x_2 - x_3)(y_1 - y_2) - (x_1 - x_2)(y_2 - y_3)} \quad (2)$$

and for the radius:

$$R = \sqrt{(x_i - x_0)^2 + (y_i - y_0)^2} \qquad i = 1, 2, 3 \quad (3)$$

We use a Hough transform which consists in two steps. As a first step, using the formulae (1), (2), we calculate the centers of the circles described by each possible set of three points. In the space of the centers, we made a two dimensional histogram (x_c, y_c, N) where x_c and y_c represent the coordinates of the centers, and N the number of possible combinations in each box. In the second step, for each found center we define a ring centered in the (x_c, y_c) and limited by R_{min}, R_{max}. For each data point situated inside this ring we use (eq. 3) to calculate the polar radius related to this center. Then, for each center we made a histogram of these radii.

In order to have better solutions (even in the case of splitting peaks or if the displacement of the points from the circle is large) we give as good value for centers and radii, not the center of the square (or the segment) defined by granularity, but a weighted average over the neighboring cells

3 Deformable Templates (Elastic Arms) Algorithm

Each event corresponds to a set of rings (circles). In order to find these rings (circles), a set $\vec{\pi} = \{\vec{\pi}_1, .., \vec{\pi}_M\}$ of M deformable templates, or elastic arms, is introduced, where an arm a is completely described by its P parameters: $\vec{\pi_a} = (\pi_a^1, .., \pi_a^P)$.

We reduce the circle finding problem to finding minimum of the following energy function:

$$E(\{S_{ia}\}; \vec{\pi}) = \sum_{i=1}^{N}\sum_{a=1}^{M} S_{ia} M_{ia}(\vec{x}, \vec{\pi}) + \lambda \sum_{i=1}^{N}\left(\sum_{a=1}^{M} S_{ia} - 1\right)^2, \quad (4)$$

where S_{ia} is a binary decision unit (neuron) $S_{ia} = 1$ if i is assigned to arm a and $S_{ia} = 0$ otherwise, $M_{ia}(\vec{x}, \vec{\pi})$ is the Euclidean distance between signal point i and arm a. The parameter λ imposes a penalty if $\sum_a S_{ia} = 0$, that means, signal i is not assigned to any arm. In this way the parameter λ governs the amount of noise the algorithm allows for.

In order to avoid local minima when minimizing E, one often introduces thermal noise into neural system. A commonly used procedure for doing this is simulated annealing where the neural system is allowed to be thermalized for a sequence of temperatures $T_n > T_{n-1} > .. > T_0$ according to the Boltzmann distribution [1] [2], :

We proceed by calculating the marginal probability distribution by summing out the neuronic degrees of freedom, S_{ia}. We obtain:

$$P_M(\vec{\pi}) = \frac{1}{Z} e^{-\beta E_{eff}(\vec{\pi})}, \tag{5}$$

where the effective energy is defined as:

$$E_{eff}(\vec{\pi}) = -\frac{1}{\beta} \sum_i \log \left(e^{-\beta \lambda} + \sum_a e^{-\beta M_{ia}} \right) \tag{6}$$

The most probable configuration according to (eq. 5) corresponds to the minimum of E_{eff}.

In our case the templates are circles with variable center and radius. The template is defined in this case by the coordinates of center and by the radius (x_a, y_a, R_a).

Using a gradient descendent method to minimize E_{eff} one gets the updating rules for the elastic circles:

$$\begin{array}{ll} \triangle x_a = -\eta_a \frac{\partial E_{eff}}{\partial x_a} = -\eta_a \sum_i \widehat{V}_{ia} \frac{\partial M_{ia}}{\partial x_a}, & \\ \triangle y_a = -\eta_a \frac{\partial E_{eff}}{\partial y_a} = -\eta_a \sum_i \widehat{V}_{ia} \frac{\partial M_{ia}}{\partial y_a}, & a = \{1, ..M\} \\ \triangle R_a = -\eta_a^r \frac{\partial E_{eff}}{\partial R_a} = -\eta_a^r \sum_i \widehat{V}_{ia} \frac{\partial M_{ia}}{\partial R_a}, & \end{array} \tag{7}$$

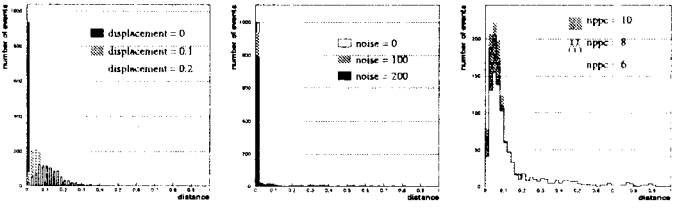
where η_a is the updating parameter for the coordinates and η_a^r for the radius. $\widehat{V}_{ia}$ is the Potts factor:

$$\widehat{V}_{ia} = \frac{e^{-\beta M_{ia}}}{e^{-\beta \lambda} + \sum_{b=1}^{M} e^{-\beta M_{ib}}} \tag{8}$$

The relative importance of the different signals is measured by the Potts factor (eq.8). Also the Potts factor implicitly contains a repulsive force between each two arms. The repulsive force, generated by the V_{ia} term, is useful when the main problem is to separate two overlapping circle (Fig.2) but it is also an important source of local minima. Many local minima appear when the sum $\sum_{b=1}^{M} e^{-\beta M_{ib}}$ is not $e^{-\beta M_{ia}}$. To avoid this we use a new simplified form of Potts factor without interaction between circles:

$$\widehat{V}_{ia} = \frac{e^{-\beta M_{ia}}}{e^{-\beta \lambda} + e^{-\beta M_{ia}}} \tag{9}$$

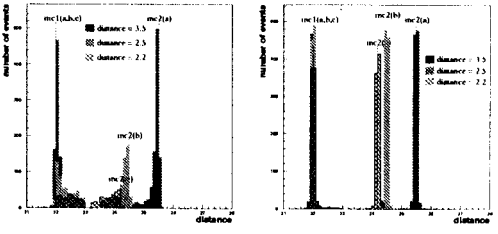
In (eq.7), the partial derivatives $\partial E_{eff}/\partial x_a$, $\partial E_{eff}/\partial y_a$, $\partial E_{eff}/\partial R_a$ haveinfluence is not important in this situation. Even for 200 % noise/signal ratio in more than 80% of cases the distance is less than 0.01. Then we can observe that the algorithm is very robust to noise. We studied the influence of number of points per circle in the range $[6, 10]$ at 0.1 displacement and 100 % noise/signal ratio. The precision of results decreases with the number of point per circle because the amount of information about the circle decreases. The histograms in this case depend strongly on displacement. In Fig. 1c we use 0.1 displacement and different number of points per circle. If the displacement increases the required NPP must be increased to maintain the same amount of information about the circle.



a b c

Fig. 1 Histogram of the distance between the center of Monte Carlo generated circle and the obtained circlefor 0, 0.1 and 0.2 displacement , 100% noise, 10 points per circle (a); for 0%, 100% and 200% noise 0.1 displacement, 10 points per circle (b) and for 6, 8 and 10 points per circle. (0.1 displacement, 100% noise)(c).

Deformable template approach performs a simultaneous fit of circles. When two circles are present, from one circle point of view, the points corresponding to other circle are noise points. If two circles are very close and have the same radius the relative disturbing effect increases. The histograms of x coordinates of centers for two crossing circles, without interaction (given by Potts factor (eq.9), are shown in Fig. 2a corresponding to three various distances between centers. The most difficult case when circles have the same radius was chosen. Each elastic arm is attracted by both sets of points. The attraction force increases when the distance decreases. Using Potts factor (eq.8) the circles will interact and as we demonstrated in subsection 4.3 a repulsive force will change the behavior of two crossing circles. Fig.5b shows a histogram represented in the same conditions as in Fig.2a, but we use (eq.8) instead of (eq.9) to compute the Potts factor. As one can see the repulsive force will generate the decrease of the attractive force of points belonging to other close circle.



a b

Fig. 2 Histograms of x coordinates of centers for two crossing circles at different distances without interaction (Potts factor eq.9)(a) and with interaction (Potts factor eq.8)(b).

This technique was applied to real data obtained in lead-gold interactions from RICH detectors used in CERES/NA45 experiment. In our example we identify pions with $p_t > 4.5\ GeV$ (asymptotic radius = 16.2 pads) in RICH 2. In Fig. 3, all results of the ring

reconstruction that can be given by pions (we use $R_{min} = 5$, $R_{max} = 17$, displacement= 0.2) and numbe: of points per circle$\geq$ 6 are shown.

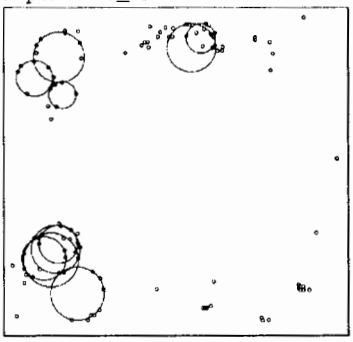


Fig. 3 Output of the algorithm (RICH 2 detectors CERES NA45 experiment).

Looking at the results we may say that the deformable templates algorithm is useful in identificatior of rings given a very good simultaneous fitting with a satisfactory efficiency. During data analysis we observed that the choice of parameters (required NNP, displacement, r nge of radii) according with the experimental setup is very important for the algorithm ability to recognize the Cherenkov rings.

4 Conclusions

We propose a circle finding method that combines the matching and the fitting problem into a single algorithm. It goes from coarse to fine resolution by using a Hough transform to initialize a set of elastic arms. A simultaneous fitting of these input arms is performed by deformable templates method finding the circles with high precision. Our use of Potts factor as a repulsive force is similar to using robust weights [4]. Therefore the algorithm is closely related to robust statistics, it ignores noise data to a required level. The elastic arms approach has the advantage of being flexible to host a variety of experimental setups. Concluding, we can say that our deformable templates algorithm is a considerable contributions to pion physics since it can be applied in cases of arbitrary ring radii where known convent onal algorithms are helpless.

References

[1] Mattias Ohlsson, Comput. Phys. Commun. **77**(1993)19.

[2] M. Ohlsson, C. Peterson, A. Yuille, Comput. Phys. Commun. **71** (1992)77.

[3] T. Altherr, J. Seixas, CERN-TH-**6133/91** July 1991

[4] N. Chernov, E. Kolganova, G. Ososkov, Communication of the Joint Institute for Nuclear Research, Dubna, **E10-95-468** 1995

TRANSFORMATION THEORY OF MULTIPLE HYPERGEOMETRIC SERIES AND COMPUTER AIDED SYMBOLIC CALCULATIONS

A.W.Niukkanen

Vernadsky Institute of Geochemistry and Analytical Chemistry, RAS,
Moscow 117 975, Kossygin street, 19. Tel. 939-70-28
e-mail: elkor@geokhi.msk.su

Abstract. Comments on a new operator factorisation theory of multiple hypergeometric series are given. Some examples relating to transformation theory of the series are presented and possible aplications to development of advanced computer algebra systems are also indicated.

Mathematical models which play the role of theoretical substitutes for underlying natural or technical phenomena including those under supervision of computer-aided control systems give us an information about the phenomena in terms of functions having specialized structures. The vast majority, up to 95%, of the **special functions** involved in description of fundamentally important processes are connected with **hypergeometric series** which thereby give us a key to a substantial part of **applied mathematics really needed by numerous users** . It suggests that along with elementary functions higher transcendental functions of hypergeometric type will inevitably turn into an **obligatory component of modern software** especially for the new computer generation with prevailing role of "intellectual" interactive symbolic analysis over purely numerical computations.

The new **non-elementary level of problems** invoked by the immensely grown number of functions of several variables with analytically comlicated coefficients cannot be coped with the help of traditional mathematical methods. First, an approach based on **first principles** rather than a collection of artificial calculational expedients is needed to provide reliable foundation for a new comprehensive theory. Second, recourse to the methods of computer algebra allowing **automatical performance of analytical transformations** is highly desirable to provide greater calculational productivity and prevent misprints and errors.

The only approach meeting the above requirements is the **operator factorization method** based on the use of a **new binary analytical operation** $< u * v|x >= u(d/ds)v(xs)|_{s=0}$ [1-5] which allows any hypergeometric series (in any number of variables) to be expressed through simpler, eventually, simplest series of the same hypergeometric type (closure property). Utilizing known properties of the simple series any property of the initial complicated series can be established with the help of convenient operator technique. In practice this very simple idea works with **striking effectiveness** .

As an illustration we present here a set of formulas relevant to the transformation theory of hypergeometric series. It gives us an indication at what the core of universal computer algebra system capable of working with multiple hypergeometric series may look like.

First we introduce, for multiple hypergeometric series, the following three canonical forms belonging to Bessel-type (B), Kummer-type (K) and Gauss-type (G), respectively

$$\mathrm{B} \equiv^{N+1} F[L^*//\langle \nu_0 \mid 1, \mathbf{m}_0\rangle; x_0, \mathbf{x}] \tag{1}$$

$$\mathrm{K} \equiv^{N+1} F[\langle \nu_1 \mid 1, \mathbf{m}_1\rangle, L^*//\langle \nu_0 \mid 1, \mathbf{m}_0\rangle; x_0, \mathbf{x}] \tag{2}$$

$$\mathrm{G} \equiv^{N+1} F[\langle \nu_1 \mid 1, \mathbf{m}_1\rangle, \langle \nu_2 \mid 1, \mathbf{m}_2\rangle, L^*//\langle \nu_0 \mid 1, \mathbf{m}_0\rangle; x_0, \mathbf{x}] \tag{3}$$

where $\mathbf{x} \equiv [x_1, \ldots, x_N], \mathbf{m}_s \equiv \left[m_s^1, \ldots, m_s^N\right]$; The complex set of parameters $\langle \nu_s \mid 1, \mathbf{m}_s\rangle$ corresponds to the coefficient $(\nu_s, i_0 + m_s^1 i_1 \cdots + m_s^N i_N)$ where (ν, n) – Pochhammer symbol; $i_1, \ldots, i_N$ – summation indicies; L^* symbolises coefficients $L(i_1, \ldots, i_N)$ **independent** of i_0. Note that any series having type B, K or G is necessarily contained in one of the above formulas. The simplest linear transformation connecting B and K series

$$F\left[\begin{array}{l} \langle \nu_1 \mid 1, \mathbf{m}_1\rangle, L^* \quad ; x_0, \mathbf{x} \\ \langle \nu_0 \mid 1, \mathbf{m}_0\rangle \end{array}\right] = L^0 F =$$
$$e^{x_0} F\left[\begin{array}{l} \langle \nu_{0\bar{1}} \mid 1, \mathbf{m}_{0\bar{1}}\rangle, \langle \nu_1 \mid 0, \mathbf{m}_1\rangle, L^* \quad ; -x_0, \mathbf{x} \\ \langle \nu_0 \mid 1, \mathbf{m}_0\rangle, \langle \nu_{0\bar{1}} \mid 0, \mathbf{m}_{0\bar{1}}\rangle \end{array}\right] \tag{4}$$

is the most general analog of the well known first Kummer-type transformation of $F_1^2(x)$.

The complete set of linear transformations of Gauss-type form is given by:

$G = L_1 G = (1 - x_0)^{-\nu_2} \times$

$$F\left[\begin{array}{l} \langle \nu_{0\bar{1}} \mid 1, \mathbf{m}_{0\bar{1}}\rangle, \langle \nu_2 \mid 1, \mathbf{m}_2\rangle, \langle \nu_1 \mid 0, \mathbf{m}_1\rangle, L^* \quad ; \frac{x_0}{x_0 - 1}, \mathbf{x}(1 - x_0)^{-\mathbf{m}_2} \\ \langle \nu_0 \mid 1, \mathbf{m}_0\rangle, \langle \nu_{0\bar{1}} \mid 0, \mathbf{m}_{0\bar{1}}\rangle \end{array}\right] \tag{5}$$

$G = L_2 G = (1 - x_0)^{-\nu_1} \times$

$$F\left[\begin{array}{l} \langle \nu_1 \mid 1, \mathbf{m}_1\rangle, \langle \nu_{0\bar{2}} \mid 1, \mathbf{m}_{0\bar{2}}\rangle, \langle \nu_2 \mid 0, \mathbf{m}_2\rangle, L^* \quad ; \frac{x_0}{x_0 - 1}, \mathbf{x}(1 - x_0)^{-\mathbf{m}_1} \\ \langle \nu_0 \mid 1, \mathbf{m}_0\rangle, \langle \nu_{0\bar{2}} \mid 0, \mathbf{m}_{0\bar{2}}\rangle \end{array}\right] \tag{6}$$

$G = L_0 G = (1 - x_0)^{-\nu_{0\bar{1}\bar{2}}} \times$

$$F\left[\begin{array}{l} \langle \nu_{0\bar{1}} \mid 1, \mathbf{m}_{0\bar{1}}\rangle, \langle \nu_{0\bar{2}} \mid 1, \mathbf{m}_{0\bar{2}}\rangle, \langle \nu_1 \mid 0, \mathbf{m}_1\rangle, \langle \nu_2 \mid 0, \mathbf{m}_2\rangle, L^* \quad ; x_0, \frac{\mathbf{x}}{(1 - x_0)^{\mathbf{m}_{12\bar{0}}}} \\ \langle \nu_0 \mid 1, \mathbf{m}_0\rangle, \langle \nu_{0\bar{1}} \mid 0, \mathbf{m}_{0\bar{1}}\rangle, \langle \nu_{0\bar{2}} \mid 0, \mathbf{m}_{0\bar{2}}\rangle \end{array}\right] \tag{7}$$

which are the most general analogs of the well known Euler transformations of the $F_1^2(x)$.

The simplest quadratic transformation is given by the formula

$$F\left[\begin{array}{l} L^* \quad ; x_0^2, \mathbf{x} \\ \langle \nu \mid 1, \mathbf{m}\rangle \end{array}\right] = e^{-2x_0} F\left[\begin{array}{l} \langle \nu - 1/2 \mid 1, \mathbf{m}\rangle, L^* \quad ; 4x_0, 4^{\mathbf{m}}\mathbf{x} \\ \langle 2\nu - 1 \mid 1, 2\mathbf{m}\rangle \end{array}\right] \tag{8}$$

which connects Bessel-type form with the Kummer-type series.

To formulate quadratic transformations of the canonical form G we first introduce the following three basic series

$$F_1 \equiv F_1\left[\langle \nu_1 \mid 1, \mathbf{m}_1\rangle, \langle \nu_1 + 1/2 \mid 1, \mathbf{m}_1\rangle, L^* // \langle \nu_0 \mid 1, \mathbf{m}_0\rangle; x_0, \mathbf{x}\right| \tag{9}$$

$$F_2 \equiv F_2\left[\langle \nu_1 \mid 1, \mathbf{m}_1\rangle, \langle \nu_2 \mid 1, \mathbf{m}_2\rangle, L^* // \langle 1 + \nu_{1\bar{2}} \mid 1, \mathbf{m}_{1\bar{2}}\rangle; x_0, \mathbf{x}\right. \tag{10}$$

$$F_3 \equiv F_3\left[\langle \nu_1 \mid 1, \mathbf{m}_1\rangle, \langle \nu_2 \mid 1, \mathbf{m}_2\rangle, L^* // \langle 2\nu_2 \mid 1, 2\mathbf{m}_2\rangle; x_0, \mathbf{x}\right] \tag{11}$$

Three basic quadratic transformations relating to the series have the form:

$F_1 = Q_1^3 F_1 = (1 + \sqrt{x_0})^{-2\nu_1} \times$

$$F_3\left[\begin{array}{l} \langle 2\nu_1 \mid 1, 2\mathbf{m}_1\rangle, \langle \nu_0 - 1/2 \mid 1, \mathbf{m}_0\rangle, L^* \\ \langle 2\nu_0 - 1 \mid 1, 2\mathbf{m}_0\rangle \end{array} ; \frac{2\sqrt{x_0}}{1+\sqrt{x_0}}, \frac{4^{\mathbf{m}_0\bar{1}}\mathbf{x}}{(1+\sqrt{x_0})^{2\mathbf{m}_1}}\right] \tag{12}$$

$F_3 = Q_3^1 F_3 = (2/(2 - x_0))^{\nu_1} \times$

$$F_1\left[\begin{array}{l} \langle \frac{\nu_1}{2} \mid 1, \frac{\mathbf{m}_1}{2}\rangle, \langle \frac{\nu_1+1}{2} \mid 1, \frac{\mathbf{m}_1}{2}\rangle, L^* \\ \langle \nu_2 + 1/2 \mid 1, \mathbf{m}_2\rangle \end{array} ; \frac{x_0^2}{(2-x_0)^2}, \frac{4^{\mathbf{m}_1\bar{2}}\mathbf{x}}{(2-x_0)^{\mathbf{m}_1}}\right] \tag{13}$$

$F_1 = Q_1^2 F_1 = (2/(1 + \sqrt{1 - x_0}))^{2\nu_1} \times$

$$F_2\left[\begin{array}{l} \langle 2\nu_1 \mid 1, 2\mathbf{m}_1\rangle, \langle 2\nu_1 - \nu_0 + 1 \mid 1, 2\mathbf{m}_1 - \mathbf{m}_0\rangle, L^* \\ \langle \nu_0 \mid 1, \mathbf{m}_0\rangle, \langle 2\nu_1 - \nu_0 + 1 \mid 0, 2\mathbf{m}_1 - \mathbf{m}_0\rangle \end{array} ; \frac{1-\sqrt{1-x_0}}{1+\sqrt{1-x_0}}, \frac{\mathbf{x}}{(1+\sqrt{1-x_0})^{2\mathbf{m}_1}}\right] \tag{14}$$

$F_2 = Q_2^1 F_2 = (1 + x_0)^{-\nu_1} \times$

$$F_1\left[\begin{array}{l} \langle \frac{\nu_1}{2} \mid 1, \frac{\mathbf{m}_1}{2}\rangle, \langle \frac{\nu_1+1}{2} \mid 1, \frac{\mathbf{m}_1}{2}\rangle, \langle \nu_2 \mid 0, \mathbf{m}_2\rangle, L^* \\ \langle 1 + \nu_1 - \nu_2 \mid 1, \mathbf{m}_1 - \mathbf{m}_2\rangle \end{array} ; \frac{4x_0}{1+x_0^2}, \frac{2^{\mathbf{m}_1}\mathbf{x}}{(1+x_0)^{\mathbf{m}_1}}\right] \tag{15}$$

$F_3 = Q_3^2 F_3 = (2/(1 + \sqrt{1 - x_0}))^{2\nu_1} \times$

$$F_2\left[\begin{array}{l} \langle \nu_1 \mid 1, \mathbf{m}_1\rangle, \langle \nu_{1\bar{2}} + 1/2 \mid 1, \mathbf{m}_{1\bar{2}}\rangle, L^* \\ \langle \nu_2 + 1/2 \mid 1, \mathbf{m}_2\rangle, \langle \nu_{1\bar{2}} + 1/2 \mid 0, \mathbf{m}_{1\bar{2}}\rangle \end{array} ; \frac{1-\sqrt{1-x_0}}{1+\sqrt{1-x_0}}, \frac{4^{\mathbf{m}_1\bar{2}}\mathbf{x}}{(1+\sqrt{1-x_0})^{2\mathbf{m}_1}}\right] \tag{16}$$

$F_2 = Q_2^3 F_2 = (1 + \sqrt{x_0})^{-2\nu_1} \times$

$$F_3\left[\begin{array}{l} \langle \nu_1 \mid 1, \mathbf{m}_1\rangle, \langle \nu_{1\bar{2}} + 1/2 \mid 1, \mathbf{m}_{1\bar{2}}\rangle, \langle \nu_2 \mid 0, \mathbf{m}_2\rangle, L^* \\ \langle 1 + 2\nu_{1\bar{2}} \mid 1, 2\mathbf{m}_{1\bar{2}}\rangle \end{array} ; \frac{4\sqrt{x_0}}{(1+\sqrt{x_0})^2}, \frac{4^{\mathbf{m}_1\bar{2}}\mathbf{x}}{(1+\sqrt{x_0})^{2\mathbf{m}_1}}\right] \tag{17}$$

Using linear transformations many additional induced quadratic formulas can be readily obtained. First we untroduce 6 associated series

$$F_{1,\omega} \equiv F_{1,\omega}\left[\langle \nu_1 \mid 1, \mathbf{m}_1\rangle, \langle \nu_2 \mid 1, \mathbf{m}_2\rangle, L^* // \langle \nu_{12} - \omega + 1/2 \mid 1, \mathbf{m}_{12}\rangle; x_0, \mathbf{x}\right] \tag{18}$$

$$F_{2,\omega} \equiv F_{2,\omega}\left[\langle \nu_1 \mid 1, \mathbf{m}_1\rangle, \langle 1 - \nu_1 + 2\omega(\nu_0 - 1) \mid 1, -\mathbf{m}_1 + 2\omega\mathbf{m}_0\rangle, L^* // \langle \nu_0 \mid 1, \mathbf{m}_0\rangle; x_0, \mathbf{x}\right] \tag{19}$$

$$F_{0,\omega} \equiv F_{0,\omega}\left[\langle \nu_1 \mid 1, \mathbf{m}_1\rangle, \langle \nu_2 \mid 1, \mathbf{m}_2\rangle, L^* // \langle \omega + 1/2 \mid 1, 0\rangle; x_0, \mathbf{x}\right] \tag{20}$$

where $\omega = 0, 1$. We then introduce notations

$$L(0,0) \equiv I, L(1,0) \equiv L_1, L(0,1) \equiv L_2, L(1,1) \equiv L_0, L_0^0 \equiv I, L_0^1 \equiv L_0, \tag{21}$$

Finally, the complete list of quadratic transformations takes on the following compact form:

$$F_1 = L(\omega_1, \omega_2) Q_1^3 F_1 \sim F_3, \quad F_3 = L_0^\sigma Q_3^1 F_3 \sim F_1 \tag{22}$$

$$F_1 = L_0^\sigma Q_1^2 F_1 \sim F_3, \quad F_2 = L_0^\sigma Q_2^1 F_2 \sim F_1 \tag{23}$$

$$F_3 = L_0^\sigma Q_3^2 F_3 \sim F_2, \quad F_2 = L(\omega_1, \omega_2) Q_2^3 F_2 \sim F_3 \tag{24}$$

$$F_{1,\omega} = L(\omega_1, \omega_2) Q_1^3 L_2 F_{1,\omega} \sim F_3, \quad F_3 = L_{2-\omega} Q_3^1 F_3 \sim F_{1,\omega} \tag{25}$$

$$F_1 = L_{1+\omega} Q_1^2 F_1 \sim F_{2,\omega}, \quad F_{2,\omega} = L_0^\sigma Q_2^1 L_2 F_{2,\omega} \sim F_1 \tag{26}$$

$$F_{1,\omega} = L_0^\sigma Q_1^2 L_2 F_{1,\omega} \sim F_2, \quad F_2 = L_{2-\omega} Q_2^1 F_2 \sim F_{1,\omega} \tag{27}$$

$$F_{1,\omega} = L_{1+\omega'} Q_1^2 L_2 F_{1,\omega} \sim F_{2,\omega'}, \quad F_{2,\omega'} = L_{2-\omega} Q_2^1 L_2 F_{2,\omega'} \sim F_{1,\omega'} \tag{28}$$

$$F_3 = L_{1+\omega} Q_3^2 F_3 \sim F_{2,\omega}, \quad F_{2,\omega} = L_{(\omega_1, \omega_2)} Q_2^3 L_2 F_{2,\omega} \sim F_2 \tag{29}$$

In particular case of one variable ($N = 0$) the above formulas give us a systematic and comprehensive alternative to the chaotic and incomplete Goursat list of quadratic transformations of Gauss function.

Many other properties of the multiples series including analytical continuation formulas, functional expansions, recurrence relations, linearization theorems, addition and multiplication formulas, generating functions, etc. were also investigated in detail.

We note that the whole set of relations including those presented in the report not only gives us **a complete theory of linear and quadratic transformations and analitical continuation formulas of whatever complicated series** but allows us to immediately use this relations to develop a computer program capable of handling with those hundreds and thousands of series which were mentioned by Erdely as a dramatic challenge for the theory of hypergeometric series.

Execution of the project may give us a **useful and sophisticated superstructure over existent computer algebra systems** not only in the field of transformation theory but in a number of other important classes of formulas including those popular amongst physicists and theoretical chemists. The total outcome of the project is expected to exceed in power all known means of analyzing special functions. Moreover the projected **interactive knowledge-based Special Functions Adviser** is expected to acquire essentially new faculty of synthesizing necessary relations at the moment they are needed by the user.

Co-operation efforts would be desirable to speed up realization of the rather ambitious plan outlined above. Inquiries about technical details of the project are invited to be sent to the author by the above working address or residence address: Russia 300 062, Tula, Gorky st., 11, 304 .

The author is highly indebted to Russian Foundation of Fundamental Research for Support of the author's hypergeometric project (grant 97-01-00317).

References

[1] A. W. Niukkanen, J. Phys. A: Math. Gen. **16**, 1813 (1983).

[2] A. W. Niukkanen, J. Phys. A: Math. Gen. **17**, L731 (1984).

[3] A. W. Niukkanen, J. Phys. A: Math. Gen. **18**, 1399 (1985).

[4] A. W. Niukkanen, Uspekhi Matematicheskikh Nauk, **43**, 191 (1988). See also Russian Mathematical Surveys (London), **43**, 218 (1988).

[5] A. W. Niukkanen, Matematicheskie Zametki, **50**, 65, (1991).

On Numerical Solving of the Direct Regulator Problem

A. D. Ovsyannikov

Control Theory Department, St.Petersburg State University
Bibliotechnaya pl.2, Petrodvorets, St.Petersburg, 198904, Russia
e-mail: Alex@ovs.apmath.spb.su

Abstract. Problem of computation of direct regulator matrix, providing given eigenvalues for closed loop system, is considered. System of the form

$$\frac{dx}{dt} = Ax + Bu, \quad y = Cx, \tag{1}$$

where A, B and C are matrices given, and asymptotic observer equation

$$\frac{dz}{dt} = Az + Bu + H(y - Cz), \quad u = K(z - x^*)$$

are investigated. Transient processes of (1) closed with controls of $u = K(z - x^*)$ and $u = K(x - x^*)$ are known to be equal if $A - BH$ matrix is Hurwitz and $x(0) = z(0)$. (Here x^* is a constant vector.) Matrix H is supposed to be an optimal observation problem solution. Transient process analysis problem is reduced to the analysis problem of system (1) closed with $u = K(x - x^*)$ control. Matrix K is chosen to provide eigenvalues given for $A + BK$ matrix. Theoretical bases of this type regulator construction are well known. However, practical implementation of familiar algorithm is troublesome. Especially, difficulties appear in studies of high order controlled systems. Certain modification of common algorithm of spectrum assigning for closed system based on exploitation of special orthogonal transformation instead of the ordinary one is proposed. Suggested algorithm is used for solving tokamak plasma shape stabilization problem. Plasma cord behaviour is described by 22th-order system of linear differential equations. Transient process optimization is performed by means of eigenvalue choosing in a given domain of characteristic polynomial roots plane. The corresponding computer program is composed.

The system of linear ordinary differential equations is considered:

$$dx/dt = Ax + Bu \quad (1) \quad ,$$

where x is n-dimensional vector of system state, A is $n \times n$ constant matrix, u is r-dimensional vector of control, B is $n \times r$ constant matrix.

To construct the control of the form $u = Kx$ is required; K is $r \times n$ matrix such that matrix $(A + BK)$ of closed loop system has given eigenvalues.

It is well known fact of the control theory that sometimes it is impossible to assign all spectrum. So further we must bear in mind that spectrum assigning means to assign controlable part of spectrum.

Let us transform system (1) by special orthogonal transformation $x_1 = Sx$ to block-triangular form such that diagonal blocks have almost triangular structure. The way of constructing the matrix S is following. Let take linear independent vectors

$$b_1, \ Ab_1, \ \ldots \ , \ A^{k_1 - 1} b_1,$$

where number k_1 is chosen such that vector $A^{k_1}b_1$ is a linear combination of them.(b_i is the ith column of the matrix B.)

Let column b_2 is linear independent on built vectors. Then we consider the second set of vectors

$$b_2,\ Ab_2,\ \ldots\ ,\ A^{k_2-1}b_2,$$

which have property that they and the first vector set are linear independent but vector $A^{k_2}b_2$ is a linear combination of all previous vectors.

If we continue this process to the last , we will find final series

$$b_r,\ Ab_r,\ \ldots\ ,\ A^{k_r-1}b_r,$$

that are linear independent on themselves and all earlier found vectors. And relation $k_1+k_2+\ldots+k_r = n$ is satisfied when the system (1) is all-controllable, or $k_1+k_2+\ldots+k_r = m < n$ is relevant in the other case. In the second case we must complete built vector collection to basis by any suitable vectors $t_1,\ \ldots\ ,\ t_{k_{r+1}}$; $(k_{r+1} = n - m)$.

Let all these vectors are subjected to procedure of orthogonalization and normalized. So we have orthonormal basis of whole space. Let us take these basis vectors as columns of matrix S:

$$S = (s_1,\ s_2, \ldots,\ s_{m_1};\ s_{m_1+1}, \ldots,\ s_{m_2}; \ldots\ldots; s_{m_{r-1}+1}, \ldots,\ s_{m_r};\ s_{m_r+1}, \ldots\ s_n),$$

where $m_1 = k_1,\ m_2 = k_1 + k_2,\ \ldots,\ m_r = k_1 + k_2 + \ldots + k_r$.

Some numbers among the $k_1, \ldots, k_r$ can be equal zero. It means that corresponding column b_i of matrix B and previously built vectors are linear dependent. These columns are not needed for control constructing.

Built by this way matrix S will be orthogonal :

$$S^{-1} = S^T\ ,\ S \cdot S^T = E\ .$$

After the transformation the system (1) will have the form :

$$dx_1/dt = A_1x_1 + B_1u\ . \qquad (2)$$

$$A_1 = S^TAS\ ,\ B_1 = S^TB\ . \quad (3)$$

Let us detail structure of the matrix A_1 :

$$A_1 = S^T(AS) \ \Rightarrow\ S^T \cdot (As_1,\ As_2, \ldots,\ As_n)$$

Evidently the following correlations are fulfilled:

$$As_1 = d_{11}s_1 + d_{21}s_2\ ,\ d_{21} \neq 0$$

$$As_2 = d_{12}s_1 + d_{22}s_2 + d_{32}s_3\ ,\ d_{32} \neq 0$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$As_k = d_{1k}s_1 + d_{2k}s_2 + \ldots + d_{kk}s_k + d_{k+1,k}s_{k+1}$$

$$d_{k+1,k} \neq 0, \quad k \neq m_i \ , \ i = 1, 2, \ldots, r$$

$$d_{m_i+1,m_i} = 0, \quad i = 1, 2, \ldots, r$$

It is clear that d_{ij} and corresponding elements of the matrix A_1 are equal; all the others elements of A_1 are equal to zero. Hence the matrix A_1 has a block-triangular structure. And diagonal blocks have the following dimensions: $m_1 \times m_1$, $(m_2 - m_1) \times (m_2 - m_1)$, $\ldots$ $(m_r - m_{r-1}) \times (m_r - m_{r-1})$, All diagonal blocks have almost triangular form and non-zero subdiagonal ($a^1_{ij} = 0$, $i > j + 1$; $a^1_{i+1,i} \neq 0$), except $(r + 1)$th block.

Those corre ations give us the way of the matrix S calculating:

$$s_1 = b_1/\|b_1\| \ ,$$

$$s_2 = [As_1 - (As_1, s_1)s_1]/\|[\ldots]\| \ ,$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$s_{k+1} = [As_k - \sum_{i=1}^{k}(As_k, s_i)s_i]/\|[\ldots]\| \ ,$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$till \quad As_{m_1} - \sum_{i=1}^{m_1}(As_{m_1}, s_i)s_i = 0 \ ,$$

$$then \quad s_{m_1+1} = [b_2 - \sum_{i=1}^{m_1}(b_2, s_i)s_i]/\|[\ldots]\| \ ,$$

$$s_{m_1+2} = [As_{m_1+1} - \sum_{i=1}^{m_1+1}(As_{m_1+1}, s_i)s_i]/\|[\ldots]\| \ ,$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

where s_i is ith column of the matrix S , b_i is of the matrix B correspondingly.

The characteristic polynomial of the matrix A_1 is a product of the characteristic polynomials of the diagonal blocks. Let a control u_1 acts at the first block, a u_2 acts at the second, et cetera, a u_r acts at the rth diagonal block. Thus we will find the first control as a linear combination of $x^1_1, \ldots, x^1_{m_1}$, the second as a linear combination of $x^1_{m_1+1}, \ldots, x^1_{m_2}$, $\ldots$, the rth as a linear combination of $x^1_{m_{r-1}+1}, \ldots, x^1_{m_r}$.

In accordance with configuration of the matrix B_1, any control will modify the first row of corresponding block. Is it possible to assign any eigenvalues to almost triangular matrix by fixing the values of the first row elements?

Theorem. Let us consider almost triangular $k \times k$ matrix P such that its subdiagonal does not contain zero. ($p_{i+1,i} \neq 0$, $i = 1, \ldots, k - 1$). Let given polynomial $q_k(\lambda)$ has k predesigned roots (taking into account multiplicity). (Degree of the polynomial is k, and its leading coefficient is unity.) Then there always exists such one set of numbers $c_1, c_2, \ldots, c_k$, that a matrix being the result of replacement of the matrix P first row with this set has the same roots as the given polynomial.

Proof. We denote the matrix P after the first row substitution by P_c. Let consider the determinant: $D_k = det(P_c - \lambda E)$. It has a connection with the characteristic polynomial

of the matrix P_c by formula: $\chi(\lambda) = (-1)^k D_k$. We can decompose determinant D_k via the first row:

$$D_k = -\lambda D_{k-1} + \sum_{i=1}^{k} c_i [D_{k-i}(-1)^{i+1} \prod_{j=2}^{i} p_{j,j-1}] ,$$

where D_{k-i} is determinant of a matrix that built from $(P_c - \lambda E)$ by discarding of i front rows and columns.

It is clear that

$$D_k = -\lambda D_{k-1} + \sum_{i=1}^{k} c_i v_{k-i}(\lambda) ,$$

where $v_{k-i}(\lambda)$ are $(k-i)$ degree polynomials. Let us demand that $\chi(\lambda \equiv q_k(\lambda)$ or $(-1)^k q_k(\lambda) \equiv D_k$. Hence we can draw the conditions for $c_1, \ldots, c_k$:

$$(-1)^k q_k(\lambda) + \lambda D_{k-1} \equiv \sum_{i=1}^{k} c_i v_{k-i}(\lambda) ,$$

Let denote the polynomial staying at left part of equation by $w_{k-1}(\lambda)$. Its degree is equal to $(k-1)$. Finally we obtain:

$$w_{k-1}(\lambda) \equiv \sum_{i=1}^{k} c_i v_{k-i}(\lambda) , \quad (4)$$

It means that coefficients at the same degree of left and right polynomials must be equal. So we have triangular linear system to find $c_1, \ldots, c_k$. (As coefficient at the λ^{k-1} linear depend on c_1, one at λ^{k-2} count on c_1, c_2, and at λ^{k-i} — on $c_1, c_2, \ldots, c_i$). And diagonal of the system matrix does not contain zero. There always exists only decision of such type systems. Q.E.D.

Thus the procedure of closed-loop system matrix spectrum assigning has the following framework:

1. Calculating of the matrices S, A_1, B_1.
2. Assigning of polynomials $q_{k_1}(\lambda), q_{k_2}(\lambda), \ldots, q_{k_r}(\lambda)$.
3. Deriving and solving of the system (4) for each diagonal block except maybe the last one if the system (1) has uncontrollable part. Finding of the sets $c_1^1, \ldots, c_{k_1}^1$; $c_1^2, \ldots, c_{k_2}^2$; $\ldots$; $c_1^r, \ldots, c_{k_r}^r$.
4. Matrix K *designing:*

$$x_1 = S^T x \ , \ u = K_1 x_1 = K_1 S^T x \ , \ K = K_1 S^T$$

The i*th row elements from the* $(m_{i-1}+1)$*th position to the* (m_i)*th of the* K_1 *are equal to a difference between* c_j^i *and the* j*th element of the first row of the* i*th diagonal block* ($j = 1, \ldots, k_i$), *divided by* $b^1_{i,m_{i-1}+1}$. ($m_0 = 0$; $i = 1, \ldots, r$; $b^1_{i,m_{i-1}+1} \neq 0$ *by the construction of* B_1). *The other elements of the matrix* K_1 *are equal to zero.*

References

[1] V. I. Zubov, Mathematical Methods of Research of Automatic Control Systems, 319 (1974).

[2] Yu. N. Dnestrovskij, D. P. Kostomarov, Mathematical Modelling of Plasma, 231 (1993).

[3] E. I. Verenei, Engineering Cybernetics **9**, (1986).

[4] D. K. Faddeev, V. N. Faddeeva, Computational Methods of Linear Algebra, 735 (1963).

Algorithms for selecting the form of approximation expressions and splines

Bogdan A. Popov

Institute for Phisics and Mechanics of Ukrainian Academy of Sciences
Naukova str. 5, Lviv, 290601, Ukraine
e-mail: bogdan@popov.lviv.ua

1 Introduction

Let a function $f(x)$ be approximated in the interval $[a,b]$ by one of the rational expressions

$$R_{k,l}(A,x) = \frac{\sum_{j=0}^{k} a_j x^j}{\sum_{j=0}^{l} b_j x^j}, \quad k = 0,1,\ldots,m, \quad l = m-k. \tag{1}$$

It is sometimes convenient to use different approximations on different subintervals of the interval $[a,b]$, i.e. to change to an approximations by splines. If the maximum approximation error is the same on each subinterval (spline links) then the approximation is called a "balanced" spline approximation. If each spline link is a best Chebyshev approximation then such approximation is an optimal one [2]. This means that if the boundaries of the spline links (knots of spline) are selected from the condition of the balanced approximation we get the least possible error with the fixed number of links or the least possible number of links with the fixed maximum error.

The balanced approximation usage may greatly reduce the evalution time and it makes it possible to approximate functions on large intervals. The approximation efficiency may greatly increase if different approximating expressions of fixed number $s = m+1$ are used on different subintervals.

The balanced approximation is not found in any computer algebra systems due to the great amount of calculations needed. All the more the balanced approximation with different expressions on subintervals as it increases the calculaiton amount in s times in the case of the simple sort. But the evaluation time can be decreased by using analytical transformations on the error expressions for the balanced approximation.

2 Theoretical Background

Consider a set of points $Z = \{z_i\}$, $i = 0,\ldots,r$ in $[a,b]$ that are arranged in ascending order: $a = z_0 < z_1 < \ldots < z_r = b$. If the function $f(x)$ is approximated on each subinterval (z_{i-1}, z_i), $i = 1,\ldots,r$ by an expression of the form (1) with different numerical values of approximating parameters then this function is approximated on the whole interval $[a,b]$ by the spline:

$$S(F,x) = \sum_{i=1}^{r} R_{k,l}(A_i,x)\ \Theta((x-z_{i-1})(z_i-x)), \tag{2}$$

where $\Theta(x)$ is the Heavyside function.

If approximations on each link of spline (2) are the best Chebyshev approximations, the spline is called a Chebyshev spline.

Definition. If for all $i = 1, 2, \ldots, r$ links of spline (2)

$$\mu_i = \max_{x \in [x_{i-1}, x_i]} |f(x) - S(F,x)|\ w(x) = \mu = Const\ , \tag{3}$$

where $w(x)$ is an approximation weight function, $w(x) > 0$ for all $x \in [a,b]$, then the approximation of $f(x)$ by spline $S(F,x)$ on the interval $[a,b]$ is called a balanced approximation with weight $w(x)$.

Let as define the kernel $\eta_{k,l}(f)$ of approximation of the function $f(x)$ by a rational expression (1) :

$$\eta_{k,l}(f) = \frac{(m+1)!\ \Delta_{k+1,l+1}(f)}{\Delta_{k,l}(f)}\ , \tag{4}$$

where

$$\Delta_{t,s}(f) = \begin{cases} 1 & \text{for } s \le 0 \\ \begin{vmatrix} c_{t+1-s} & c_{t+2-s} & \ldots & c_t \\ c_{t+2-s} & c_{t+3-s} & \ldots & c_{t+1} \\ \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots \\ c_t & c_{t+1} & \ldots & c_{t+s-1} \end{vmatrix} & \text{for } s = 1, 2, \ldots, \end{cases}$$

$$c_\nu = c_\nu(x) = \begin{cases} 0 & \text{for } \nu < 0, \\ \frac{f^{(\nu)}(x)}{\nu!} & \text{for } \nu = 0, 1, 2, \ldots \end{cases}$$

For the polinomial of degree m the kernel is

$$\eta_{m,0}(f) = f^{(m+1)}(x)\ . \tag{5}$$

The error of a balanced Chebyshev rational spline approximation is extablished by the following theorem [3].

Theorem. Suppose that $f(x) \in C^{m+2}[a,b], w(x) \in C[a,b], \eta_{k,l}(f) \in C[a,b], w(x)\eta_{k,l}(x) \neq 0$ for $x \in [a,b]$. Then, for $r \to \infty$, the error μ of the balanced Chebyshev approximation by spline (2) with r links of function $f(x)$ on the interval $[a,b]$ will be

$$\mu = \frac{r^{-m-1}}{2^{2m-1}(m+1)!} \left(\int_a^b |\ \eta_{k,l}(f)\ w(x)\ |^{\frac{1}{m+1}} dx \right)^{m+1} \left(1 + O\left(\frac{b-a}{r}\right) \right) \tag{6}$$

3 Finding Kernel Expressions

We can find the kernels $\eta_{k,l}(f)$ from the formula (4). This transformations are generally beyong the capabilities of pen-and-paper. But a symbolic computation system such as Maple, however, has no difficulty with these manipulations. We use Maple V Release 3 [1]. Procedure $ratc\ (F,k,l)$ was created for this purpose. In the procedure F stands for the approximated function, k and l are numerator and denominator degrees of rational expression (1) ($l = 0$ by default).

```
> #Kernals for rational expressions
> rac:=proc(F,k,l)
>    local v,c,i1,i2,a,b,A,B,j,m,eta;
>    if nargs < 3 then v:=0 else v:=l fi;
>    with(linalg,matrix,det):
>    if v=0 then a:=1 else a:=matrix(v,v) fi: m:=k+v:
>    b:=matrix(v+1,v+1):c:=0;
>    for i1 to v+1 do j:=k-v+i1:
>      if j<0 then c:=0 elif j=0
>        then c:=F(x)
>            else if c<>0 then c:=factor(simplify(diff(c,x)/j))
>                    else c:=factor(simplify(diff(F(x),x$j)/j!)) fi fi;
>    b[i1,1]:=c; od;
>    for i1 to v+1 do for i2 from 2 to v+1 do j:=k-v-1+i1+i2;
>          if i1<v+1 then b[i1,i2]:=b[i1+1,i2-1]
>              else c:=factor(simplify(diff(c,x)/j));
>                       b[i1,i2]:=c fi od od;
>    for i1 to v do for i2 to v do a[i1,i2]:=b[i1,i2] od od;
>    if v=0 then A:=1 else A:=factor(simplify(det(a))) fi:
>    B:=factor(simplify(det(b))):
>    eta:=factor(sort(simplify((m+1)!*B/A)));
> end:
```

Let us consider some examples of kernels finding :

Example 1. Find the approximating kernels for $F = e^x$; $m = 9$; $l = 0, 1, \ldots, 9$; $k = m - l$.

```
> m:=9:for l from 0 to m do k:=m-l: a[l+1]:=rac(exp,k,l); od:
> print(seq(a[l],l=1..m+1));
```

$$e^x, -\frac{1}{9}e^x, \frac{1}{36}e^x, -\frac{1}{84}e^x, \frac{1}{126}e^x, -\frac{1}{126}e^x, \frac{1}{84}e^x, -\frac{1}{36}e^x, \frac{1}{9}e^x, -e^x$$

The results can be obtained from the general expresion for $\eta_{k,l}\,(e^x)$ [3].

Example 2. $F = \ln(x)$; $m = 4$; $l = 0, 1, 2, 3, 4$; $k = m - l$.

```
> m:=4:for l from 0 to m do k:=m-l: a[l+1]:=rac(ln,k,l); od:
> print(seq(a[l],l=1..m+1));
```

$$24\,\frac{1}{x^5}, \frac{3}{2}\,\frac{1}{x^5}, \frac{2}{3}\,\frac{1}{x^5}, \frac{1}{2}\,\frac{3\ln(x)^2+14\ln(x)+19}{(\ln(x)^2+3\ln(x)+3)\,x^5},$$
$$2\,\frac{12\ln(x)^4+50\ln(x)^3+105\ln(x)^2+120\ln(x)+60}{x^5\ln(x)^4}$$

The first three expressions can also be obtained from the general expression for $\eta_{k,l}\,(\ln x)$, which is valid for $k \geq l$ [3].

Example 3. $F = \Gamma(x)$; $m = 2$; $l = 0, 1, 2$; $k = m - l$.

```
> m:=2:for l from 0 to m do k:=m-l;a[l+1]:=rac(GAMMA,k,l) od:
> print(seq(a[l],l=1..m+1));
```

$$\left(\Psi(x)^3+3\,\Psi(x)\,\Psi(1,x)+\Psi(2,x)\right)\Gamma(x),$$
$$-\frac{1}{2}\,\frac{\left(\Psi(x)^4-2\,\Psi(x)\,\Psi(2,x)+3\,\Psi(1,x)^2\right)\Gamma(x)}{\Psi(x)},$$
$$-\left(-\Psi(x)^3+3\,\Psi(x)\,\Psi(1,x)-\Psi(2,x)\right)\Gamma(x)$$

Example 4. $F=x^v$; $m=2$; $l=0,1,2$; $k=m-l$.

```
> m:=2:for l from 0 to m do k:=m-l:a[l+1]:=rac(x->x^v,k,l); od:
> print(seq(a[l],l=1..m+1));
```

$$(v-1)(v-2)\,v\,x^{(v-3)},\ -\frac{1}{2}\,\frac{(v-1)(v+1)\,v\left(\frac{1}{x}\right)^{(-v)}}{x^3},\ \frac{(v+2)(v+1)\,v\left(\frac{1}{x}\right)^{(-v)}}{x^3}$$

The results can be obtained from the general expression for $\eta_{k,l}(x^v)$ [3].

4 Selecting Approximating Expressions

The numerator and denominator degrees k and l of the rational expression that give the least approximation error on the interval $[a,b]$,with the weight $w(x)$ and $m=k+l=Const$, can be found from formula (6). This is done in the procedure $selection(F,m,a,b,w)$. If w is omitted then $w(x)=1$ is used (approximation with absolute error). To speed up the calculations the series expansion of the integrated expression in the neighbourhood of the point $x=(b+a)/2$ is used. If b is omitted the values of the functions $\Phi_{k,l}(x)=|\eta_{k,l}(f)\,w(x)|$ at the point $x=a$ are compared. If a is omitted then the point $x=0$ is taken. The examples from the previous section show that this is often sufficient. The results of the procedure are the optimal values of k and l and corresponding kernel expression multiplied by the weight function. Let us consider several examples:

Example 5. Find the optimal k,l and the kernel for $F=e^x$, $m=9$, $w(x)=1$

```
> selection(exp,9);
```

$$k=5, l=4, kernel=\frac{1}{126}\,e^x$$

The result is the same as follows from example 1.

Example 6. $F=\ln(x)$, $m=4$, $w(x)=1$ in the neighbourhood of the point $x=1$

```
> selection(ln,4,1);
```

$$k=2, l=2, kernel=\frac{2}{3}\,\frac{1}{x^5}$$

The result is the same as follows from example 2.

Example 7. Find the optimal k,l and the kernel multiplied by $w(x)$ for $F=\Gamma(x), m=2$, for approximation with relative error on the intervals $[1,2]$ and $[2,3]$.

```
> selection(GAMMA,2,1,2,1/GAMMA);
```

$$k = 2, l = 0, kernel = \Psi(x)^3 + 3\,\Psi(x)\,\Psi(1,x) + \Psi(2,x)$$

```
> selection(GAMMA,2,2,3,1/GAMMA);
```

$$k = 0, l = 2, kernel = \Psi(x)^3 - 3\,\Psi(x)\,\Psi(1,x) + \Psi(2,x)$$

Here $\Psi(x) = \Gamma'(x)/\Gamma(x)$, $\Psi(k,x) = d^k\Psi(x)/dx^k$.

Example shows that the same function can have different optimal approximating expressions on different subintervals.

Example 8. It is impossible to find the optimal k and l for $F = x^v$, if parametr v have no concret numerical meaning. But if, for example, $v = \pm 1/3$, $m = 5$, $x \in [1,2]$, $w(x) = 1/F$ it is quite possible.

```
> selection(x->x^(1/3),5,1,2,x->x^(-1/3));
```

$$k = 3, l = 2, kernel = -\frac{224}{729}\frac{1}{x^6}$$

```
> selection(x->x^(-1/3),5,1,2,x->x^(1/3));
```

$$k = 2, l = 3, kernel = \frac{224}{729}\frac{1}{x^6}$$

5 Conclusions

It is possible to extend the set of valide approximations to non-linear for parameters expressions (sum of polynomials and exponent, sum of logarithm and polynomials etc.). Analytical expressions for corresponding kernels can be found in [3,4]. It is possible to build procedures that derives an optimal balanced approximation on the base of selecting the optimal expression on each link.

References

[1] A. Heck, Introduction to Maple: Springer-Verlag, N.Y., 1993.

[2] J. Meinargus, Approximation of functions. Theorie and numerical methods: Springer, Berlin, 1967.

[3] B.A. Popov, Balanced Approximation by Splines, Naukova Dumka, Kyiv, 1989 [Russian].

[4] B.A. Popov, Information Technology of Formulaic Representation of Physical Relationships, Pattern Recognition and Image Analysis, vol 4, N 3, pp.306-312, 1994.

Simulation of a boundary layer between the high density non-ideal plasma and condenced medium [1]

Poritsky P.V., Kononov A.V.,Starchyk P.D., and Voitenko L.M.

Department of Plasma Physics,
Scientific Centre "Institute for Nuclear Research",
pr. Nauky 47, Kyiv 252022
UKRAINE
Phone: +380-(44)-265-3839
Fax: +380-(44)-265-4463
E-mail: kinr@sovam.com

Abstract. Electrical breakdown of gases, liquids and insulators frequently occurs by means of the narrow plasma streamer channel that inhibits a strong tendency to branching into complicated stochastic pattern . In order to describe these structures it is usefull on to calculate the fractal dimension of channel pattern. It is shown that the process of electrical breakdown is more intensive by means of photoinisation mechanism. Also, it is shown that the growth of a leader channel should be non-regular and the velocity of its spreading is variable in time.

Electrical breakdown of gases, liquids and insulators frequently occurs by means of the narrow plasma streamer channel that inhibits a strong tendency to branching into complicated stochastic pattern [1, 2]. In order to describe these structures it is usefull to take into consideration the fractal formalism. In that way a number of interesting results have been obtained with the using of computational simulation on the basis of the equation of potential [3, 4]. However, an analytical approach allows us to obtain more detailed information about the space properties of branching structures and a physical processes under their forming. A characteristic process of development of electrical breakdown by means of the forming of the high conductive plasma streamer has some stages. This process initiates with the motion of weak-luminous pilot-streamer, which moves forward continuously from cathode to anode. The weak-ionised channel produced by pilot-streamer becomes unstable. A secondary ionisation takes place in the channel for a some time, and as a consequence of that more bright radiation and more high conduction are appeared. That process represents motion of the branching leader which spreads fast and stagly inside the channel produced by pilot-streamer. The velocity of single stage can be reached 10^9 cm/s. Motion of a leader is due to an additional ionisation which is produced by the increasing of electric intensity or by the amplification of photoinisation as a consequence of irradiation of channel plasma. The problem is to determine types of elementary processes which cause the high conduction. Thus, a leader type of streamer is characterised by bright radiation and non-regular motion. In this paper we propose an analytical approach to the description of the branching properties of the leader type of streamer channel on base of the theory for the reaction-diffusion growth on fractal clasters [5, 6]. Consider a radial discharge in continuum. In that case the spreading of a leader is described by the equation of a reaction-diffusion growth in fractal geometry where the diffusion coefficient

[1]This work was supported in part by Ministry of Science and Technology Policy of Ukraine

is used in the form of the diffusion coefficient of the channel plasma and the reaction order parameter (corresponds to the order of ionisation reaction:

$\frac{\partial \rho}{\partial t} = \frac{1}{x^{p-1}} \frac{\partial}{\partial x} \left[k \cdot x^{-\sigma} \cdot x^{p-1} \cdot \frac{\partial p}{\partial x} \right] + R \cdot p^{r}, (1)$

where is the $\rho = (x,t)$ density of probability, which represents the belonging of the space point to the channel; x and t are radial and time dimensionless coordinates; D is fractal dimension, σ is the anomalous diffusion parameter ($\sigma = D - d_e$; d_e is the dimension of Euclidean embedding space); R is parameter which represents the part of ionisation in the process of structure growth and which has a different value for the different types of discharge; (γ is parameter corresponded to the order of ionisation reaction, and is determined via spectral fractal dimension : (γ=1 + $d/2$, where $d_s = 2D/(\sigma + 2)$. It is considered the two ionisation processes: both the ionisation by means of electron strike and photoionisation. Spectral fractal dimension corresponds to an effective number of degrees of freedom in physical system. It should be have in mind that for the case of electron strike ionisation it is essential the influence of external electric field because of that the derived direction is appeared . According to that an effective spectral dimension is introduced in the following form $d_{eff} = d_s$ -1, where (-1) points the derived direction. Thus, for the case of prevail ionisation by electron strike we have , and for the photoinisation . The eq. (1) has the solution at the form $\rho(x,t) \propto t^{\alpha} \cdot x^{\beta}$, where α, β are parameters depended on D, d_e, γ. It should be pointed that the evolution of channel have a dynamical character due to the density of probability depends obvious on time. For the fractal cluster should be valid the dependence $\rho(x,t) \propto x^{D-d_e}$(according to such rule the density of particles numbers on fractal is distributed in a space [4]), wherefrom we obtain that the fractal dimension D is the certain function of $\gamma^3 \cdot d_e$. The value of D may be obtained from algebraic equation:

$\frac{2 \cdot D^2 - 3 \cdot D \cdot (d_e - 1) - 1 + d_e \cdot (d_e - 2)}{2 \cdot (D - d_e + 2)} + \frac{1}{\gamma - 1} = 0, (2)$

Substituting into eq. (2) the values of γ for different cases of ionisation, it may be obtained the corresponding values of D. In that way, in three-dimensional space for the case of prevail ionisation by electron strike we obtain $D = D_e = 2.15$, and for the case of prevail photoinisation $D = D_f = 2.25$. It should be pointed that the global space properties of fractal structure are characterised by the fractal dimension D. The sense of this result in that the process of electrical breakdown is more intensive in a case of photoinisation mechanism. Also, it is shown that the growth of a leader channel should be non-regular and the velocity of its spreading is variable in time. Thus, the character of evolution and the parameters of the fractal structure of leader channel are determined by the prevail type of ionisation. The parameters of fractal structure (fractal dimension and spectral fractal dimension) have the different values in dependence on the mechanisms of channel growth by means of the ionisation by electron strike or photoinisation. The process of electrical breakdown is more intensive under the prevail photoionitation growth and channel have a more strong tendency to branching by comparison with the case of ionisation by electron strike. Also, it is shown that the growth of a leader channel should be non-regular and the velocity of its spreading is variable in time. It should be pointed that the considered theoretical model of electrical breakdown in continuum may be modified and improved with the involving of the others elementary processes: ion diffusion, recombination processes, secondary order ionisation etc. Also the considered theory may be applied to describe an initial stage of arc ignition in the controlled fusion reactors.

References

[1] Meek J.M., Craggs J.D. Electrical Breakdown of Gases (Wiley,New York,1978).

[2] Feder J. Fractals (Plenum,New York,1992).

[3] Niemeyer L. , Pietronero L., Wiesmann H.J. Phys.Rev. Lett. (1984) -82.-p.1033.

[4] Kupershtokh A.L. Pisma v Zhurn. Tekhn. Fiz. (1992) -18.-p.91.

[5] Clement E , Sander L.M., Kopelman R. Phys.Rev. (1984) -A39.-p.2305.

[6] Milovanov A.V., Zelenyi L.M. Phys.Fluids (1993) -B5.-p.2609.

Simulation of a boundary layer between the high density non-ideal plasma and condenced medium [1]

Poritsky P.V., Kononov A.V.,Starchyk P.D., and Voitenko L.M.

Department of Plasma Physics,
Scientific Centre "Institute for Nuclear Research",
pr. Nauky 47, Kyiv 252022
UKRAINE
Phone: +380-(44)-265-3839
Fax: +380-(44)-265-4463
E-mail: kinr@sovam.com

Abstract. It is considered the numerical model for the simulation of a thin high-density layer formed between the high-density non-ideal plasma and bounded medium that occurs in an electrical pulse discharge in condenced media. To study the basic system of the equations in one-dimensional approximation the method of large particles has been used with the separated solving of the energy equation to carry out the conservation of energy in the studied cell.

The electrical pulse discharges in liquids and high pressure discharges in capillary tubes has been thoroughly studied to use in a number of technological applications. In that discharges a thin high-density layer is formed between the high-density plasma and bounded medium. In this layer gas-kinetical parameters varied mostly into the direction of normal to ablation surface and the large flows of energy and matter has been observed. The behavior of plasma and gas in this layer is important for the dynamics and for the all of discharge characteristics. The similar layer is formed also under optical breakdown of transparent matters at the focusing of layer irradiation. Experimental studying of temperature and density in boundary layer have a number of difficulties. Because of that the numerical simulation allows to obtain a number of the essential results. The basic system of the equations of one-dimensional model for boundary layer have the form

$$\rho \cdot v = -\frac{\partial P}{\partial x} + f_i, \ (f_i = 0) \ (1.1)$$

$$\frac{\partial(\rho \cdot v)}{\partial x} = 0 \ (1.2)$$

$$\frac{\partial\left(\frac{1}{2}\cdot\rho\cdot v^2\right)}{\partial x} + \frac{\partial(E\cdot v)}{\partial x} - \frac{\partial}{\partial x}\left(\lambda \cdot \frac{\partial}{\partial x} T\right) = 0; \left(= W_{\sum}\right) \ (1.3)$$

$$S(P, \rho, T) = 0 \ (1.4)$$

where eq.(1.1) is the hydrodynamical equation of motion, eq.(1.2) equation of continuity, eq.(1.3) is the energy conservation law, eq.(1.4) is the equation of state for potential in non-explicit form, ρ is density, v is velocity , P is pressure, T is temperature, $E(\rho, T)$ is the internal energy of gas, $\lambda(\rho, T)$ is heat conductivity, f_i is the density of inertial forces in accelerated frame of reference (as a rule $f_i = 0$), W is a total density of energy sources $= 0$ under an absence of ionization loses and Joule heating that is valid in this approximation. The x–axis is directed on normal to the ablation surface (the external surface of layer). The studied system

[1]This work was supported in part by Ministry of Science and Technology Policy of Ukraine

(1)-(4) having reduced and solved by the numerical methods. For energy equation the non-local subtraction scheme is used with the calculation of. 1st, 2nd, and 3rd terms in different space points under averaging step by step in following iterations. For this scheme the non-linear coefficients $E(\rho,T)$ and $\lambda(\rho,T)$ are calculated with solving the equation of state which having solved by the method of large particles in plasma cubic cell with the radii of 2-4 Debay radius when the number of particles no more then 50-100 and averaging at center. In that way the spatial distributions of temperature, pressure and internal energy were calculated in boundary layer for the various initial conditions which are correspond to the physical experiments. The obtained results have a good agreement with the experimental data. The equations (1.1) and (1.2) have analytical solutions that allows to reduce the system (1) to the following form

$C \cdot \rho \cdot V = P_0 - P$ (2.1)

$\rho \cdot V = C$ (2.2)

$C \cdot \frac{\partial\left(\frac{1}{2}V^2\right)}{\partial x} + \frac{\partial(E \cdot V)}{\partial x} - \frac{\partial}{\partial x}\left(\lambda \cdot \frac{\partial}{\partial x}T\right) = 0$ (2.3)

where C and P_0 are the certain constants. The system (2) is studied with using the numerical methods. For energy equation (2.3) the non-local subtraction scheme is used with the calculation of 1st, 2nd, and 3rd terms in different space points under averaging step by step in the following iterations:

$C \cdot \frac{V_+^2 - V_0^2}{2 \cdot h} + \frac{E_0 \cdot V_0 - E_{-1} \cdot V_{-1}}{h} - \frac{(\lambda_0 + \lambda_{-1}) \cdot T_0 - (\lambda_0 - 2 \cdot \lambda_{-1} + \lambda_{-2}) \cdot T_{-1} + (\lambda_{-1} + \lambda_{-2}) \cdot T_{-2}}{2 \cdot h^2} = 0$ (3.3)

For this scheme the non-linear coefficients internal energy $E(\rho,T)$ and heat conductivity $\lambda(\rho,T)$ are calculated with the combined solving of the equation of state (2.4) which having solved by the method of large particles in plasma cell with the radii of 2-4 Debay radius r_D when the number of particles no more then 50-100 and averaging at center.

The choice of scheme (3.3) is determined to find in each spatial point x the routs of algebraic subsystem (2.1), (2.2), (2.4) in explicit form as a result of a single iterated evaluation of numerical model (2.4). The using of standard subtraction schemes for eq.(2.3) makes the non-explicit form of system (2.1), (2.2), (2.4) with reference to the variables P, ρ and v that having caused to the re-iterated evaluations at the calculating of routs and lead to the unwarranted expenses of computer time.

Then, if it is needed, the obtained in this way the first approximation may be made more precise by the using of local subtraction scheme which is the analogue of the standard scheme for the heat conduction equation:

$C \cdot \frac{V_+^2 - V_-^2}{2 \cdot h} + \frac{E_+ \cdot V_+ - E_{-1} \cdot V_{-1}}{h} - \frac{(\lambda_+ + \lambda_0) \cdot T_+ - (\lambda_+ - 2 \cdot \lambda_0 + \lambda_-) \cdot T_0 + (\lambda_0 + \lambda_-) \cdot T_-}{2 \cdot (h^2)} = 0$

In plasma the interaction between particles is small and may be neglected for due to the Debay shielding (decreasing as $exp(-R/x)$. Hence, having used the method of particles to simulate a plasma, it is enough to calculate the interactive forces between charges placed on a relative distance no large than (1,5-2). In the case of plasma with non-ideality parameter 0.1-1 (the ratio of Coulomb energy to heat one) the number of particles is no more than 10-15. Because of that the number of particles in calculation cell measuring (2-4) should be no more than 50-100, that it allows to obtain the solution of eq.(2.4) for acceptable time. To eliminate the edge effects a calculation cell is bounded by "blanket" having the thickness about , with taking

into account in a dipole approximation. The statistical distribution of particles in "blanket" is given by the chance translation of central cell in the region of "blanket" after the certain iterations. To carrying out a several conservation of energy after the each 4-10 steps the multiply normalization of a kinetic energy of all particles is done to conservate a total energy of a cell. The calculation of the statistically averaging values is done on a central region of cell measuring . To calculate the heat conduction coefficient a central cell is redoubled. Between the both of its halfs by means the normalization of energy the small overfall is supported, and the energy flow is integrated over a separating surface. Because of the smooth dependence of gas-kinetic parameters on space coordinate in a boundary layer the simulation of cell is carried out in the space point x from its established state in the previous point.

In that way the spatial distributions of temperature, pressure and internal energy were calculated in boundary layer for the various initial conditions which are correspond to the physical experiments (Fig.1). The obtained results have a good agreement with the experimental data.

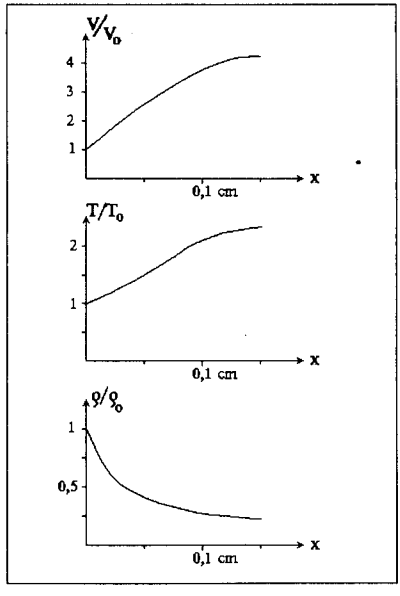


Fig.1.

The results of calculation for the typical conditions under the pulse electrical discharge in a water (energy power is $300J$ per a cm of the channe length). Here $V_0 = 10^4 cm/s$, $T_0 = 4eV$, $\rho = 0.5g/cm^3$.

High accuracy Newton iteration scheme for solving the multichannel eigenvalue problem of some exotic few-body systems [1]

Puzynin I.V., Puzynina T.P., Tyukhtyaev A.Yu., Vinitsky S.I.

Joint Institute for Nuclear Research, Dubna, Russia

e-mail: vinitsky@thsun1.jinr.dubna.su

Abstract. An iterative algorithm for numerical solving the partial Sturm-Liouville problem for the system of second-order linear differential equations with homogeneous boundary conditions is described. The discrete scheme which has been realized by means of the continuous analog of Newton method provides an accuracy of the four or six order with respect to mesh steps of the quasi-uniform grid of an independent variable. The method allows to inverse a conventional three-point difference operator during iteration process. The efficiency of the implemented algorithm is demonstrated by solving the eigenvalue problem for the system of the radial equations with the rotational coupling describing $\bar{p}He^+$ atomcule and verified on the exact solvable problem with the Coulomb potential.

1 Introduction

In the recent work [1] the energy level scheme of the $\bar{p}He^+$ atomcule has been studded in an adiabatic approximation with help of "SLIPH4" program [2] for the numerical solving of Sturm-Liouville problem for the radial Shroedinger equation. For the further study of the rotational-vibrational spectrum of this exotic Coulomb system with the values of the total orbital momentum $J = 30 - 40$ the high-accuracy program for solving the eigenvalue problem for N radial equations with the strong rotational coupling should be applied.

In present work we describe an iterative algorithm for numerical solving the partial Sturm-Liouville problem for the system of N-th radial differential equations

$$\phi_1(\lambda, y) = Iy''(x) + 2p(x)y'(x) + (q(x) - \lambda r(x))y(x) = 0 \,, \tag{1}$$

with homogeneous boundary conditions.

$$\phi_2(\lambda, y) = d_1(\lambda, a)y'(a) + f_1(\lambda, a)y(a) = 0 \,, \tag{2}$$

$$\phi_3(\lambda, y) = d_2(\lambda, b)y'(b) + f_2(\lambda, b)y(b) = 0 \,. \tag{3}$$

Here I is an unit matrix of size $N \times N$, $p(x)$, $q(x)$, $r(x)$ are square matrixes $N \times N$ which are given as a table of numerical values on the quasi-uniform grid [1,3], $y(x)$ is an eigenvector and λ is an eigenvalue. To find an isolate solution $\mathbf{z}^* = \{\lambda^*, y(x)^*\}$ on an interval $[a, b]$ we will specify matrixes d_i, f_i, $i = 1, 2$ and fixed a normalization condition

[1] The talk has been presented on International Conference Computational Modeling and Computing Physics, Dubna, September, 16-21, 1996, and submitted to Proceedings CMCP-96

$$\phi_4(\lambda, y) = \int_a^b y^T(x) r(x) y(x) dx - 1 = 0 \ . \tag{4}$$

In the section two we present the iteration scheme that has been realized in the continuous analog of Newton method. Using the perturbation function allows us to inverse a conventional three-point difference operator during iteration process. The proposed approach provides an accuracy of the four or six order with respect to mesh steps of the quasi-uniform grid of an independent variable.

In the section three we demonstrate the efficiency of the implemented algorithm by solving the eigenvalue problem for the system of the radial equations with the rotational coupling describing $\bar{p}He^+$ atomcule and verified on the exact solvable problem with the Coulomb potential describing $\bar{p}He^{++}$ atom.

2 Iteration scheme

We can treat the problem (1)-(4) as a nonlinear equation with respect to some isolate solution $\mathbf{z}$ following to paper[4]

$$\phi(\mathbf{z}) = 0 \ . \tag{5}$$

We suppose that a conventional finite difference approximation of n-th order accuracy with respect to the mesh step h of an appropriate quasi-uniform grid $\omega_l(x)$ [5] may be applied to reduce nonlinear equation (5) to the following one

$$\phi_h^{(n)}(\mathbf{z}_h) = 0 \ . \tag{6}$$

We suppose also that a perturbation parametrization of the reduced equation may be implemented by such a way

$$\phi_h^{(n)}(g, \mathbf{z}_h) = \phi_h^{(2)}(\mathbf{z}_h) + g(\phi_h^{(n)}(\mathbf{z}_h) - \phi_h^{(2)}(\mathbf{z}_h)) \ , \tag{7}$$

that a real value parameter $g \in [0, 1]$ will specify a finite difference approximation of n(g)-th order mentioned above. In further we will use the short notation $\varphi(g, z) \equiv \phi_h^{(n)}(g, \mathbf{z}_h)$, $z \equiv \mathbf{z}_h$. to rewrite equation (7) without subscript h :

$$\varphi(g, z) = \varphi_0(z) + g\Delta\varphi(z) \ , \quad z \equiv \mathbf{z}_h \ . \tag{8}$$

Now we introduce a continuous parameter $t \in [0, \infty)$ and consider equation (8) as a parametric one with respect to t:

$$\begin{aligned} \varphi(t, z(t)) &= \varphi_0(z(t)) + g(t)\Delta\varphi(z(t)) \ , \\ g(t) &= 1 - e^{-\alpha t} \ , \quad 0 < \alpha \leq 1 \ . \end{aligned} \tag{9}$$

To solve nonlinear equation (6) we will consider the abstract Cauchy problem in the framework of the evolution the Newton procedure following by paper [6]

$$\frac{d}{dt}\varphi(t, z(t)) = -\varphi(t, z(t)) \ , \quad z(0) = z_0 \ , \quad \varphi(0, z(0)) = \varphi_0(z_0) \ . \tag{10}$$

We can rewrite equation (10) in the conventional form

$$\varphi'_z \frac{dz}{dt} = -\varphi(t, z(t)) - g'(t)\Delta\varphi(z(t)) \ , \qquad (11)$$
$$z(0) = z_0 \ , \quad \varphi(0, z(0)) = \varphi_0(z_0) \ .$$

Here φ'_z means Fr'eshet derivative [4]. In particular case, when $\alpha = 1$ one can see that $g(t) + g'(t) = 1$ and one can rewrite (11) as follows

$$(\varphi'_{0z}(z(t)) + g(t)\Delta\varphi'_z(z))\frac{dz}{dt} = -\varphi(z(t)) \ , \qquad (12)$$
$$z(0) = z_0 \ , \quad g(0) = 0 \ .$$

To solve the equation (11) we use Euler's method

$$\varphi'_{0z} v_k = -\varphi_0(z_k) - (g_k + g'_k)\Delta\varphi(z_k) - g_k\Delta\varphi'_z v_{k-1} \ , \qquad (13)$$
$$z_{k+1} = z_k + \tau_k v_k \ , \quad g_{k+1} = g_k + \tau_k g'_k \ ,$$
$$k = 0, \ 1, \cdots \ , 0 < \tau_o \le \tau_k \le 1 \ .$$

Let the correction $v_k = \{\mu_k, \nu_k\}$ to the solution $z_k = \{\lambda_k, y_k\}$ be defined as follows

$$\nu_k = -y_k + \nu_{1k} + \mu_k\nu_{2k} \ . \qquad (14)$$

Then we have two boundary problems for determining the corrections ν_{1k} and ν_{2k}:

$$\varphi'_{0y}\nu_{1k} = -(g_k + g'_k)\Delta\varphi(z_k) - g_k\Delta\varphi'_z v_{k-1} \ , \qquad (15)$$
$$\varphi'_{0y}(\nu_{2k}) = -\varphi'_{0\lambda}(z_k) \ , \qquad (16)$$

and the following equation for determining the correction μ_k:

$$\mu_k = -\frac{\phi_4(y_k) - \phi'_{4y}\nu_{1k}}{\phi'_{4y}\nu_{2k}} \ . \qquad (17)$$

Thus, we can find the iteration corrections v_k at the k-th step and go to next layer $k+1$ with help of the following equations:

$$\lambda_{k+1} = \lambda_k + \tau_k\mu_k \ , \quad y_{k+1} = y_k + \tau_k\nu_k \ , \quad g_{k+1} = g_k + \tau_k g'_k \ . \qquad (18)$$

The iteration process (13)-(18) will be finished at such $k = k_{max}$ if for a given and sufficient small value of $\varepsilon > 0$ the discrepancy $\delta_k = ||\phi(z_k)||_C \le \varepsilon$.

3 Numerical results and discussion

The $\bar{p}He^+$ atom cule has reach rotational-vibrational spectrum. The most interesting for the experimental observing part of the spectrum lie between $J = 30\cdots43$ [7,8]. In a general approach the corresponding states will have the widths, which are caused by the radiative transitions and Auger's decay $\bar{p}He^+ = \bar{p}He^{++} + e^-$. In the latter the exit channel is described by the product of the wave function of a bound state of the $\bar{p}He^{++}$

atom and the wave function of Auger's electron. As an example of implementation and efficiency of the proposed iteration scheme, in Tables 1,2 we present the results of our numerical calculations of the eigenvalue problem for the radial Shroedinger equation with Coulomb potential describing the bound state with $J = 31$, $v = 0$. In tab e 3 we present our results for the problem (1)-(4) describing the bound state of the $\bar{p}He^+$ atomcule with $J = 31$, $v = 0$. Here we restrict our attention by the system of 7-th radial equations with the rotational coupling. As an initial approximation $z_0 = \{\lambda_0,\ y_0(x)\}$ we have used the numerical solution $z^{ad} = \{\lambda^{ad},\ y^{ad}(x)\}$ of the adiabatic approximation obtained with help of SLIPH4 program.

Table 1. Numerical solution of the problem for the radial equation with Coulomb potential describing $\bar{p}He^{++}$ atom ($J = 31$, $v = 0$) on interval $x \in [0.1, 1.0]$ with the mesh step $h = 0.00625$, by means of 4-th order difference scheme.

	λ_h	$\lambda_{h/2}$	$\lambda_{h/4}$	λ^*	σ
	.840543622E+04	.840543560E+04	.840543556E+04	.84054355[illegible]E+04	15.95
x	y_h	$y_{h/2}$	$y_{h/4}$	y^*	σ
0.1	.102908995E-06	.105560356E-06	.105721635E-06	.105732313E-06	16.44
0.2	.473713820E-01	.473702828E-01	.473702134E-01	.473702088E-01	15.84
0.3	.213188499E+01	.213188278E+01	.213188264E+01	.213188263E+01	15.88
0.4	.221380394E+01	.221381068E+01	.221381110E+01	.221381113E+01	15.93
0.5	.291472399E+00	.291473186E+00	.291473235E+00	.291473233E+00	15.98
0.6	.103928520E-01	.103928860E-01	.103928882E-01	.103928883E-01	15.83
0.7	.150433327E-03	.150442152E-03	.150442703E-03	.150442740E-03	16.01
0.8	.112551101E-05	.112575768E-05	.112577304E-05	.112577406E-05	16.05
0.9	.508662278E-08	.508926106E-08	.508942513E-08	.508943606E-08	16.08
1.0	.154452575E-10	.154603385E-10	.154612752E-10	.154613376E-10	16.10

Table 2. Numerical solution of the problem for the radial equation with Coulomb potential describing $\bar{p}He^{++}$ atom ($J = 31$, $v = 0$) on interval $x \in [0.1, 1.0$ with the mesh step $h = 0.005$, by means of 6-th order difference scheme.

	λ_h	$\lambda_{h/2}$	$\lambda_{h/4}$	λ^*	σ
	8405.435558477	8405.435557369	8405.435557352	8405.435[illegible]57351	63.72
x	y_h	$y_{h/2}$	$y_{h/4}$	y^*	σ
0.1	.105776617E-06	.105733004E-06	.105732324E-06	.105732[illegible]4E-06	64.16
0.2	.473702158E-01	.473702090E-01	.473702089E-01	.473702[illegible]9E-01	63.38
0.3	.213188265E+01	.213188264E+01	.213188264E+01	.213188264E+01	63.32
0.4	.221381111E+01	.221381113E+01	.221381113E+01	.221381113E+01	63.69
0.5	.291473238E+00	.291473239E+00	.291473239E+00	.291473239E+00	63.06
0.6	.103928882E-01	.103928884E-01	.103928884E-01	.103928884E-01	63.86
0.7	.150442744E-03	.150442740E-03	.150442740E-03	.150442740E-03	63.71
0.8	.112577460E-05	.112577408E-05	.112577407E-05	.112577407E-05	64.18
0.9	.508944463E-08	.508943619E-08	.508943606E-08	.508943606E-08	64.33
1.0	.154613995E-10	.154613386E-10	.154613376E-10	.154613376E-10	64.43

The calculations are realized on a sequences of three uniform grids $\omega_h(x)$ of the interval [0.1, 1.0] with steps h, $h/2$, $h/4$. It needs to calculate the relation σ:

$$\sigma = \frac{z_h - z_{h/2}}{z_{h/2} - z_{h/4}} . \tag{19}$$

It allows us to show the order n of approximation of the implemented difference scheme. Note, that the theoretical value of σ for the scheme of accuracy $O(h^n)$ is defined as 2^n [9].

Table 3. Numerical solution of the problem for system of 7-radial equations describing $\bar{p}He^+$ atomcule ($J = 31$, $v = 0$) on interval $x \in [0.1, 1.0]$ with the mesh step $h = 0.025$, by means of 6-th order difference scheme.

	λ_h	$\lambda_{h/2}$	$\lambda_{h/4}$	σ	$\delta\lambda$
	.102903[illegible]38895E+5	.102903753093E+5	.102903751659E+5	59.87	0.14E-03
x	$y_{1,h}$	$y_{1,h/2}$	$y_{1,h/4}$	σ	δy_1
0.1	0.4396[illegible]4387E-06	0.623047062E-07	0.559552734E-07	59.4	0.20E-01
0.2	0.2854[illegible]9112E-01	0.284966990E-01	0.284958277E-01	53.0	0.45E-02
0.3	0.1655[illegible]1273E+01	0.165537380E+01	0.165537427E+01	66.7	0.14E-02
0.4	0.2568[illegible]2480E+01	0.256884048E+01	0.256884258E+01	55.0	0.11E-02
0.5	0.5920[illegible]3730E+00	0.592079636E+00	0.592080012E+00	61.0	0.34E-02
0.6	0.4338[illegible]9716E-01	0.433890427E-01	0.433891141E-01	57.0	0.54E-02
0.7	0.1506[illegible]5918E-02	0.150714972E-02	0.150715414E-02	65.7	0.72E-02
0.8	0.3129[illegible]5709E-04	0.312680472E-04	0.312676708E-04	65.1	0.88E-02
0.9	0.4488[illegible]1116E-06	0.446973317E-06	0.446946017E-06	68.4	0.10E-01
1.0	0.4884[illegible]0410E-08	0.483071330E-08	0.482995616E-08	70.3	0.11E-01

In the last column of the Table 3 we point out also the value of $\delta z = \{\delta\lambda,\ \delta y_1\}$ which shows relative difference between the adiabatic approximation $z_0 = \{\lambda^{ad},\ y^{ad}(x)\}$ and solution $z = \{\lambda,\ y_1(x)\}$ of the problem for 7 radial equations with the rotational coupling. The obtained values δz confirms the theoretical estimation of [1]. In calculations we have used the following values of masses of the particles $M_{^4He^{++}} = 7294.299\ m_e$ and $M_{\bar{p}} = 1836.1527\ m_e$.

The obtained results indicate that the implemented high-accuracy Newton iteration scheme may be applied for further study of physical characteristics of the exotic Coulomb few-body systems.

Acknowledgements. We would like to thank Professor P.G. Akishin for helpful discussions. The authors are grateful to RFBR and INTAS-RFBR under Grants No. 96-02-17715 and No. 95-0512 for support.

References

[1] I.V. Puzynin, T.P. Puzynina, S.I. Vinitsky and V.I. Puzynin, Hyperfine Interection **101/102**, 493 (1996).

[2] I.V. Puzynin, T.P. Puzynina, and T.A. Strizh, Preprint JINR P1-87-332 (Dubna, 1987).

[3] L.I. Ponomarev and T.P. Puzynina, Zh.Vychisl.Math. **8**, 1256 (1968).
[4] L.V. Kantorovich and G.P. Akilov, Funchional Analysis, (Nauka, Moscow, 1977).
[5] A.A. Samarskii, Theory of Difference Schemes, (Nauka, Moscow, 1983).
[6] T. Zhanlav and I.V. Puzynin, Comput.Maths.Math.Phys. **32**, 1 (1992).
[7] T. Yamazaki, Hyperfine Interection **101/102**, 451 (1996).
[8] L.A. Melnikov, V.I. Derbov, I.V. Umansky and S.I. Vinitsky, Hyperfine Interection **101/102**, 471 (1996).
[9] G.I. Marchuk and V.V. Shaidurov, Improving the Accuracy of the Solution of Difference Schemes, (Nauka, Moscow, 1979).

ANALYTICAL STUDY OF THE GARY SCHEME ON PC. DISPLACEMENT OF OVERLAPPING GRIDS GIVES THE STABILITY

Svetlana Serdyukova [1]

JINR, Dubna, Russia
e-mail:sis@lcta36.jinr.dubna.su
Michael Thuné [2]

Uppsala University, Sweden
e-mail:michael@tdb.uu.se

Abstract. By using CAS REDUCE and MAPLE the Gary scheme with initial, complementary boundary and overlap conditions is studied. In the case of two different grids without displacement a power instability in space L_2 is proved: $\|\mathcal{G}^n\|$ grows at least as $\sqrt{n}$, when $n \to \infty$. It is an example of interesting instability phenomenon. The instability is observed in calculations, if only the overlap interval has an even number of full steps for the outflow problem. We state the power instability for $\Delta = 2h_1$, Δ is the length of the overlap interval. In the case of $\Delta = 2.5h_1$ the stability is proved.

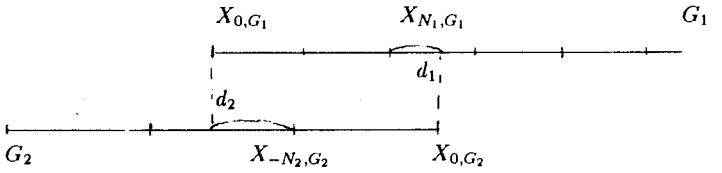
1 Introduction

We developed and realized on PC by using CAS an algorithm for stability verification of difference boundary problems. In this report we demonstrate possibility of this algorithm by searching the Gary scheme [1]. This scheme is used in airflow simulations.

The one-stage form of the Gary scheme for $u_t = u_x$ is

$$u_\nu^{n+1} = u_\nu^n + \frac{\alpha}{2}(u_{\nu+1}^n - u_{\nu-1}^n) + \frac{\alpha^2}{8}(u_{\nu+2}^n - 2u_\nu^n + u_{\nu-2}^n) + \frac{\alpha^3}{32}(u_{\nu+3}^n - 3u_{\nu+1}^n + 3u_{\nu-1}^n - u_{\nu-3}^n).$$

The corresponding Cauchy problem is stable in space L_2, when [2] $\alpha = \tau/h \leq 2$. Following to [2], we consider model problem on substructured domain consisting of two semiinfinite intervals with different grids G_1 and G_2, having h_1, h_2 steps in x.



N_1, N_2 are supposed independent of h_1, h_2. A problem with initial data is solved: $w^{n+1} = \mathcal{G}w^n$, $u = \{u_\nu\} \cup \{v_\nu\}$, $\nu \geq 0$. $\mathcal{G}$ is the operator of transition from the layer t to the layer $t + \tau$. What is $\mathcal{G}$ here?

[1] The research of this author was supported in part by the International Science Foundation and the Russian Government under Grant No. NKC000 and Grant No.NKC300

[2] The research of this author was funded through a faculty grant from Uppsala University

On grid G_1 the left boundary problem (the outflow problem) is solved u_ν^{n+1}, $\nu \geq 3$ are calculated by the Gary scheme with $\alpha = \tau/(h_1)$. The Gary scheme

requires two complementary boundary conditions to define u_1^{n+1}, u_2^{n+1}. We use [2]

$$u_1^{n+1} = u_1^n + \frac{\alpha}{4}[2u_2^n - u_0^n - u_0^{n+1} + \frac{\alpha}{2}(u_3^n - u_1^n) + \frac{\alpha^2}{8}(u_4^n - 2u_2^n + u_0^n)],$$

$$u_2^{n+1} = u_2^n + \frac{\alpha}{4}[2(u_3^n - u_1^n) + \frac{\alpha}{4}(2u_4^n - 4u_2^n + u_0^n + u_0^{n+1} + \frac{\alpha}{2}(u_5^n - 3u_3^n + 2u_1^n))].$$

The value of u_0^{n+1} is calculated by the interpolation of the solution values in the points of G_2 closest to X_{0G_1}.

Solution on G_2? On grid G_2 the right boundary problem is solved for $u_t = u_x$. This inflow problem can be interpreted as the left boundary problem for $v_t = -v_x$. The values of v_ν^{n+1}, $\nu \geq 1$, are found by the formulas above with u replaced by v and α replaced by $\hat{\alpha} = -\tau/h_2 = -\phi\alpha$, $\phi = h_1/h_2$. The value of v_0^{n+1} is found by the interpolation of the solution values in the points of grid G_1 closest to the considered boundary point X_{0,G_2}. Put $\beta_1 = d_1/h_1 < 1$, $\beta_2 = d_2/h_2 < 1$. In the case of the linear interpolation the overlap conditions are

$$u_0^{n+1} = (1-\beta_2)v_{N_2}^{n+1} + \beta_2 v_{N_2+1}^{n+1}, \quad v_0^{n+1} = (1-\beta_1)u_{N_1}^{n+1} + \beta_1 u_{N_1+1}^{n+1}.$$

The operator $\mathcal{G}$ is presented.

The considered problem with initial, complementary boundary and overlap conditions (i.b.o. problem) is stable if there is a positive constant c such that $||\mathcal{G}^n|| \leq c$ for all $n \geq 0$.

It is necessary for the stability, that the spectrum of $\mathcal{G}$ lies in the unit disk $|z| \leq 1$. The classic definition is used: a point of the complex plain z_0 is a spectrum point of $\mathcal{G}$, if there is found a nonzero sequence w_0 with $\mathcal{G}w_0 = z_0 w_0$. Here, w consists of two semi-infinite sequences $\{u_\nu\}, \{v_\nu\}, \nu \geq 0$, from L_2 satisfying the complementary boundary and overlap conditions.

2 Spectral problem

The spectral problem for the considered i.b.o. problem is a system of ordinary difference equations with a spectral parameter z, which is obtained from the original one by substituting $zu_\nu, zv_\nu, u_\nu, v_\nu$ instead of $u_\nu^{n+1}, v_\nu^{n+1}, u_\nu^n, v_\nu^n$, respectively. The solution of the spectral problem can be represented by the roots of the characteristic polynomials:

$$P_1(z,\kappa) = (1-z)\kappa^3 + \frac{\alpha}{2}(\kappa^4 - \kappa^2) + \frac{\alpha^2}{8}(\kappa^5 - 2\kappa^3 + \kappa) + \frac{\alpha^3}{32}(\kappa^6 - 3\kappa^4 + 3\kappa^2 - 1),$$

$$P_2(z,\hat{\kappa}) = (1-z)\hat{\kappa}^3 - \frac{\phi\alpha}{2}(\hat{\kappa}^4 - \hat{\kappa}^2) + \frac{(\phi\alpha)^2}{8}(\hat{\kappa}^5 - 2\hat{\kappa}^3 + \hat{\kappa}) - \frac{(\phi\alpha)^3}{32}(\hat{\kappa}^6 - 3\hat{\kappa}^4 + 3\hat{\kappa}^2 - 1).$$

It is necessary for the stability of the considered i.b.o. problem that the corresponding Cauchy problems on the infinite grids with the steps h_1, h_2 are stable. So we suppose,

that $\alpha^2 \le 4$, $\varphi^2\alpha^2 \le 4$. Then [3] for z with $|z| > 1$ each of the characteristic equations $P_1(z,\kappa) = 0$, $P_2(z,\hat{\kappa}) = 0$ has exactly three solutions less than one in the absolute value: $\kappa_1, \kappa_2, \kappa_3$ and $\hat{\kappa}_1, \hat{\kappa}_2, \hat{\kappa}_3$, respectively. In the general case (when $\kappa_1, \kappa_2, \kappa_3$ are different and $\hat{\kappa}_1, \hat{\kappa}_2, \hat{\kappa}_3$ are different) the solution of the considered spectral problem has the following form: $u_\nu = c_1\kappa_1^\nu + c_2\kappa_2^\nu + c_3\kappa_3^\nu$, $v_\nu = c_4\hat{\kappa}_1^\nu + c_5\hat{\kappa}_2^\nu + c_6\hat{\kappa}_3^\nu$, $\nu \ge 0$. On introducing these functions into the boundary and overlap conditions we get a linear homogeneous system for the definition of $c_1, c_2, c_3, c_4, c_5, c_6$. This system has nonzero solution, if

$$D(z) = det(E) = \det \begin{bmatrix} -1 & -1 & -1 & e_1(\hat{\kappa}_1) & e_1(\hat{\kappa}_2) & e_1(\hat{\kappa}_3) \\ e_2(\kappa_1) & e_2(\kappa_2) & e_2(\kappa_3) & -1 & -1 & -1 \\ e_3(\kappa_1) & e_3(\kappa_2) & e_3(\kappa_3) & 0 & 0 & 0 \\ 0 & 0 & 0 & e_4(\hat{\kappa}_1) & e_4(\hat{\kappa}_2) & e_4(\hat{\kappa}_3) \\ e_5(\kappa_1) & e_5(\kappa_2) & e_5(\kappa_3) & 0 & 0 & 0 \\ 0 & 0 & 0 & e_6(\hat{\kappa}_1) & e_6(\hat{\kappa}_2) & e_6(\hat{\kappa}_3) \end{bmatrix} = 0.$$

Here e_i are the polynomials determined by the boundary and overlap conditions. The coefficients of e_3, e_4, e_5, e_6 are linear functions of z. We have shown that the spectrum of the considered i.b.o. problem is described by a system of seven polynomial equations with seven variables

$$D(z, \kappa_1, \kappa_2, \kappa_3, \hat{\kappa}_1, \hat{\kappa}_2\hat{\kappa}_3) = 0, \quad |z| \ge 1,$$
$$P_1(z,\kappa_1) = 0, \quad P_1(z,\kappa_2) = 0, \quad P_1(z,\kappa_3) = 0,$$
$$P_2(z,\hat{\kappa}_1) = 0, \quad P_2(z,\hat{\kappa}_2) = 0, \quad P_2(z,\hat{\kappa}_3) = 0.$$

We reduce this system to a system of two polynomial equations in two variables $x = \kappa_1\kappa_2\kappa_3$, $y = \hat{\kappa}_1\hat{\kappa}_2\hat{\kappa}_3$. At first the determinant equation is transformed by a number of elementary manipulations with the columns of E into

$$det(\hat{E}(x_1, x_2, x, \hat{x}_1, \hat{x}_2, y, z)) = 0,$$

$x_1 = \kappa_1 + \kappa_2 + \kappa_3$, $\hat{x}_1 = \hat{\kappa}_1 + \hat{\kappa}_2 + \hat{\kappa}_3$, $x_2 = \kappa_1(\kappa_2 + \kappa_3) + \kappa_2\kappa_3$, $\hat{x}_2 = \hat{\kappa}_1(\hat{\kappa}_2 + \hat{\kappa}_3) + \hat{\kappa}_2\hat{\kappa}_3$.

Then by using the Vieta relations for the characteristic equations we get $x_1, \hat{x}_1, x_2, \hat{x}_2, z$ as functions of x, y. All these functions are simple ratios of polynomials. After substituting them into the transformed determinant equation we get one of the polynomial equations, another equation is obtained from one of the Vieta relations:

$$P(x,y) = num(\alpha^3(x_1^2(x,y) - x_2^2(x,y) + x^2 - 1) - x(8\alpha^2 + 32(z(x,y) - 1))) = 0,$$

$$Q(x,y) = num(det(\hat{E}(x_1(x,y), x_2(x,y), x, \hat{x}_1(x,y), \hat{x}_2(x,y), y, z(x,y)) = 0.$$

Remark, that this system automatically describes the spectrum points $z, |z| \ge 1$, when there are coincident roots among $\kappa_1, \kappa_2, \kappa_3$ or among $\hat{\kappa}_1, \hat{\kappa}_2, \hat{\kappa}_3$.

3 Separation of the false spectrum points

The attemp to solve this system by using REDUCE: $solve(\{P,Q\}, \{x,y\})$; have given no result. After this we used the resultant method. This leads to the false solutions. Another reason why the false solutions arise is the following. We can't say a priori in the

program, that the new variables $x_1, x_2, x, \hat{x}_1, \hat{x}_2, y$ are symmetric functions of κ_j and $\hat{\kappa}_j$ with the absolute value less than one only. Therefore in the program x_1, x_2, x depend on three arbitrary κ_j and $\hat{x}_1, \hat{x}_2, y$ depend on three arbitrary $\hat{\kappa}_j$. An algorithm for separating the false spectrum points was developed and implemented on a PC using REDUCE and MAPLE systems. The separation has two steps.

In the first step, for each $x, |x| \leq 1$, (the solution of the resultant) we find all solutions $y, |y| \leq 1$, of $P(x, y) = 0$ (the first polynomial equation). For each of the selected pairs (x, y) the values $x_1, x_2, \hat{x}_1, \hat{x}_2, z$ are calculated by the simple explicit formulas. For the following analysis we kept only the sets $(z, x_1, x_2, \hat{x}_1, \hat{x}_2, x, y)$ satifying the natural inequalities $|x_1| \leq 3,\ |x_2| \leq 3,\ |\hat{x}_1| \leq 3,\ |\hat{x}_2| \leq 3,\ |z| \geq 1 - \epsilon$.

In our program $\epsilon = 10^{-10}$. The points $z, 1 - \epsilon \leq |z| \leq 1 + \epsilon$, from the $\epsilon-$ vicinity of the unit circle are replaced by $z/|z|$, the closest points of the unit circle. After this we separate the false spectrum points, arising in the reduction to the resultant. We must simply verify if the selected (x, y) satisfy $Q(x, y) = 0$ (the second polynomial equation). But x, y were found numerically and $Q(x, y)$ has many terms with rather big coefficients. In order to avoid numerical difficulties we solve the resultant equation with high accuracy: in our REDUCE calculations PRECISION 15 was. Consequently, the solutions $x, |x| \leq 1$, lie at a distance less than 10^{-15} from the corresponding exact solutions. In our program x_0, y_0 are assumed to satisfy the second equation if the following inequality holds

$$|Q(x_0, y_0)| < 2 \times 10^{-15} \left| Q_x(x_0, y_0) + Q_y(x_0, y_0) \frac{P_x(x_0, y_0)}{P_y(x_0, y_0)} \right|.$$

This completes the first step of the separation.

In the second step the selected sets $(z, x_1, x_2, \hat{x}_1, \hat{x}_2, x, y)$ are analysed. For each set we find $\kappa_1, \kappa_2, \kappa_3, \hat{\kappa}_1, \hat{\kappa}_2, \hat{\kappa}_3$ less than one in absolute value (the roots of $P_1(\kappa, z), P_2(\hat{\kappa}, z)$ characteristic polynomials with z from the considered set). We check if $x_1, \hat{x}_1$ from the considered set are really the sums of the computed $\kappa_1, \kappa_2, \kappa_3$ and $\hat{\kappa}_1, \hat{\kappa}_2, \hat{\kappa}_3$ respectively. In our implementation, this is considered to be the case if the following inequalities hold:

$$|x_1 - (\kappa_1 + \kappa_2 + \kappa_3)| < 10^{-3}|x_1|,\ \ |\hat{x}_1 - (\hat{\kappa}_1 + \hat{\kappa}_2 + \hat{\kappa}_3)| < 10^{-3}|\hat{x}_1|.$$

All constants and inequalities were chosen in the process of numerical experiments. Further we present the results of concrete studies produced on a PC, using REDUCE and MAPLE.

4 Example I. Power instability.

The Gary scheme with $\alpha = 2,\ N_1 = 2,\ N_2 = 1,\ \phi = 1/2,\ \beta_1 = \beta_2 = 0$ is considered. There is no displacement here. The overlap conditions are $u_0^{n+1} = v_1^{n+1},\ v_0^{n+1} = u_2^{n+1}$ simply. The resultant $R(x)$ is a polynomial of 900th degree. REDUCE was not able to find it. MAPLE did this and factored $R(x)$ in approximately 4 hours on PC486-66/4MB. The factored $R(x)$ has 10 different factors-polynomials of maximal degree 68 with sufficiently small coefficients. The resultant has 96 nonzero solutions $x, |x| \leq 1$. For each computed x the polynomial equation $P(x, y) = 0$ of the fifteenth order is solved with respect to y. Thus we have to solve a lot of polynomial equations with the complex coefficients. All polynomials were solved and all the selected solutions were analysed. As the result the

unique (in the region $|z| \geq 1$) spectrum point $z = 1$ was found. Here $\kappa 1(1) = -1$. This leads to a power instability in space L_2. For the considered problem with initial, complementary boundary and overlap conditions $||\mathcal{G}^n||$ grows as $\sqrt{n}$ when $n \to \infty$. Such instability is given [4] by

$$\gamma_\nu^n = \frac{1}{2\pi i}\int_\Gamma \kappa_1^\nu(z) z^n \frac{dz}{(z-1)}, \ \nu \geq 1.$$

For $1 \leq \nu \leq cn$ the principal part of γ_ν^n is equal to $(-1)^\nu$ residual with respect to the pole $z = 1$: $\sqrt{\Sigma_1^{cn}|\gamma_\nu^n|^2}$ grows as $\sqrt{n}$ when $n \to \infty$. This is in agreement with the results of numerical experiments [2]. It is an example of an interesting phenomenon: the instability is observed in calculations only in the case of an even number (N_1) of full steps (h_1) on the overlap interval for the outflow problem.

5 Example II. Displacement gives the stability.

The Gary scheme with $\alpha = 2$, $N_1 = 2$, $N_2 = 1$, $\phi = 1/2$, $\beta_1 = 1/2$, $\beta_2 = 1/4$ is considered. This time there is a displacement and we get stability. u_0^{n+1} is found by the linear interpolation of v_2^{n+1}, v_1^{n+1} values. And v_0^{n+1} is found by the interpolation of u_2^{n+1}, u_3^{n+1} solution values in the inner points of G_1. The resultant is a polynomial of 924th degree. MAPLE determined and factored it in 20 min. approximately on PC486-66/16MB:

$$R(x) = 6947336878393197988213864912235291457122369204750211575216332097$$

$$79402989590170262629080651157247830044603970766828675594253762560 * x^{188} *$$

$$*R_2(x) * R_3(x) * (x^2+1)^{24} * (x^2+2x-1)^{96} * (4x^8 - 32x^7 + 89x^6 - 96x^5 + 152x^4 -$$

$$-96x^3 + 80x^2 - 32x + 4)^2 * (4x^8 + 32x^7 + 80x^6 + 96x^5 + 152x^4 + 96x^3 + 89x^2 + 32x + 4)^2 *$$

$$*(x^2 - 4x + 1)^8 * (21 + 88x + 80x^2 - 1080x^3 - 1524x^4 - 3960x^5 + 1072x^6 - 2792x^7 + 894x^8 +$$

$$+2792x^9 + 1072x^{10} + 3960x^{11} - 1524x^{12} + 1080x^{13} + 80x^{14} - 88x^{15} + 21x^{16})^{16}.$$

Here $R_2(x), R_3(x)$ are polynomials of 92 and 100 orders respectively with big integer coefficients. So we have 9 different factors of degrees not more than 100. These polynomials were solved in REDUCE by using the special patch of Stan Kameny. All solutions were analysed. The characteristic equations have not zero solutions. So only $x, |x| \leq 1$, are interesting. The resultant has 113 such solutions among 231 distinct solutions. We demonstrate below, how the real false spectrum points are separated. After the first step of separation only four sets were kept. They originate from two pairs of complex-conjugate solutions of the polynomial $R_3(x)$. We present sets corresponding to solutions with positive imaginary parts. The other two give complex-conjugite sets.

$$[z, \ x_1, \ \hat{x}_1, \ x_2, \ \hat{x}_2, \ x, \ y]$$

$$[-4.06441888993069 + 3.51812432083375 * i, 0.158325125400724 -$$

$$2.22203591880914 * i, -0.0554995523291884 - 0.0412152562155017 * i,$$

$$1.0 + 9.40653804848424E - 16 * i, 0.017064546580574 9 + 0.0126272932 92366 * i,$$

$$0.22492368095103 + 0.231563119097344 * i,$$

$$-0.00423895984976586 - 0.00307054241832905 * i],$$

$$[2.23261162700076 + 1.87010672569856 * i, -1.58409955347499 -$$

$$1.76688689793465 * i, -2.61811328317628 - 1.16678863932354 * i,$$

$$0.999999999999995 + 6.72205346941013E - 17 * i, 0.618777428864''25+$$

$$0.148165305765429 * i, 0.126226951078023 + 0.536255125209603 \cdot i,$$

$$-0.0945070704296629 - 0.00572603218332348 * i].$$

On the second step all sets were rejected. Let us consider, why the sets, pr sented above were rejected. The polynomial $P_1(\kappa, z)$ with $z = -4.06441888993069+3.51812432083375*i$ from the first set has the roots

$$\kappa_1 = -0.0155300629675805 - 0.379309688698518 * i,$$

$$\kappa_2 = -0.232113823750295 + 0.357542143667048 * i,$$

$$\kappa_3 = 0.282678861784193 + 0.0523781168748906 * i,$$

$$\kappa_4 = -3.42015856072792 + 0.633727841175572 * i,$$

$$\kappa_5 = 1.27736349829478 + 1.96761776633209 * i,$$

$$\kappa_6 = 0.107760087366824 - 2.63195617935108 * i.$$

PC found $\kappa_1 + \kappa_2 + \kappa_3 = 0.0350349750663175 + 0.0306105718434206 * i,$

$$\kappa_1 * (\kappa_2 + \kappa_3) + \kappa_2 * \kappa_3 = 0.0703603925019518 + 0.063366024929440 9 * i,$$

$$\kappa_1 * \kappa_2 * \kappa_3 = 0.0350349750663178 + 0.0306105718434205 * i.$$

These values don't equal x_1, x_2, x from the considered first set:

$$x_1 = 0.158325125400724 - 2.22203591880914 * i,$$

$$x_2 = 1.0 + 9.40653804848424E - 16 * i,$$

$$x = 0.22492368095103 + 0.231563119097344 * i.$$

But you can see, that the last ones are close to the values of

$$\kappa_2 + \kappa_3 + \kappa_6 = 0.158325125400722 - 2.22203591880914 * i,$$

$$\kappa_2(\kappa_3 + \kappa_6) + \kappa_3 * \kappa_6 = 0.999999999999998 + 1.16399945238044E - 15 * i,$$

$$\kappa_2 * \kappa_3 * \kappa_6 = 0.224923680951029 + 0.231563119097343 * i$$

respectively. As a matter of fact they coincide in 14 digits printed, which g uarantees the correctness of the results computed indepedently on the PC. The values z_1 in the sets were obtained at the end of long analytical and numerical calculations. So t ie considered z is a false spectrum point.

The polynomial $P_1(\kappa, 2.23261162700076 + 1.87010672569856 * i)$ has solutions

$$\kappa_1 = -0.56769819895633 + 0.180684005867448 * i,$$

$$\kappa_2 = 0.296965367383044 + 0.213386989838339 * i,$$

$$\kappa_3 = 0.20538455976233 - 0.337931042601437 * i,$$

$$\kappa_4 = -2.22075946537033 + 1.5957455970249 * i,$$

$$\kappa_5 = -1.31336672190171 - 2.16095789364045 * i,$$

$$\kappa_6 = 1.599474459083 + 0.509072343511198 * i.$$

The values of z_1, z_2, z from the second set are close respectively to

$$\kappa_1 + \kappa_2 + \kappa_5, \quad \kappa_1(\kappa_2 + \kappa_5) + \kappa_2 * \kappa_5, \quad \kappa_1 * \kappa_2 * \kappa_5.$$

They coincide in 14 digits printed again. So the considered z is a false spectrum point.

No spectrum point were found in $|z| \geq 1$. This and the stability of the corresponding Cauchy problems gives [5] the stability of the considered i.b.o. problem. Previous attempts to study the stability of the Gary scheme, under such nontrivial conditions, have been based on numerical approaches [2].

Preliminarily i.b.o. problem for the Rusanov scheme [6] had been searched [7]. The stability in L_2 and in uniform metric had been proved. The Rusanov scheme



has only five points on the low layer. Calculations here are not so complicated as in the case of the Gary scheme. The results of searching the Rusanov scheme had been reported at the 4th International Conference on Advance Numerical Analysis, August 15-18, 1995, Moscow, Russia.

References

[1] J.Gary, *On certain finite difference schemes for hyperbolic systems*, Math.Comp., Vol.18, 1-18 (1964).

[2] K.Otto and M.Thune. *Stability of a Runge-Kutta method for the Euler equations on a substructured domain*, SIAM J. Sci. Stat. Comput., Vol.10, No.1, 154-174 (1989).

[3] H.-O.Kreiss *Stability theory for mixed initial boundary value problems.I*, Math. Comp., Vol.22, No.104, 703-714(1968).

[4] . S.I.Serdyukova *On stability in C of difference boundary-value problems.*In:Vychisl. Protsessy i Sistemy, No.8, Nauka, Moscow, 292-327 (in Russian) (1991).

[5] B.Gustafsson, H.-O.Kreiss and A.Sundström *Stability theory of difference approximations for mixed initial boundary-value problems.II* Math. Comp. , V.26, No.119, 649-686(1972).

[6] V.V.Rusanov *Difference scheme of third-order accuracy for "across" computation of discontinuous solutions.* Fluid Dynam. Trans., V.4, 285-294(1969).

[7] S.I.Serdyukova, M.Thune *Studying the stability of difference problems on substructured domains* Preprint of JINR,E5-95-381, Dubna, 1995.

GRID APPROXIMATION OF THE SOLUTION AND DIFFUSION FLUX FOR SINGULARLY PERTURBED PARABOLIC EQUATIONS WITH DISCONTINUOUS COEFFICIENTS AND CONCENTRATED FACTORS [1]

G. I. Shishkin

Institute of Mathematics and Mechanics, Ural Branch of Russian Acad. Sci.
S. Kovalevskaya Str., 16, Ekaterinburg, 620219, Russia
e-mail: Grigorii@shishkin.ural.ru

Abstract. One-dimensional singularly perturbed parabolic equations, defined on composite type domains, with mixed boundary conditions are considered. The parameter ε^2 multiplying the highest derivative takes any values in the half-interval $(0, 1]$; for $\varepsilon = 0$, the reduce equation involves the time derivative only. Coefficients of the equation and the right-hand side have discontinuities of the first kind on the interfaces of the subdomains. These problems describe heat and mass transfer in inhomogeneous media and with complex exchange conditions on the interface of medium. By using special piecewise uniform grids which condense in the boundary and interior layers, we construct difference schemes that converge uniformly in ε. The approximations of normalized diffusion fluxes, i.e. products of the parameter ε and the first-order derivatives of the solution with respect to the space variable x, are studied.

1. Problem formulations

1. In the domain $\overline{G}$ with the boundary $S = \overline{G} \setminus G$ where

$$G = D \times (0, T], \quad D = \{x : d_* < x < d^*\}, \tag{1.1}$$

we consider a boundary value problem for the parabolic equation

$$L_{(1.2)} u(x,t) = f(x,t), \quad (x,t) \in G. \tag{1.2a}$$

Here

$$L_{(1.2)} \equiv \varepsilon^2 \left\{ \frac{\partial}{\partial x} a(x,t) \frac{\partial}{\partial x} + b(x,t) \frac{\partial}{\partial x} \right\} - c(x,t) - p(x,t) \frac{\partial}{\partial t}, \quad \varepsilon \in (0, 1],$$

the functions $a(x,t)$, $b(x,t)$, $c(x,t)$, $p(x,t)$, $f(x,t)$ are sufficiently smooth on the set $\overline{G}$, $a(x,t)$, $p(x,t) \geq \alpha_0 > 0$, $c(x,t) \geq 0$, $(x,t) \in \overline{G}$. The boundary S is represented as a sum of sets: $S = S^0 \cup S^D \cup S^N = S^0 \cup S^L$, where $S^0 = \{(x,t) : x \in \overline{D}, t = 0\}$ is the lower base of G. The lateral boundary consists of two parts: $S^L = S^D \cup S^N$; Dirichlet and Neumann boundary conditions are given on the sets S^D and S^N, respectively. Assume that the type of the boundary conditions does not vary in time. One of the sets S^D or S^N can be empty. The function $u(x,t)$ on S satisfies the boundary conditions

$$\varepsilon \frac{\partial}{\partial n} u(x,t) = \psi(x,t), \quad (x,t) \in S^N, \quad u(x,t) = \varphi(x,t), \quad (x,t) \in S^0 \cup S^D. \tag{1.2b}$$

[1] This work was supported in part by the Russian Foundation for Basic Research under Grant No. 95-01-00039

Here $\partial/\partial n$ is a derivative with respect to the outward normal to the set G, the functions $\varphi(x,t)$, $(x,t) \in S^0 \cup S^D$ and $\psi(x,t)$, $(x,t) \in S^N$ are sufficiently smooth on the set S^0, S^D and S^N respectively, $\varphi(x,t)$ is continuous on $S^0 \cup S^D$.

The solution of problem (1.2), (1.1) is regarded as a function $u \in C^{2,1}(G) \cap C(\overline{G}) \cap C^{1,0}(G \cup S^N)$, which satisfies the differential equation (1.2a) on G and the boundary conditions (1.2b) on S. For simplicity, we assume that the compatibility conditions [1] are fulfilled on the set $\gamma = S^0 \cap \overline{S}^L$ that ensure the sufficient smoothness of the solution on $\overline{G}$, in particular, $u \in C^{1,0}(\overline{G})$, for each fixed value of the parameter ε.

When the parameter ε tends to zero, parabolic boundary layers appear in the neighbourhood of the set S^L that are described by a parabolic equation.

We are interested in the function $u(x,t)$ and its gradient $(\partial/\partial x)u(x,t)$, $(x,t) \in \overline{G}$. Notice that the solution of problem (1.2), (1.1) depending on ε is bounded uniformly in the parameter ε. However, as $\varepsilon \to 0$, the derivative $(\partial/\partial x)u(x,t)$ increases unboundedly in the boundary layers. It is convenient to consider instead of $(\partial/\partial x)u(x,t)$ the function $P(x,t) = \varepsilon(\partial/\partial x)u(x,t)$ $(x,t) \in \overline{G}$, which is bounded uniformly in ε on the set $\overline{G}$. The function $P(x,t)$ is called the normalized diffusion flux or, in short, the flux.

We are to find the functions $u(x,t)$ and $P(x,t)$, $(x,t) \in \overline{G}$ for problem (1.2), (1.1).

2. We consider also the parabolic equation with discontinuous coefficients and right-hand side. Let the domain G be divided into several subdomains $G^{(j)}$, $j = 1, 2, \ldots, J$:

$$\overline{G} = \bigcup_j \overline{G}^{(j)}, \quad G^{(j)} = D^{(j)} \times (0,T], \quad D^{(j)} = \{x : \ d_{j-1} < x < d_j\}, \tag{1.3}$$

where $d_0 = d_*$, $d_J = d^*$, $d = d_{(1.3)} = \min_j (d_j - d_{j-1}) \geq m$; on each of $G^{(j)}$ we give the equation

$$L^{(j)}u(x,t) \equiv \left\{\varepsilon^2 \left\{\frac{\partial}{\partial x}(a^{(j)}(x,t)\frac{\partial}{\partial x}) + b^{(j)}(x,t)\frac{\partial}{\partial x}\right\} - \right. \tag{1.4a}$$

$$\left. - c^{(j)}(x,t) - p^{(j)}(x,t)\frac{\partial}{\partial t}\right\} u(x,t) = f^{(j)}(x,t), \quad (x,t) \in G^{(j)}.$$

On the every subdomain $\overline{G}^{(j)}$, the coefficients and right-hand side of the equation are sufficiently smooth, moreover, $a^{(j)}(x,t)$, $p^{(j)}(x,t) \geq \alpha_0 > 0$, $c^{(j)}(x,t) \geq 0$, $(x,t) \in \overline{G}^{(j)}$. On the boundary S the function $u(x,t)$ satisfies the boundary conditions

$$\varepsilon\frac{\partial}{\partial n}u(x,t) = \psi(x,t), \quad (x,t) \in S^N, \quad u(x,t) = \varphi(x,t), \quad (x,t) \in S^0 \cup S^D, \tag{1.4b}$$

and, on the boundaries $S^{[j]}$, $S^{[j]} = \{\overline{G}^{(j)} \cap \overline{G}^{(j+1)}\} \setminus S^0$ for the interface of the subdomains $G^{(j)}$, $G^{(j+1)}$, $j = 1, 2, \ldots, J-1$, it satisfies the conjugation conditions

$$\left[a(x,t)\frac{\partial}{\partial x}u(x,t)\right] = q(x,t), \quad [u(x,t)] = 0, \quad (x,t) \in S^*. \tag{1.4c}$$

Here $[v(x,t)]$, $(x,t) \in S^{[j]}$ is a jump which the function $v(x,t)$ has on the interface $S^{[j]}$; $[a(x,t)(\partial/\partial x)u(x,t)] = a^{(j+1)}(x+0,t)(\partial/\partial x)u(x+0,t) - a^{(j)}(x-0,t)(\partial/\partial x)u(x-0,t)$, $(x,t) \in S^{[j]}$; $S^* = \cup_j S^{[j]}$, $j = 1, 2, \ldots, J-1$. The functions $\varphi(x,t)$ and $\psi(x,t)$, $q(x,t)$ are

sufficiently smooth on the sets $S^0 \cap S^{(j)}$, S^D and S^N, S^* respectively, $S^{(j)} = \overline{G}^{(j)} \setminus G^{(j)}$, $\varphi(x,t)$ is continuous on $S^0 \cup S^D$. We suppose $G^* = G \setminus S^*$.

The solution of problem (1.4), (1.3) is regarded as $u \in C^{2,1}(G^*) \cap C(\overline{G}) \cap C^{1,0}(\overline{G}^{(j)} \setminus \{S^0 \cup S^D\})$, $j = 1, 2, \ldots, J$, which satisfies equation (1.4a) on $G^{(j)}$, boundary conditions (1.4b) on S and conjugation conditions (1.4c) on S^*. Notice that, for a fixed ε, $u \notin C^{2,1}(G)$; the derivative $(\partial/\partial x)u(x,t)$ has a discontinuity on the set S^*.

We suppose that, on the set $\gamma = S^0 \cap \{\overline{S}^L \cup \overline{S}^*\}$, the compatibility conditions are satisfied that ensure the sufficient smoothness of the solution on $\overline{G}^{(j)}$, in particular, $u \in C^{1,0}(\overline{G}^{(j)})$ for each value of the parameter ε. The solution of the reduced problem with $\varepsilon = 0$ is discontinuous on S^* and does not satisfy the conjugation conditions (1.4c), and conditions (1.4b) on S^L. When $\varepsilon \to 0$, parabolic boundary and interior layers appear in the neighbourhood of the sets S^L and S^*, respectively.

In the case of problem (1.4), (1.3) we are to find the functions $u(x,t)$, $(x,t) \in \overline{G}$ and $P^j(x,t) \equiv \varepsilon(\partial/\partial x)u(x,t)$, $(x,t) \in \overline{G}^{(j)}$, $j = 1, 2, \ldots, J$.

2. A priori estimates

By the analog of the maximum principle [1] and the technique of barrier functions [2], we establish the ε-uniform boundedness of the solution for problem (1.4), (1.3)[2]

$$| u(x,t) | \le M, \quad (x,t) \in \overline{G}. \tag{2.1}$$

To estimate the derivatives, we assume that the following condition is satisfied:

$$\text{either } a(x,t) = a(x),\ b(x,t) = b(x) \text{ or } a(x,t) = a(t),\ b(x,t) \equiv 0, \quad (x,t) \in \overline{G}. \tag{2.2}$$

Let us find the derivatives $(\partial^{k_0}/\partial t^{k_0})u(x,t)$ for $t = 0$, $k_0 = 1, 2, 3$ according to (1.4a). Under conditions imposed on the problem data, these derivatives are bounded ε-uniformly. Applying the operators $\partial^{k_0}/\partial t^{k_0}$, $k_0 = 1, 2, 3$ to the data of problem (1.4) for $t > 0$ and using the maximum principle, we find the estimates

$$| \frac{\partial^{k_0}}{\partial t^{k_0}} u(x,t) | \le M, \quad (x,t) \in \overline{G}, \quad k_0 \le 3. \tag{2.3}$$

The solution of problem (1.4) in each $\overline{G}^{(j)}$ can be represented as a sum

$$u(x,t) = U(x,t) + V(x,t), \quad (x,t) \in \overline{G}^{(j)}, \tag{2.4}$$

where $U(x,t)$ and $V(x,t)$ are the regular and singular parts of the solution. Using (2.3), we can get the estimates

$$| \frac{\partial^{k+k_0}}{\partial x^k \partial t^{k_0}} u(x,t) | \le M\varepsilon^{-k}, \quad | \frac{\partial^{k+k_0}}{\partial x^k \partial t^{k_0}} U(x,t) | \le M[1 + \varepsilon^{3-k}] \tag{2.5}$$

$$| \frac{\partial^{k+k_0}}{\partial x^k \partial t^{k_0}} V(x,t) | \le M\varepsilon^{-k} \exp(-m\varepsilon^{-1} r(x, \Gamma^{(j)})), \ (x,t) \in \overline{G}^{(j)},\ k + 2k_0 \le 6,\ j = 1, 2, \ldots, J$$

where $\Gamma^{(j)} = \overline{D}^{(j)} \setminus D^{(j)}$. Here we used the estimation technique like that in [3].

[2] We denote by M different positive (sufficiently large) constants which are independent of ε.

Theorem 1. *Let condition* (2.2) *be satisfied. Then, for the solution of problem* (1.4), (1.3) *and for its components from* (2.4), *estimates* (2.1) *and* (2.5) *are valid.*

Remark. We assume that the solution of problem (1.4), (1.3) satisfies the estimates of Theorem 1 even if condition (2.2) is violated.

3. Finite difference schemes

1. Classical difference scheme. On the set $\overline{G}_{(1.3)}$ we introduce the grid

$$\overline{G}_h = \overline{D}_h \times \overline{\omega}_0 = \overline{\omega}_1 \times \overline{\omega}_0, \tag{3.1}$$

where $\overline{\omega}_0$ is a uniform grid on $[0,T]$ with the step-size $\tau = TN_0^{-1}$; N_0+1 is the number of nodes in the grid $\overline{\omega}_0$. The grid $\overline{\omega}_1$ is assumed to be a grid on the interval $[d_*, d^*]$ with an arbitrary allocation of the nodes, and $d_j \in \overline{\omega}_1$, $j = 1,2,\ldots,J-1$; $N+1$ is the number of nodes in the grid $\overline{\omega}_1$. On the set $\overline{G}_h$, to problem (1.4) we assign the difference scheme

$$\begin{gathered}\Lambda_{(3.2)} z(x,t) = f(x,t), \quad (x,t) \in G_h^*, \\ \varepsilon[\hat{a}^h(x,t)\delta_x z(x,t) - a^h(x,t)\delta_{\overline{x}} z(x,t)] = q(x,t), \quad (x,t) \in S_h^*, \\ \varepsilon\delta_n z(x,t) = \psi(x,t), \ (x,t) \in S_h^N, \quad z(x,t) = \varphi(x,t), \ (x,t) \in S_h^0 \cup S_h^D.\end{gathered} \tag{3.2}$$

Here $G_h^* = G^* \cap \overline{G}_h$, $S_h^* = S^* \cap \overline{G}_h$, $\delta_n z(x,t) = \left\{ \begin{array}{c} -\delta_x z(x,t), \ x = d_*, \\ \delta_{\overline{x}} z(x,t), \ x = d^* \end{array} \right\}$, $(x,t) \in S_h^L$,

$$\Lambda_{(3.2)} \equiv \varepsilon^2 \left\{ \delta_{\hat{x}}(a^h(x,t)\delta_{\overline{x}}) + b^+(x,t)\delta_x + b^-(x,t)\delta_{\overline{x}} \right\} - c(x,t) - p(x,t)\delta_{\overline{t}}, \ (x,t) \in G_h^{(j)},$$

where $G_h^{(j)} = G^{(j)} \cap \overline{G}_h$, $b^+(x,t) = 2^{-1}(b(x,t) + |b(x,t)|)$, $b^-(x,t) = 2^{-1}(b(x,t) - |b(x,t)|)$, $\delta_{\hat{x}}(a^h(x,t)\,\delta_{\overline{x}} z(x,t)) = (a^h(x,t)\, z_{\overline{x}})_{\hat{x}}$, cf. [4], $a^h(x,t) = a(2^{-1}(x^{i-1}+x^i),t)$, $x = x^i$, $\hat{a}^h(x^i,t) = a^h(x^{i+1},t)$ for $(x^i,t), (x^{i+1},t) \in \overline{G}_h$. For the approximation of the diffusion flux $P^j(x,t)$, $(x,t) \in \overline{G}^{(j)}$ we use the grid function

$$P^{jh+}(x,t) \equiv \varepsilon\delta_x z(x,t), \quad (x,t) \in G_h^{(j)-}, \ j = 1,2,\ldots,J, \tag{3.3a}$$

where

$$G_h^{(j)-} = D_h^{(j)-} \times \overline{\omega}_0, \quad D_h^{(j)-} = \{\omega_1 \cap D^{(j)}\} \cup \{x = d_j\}. \tag{3.3b}$$

Taking into account the estimates of Theorem 1 and Remark, we establish the convergence of the solution of scheme (3.2), (3.1) and the functions $P^{jh+}(x,t)$ to the solution $u(x,t)$ and the functions $P^j(x,t)$, respectively, for a fixed value of the parameter

$$|u(x,t) - z(x,t)| \le M[\varepsilon^{-1}N^{-1} + N_0^{-1}], \quad (x,t) \in \overline{G}_h, \tag{3.4a}$$

$$|P^j(x,t) - P^{jh+}(x,t)| \le M\varepsilon^{-1}[\varepsilon^{-1}N^{-1} + N_0^{-1}], \quad (x,t) \in G_h^{(j)-}, \tag{3.4b}$$

$j = 1,2,\ldots,J$. To obtain (3.4b), we assume the fulfilment of the condition

$$b(x,t) \equiv 0, \ (x,t) \in \overline{G}; \quad a(x,t) = a(x) \ \text{ or } \ a(x,t) = a(t), \ (x,t) \in \overline{G}^{(j)}, \ j = 1,\ldots,J. \tag{3.5}$$

If condition (3.5) is violated, for the approximation of the fluxes $P^j(x\ t)$ we use the functions $P^{0jh+}(x,t)$

$$P^{0jh+}(x,t) \equiv \varepsilon\delta_x^0 z(x,t), \quad (x,t) \in G_h^{(j)}, \quad x \le d_j - h^0, \tag{3.6a}$$

where

$$\delta_x^0 z(x,t) = (h^0)^{-1}[\tilde{z}(x+h^0,t) - z(x,t)], \quad (x,t) \in \overline{G}_h^{(j)}, \quad x \le d_j - h^0; \tag{3.6b}$$

the function $\tilde{z}(x,t)$, $x \in \overline{D}$, $t \in \overline{\omega}_0$ is continuous in x and discrete in t, $\tilde{z}(x.t)$ is the linear interpolation in x from the grid function $z(x,t)$, $(x,t) \in \overline{G}_h$. We define h^0 by

$$h^0 = h^0(\varepsilon, \beta(N, N_0, \varepsilon)) = \min[4^{-1}d,\ M\varepsilon\beta^{1/2}(N, N_0, \varepsilon)], \tag{3.6c}$$

where $d = d_{(1.3)}$, M is any number, the function $\beta(N, N_0, \varepsilon)$ is the right-hanc side in (3.4a).

The modified flux $P^{0jh+}_{(3.6)}(x,t)$, $j = 1, 2, \ldots, J$ satisfies the estimate

$$|P^j(x,t) - P^{0jh+}(x,t)| \le M[\varepsilon^{-1}N^{-1} + N_0^{-1}]^{1/2}, \quad (x,t) \in G_h^{(j)}, \quad x \le d_j - h^0. \tag{3.7}$$

2. Special difference scheme. On the set $\overline{G}$ we construct a grid cor densing in the neighbourhood of the boundary and interior layers

$$\overline{G}_h = \overline{D}_h^* \times \overline{\omega}_0 = \overline{\omega}_1^* \times \overline{\omega}_0. \tag{3.8}$$

Here $\overline{\omega}_0 = \overline{\omega}_{0(3.1)}$, $\overline{\omega}_1^*$ is a piecewise uniform grid on $\overline{D}$. We divide the inte·val $\overline{D}$ into $3J$ parts: $[d_{j-1},\ d_{j-1} + \sigma]$, $[d_{j-1} + \sigma,\ d_j - \sigma]$, $[d_j - \sigma, d_j]$, $j = 1, 2, \ldots, J$. The step-size of the grid $\overline{\omega}_1^*$ equals $h^{(1)} = 4\sigma JN^{-1}$ on the intervals $[d_{j-1},\ d_{j-1} + \sigma]$, $[d_j - \sigma, d_j]$, and it is equal to $h^{(2)} = 2(d_j - d_{j-1} - 2\sigma)JN^{-1}$ on the interval $[d_{j-1} + \sigma,\ d_j - \sigma]$. The value σ is chosen to satisfy $\sigma = \min[4^{-1}d, m^{-1}\varepsilon \ln N]$, where $d = d_{(1.3)}$, m is an arbit·ary constant.

On the grid $\overline{G}_{h(3.8)}$ the solution of difference scheme (3.2) converges un formly in ε

$$|u(x,t) - z(x,t)| \le M[N^{-1}\ln N + N_0^{-1}], \quad (x.t) \in \overline{G}_h. \tag{3.9}$$

The modified flux $P^{0jh+}(x,t)$ converges ε-uniformly to the exact flux $P^j(x.t)$

$$|\ P^j(x,t) - P^{0jh+}(x,t)\ | \le\ M[N^{-1}\ln N + N_0^{-1}]^{1/2}, \quad (x,t) \in G_h^{(j)}, \quad x \le d_j - h^0, \tag{3.10}$$

$j = 1, 2, \ldots, J$. Here the set $G_h^{(j)}$ is generated by the grid $\overline{G}_{h(3.8)}$. $h^0 = h^0_{(3.6)}$ where the function $\beta = \beta(N, N_0)$ is the right-hand side in (3.9).

In the case of condition (3.5), for the computed flux $P^{jh+}(x,t)$ we get the estimate

$$|\ P^j(x,t) - P^{jh+}(x,t)\ | \le\ M[N^{-1}\ln N + N_0^{-1}], \quad (x.t) \in G_h^{(j)-}, \tag{3.11}$$

where $G_h^{(j)-} = G^{(j)-}_{h(3.3)}(\overline{G}_{h(3.8)})$, $j = 1, 2, \ldots, J$.

Theorem 2. *Let the solution of problem* (1.4), (1.3) *satisfy the estin ates of Theorem* 1 *and Remark. Then the solution of scheme* (3.2), (3.8) *(or scheme* (3 2), (3.1)*) and the function* $P^{0jh+}(x,t)$ *converge to the solution of the boundary value pr·blem and the function* $P^j(x,t)$ ε*-uniformly (for a fixed value of the parameter* ε*). Estimates* (3.9) *and* (3.10) *(or* (3.4a) *and* (3.7)*) are valid for the discrete solution and for the con puted flux. In the case of condition* (3.5), *estimates* (3.4b), (3.11) *are fulfilled for the function* $P^{jh+}(x,t)$.

All the considerations of sections 2, 3 can be easily modified for problen (1.2), (1.1).

References

[1] O. A. Ladyzhenskaya, V. A. Solonnikov and N. N. Ural'tseva, *Linear and Quasi-linear Equations of Parabolic Type*, Nauka, Moscow (1968) (in Russian).

[2] A. M. Il'in, A. S. Kalashnikov, O. A. Oleinik, Uspekhi matem. nauk, **17** (3), 3–146 (1962) (in Russian).

[3] G. I. Shishkin, *Grid Approximations of Singularly Perturbed Elliptic and Parabolic Equations* Ural Branch of Russian Acad. Sci., Ekaterinburg (1992) (in Russian).

[4] A. A. Samarskii, *Theory of Difference Schemes*, Nauka, Moscow (1989) (in Russian).

Pixel Decimation Motion Estimation using Wavelet Transform

T.M.Y.Wong and J.F.Zhang
School of Electrical Engineering, Electronic and Physics, Liverpool John Moores University, Byrom Street, Liverpool, UK L3 3AF

Abstract

Traditional exhaustive Block Matching motion estimation (BMME) is difficult to achieve in real time. Several methods had been proposed in order to simplify the process while at the same time maintaining good matching performance. Uniform pixel decimation [1] is one of the methods. The disadvantage of it is some crucial pixels in a block may not be selected for the motion matching criterion, resulting in inaccurate motion vectors. To overcome this problem, Chan and Siu [2] proposed an adaptive pixel decimation method. However, this method still contains inactive pixels at the beginning of the pixel selection stage. It also employs different number of pixels in each block, which makes it difficult for hardware implementation. In this paper, we use Wavelet Transform to extract the high activity pixels and adaptively select a fixed number of the most representative pixels for the matching criterion. The simulation result shows that the performance of this adaptive pixel decimation is close to that by the traditional exhaustive search with a significant reduction in computation.

Introduction

BMME is widely used in video coding. Exhaustive full search BMME (FSBMME) is one of the most popular methods because of its optimal solution and low control overhead. The basic idea of FSBMME is to search all the possible displaced candidate blocks within the search area in the previous frame to find the best match block using either Mean Square Difference (MSD) or Mean Absolute Difference (MAD) matching criterion. Obviously, the FSBMME requires a large amount of computations and cannot be performed in real time by a single processor. To reduce the computation time, Koga et al [1] proposed to use a uniform subsampled pixels for the matching. However, the matching is mainly controlled by the highly active pixels, such as those representing edges and textures. These active pixels may not be selected during uniform subsampling, which leads to a poor accuracy of the motion vectors.

To overcome this problem, adaptive pixel decimation had been proposed by Chan and Siu[2]. However, their method still contains inactive pixels at the beginning of the pixel selection stage. Their method also uses different number of pixels in each block, which increases the difficulty in hardware implementation. In this paper, we propose to detect the active pixels by means of wavelet transform. S.Mallat [3] proved that for particular class of wavelets, the zero crossing or the local extrema of a wavelet transform provided the location of the signal sharp variation points (edges or textures).

High variation point extraction by wavelet transform

If a signal f(x) is smoothed (low passed), the extrema of their first derivative and the zero crossings of the second derivative provide the location of high variation points. A smoothing function is defined here as a function, $\theta(x)$, whose integral is equal to 1 and that converges to 0 at infinity. For example, one can choose $\theta(x)$ equal to a Gaussian. Supposing this smoothing function is twice differentiable, we can define, $\psi^a(x)$ and $\psi^b(x)$ as the first and second order derivative of $\theta(x)$, respectively,

$$\psi^a(x) = \frac{d\theta(x)}{dx} \qquad \psi^b(x) = \frac{d^2\theta(x)}{dx^2} \tag{1}$$

These two functions can be considered to be wavelets because their integral is equal to 0.

The wavelet transform of f(x) at the scale s and position x, computed with respect to the wavelet $\psi^a(x)$, is

$$W_s^a f(x) = f * \psi_s^a(x) \tag{2}$$

and the wavelet transform of f(x) with respect to $\psi^b(x)$ is

$$W_s^b f(x) = f * \psi_s^b(x) \tag{3}$$

The wavelet transform $W_s^a f(x)$ and $W_s^b f(x)$ are, respectively, the first and second derivative of the signal smoothed at the scale s. The local extrema of $W_s^a f(x)$ thus corresponding to the zero crossings of $W_s^b f(x)$ and to the inflection points of $f * \theta_s(x)$. In particular case where θ(x) is a Gaussian, the zero-crossing detection is equivalent to a Marr-Hildreth [4] edge detection, whereas the extrema detection corresponds to a Canny [5] edge detection. And generally, the local extrema method is preferable to the zero-crossing due to the simplicity of implementation. It also generates more information about the edge, such as if the edge is rising or falling. The zero crossing method only tells whether the pixels are edges or not. Therefore, we chose the local extrema method for the high activity pixel selection.

In this paper, we chose the Daubechies' 4 tap wavelet filter to extract the edge information. This filter is better than the others in the same Daubechies' filter family due to its compacted format. It also generates less ringing effect on the edge map. Since images are two dimensional, equation (2) has to be applied on both the horizontal and vertical directions in order to obtain the final edge map as shown in Fig.1. W_{s1} represents the filter in horizontal direction, while W_{s2} in the vertical direction. After filtering, all the output coefficients will be retained without subsampling, i.e. s=0.

By adding these two edge maps together and taking the absolute values of the coefficients, we obtained the desired edge map. Thresholding will then be applied on the coefficients to remove the unwanted noise. A typical edge map for the image of Fig.2 is shown in Fig. 3.

Adaptive Pixel Decimation

In order to select the pixels with highest activity for motion estimation, the coefficients of each block from the edge map are sorted in descending order. The high value coefficients represents highly active pixels. Depending on the requirement of computational reduction, a certain number of pixels (with the highest activity within a block) will be selected to enter the MAD matching criterion. For instance, a four to one decimation, which gives a 75% reduction in computation time, for a 16x16 block, results in only 64 pixels for motion estimation (Fig. 4).

Computational Complexity

The computing time for detecting the activity pixels is negligible as compared to the exhaustive full search process. The average number of operations for motion estimation using the adaptive pixel decimation per block (F) is,

$$F = \frac{SKN^2}{PL} \sum_{x=0}^{\frac{P}{N}-1} \sum_{y=0}^{\frac{L}{N}-1} H(x,y)$$

S=Number of location searched (Searching window size)
K=MAD matching operation
N=Square block width
P=Width of the image
L=Height of the image
H(x,y)=Number of pixels selected at block location (x,y) in the frame

The computational complexity of the adaptive pixel decimation is determined by the following factors:

- Searching window size (S)
- Image size (P & L)
- Square block width (N)
- Number of pixels selected for each MAD matching criterion (H)

Results and Discussions

A series of experiments were carried out to evaluate the performance of the adaptive pixel decimation technique. The Table Tennis (720x486) video sequence is used as the test sequence and it follows the MPEG 2-frame interpolation structure (i.e. IBBPBBI...). The public domain PVRG-MPEG Codec 1.1 package has been adapted. Fig.5 and Fig.6 show the results of B frames. It can be seen that the adaptive wavelet pixel decimation produces better peak signal to noise ratio and, it employed less bits to encode the sequence at the same time. Fig. 7 shows that the Displacement Frame Difference (DFD) of our method is always less than the uniform decimation, which means more accurate motion estimation. The results show the performance of the this motion estimation technique is close to that of the traditional exhaustive full search method, while the computational time has been reduced by 75%.

Conclusions

Wavelet Transform has been successfully used to detect the active pixels in a block. These active pixels are then selected for motion estimation. Results show that this new pixel decimation technique is superior to the uniform decimation, in terms of the peak signal to noise ratio, bits per pixel and the DFD values. Its performance is close to the traditional exhaustive full search method, with a 75 % reduction in the computation time.

Acknowledgment

This project is supported by the Higher Education Funding Council of England and T.M.Y. is grateful for a studentship.

References:

[1] T.Koga, K.Iinuma, A.Hirano, Y.Iijima, and T.Ishiguro, "Motion Compensated Interframe Coding for Video Conferencing", in Proc. Nat. Telecommun. Conf., Nov. 29 - Dec. 3, 1981, pp.G.5.3.1-5.3.5

[2] Y.L.Chan and W.C.Siu, "A New Block Motion Vector Estimation using Adaptive Pixel Decimation", 1995 IEEE 0-7803-2431-5/95, pp2257-pp.2260

[3] S.Mallat, "Zero-Crossings of a Wavelet Transform", IEEE Transaction on Information Theory, Vol. 37, No. 4, July 1991, pp.1019-pp.1033

[4] S.Mallat and S.Zhong, "Characterization of Signals from Multiscale Edges", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 14, No. 7, July 1992, pp.710-pp.732

[5] J.Canny, "A Computational Approach to Edge Detection", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. PAMI-8, No. 6, Nov., 1986, pp679-698

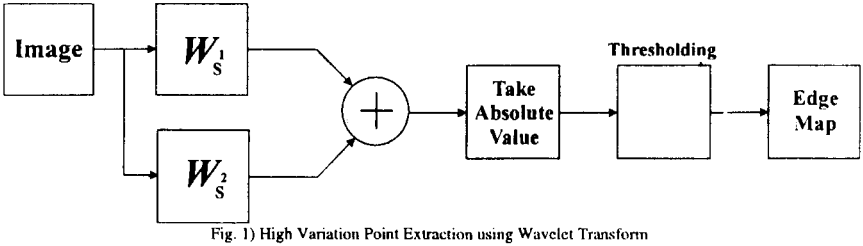


Fig. 1) High Variation Point Extraction using Wavelet Transform

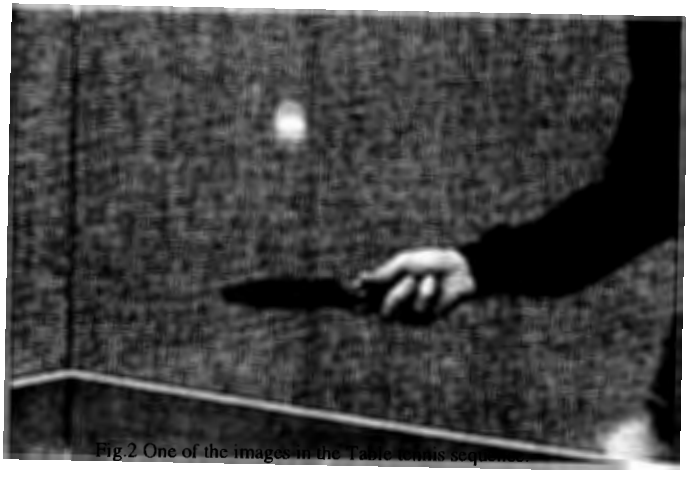
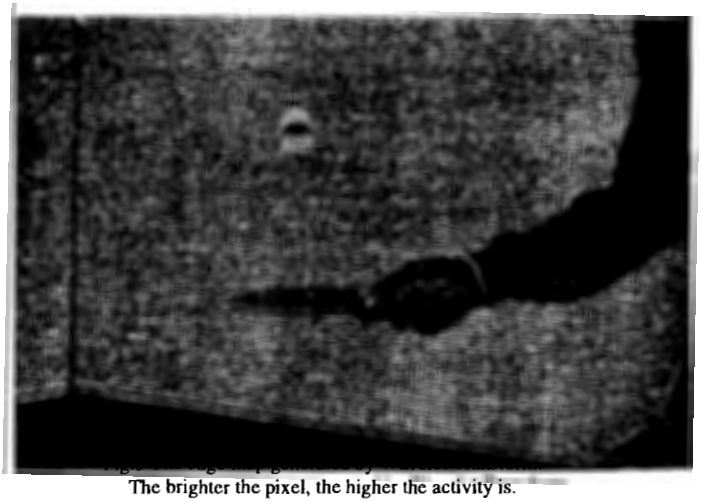


Fig.2 One of the images in the Table tennis sequence.



The brighter the pixel, the higher the activity is.

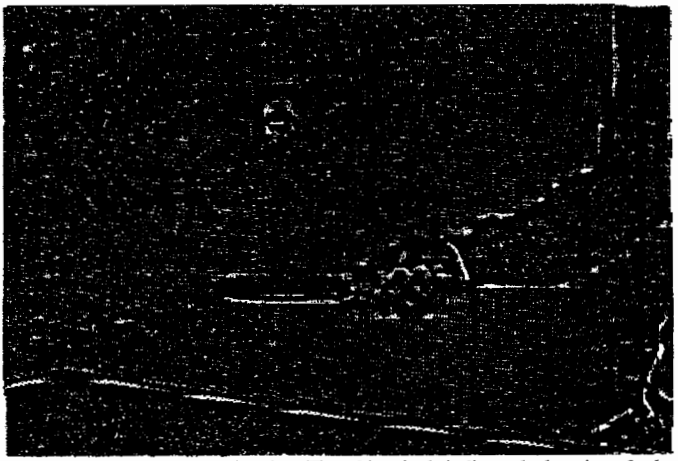


Fig.4 The selected pixels for motion estimation. The white pixels indicate the location of selected pixels.

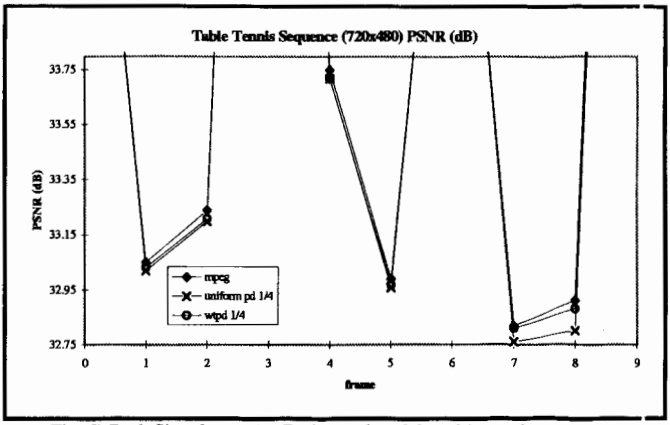


Fig. 5) Peak Signal to Noise Ratio results of the table tennis sequence.
mpeg- MPEG-1 exhaustive full search with one pixel accuracy motion estimation
uniform pd ¼ - Uniform four to one pixel decimation with one pixel accuracy motion estimation
wtpd ¼ - Adaptive four to one Wavelet pixel decimation with one pixel accuracy motion estimation

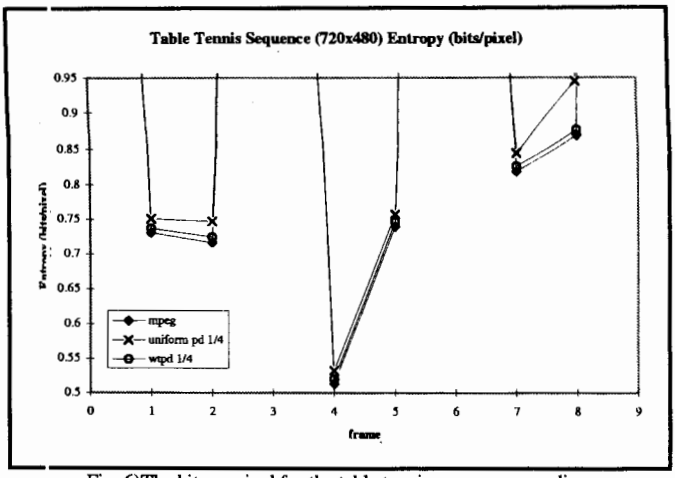


Fig. 6)The bit per pixel for the table tennis sequence encoding.

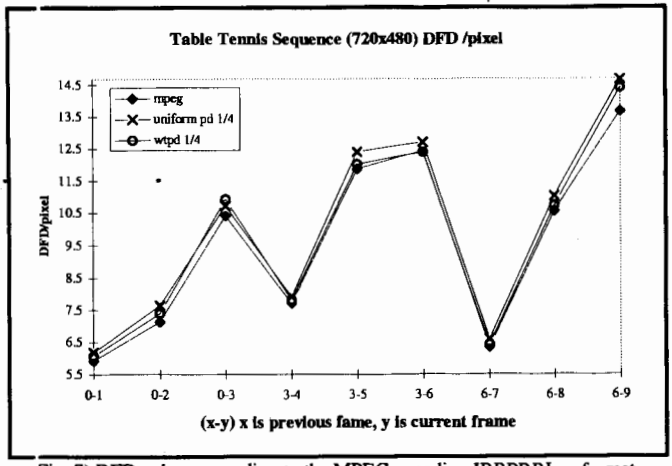


Fig. 7) DFD values according to the MPEG encoding IBBPBBI.... format.

Execution Time Analysis for Least Squares Problems on Massively Parallel Distributed Memory Computers

Tianruo Yang

Department of Computer Science, Linköping University
S-581 83, Linköping, Sweden

Abstract. In this paper we mainly focus on the study of the parallelization of PCGLS, a basic iterative method whose main idea is to organize the computation of conjugate gradient method with preconditioner applied to normal equations. Based on the data distribution model, we analyze fully the most suitable communication network topology for solving least squares problems on massively distributed memory computers. A theoretical model of communication phases is presented which allows us to give a detail execution time complexity analysis and investigates its usefulness. It is shown that the implementation of PCGLS, with a row-block decomposition of the coefficient matrix, on a ring of communication structure is the most efficient choice. Performance tests of the developed parallel PCGLS algorithm have been carried out on the massively distributed memory system ParsytecGC/PowerPlus and experimental timing results are compared with the theoretical execution time complexity analysis.

1 Introduction

Many scientific applications such as tomography, geodetic survey and structural analysis give rise to the least squares problems

$$\min_x \| Ax - b \|_2, \tag{1}$$

with a large and sparse coefficient matrix A. Minimizing by solving the normal equations is a common and often efficient approach, because $A^T A$ is symmetric and positive definite and it can be solved by using conjugate gradient method with preconditioner which was developed in the early 1950's by Hestenes and Stiefel [8]. The resulting method, PCGLS, is used as the basic iterative method to solve the least squares problems. Although this method does not explicitly form the normal equations, the rate of convergence still relies on the spectrum of $A^T A$. Preconditioning techniques that accelerate the convergence of these methods have received extensive attention in the literature.

In many situations, especially when matrix operations are well-structured, these operations are suitable for implementation on vector and share memory parallel computers [6, 5]. But for parallel distributed memory machines the picture is entirely different. In general the matrices and the vectors are distributed over the processors, so that even when the matrix operations can be implemented efficiently by parallel operations, we can not avoid the communication required for information exchange. About discussions on how to reduce the communication costs on distributed memory systems, see [2, 3, 4, 13]. In a word, these communication costs become relatively more and more important when the number of the parallel processors is increased and thus they have the potential to affect the scalability of the algorithm in a negative way [3].

We are also highly concerned with preconditioners for massively parallel distributed memory computers because the preconditioning part is often the most problematic part in a parallel environment. The preconditioning techniques have been studied by many authors. In our paper we will focus on the Incomplete Modified Gram-Schmidt(IMGS) preconditioner, a class of incomplete QR factorization preconditioners, because their existences can be guaranteed [15].

Here, we present the functional aspects of parallelizing the PCGLS method for solving least squares problems. Parallelizing a problem basically implies answering the following questions: what is the best decomposition for problem or data distribution and which communication scheme is best suited. Since the first one has been analyzed in [18], we will foucs on the second issue.

A suitable metric is needed to decide which data distribution and communication scheme should be selected. We will follow the definition and the abstract model of the parallel computing system proposed in [10, 11]. We also choose the total execution time of the parallel program T_{par} as our metric. The decomposition and interconnection that minimize the total execution-time of the parallel program T_{par} must be chosen. For data parallelism, the total execution time T_{par} depends on two variables and system parameters

$$T_{par} \equiv T_{par}(p, n, system\ parameters),$$

where p is the number of processing elements and n is the data size.

In general the communication is described by a startup time $\tau_{startup}$ plus a pure sending time $n\tau_{comm}$, where n is the number of bytes that is sent, and τ_{comm} the time to send one byte. Now we restrict ourselves first that our parallel computing system is Hoare's Communicating Sequential Processes [9]. In other words, this means that the system consists of a set of sequential processes, running in parallel and exchanging data by means of synchronous point to point communication. This kind of model needs to be specified how the parallel processors are connected, the time needed to send a certain amount of data over a link connecting two processors, and execution times on a single processor. We will describe our platform Parsytec GC/PowerPlus's communication model later which suits our assumption well. The next assumption is that the action within the sequential processes mainly consists of manipulating floating point numbers, and that the time needed to perform one floating point operation can be described with a single parameter τ_{calc}. So based on above assumptions, T_{par} can be specified as follows:

$$T_{par} \equiv T_{par}(p, n, \tau_{calc}, \tau_{startup}, \tau_{comm}, topology).$$

Finally we assume that the parallel program consists of a sequence of cycle [1]. During a cycle all processors first perform some computations and after finishing the computation they all synchronize and exchange data. This program execution model is valid if parallelism is obtained through data decomposition. With the final assumption, T_{par} can be expressed as follows:

$$T_{par} = \frac{T_{seq}}{p} + T_{calc.np} + T_{comm},$$

where $T_{seq} \equiv T_{seq}(p = 1, n)$, is the execution time of the sequential algorithm or the execution time of the parallel algorithm on 1 processor; $T_{calc.np} \equiv T_{calc.np}(p, n, \tau_{calc}, topology)$,

is the part that can not be performed completely in parallel; $T_{comm} \equiv T_{comm}(p, n, \tau_{startup}, \tau_{comm}, topology)$ is the total communication time of all cycles of the paralle program.

In this paper, we will focus on how to minimize T_{par} as a function of p and n and decide the corresponding data distribution and communication network topology combination for PCGLS method. Here we assume throughout that no preconditioner is used to verify the effect different parallelization strategies do.

Several timing experiments are carried out on Parsytec GC/PowerPlus at the National Supercomputing Center, Linköping University, in order to verify the validity of our theoretical expectations. The Parsytec GC/PowerPlus consists of 64 nodes, Each node contains two PowerPC601, four T805 transputers and 32Mbyte memory. The machine can be partitioned so that users can have their partitions when running a program. In addition to this there is a parallel file system for input and output. Parix provides a set of easy to use functions to build and use a communication network matching our problem and define suitable routines manage data operation. This enables us to write efficient, reliable, scalable programs for a large variety of applications. Parix's partition management, the NRM software, assures flexibility and versatility in using a Parsytec GC/PowerPlus's capabilities. Depending on your request you are granted sets of processors according to your requirement. The basic communication facility of Parix are virtual links. A virtual link is a bidirection, synchronizing, non-buffering, point to point communication line between two threads on different or same processor. The links are called 'virtual' because there is no need for a direct hardware connection or link between the communication processors. Communication is modeled based on virtual topologies, which build point to point connections between arbitrary threads. A well defined set of virtual links can be combined to build a virtual topology such as grid, tree, ring. We will introduce the communication model in detail in our experimental part.

The left of paper is organized as follows. In section 2, we will present shortly our introduction for the PCGLS method with IMGS preconditioner. And an abstract model for data distribution or decomposition and communication schemes will be described shortly based on [18] in section 3. In section 4, the comparison of actual implementation and theoretical total execution time complexity analysis will be described fully. Finally, performance tests of the developed parallel PCGLS have been carried out and experimental timing result are compared with the theoretical time execution analysis.

2 PCGLS with IMGS Preconditioner

PCGLS [16] has been used widely for large and sparse least squares problems whose main idea to organize the computation of conjugate gradient method with preconditioner applied to the normal equations.

The PCGLS algorithm contains three distinct computational tasks per iteration namely, two matrix vector products, At_k and $A^T r_{k+1}$, where $T_{seq}^{m-v} = 8n\tau_{calc}$; two inner products, (q_k, q_k) and (s_{k+1}, s_{k+1}), where $T_{seq}^{inn} = (8n-2)\tau_{calc}$; three vector updates, x_{k+1}, r_{k+1} and p_{k+1}, where $T_{seq}^{v-u} = (8n^2 - 2n)\tau_{calc}$.

IMGS, is an incomplete modified version of Gram-Schmidt orthogonalization to obtain an incomplete orthogonal factorization preconditioner $M = \bar{M}$, where $A = \bar{Q}\bar{R} + E$ is an approximation of QR factorization, $\bar{Q}$ is an orthogonal matrix and $\bar{R}$ is a upper triangular

matrix respectively. The main advantage of the preconditioner is that its existence can be fully guaranteed [16]. The IMGS preconditioning procedure requires approximately $\frac{1}{4}(n^2+5n)$ inner products, and $\frac{1}{2}(n^2+n)$ vector updates, and n matrix-vector multiplications [17].

3 Data Distribution and Communication Schemes

In [18], three different data decompositions of the matrix have been discussed: row-block decomposition, column-block decomposition and grid decomposition. First of all, we will present briefly the basic parallel vector operations and their execution time complexity based on [10, 1 ., 18].

3.1 Vector update operations

The parallel vector update operations can be performed in parallel. The computing time for the parallel vector update operations

$$T_{par}^{v-u}(p,n) = 8\lceil\frac{n}{p}\rceil\tau_{calc} = \frac{T_{seq}^{v-u}(n)}{p} + 8(\lceil\frac{n}{p}\rceil - \frac{n}{p})\tau_{calc}, \tag{2}$$

where $\lceil x\rceil$ is the ceiling function of x. It does not have to communicate data, so $T_{comm} = 0$. And non-parallel part is a pure load balancing effect. From (2), we can know that

$$T_{np.calc}^{v-u}(p,n) = 8(\lceil\frac{n}{p}\rceil - \frac{n}{p})\tau_{calc}, \tag{3}$$

if the length of the vector is not divisible by the number of processor p, then some processors receive pieces of the vector with size $\lceil\frac{n}{p}\rceil$, and some with size $\lceil\frac{n}{p}\rceil - 1$. This will lead to a load imbalance, of which the effect is presented by (3). In general, for large grain size, $n/pT_{np.calc}$ is very small compared with T_{seq}/p.

3.2 Inner product operations

The total computing time for the parallel inner production operations is

$$\begin{aligned} T_{par}^{inn}(p,n) &= (8\lceil\frac{n}{p}\rceil - 2)\tau_{calc} + t_{ca.calc}^{inn} + t_{ca} \\ &= \frac{T_{seq}^{inn}(n)}{p} + (t_{ca.calc}^{inn} - 2(1-\frac{1}{p})\tau_{calc}) + 8(\lceil\frac{n}{p}\rceil - \frac{n}{p})\tau_{calc} + t_{ca}, \end{aligned}$$

where t_{ca} is the time for a complex accumulate, the total communication time to send the complex partial inner products resident on each processor to all processors. The $t_{ca.np}$ is a computing time introduced by summing the partial inner products after or during the scalar accumulation. It is quite easy to know that

$$T_{comm}^{inn}(p,n) = t_{ca}, \qquad T_{calc.np}^{inn} = (t_{ca.calc}^{inn} - 2(1-\frac{1}{p})\tau_{calc}) + 8(\lceil\frac{n}{p}\rceil - \frac{n}{p})\tau_{calc}.$$

This part consists of two terms. The second term is the load imbalance term and the first term describes the summations in the parallel vector inner product that can not be performed completely in parallel.

3.3 Matrix vector multiplication operations

The execution times for vector updates and inner products are independent of the decomposition of the coefficient matrix. But the matrix vector multiplication operations are the only matrix dependent parts of the algorithm. Full study has been conducted in [18]. Here Table 1 summarizes T_{comm} and T_{np} of the matrix vector product for the three decompositions, namely row-block, column-block and grid decompositions, where T_{np} is further subdivided in the load balance and the other non-parallel floating-point operations. Here

Table 1: T_{np} and T_{comm} of the matrix vector product for different decompositions

Decomposition	T_{np} Load balance	T_{np} other non-parallel floating-point operations	T_{comm}
Row-block	$(8n-2)(\lceil \frac{n}{p} \rceil)\tau_{calc}$	none	t_{v-g}
Column-block	$8n(\lceil \frac{n}{p} \rceil - \frac{n}{p})\tau_{calc}$	$t_{va.calc} - 2n(1-\frac{1}{p})\tau_{calc}$	t_{va}
Grid	$8(\lceil \frac{n}{p} \rceil^2 - \frac{n^2}{p})\tau_{calc}$	$t_{pva.calc} - 2n(\lceil \frac{n}{\sqrt{p}} \rceil \frac{1}{n} - \frac{1}{p})\tau_{calc}$	$t_{v-g} + t_{pva}$

t_{v-g} is the time needed for the vector gather operation in the row-block decomposition. t_{va} is the time to accumulate and scatter the resulting vector, and $t_{va.calc}$ is the time to evaluate the partial sums in the column-block decomposition. t_{p-g} is the time for the partial vector gather operation, t_{pva} is the time for the partial vector accumulate, and $t_{pva.calc}$ the time to evaluate the partial sums after or during the vector accumulation in the grid decomposition.

Based on the theoretical analysis, it is quite clear that the row-block decomposition is much more efficient than the column and grid decompositions of A for any parallel computers.

4 Theoretical Time Complexity Analysis

We will present the execution time complexity analysis of the row-block decomposition on a ring. We also have analyzed the row-block decomposition on a binary tree and a torus. The total time on the torus is higher than on the ring and the binary tree topologies. The last two are comparable. Furthermore we will focus on the grid decomposition on a torus topology.

4.1 The row-block decomposition on ring topology

The row-block decomposition gives rise to a vector operation with associated overhead computations. During the vector gather operation, each processor receives from all other processors in the network a piece of the vector, which is subsequently stored. After the gather, every processor has a copy of the total vector in its local memory. On the ring topology it can be achieved as follows:

1. First step each processor sends its private part of the vector to the left and the right and, at the same time, receives from the left and right processor their private part.

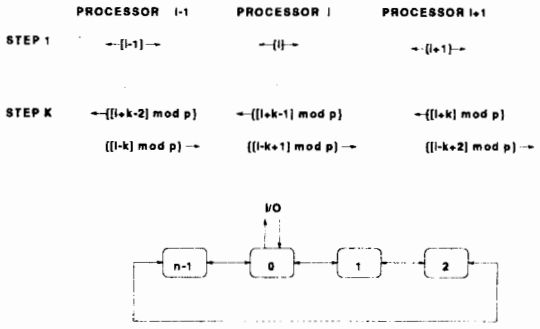


Figure 1: The vector gather operation on the ring topology

2. Next, the parts received in the previous step, are passed on from left to right and vice versa, and in parallel, parts from left and right are received.

So, after $\lfloor p/2 \rfloor$ steps, each processor in the ring receive the total vector where $\lfloor x \rfloor$ is the floor function of x.

If we assume that point to point communication of n floating points takes a time

$$t_{point-to-point} = \tau_{startup} + n\tau_{comm}, \tag{4}$$

then $t_{v-g}(p, N)$ can expressed as follows:

$$t_{v-g}^{ring}(p, N) = \lfloor \frac{p}{2} \rfloor (\tau_{startup} + 16\frac{N}{P}\tau_{comm}) = \lfloor \frac{p}{2} \rfloor \tau_{startup} + \lfloor \frac{p}{2} \rfloor \frac{16N}{p}\tau_{comm}.$$

During every communication step each processor simultaneously sends 2 messages and receives 2 messages of N/P complex numbers. A total of $\lfloor p/2 \rfloor$ steps are performed.

In Figure 1, each processor simultaneously sends to and receives from both neighboring processors a part of the vector. The $\rightarrow$ symbol gives the communication direction, the number between the brackets is the processor number where the package originally comes from.

The time for the scalar accumulate is easily derived from the previous analysis. Instead of a complex vector of N/P elements, a real number is sent, the communication pattern is the same as the vector gather. After each communication step, the newly received partial inner products are added to the partial inner products of the receiving processor. In total p real numbers are added, This leads to the following expressions for t_{ca} and $t_{ca.np}$

$$t_{ca}^{ring} = \lfloor \frac{p}{2} \rfloor (\tau_{startup} + 16\tau_{comm}), \qquad t_{ca.calc}^{ring} = 2(p-1)\tau_{calc}.$$

4.2 The grid decomposition on a torus topology

The parallel PCGLS with grid decomposition of the matrix contains a partial vector gather operation, a partial vector accumulate operation, and a complex accumulate operation. Derivation of the execution time of these operations on a torus topology is not easy.

We will analyze the partial vector gather operation first. The partia vector gather operation is implemented in this way, consists of $p^{1/2}$ simultaneous vecto gathers along the columns of torus topology. Every processor in the column must receive a package of $\lfloor n/p \rfloor$ complex words of every other processor in the column.

The vector accumulate can be separated into two communication steps The first step is a communication in the horizontal direction. The vector, which is sum decomposed in the horizontal direction is accumulated in the diagonal processors of the torus. The vector is now decomposed over the diagonal of the torus in parts of $\lceil n/(p^{1/2}) \rceil$ elements. These parts are scattered in the second step in vertical direction, resulting in the column wise scattering of the original vector. The first step implies that every processor on the diagonal must receive packages of $\lceil n/(p^{1/2}) \rceil$ complex numbers from every processor in the row of the grid. In the second step, the diagonal processors must send packages of $\lceil n/p \rceil$ complex words to all processors in their columns.

Here, we note that both in the partial vector gather and in the vector accumulate the grid can be viewed as $p^{1/2}$ independently operating pipelines or rings containing $p^{1/2}$ processors. In the first step of the vector accumulate the largest packages must be sent. Therefore it is most advantageous to create rings in row wise direction.

The partial vector gather is a gather operation of packages of $\lceil n/p \rceil$ complex words on $p^{1/2}$ simultaneously operating bidirectional pipeline each containing $p^{1/2}$ processors. This can be implemented in the same way as the vector gather operation on the ring, but now a total of $p^{1/2}-1$ steps are required to get the messages from every processor in the pipeline to every other processor. So the execution time of the partial vector gather operation on the torus topology is

$$t_{p-g}^{torus} = (\sqrt{p}-1)(\tau_{startup} + 16\lfloor \frac{n}{p} \rfloor \tau_{comm}).$$

The vector accumulate operation is an accumulation to the processors on the diagonal of the torus of a sum decomposed vector of $\lceil n/(p^{1/2}) \rceil$ complex words on $p^{1/2}$ simultaneously operating bidirectional rings of $p^{1/2}$ processors, followed by scattering of packages of $\lceil n/p \rceil$ complex words from a processor to all other processors in $p^{1/2}$ simultaneously operating pipeline containing $p^{1/2}$ processors, therefore

$$t_{pva}^{torus} = \lfloor \frac{\sqrt{p}}{2} \rfloor (\tau_{startup} + 16\lceil \frac{n}{\sqrt{p}} \rceil \tau_{comm}) + (\sqrt{p}-1)(\tau_{startup} + 16\lceil \frac{n}{p} \rceil \tau_{comm}).$$

The first step of the partial vector accumulate operation introduces the computations that should be accounted for by $t_{pva.calc}$. After every communication step the received vectors of $\lceil n/p^{1/2} \rceil$ complex words are added to the partially accumulated vector in the diagonal processor. A total of $p^{1/2}$ vectors must be added, therefore

$$t_{pva.calc}^{torus} = 2(\sqrt{p}-1)\lceil \frac{n}{\sqrt{p}} \rceil \tau_{calc}.$$

The complex accumulate is quite simple, the partial inner products are the first accumulated in the horizontal direction, followed by an accumulation in the vertical direction. In both sweeps $p^{1/2}$ complex number are added. So, this leads to the following expressions

$$t_{ca}^{torus} = \lfloor \frac{\sqrt{p}}{2} \rfloor (\tau_{startup} + 16\tau_{comm}) + (\sqrt{p}-1)(\tau_{startup} + 16\tau_{comm}),$$

$$t_{ca.calc}^{torus} = 4(\sqrt{p}-1)\tau_{calc}.$$

The derivation of the expression for the communication and non-parallel routines, as a function of the system parameters, allows us to compare T_{par} of the PCGLS with IMGS preconditioner for two different parallelization strategies.

Here we will compare T_{np} and T_{comm} of both parallelization strategies in some limited cases. From the previous parts, we know that

$$T_{np}^{PCGLS.IMGS} = \frac{1}{4}(n^2+5n)T_{np}^{inn} + \frac{1}{2}(n^2+n)T_{np}^{v-u} + nT_{np}^{m-v}, \tag{5}$$

and

$$T_{comm}^{PCGLS.IMGS} = \frac{1}{4}(n^2+5n)T_{comm}^{inn} + \frac{1}{2}(n^2+n)T_{comm}^{v-u} + nT_{comm}^{m-v}. \tag{6}$$

Combining the expression of the previous section with (5) and (6), we get

$$(T_{np}^{PCGLS.IMGS})_{ring}^{row-block} = (14n^2+12n)(\lceil\frac{n}{p}\rceil - \frac{n}{p})\tau_{calc} + \frac{1}{2}(n^2+5n)(p+\frac{1}{p}-2)\tau_{calc}, \tag{7}$$

$$(T_{comm}^{PCGLS.IMGS})_{ring}^{row-block} = (\frac{1}{4}n^2+\frac{9}{4}n)\lfloor\frac{p}{2}\rfloor\tau_{startup} + 16\lfloor\frac{p}{2}\rfloor(n\lceil\frac{n}{p}\rceil + \frac{1}{4}n^2 + \frac{5}{4}n)\tau_{comm}, \tag{8}$$

and

$$\begin{aligned}(T_{np}^{PCGLS.IMGS})_{torus}^{grid} &= (6n^2+14n)(\lceil\frac{n}{p}\rceil - \frac{n}{p})\tau_{calc} + \frac{1}{2}(n^2+5n)(2\sqrt{p}+\frac{1}{p}-3)\tau_{calc} + \\ &\quad n(\lceil\frac{n}{\sqrt{p}}\rceil^2 - \frac{n^2}{p})\tau_{calc} + n((2\sqrt{p}-4)\lceil\frac{n}{\sqrt{p}}\rceil + 2\frac{n}{p})\tau_{calc},\end{aligned} \tag{9}$$

$$\begin{aligned}(T_{comm}^{PCGLS.IMGS})_{torus}^{grid} &= (32n(\sqrt{p}-1)\lceil\frac{n}{p}\rceil + 4(n^2+5n)(\lfloor\frac{\sqrt{p}}{2}\rfloor + \sqrt{p}-1))\tau_{comm} + \\ &\quad 16n\lfloor\frac{\sqrt{p}}{2}\rfloor\lceil\frac{n}{p}\rceil\tau_{comm} + ((\frac{1}{4}n^2+\frac{9}{4}n)\lfloor\frac{\sqrt{p}}{2}\rfloor)\tau_{startup} + \\ &\quad (\frac{1}{4}n^2+\frac{13}{4}n)(\sqrt{p}-1)\tau_{startup}.\end{aligned} \tag{10}$$

If we assume that $n \gg p$ in case of a perfect load balance, we know from (7) and (9) that

$$(T_{np}^{PCGLS.IMGS})_{ring}^{row-block} = \Theta(p)\tau_{calc}, \quad (T_{np}^{PCGLS.IMGS})_{torus}^{grid} = \Theta(n)\tau_{calc}. \tag{11}$$

From (11) with a load imbalance in the system, T_{np} for the grid-torus combination has the same order of magnitude as with a perfect load balance. However, a load imbalance has a strong impact on the row-block-ring combination; Now, T_{np} becomes

$$(T_{np}^{PCGLS.IMGS})_{ring}^{row-block} = \Theta(n)\tau_{calc}, \tag{12}$$

which is the same order of magnitude as the grid-torus combination. The corresponding order of magnitude of T_{comm} follows from (8) and (10) is

$$(T_{comm}^{PCGLS.IMGS})_{ring}^{row-block} = \Theta(p)\tau_{startup} + \Theta(n)\tau_{comm}, \tag{13}$$

and

$$(T_{comm}^{PCGLS.IMGS})_{torus}^{grid} = \Theta(\sqrt{p})\tau_{startup} + \Theta(n)\tau_{comm}. \tag{14}$$

From (11) and (14), it is easy to show that T_{up} and T_{comm} for the row-block-ring combination and the grid-torus combination are comparable. We need to decide the numerical values of the system parameters. Furthermore, we must remember that T_{par} for PCGLS with IMGS preconditioner is mainly determined by T_{seq}/p, with an order of magnitude

$$\frac{T_{seq}}{p} = \Theta(\frac{n^2}{p})\tau_{calc}. \tag{15}$$

Now, we can use the crude expression suggested in [10, 11] for the efficiency ε of the parallel PCGLS with IMGS preconditioner, where we have taken the worst case T_{comm} of (13)

$$\varepsilon = \frac{T_{seq}}{pT_{par}} \approx \frac{1}{1 + \Theta(\frac{p}{n})(1 + \frac{\tau_{comm}}{\tau_{calc}}) + \Theta(\frac{p^2}{n^2})\frac{\tau_{startup}}{\tau_{calc}}}. \tag{16}$$

We know from (16) that deviations from $\varepsilon = 1$ is due to three reasons. We will first focus on the term related to τ_{comm}. This term can be viewed as the communication overhead f_c defined by Fox in [7]. For a problem dimension d_p and a dimension of the complex computer d_c, he shows that

$$f_c = \begin{cases} \frac{\text{constant}}{(n^2/p)^{1/d_p}} \frac{\tau_{comm}}{\tau_{calc}} & \text{if } d_c \geq d_p \\ p^{(1/d_c - 1/d_p)} \frac{\text{constant}}{(n^2/p)^{1/d_p}} \frac{\tau_{comm}}{\tau_{calc}} & \text{if } d_c < d_p \end{cases}$$

Thus, for the row-block-ring combination $d_p = 2$ and $d_c = 1$, the communication overhead f_c is

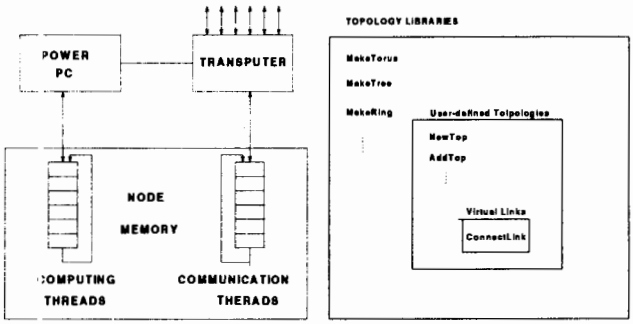
$$f_c = \text{constant} \cdot \frac{p}{n} \frac{\tau_{comm}}{\tau_{calc}},$$

which is in agreement with (16). However, for the grid-torus combination $d_c = 2$ one would expect

$$f_c = \text{constant} \cdot \frac{\sqrt{p}}{n} \frac{\tau_{comm}}{\tau_{calc}},$$

which disagrees with our results. The reason for this is the implementation of the partial vector accumulate. The first step of this operation can not exploit the full dimensionality of the network, giving rise to an order of magnitude of T_{comm} as in (14). The second term of the denominator in (16) is due to non-parallel computation (the grid-torus combination) or due to load imbalance (the row-block-ring combination) and has the same order of magnitude as the communication overhead. The last term in the denominator describes the efficiency reduction due to communication startup times. This term however is an order of magnitude smaller than the previous two.

As long as τ_{calc}, τ_{comm} and $\tau_{startup}$ have the same order of magnitude, and $n \gg p$, the efficiency of the parallel PCGLS with IMGS preconditioner can be very close to unity. In the next section we will investigate this in detail.



(a) The architecture (b) Hierarchy of topologies

5 Numerical Experiments

The Parsytec GC/PowerPlus combines the state-of-the-art RISC processor PowerPC-601 and advanced transputer communication technology and achieves top-end supercomputer performance with a scalable architecture supporting thousands of processor and offers a clear future upgrade path to PowerPC-604 and PowerPC-620 microprocessors. About the test matrices, we use the matrices from the simulation of the Elastic Light Scattering from arbitrary shaped micron sized biological particles by using the Couple Dipole method described in [11, 14]. In this case, A is symmetric $3N$ by $3N$ complex matrices with $N \sim 10^5$ and we generate the right-hand side vector consistent with a solution vector whose components are equal to 1.

5.1 Parsytec architecture

We will introduce briefly the architecture of Parsytec GC/PowerPlus. It employs a scalable routing network and a node design featuring a scalable number of application and communication processors. The overall architecture is optimized for most efficient execution of parallel programs composed of compute and communication threads. The Parsytec GC/Powerplus architecture supports parallelism at all three levels:

- Per-chip parallelism: the PowerPC-601 executes 3 instructions per cycle, the T805 transputer supports 8 independent concurrent communications.
- Per-node parallelism: 4 communication and 2 application processors forms a single GC/PowerPlus node and allows 6 parallel threads.
- Per-application parallelism: a large number of physical nodes is connected through a bidirectional fat-grid network with packet message routing.

The node architecture uses a 'variable balance architecture' and supports a variable number of application and communication processors and memory modules. Parix

gives transparent access to any node configuration and allows steamless programming of both compute and communication tasks using the same programming environment. The GC/PowerPlus node design has been selected after extensive simulation as the best compromise between cost and scalability. The use of intelligent transputer communication processors has already triggered the development of advanced communication firmware to support monitoring, fault-tolerant, global operations, high level protocols and scatter/gather communication. Scalable I/O performance is achieved via the GC/PowerPlus boundary links to the parallel storage system.

5.2 Parix operating system

Parix is designed as an open development platform for super massively parallel processors. To a great extend Parix is hardware independent. Concurrently it has been implemented for transputer based systems(T800/T9000) and systems based on the PowerPC. Standard UNIX systems serve as the embedded host machines.

There are four types of communication available in Parix [12]:

- Synchronous virtual link bound communication, is the GC's and Xplorer's native form of communication. The communicating processes have to be connected with a virtual link and are synchronized upon communication. The process, for example, which reaches the point of communication first has to wait for the other one, and data transfer takes place when both process are ready to communicate. A process can await communication from one of several links simultaneously.

- Synchronous random communication, does not require the definition of virtual links. For long messages, however, the library calls internally establish a virtual link due to performance reasons.

- Asynchronous link bound communication, can be carried out concurrently with computation. It therefore optimizes processor usage, as sending and receiving is performed in the background while the processor continues working.

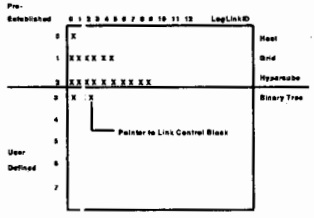
- Asynchronous random communication, based on mailboxes is the most general form of data exchange. There is no need for a link between the two communicating processes, but the size of messages is limited due to the store-and-forward principle of the router and the mailbox.

As we mentioned before, the first communication model suit our assumption well. Communication via virtual link can be performed by the following library routines of the Parsytec system:

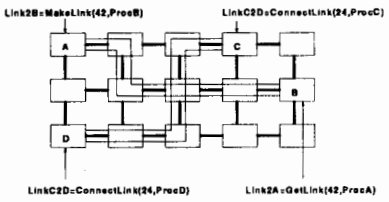
Send, Recv, SendLink, RecvLink, Select, ASend, ARecv

The first four operates on virtual topologies, specifying the relative orientation within the topology, whereas the last couple refers to a link control block directly.

To ease handling of the theoretically thousands of virtual links per processor, link can be ordered in virtual topologies. Virtual topology libraries relieve the programmer from having to building communication networks for some standard problems. The libraries provide calls for building grids, trees, hypercubes etc. The number of topologies defined on a single processor is practically unlimited, and any number of topologies can coexist



(c) Virtual links and topologies



(d) Make virtual links

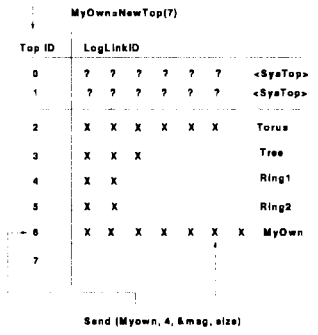
on a processor without interfering witch each other. The concept of virtual topologies offers a maximum of flexibility to modeling parallel algorithms. The hierarchy of virtual links, user defined topologies and topology libraries including the corresponding calls for handling links/topologies can be structured as shown in Figure b. Virtual links and virtual topologies are managed on the base of some data structure kept locally by each Parix kernel. Figure c describes this in more detail.

Building some special virtual topologies is done in the following three steps:

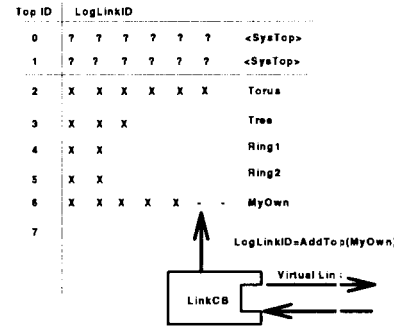
- Make virtual links. A virtual link is made through 'wormhole routing', during which a route to the other side is searched and buffer space is allocated at all intermediate nodes. Use unique ReqID's to avoid incorrect virtual link connections to be made.
- Allocate the new topologies. All virtual topologies are administrated by the Parix runtime system; The space reserved for this must be allocated from the system's heap. A unique topology ID is returned as handle to be used in all further actions on this topology.
- Add links to the topologies. The Link Control Block that is allocated as the local endpoint of a virtual link is added to the new topology by placing a LinCB-t pointer in the correct slot of the topology. The index of the LinkCB-t pointer in the topology is its LogLinkID.

Now, we will describe the structure of the program to create a ring by using another way independent of the physical network where each processor runs the same main program and determines from its own ID what to do.

1. Create ring
   If MyID even
      1. Connect link to the next processor.
      2. Connect link to the previous processor.
   Else
      1. Connect link to previous processor.
      2. Connect link to next processor.
2. Communication
   If processor 0



(e) Allocate the new algorithm



(f) Add links to topologies

1. Read line from keyword.
2. Send line to next processor.
3. Read line from previous processor.
4. Print line to screen.

Else

1. Read line from previous processor.
2. Send line to next processor.

5.3 Hardware parameters

The model parameters τ_{calc}, $\tau_{startup}$ and τ_{comm} depend on the hardware, on runtime environment and on compilers. Direct derivation of these parameters from hardware specifications alone gives always too optimistic values. Therefore reliable numerical values of the model parameters can only be obtained by actual measurements on the parallel computers. The corresponding values are

$$\tau_{calc} = 2.5\mu s, \quad \tau_{startup} = 5\mu s, \quad \tau_{comm} = 7.2\mu s.$$

With the experimental values for the system parameters a numerical calculation of T_{par} is possible. To compare the row-block-ring and the grid-torus combination relative difference function $rd(p,n)$ is defined in [10, 11]

$$rd(p,n) = \frac{(T_{par}^{PCGLS.IMGS})_{ring}^{row-block} - (T_{par}^{PCGLS.IMG})_{torus}^{grid}}{(T_{par}^{PCGLS.IMGS})_{ring}^{row-block} + (T_{par}^{PCGLS.IMG})_{torus}^{grid}}. \tag{17}$$

The sign of $rd(p,n)$ tells which parallelization strategy is faster, the amplitude tells how big the relative difference between the two is.

Since the execution time T_{par} can not justify a choice between the row-block-ring and grid-torus combinations, other criteria such as T_{loss} enter the discussion. Now we can

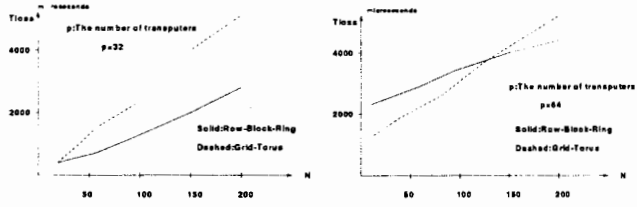


Figure 2: T_{loss} with row-block-ring and grid-torus for $p = 32$ and $p = 64$

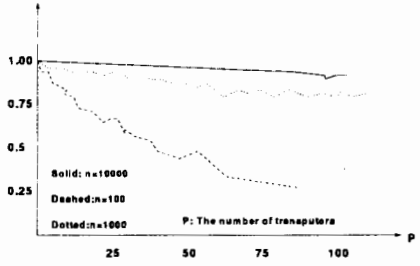


Figure 3: The theoretical efficiency with a row-block-ring combination

compute T_{loss} as function of p and N. Figure 2 gives T_{loss} as a function of N, for p equals 32 and 64, both for the row-block-ring and the grid-torus combinations. For very small N the grid-torus has a smaller T_{loss}. However, if N grows the T_{loss} of the row-block-ring combination is always smaller than for the grid-torus combination. This was concluded in the previous section. For realistic problems the row-block decomposition implemented on a ring topology is the best choice for PCGLS with IMGS preconditioner.

From an implementation point of view the row-block-ring combination is preferable. The row-block decomposition introduces just one communication routine, the gather operation. PCGLS with IMGS preconditioner via a grid-decomposition contains more and more complex communication routines. Furthermore, from a user point of view the row-block-ring combination has one important advantage. The ring can have any number of processors, and therefore the maximum number of free processors can always be used during production runs. This is quite important in parallel computing where the users can request any number of processors. These considerations are in favor of the row-block-ring combination. Figure 3 shows that the theoretical efficiency of this parallel PCGLS without preconditioner. If the number of rows per processor is large enough the efficiency will be very close to unity. So parallel PCGLS with IMGS preconditioner with row-block decomposition, implemented on a ring is very well suited for the Parsytec system.

5.4 Performance measurements

We have implemented the parallel PCGLS with IMGS preconditioner, with row-block decomposition of A, on a ring topology. This section we will describe some implementation issues and present preliminary measurements of $T_{par}(p, N)$. These performance measurements are compared with the theoretical expectations for this derived in the previous section. The preliminary results of the measurement of $T_{par}(p, N)$, the time per iteration as a function of p and N, and comparison of the results with theory is described in Figure 4. It gives the results for N equals 64, 250, 496, 918 and $1 \leq p \leq 64$. The relative difference between theory and measurement is almost 3.5%.

6 Conclusions

Our approach to parallelization of PCGLS with IMGS preconditioner is to start as general as possible, that is without any special knowledge of the computing system, and try to minimize T_{par}. Next limited characteristics of the computing system are specified, trying to find a good parallel implementation without detailed knowledge of the parallel computer. Finally numerical values of the model parameters which of course strongly depend on the particular parallel computer are needed to make an ultimate choice, if necessary, between different parallel implementations, and to compute theoretical performance figures. This allows a generic approach to the parallelization of the application. The decomposition is the only property of a parallel application that is independent of the computing system.

Here three data decompositions are considered: the row-block , column-block and grid decomposition. The resulting expressions for T_{par} contain the only parameter of the abstract model of the computing system, namely τ_{calc}. And we have proved that an implementation of PCGLS with IMGS preconditioner with the column-block decomposition is always slower than with the row-block decomposition [18]. This conclusion holds for all parallel computers which fall in the range of the abstract computing system. Although the difference in execution-time will be small in practice, it is an important point of principle that these conclusions can be drawn at such a high level of abstraction.

We have proceeded the row-block decomposition on a ring topology and the grid-torus combination. We also have derived some theoretical expressions for T_{par} as a function of the decomposition, topology and time constants. These theoretical detailed expressions are compared by an order of magnitude analysis, where in case of assuming $n \gg p$. The first striking feature is the sensitivity of the row-block-ring combination to a load imbalance, which is completely absent in the grid-torus combination. Actually the load imbalance in the grid-torus combination is $\Theta(1)\tau_{calc}$, which is negligible to the $\Theta(n)\tau_{calc}$ of the row-block-ring combination due to the computations in the partial vector accumulate. The total number of communication startups is smaller for the grid-torus combination than the row-block-ring combination, but the total amount of transmitted data is the same of $\Theta(n)$. If $\tau_{startup}$ and τ_{comm} have the same order of magnitude, the startup time for communications is negligible compared with the pure sending time. In that case the communication overhead for the torus and the less rich ring topology have the same order of magnitude. These above considerations show that the row-block-ring and the grid-torus combination have comparable execution times. As long as the time constants have the

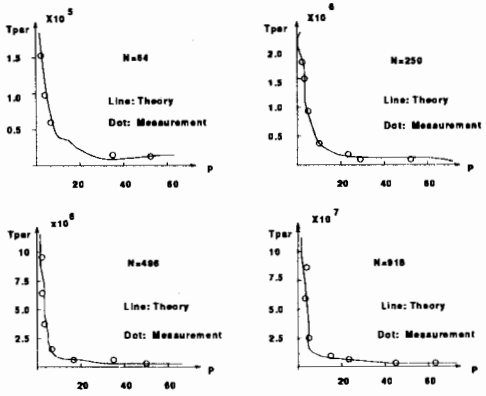


Figure 4: Comparison between experimental and theoretical results of T_{par}

same order of magnitude and $n \gg p$, both parallel implementation will have an efficiency close to 1.

Next, we measure the time constants and introduce them in the expression for T_{par}. It is shown that the row-block-ring and the grid-torus combinations are almost indistinguishable if one looks at their execution time. This means that even on a low connectivity network as a ring it is possible to implement a parallel PCGLS with IMGS preconditioner with a very high efficiency, comparable with implementations on a torus topology. We will consider this problem based on two practical consideration: programming effort and system resources.

The numerical experiments are carried out with relatively small matrices, which could be kept in memory. For very large problems, both combinations will have a comparable efficiency. We have implemented the row-block-ring combination because it contains less different communication routines, we can simplify the implementation.

Theoretical predictions of the execution-time for the parallel implementation agree very well with the experimental results and the abstract model we assumed of the computing system is well suited to present the numerical behavior of regular numerical applications, for example, solving the least squares problems using PCGLS with IMGS preconditioner, on parallel distributed memory computers.

References

[1] A. Basu, S. Srinivas, K. G. Kumar, and A. Paulray. A model for performance prediction of nessage passing multiprocessors achieving concurrency by domain decomposition. In *Proceedings of CONPAR'90*, pages 75–85, 1990.

[2] E. de Sturler. A parallel restructured version of GMRES(m). Technical Report 91-85, Delft University of Technology, Delft, The Netheland, 1991.

[3] E. de Sturler. A parallel variant of GMRES(m). In *Proceedings of the 13th IMACS World Congress on Computational and Applied Mathematics*. IMACS, Criterion Press, 1991.

[4] E. de Sturler and H. A. van der Vorst. Reducing the effect of the global communication in GMRES(m) and CG on parallel distributed memory computers. Technical Report 832, Mathematical Institute, University of Utrecht, Utrecht, The Netherland, 1994.

[5] J. W. Demmel, M. T. Heath, and H. A. van der Vorst. Parallel numerical algebra. *Acta Numerica*, 1993. Cambridge Press, New York.

[6] J. J. Dongarra, I. S. Duff, D. C. Sorensen, and H. A. van der Vorst. *Solving Linear Systems on Vector and Shared Memory Computers*. SIAM, Philadelphia, PA, 1991.

[7] G. Fox, M. Johnson, G. Lyzenga, S. Otto, J. Salmon, and D. Walker. Solving problems on concurrent processors. In *General Techniques and Regular Problems.*, volume 1. Prentice-Hall International Editions, 1988.

[8] M. R. Hestenes and E. Stiefel. Methods of conjugate gradients for solving linear system. *J. Res. Nat. Bur. Stds*, B49:409–436, 1952.

[9] C. A. R. Hoare. Communicating sequential processes. *Communications of ACM*, 21(666-677), 1978.

[10] A. G. Hoekstra, P. M. A. Sloot, M. J. de Haan, and L. O. Hertzberger. Time complexity analysis for distributed memory computers, implementation of a parallel conjugate gradient method. In J. van. Leeuwen, editor, *Computing Science in the Netherlands*, pages 249–266. Elsevier Science, 1991.

[11] A. G. Hoekstra, P. M. A. Sloot, W. Hoffmann, and L. O. Hertzberger. Time complexity of a parallel conjugate gradient solver for light scattering simulations: theory and SPMD implementation. Technical report, Department of Computer System, University of Amsterdam, 1992.

[12] Parsytec GmBH. *Parix manual version 1.3.1-PPC*, 1995.

[13] C. Pommerell. *Solution of large unsymmetric systems of linear equations*. PhD thesis, ETH, 1992.

[14] E. M. Purcell and C. R. Pennypacker. Scattering and absorption of light by nonspherical dielectric grains. *The Astrophysical Journal*, 186:705–714, 1973.

[15] T. Yang. Error analysis for incomplete modified Gram-Schmidt preconditioner. In Proceedings of *Prague Mathematical Conference (PMC-96)*, July 1996. Mathematical Institute of the Academy of Sciences, Zitza 25, CZ-115 67 Praha, Czech Republic.

[16] T. Yang. Iterative methods for least squares and total least squares problems. Licentiate Thesis LiU-TEK-LIC-1996:25, 1996. Linköping University, 581 83, Linköping, Sweden.

[17] T. Yang. Parallel least squares problems on massively distributed memory computers. In Proceedings of *The Eleventh International Conference on Computer and Information Science (ISCIS-XI)*, November 1996. Middle East Technical University, Antalya, Turkey.

[18] T. Yang. Best data decomposition model for least squares problems on massively parallel distributed memory computers. Submitted to *Second Workshop on Advanced Parallel Processing Technologies (APPT-97)*, May 1997. Koblenz, Germany.

LINEAR OPERATOR INTERPOLATION IN A SPACE OF FUNCTIONAL MATRICES

L.A. Yanovich, E.M. Manyuk

Institute of Mathematics, Academy of Sciences of Belarus
Surganov str., 11 Minsk 220072, Belarus
e-mail: imanb%imanb.belpak.minsk.by@demos.su

Abstract. Formulae of linear interpolation for operators defined on the set continuous and continuously differentiable matrices have been obtained.

Introduction

Let X and Y are some sets of matrices and $F : X \to Y$. One of the problems of interpolation consists in the following: for given function F and knots $x_k \in X$ to construct another function $\Phi : X \to Y$ such that

$$\Phi(x_k) = F(x_k).$$

Usually Φ is constructed in the form of a matrix polynomial of the fixed order. For example, when $X = Y$ is a set of square matrices of the dimension n and inverse matrices for $A_i - A_j$ for $(i \neq j \;\; i, j = 0, 1, ..., m)$ exist then one can take $\Phi(A)$ as

$$L_m(A) = \sum_{k=0}^{m} F(A_k) l_k^{-1}(A_k) l_k(A), \tag{1}$$

where

$$l_k(A) = (A - A_0)(A - A_1) \cdots (A - A_{k-1})(A - A_{k+1}) \cdots (A - A_m).$$

In the general case, the matrix polynomials (1) will not coincide identically with $F(A)$ even if $F(z)$ is an algebraic polynomial of the order $\leq m$. This will take place, e.g., when $A_k = \alpha_k E$ $(\alpha_0 < \alpha_1 < ... < \alpha_m)$, E is an identy matrix.

If $A_k = \lambda_k^3$ $(k = 0, 1, ..., n-1; \quad m = n-1)$, where λ_k are different eigenvalues of matrix A, then $L_n(A) \equiv F(A)$ for any analytic function $F(z)$ (Lagrange -Sylvester's formula, see [1] .

It is important for applications to construct interpolation formulae for functions of matrices that would be invariant with respect to matrix polynomials with matrix coefficients.

Some such formulae of linear interpolation in the space of continuous and continuously differentiable matrices will be given below.

Formulae of interpolation in the space of continuous matrices

Let $X = C(T)$ be the space of continuous on $T = [a, b]$ matrices of dimension n and Y be some another space of functional matrices of the same dimension. A matrix polynomial of the first order $\mathcal{P}_1(A) \in Y$ in this case has of the form

$$\mathcal{P}_1(A) = K_0 + \sum_{i=1}^{m} \int_T K_{1i}(t) A(t) K_{2i}(t)\, dt, \tag{2}$$

where K_0, $K_{1i}(t)$ and $K_{2i}(t)$ are given matrices, $A(t) \in C(T)$.

Introduce the number function

$$\chi_\tau(t) = \begin{cases} 1, & \text{for } \tau > t \\ 0, & \text{for } \tau < t, \end{cases}$$

$$\chi_a(t) \equiv 0, \quad \chi_b(t) \equiv 1.$$

Let $A_0(t)$ and $A_1(t)$ be matrices such that there exists on $[a, b]$ an inverse matrix $[A_1(t) - A_0(t)]^{-1}$.

Then

$$L_{10}(A) = F[A_0(\cdot)]+$$

$$+ \int_T [A(\tau) - A_0(\tau)][A_1(\tau) - A_0(\tau)]^{-1} \, d_\tau F[A_0(\cdot) + \chi_\tau(\cdot)(A_1(\cdot) - A_0(\cdot))]$$

is linear interpolation matrix polynomial for the function $F(A)$ relatively the knots $A_0(t)$ and $A_1(t)$.

This formula is invariant with respect to polynomials of the form (2) with $K_{1i}(t) \equiv E$ for $i = 1, 2, ..., m$.

Let the operator $F(A)$ be differentiable in the sense of the Gateaux. Then the matrix

$$L_1(A) = F(A_0) + \frac{1}{b-a} \int_T \delta F[A_0 + \frac{\tau - a}{b-a}(A_1 - A_0); A - A_0] \, d\tau,$$

where $\delta F[A; h]$ is the Gateaux differential at the point A in the direction h, be a linear interpolation matrix polynomial for the operator $F(A)$ with respect to knots A_0 and A_1. If $F(A)$ is a matrix polynomial of the form (2), then $L_1(A) \equiv F(A)$.

Example. Let

$$F(A) = \int_T N(t) A^2(t) M(t) \, dt,$$

where $N(t)$ and $M(t)$ are a given matrices. Then

$$L_1(A) = \frac{1}{2} \int_T N(t)[A_{01}(t)A(t) + A(t)A_{01}(t)]M(t) \, dt -$$

$$- \frac{1}{2} \int_T N(t)[A_0(t)A_1(t) + A_1(t)A_0(t)]M(t) \, dt,$$

where $A_{01}(t) = A_0(t) + A_1(t)$.

Formulae of interpolation in the space of continuously differentiable matrices

Let $C^p(T)$ be a space of matrices that have continuous derivatives up to p-th order ($p \geq 1$) on T and $F(x)$ be an operator mapping $C^p(T)$ into another space of functional matrices.

An operator (matrix) polynomial of the first order $\mathcal{P}_1(A)$ has, in the space under consideration, the form

$$\mathcal{P}_1(A) = C + \sum_{\nu=0}^{p-1} C_\nu A^{(\nu)}(a) D_\nu + \int_T Q(t) A^{(p)}(t) G(t) \, dt, \tag{3}$$

where C, C_ν, D_ν, $Q(t)$ and $G(t)$ are given matrices.

Let's assume that the inverse matrices and integrals, which are contained in the consequent interpolation formulae, exist.

Introduce the designations

$$l_{01}(A) = F(A_0) + \sum_{k=0}^{p-1}[F(\sigma_k) - F(\sigma_{k-1})][A_1^{(k)}(a) - A_0^{(k)}(a)]^{-1}(A^{(k)}(a) - A_0^{(k)}(a)),$$

$$l_{10}(A) = F(A_0) + \sum_{k=0}^{p-1}(A^{(k)}(a) - A_0^{(k)}(a))[A_1^{(k)}(a) - A_0^{(k)}(a)]^{-1}[F(\sigma_k) - F(\sigma_{k-1})],$$

where

$$\sigma_{-1}(t) = A_0(t), \quad \sigma_k(t) = \sigma_{k-1}(t) + \frac{(t-a)^k}{k!}[A_1^{(k)}(a) - A_0^{(k)}(a)] \quad (k = 0, 1, ..., p-1).$$

Theorem 1 *The matrix polynomials*

$$L_{01}(A) = l_{01}(A) + \int_T d_\tau F[g(\tau, \cdot)] \cdot [A_1^{(p)}(\tau) - A_0^{(p)}(\tau)]^{-1}[A^{(p)}(\tau) - A_0^{(p)}(\tau)] \qquad (4)$$

and

$$L_{10}(A) = l_{10}(A) + \int_T [A^{(p)}(\tau) - A_0^{(p)}(\tau)][A_1^{(p)}(\tau) - A_0^{(p)}(\tau)]^{-1} d_\tau F[g(\tau, \cdot)], \qquad (5)$$

where

$$g(\tau, t) = \sigma_{p-1} + A_1(\min(\tau, t)) - \sigma_{p-1}(\min(\tau, t))$$

are linear interpolation polynomials on $C^{(p)}(T)$ with respect to knots $A_0(t)$ and $A_1(t)$ for the matrix-valued function $F(A)$. Formulae (4) and (5) are invariant with respect to matrix polynomials of the form (3) correspondingly for

$D_\nu = G(t) \equiv E$ *and* $C_\nu = Q(t) \equiv E$ $(\nu = 0, 1, ..., p-1)$.

Proof. The coincidence $L_{01}(A)$ and $L_{10}(A)$ with $F(A)$ at knots A_0 and A_1 is verified simply. Let

$$F(A) = C_\nu A^{(\nu)}(a) D_\nu. \qquad (6)$$

Then the integrals in formulae (4) and (5) are equal to zero since in this case $F[g(\tau, \cdot)]$ does not depend upon τ: $F[g(\tau, \cdot)] = C_\nu A_1^{(\nu)}(a) D_\nu$ $(\nu = 0, 1, ..., p-1)$.

Using the correlation

$$\sigma_k^{(\nu)}(a) = \begin{cases} A_1^{(\nu)}(a), & \nu = 0, 1, ..., k \\ A_0^{(\nu)}(a), & \nu = k+1, k+2, ..., p-1, \end{cases}$$

we obtain

$$F[\sigma_k(\cdot)] - F[\sigma_{k-1}(\cdot)] = \begin{cases} C_\nu[A_1^{(\nu)}(a) - A_0^{(\nu)}(a)]D_\nu, & \nu = k \\ 0, & \nu \neq k \end{cases}$$

and, consequently, $L_{01}(A)$ and $L_{10}(A)$ coincide with (6) for $D_\nu = E$ and for $C_\nu = E$ correspondingly.

If $F(A) = Q(t)A^{(p)}(t)G(t)$ $(t \in T)$, then the difference $F[\sigma_k(\cdot)] - F[\sigma_{k-1}(\cdot)] = 0$ $(k = 0, 1, ..., p-1)$, and

$$F[g(\tau, \cdot)] = \begin{cases} Q(t)A_1^{(p)}(t)G(t), & \tau \geq t \\ Q(t)A_0^{(p)}(t)G(t), & \tau < t. \end{cases}$$

Hence it follows, that $L_{01}(A) = F(A)$ for $G(t) \equiv E$, and $L_{10}(A) = F(A)$ for $Q(t) \equiv E$. The theorem is true.

Theorem 2 *The matrix polynomials*

$$L_{01}(A) = l_{01}(A) + \frac{1}{b-a}\int_T \delta F[\sigma_{p-1}(\cdot) + \frac{\tau - a}{b-a}(A_1(\cdot) - \sigma_{p-1}(\cdot)); h(\cdot)]\, d\tau, \tag{7}$$

$$L_{10}(A) = l_{10}(A) + \frac{1}{b-a}\int_T \delta F[\sigma_{p-1}(\cdot) + \frac{\tau - a}{b-a}(A_1(\cdot) - \sigma_{p-1}(\cdot)); h(\cdot)]\, d\tau, \tag{8}$$

where $\delta F[y; h]$ is the Gateaux differential of operator F at the point y in the direction h,

$$h(t) = A(t) - A_0(t) - \sum_{k=0}^{p-1} \frac{(t-a)^k}{k!}[A^{(k)}(a) - A_0^{(k)}(a)], \tag{9}$$

will be a linear interpolation matrix polynomials of the first order for the operator $F(A)$ with respect to knots $A_0(t)$ and $A_1(t)$.

The first from these polynomials is invariant with respect to matrix polynomials of the form(3) for $D_\nu = E$, the second one with respect to that kind of polynomials when $C_\nu = E$.

Proof. Note, that $l_{01}(A_0) = l_{10}(A_0) = F(A_0)$ and $l_{01}(A_1) = l_{10}(A_1) = F(\tau_{p-1})$. According to the fact, that $h(t) = A_1(t) - \sigma_{p-1}(t)$ for $A(t) = A_1(t)$, the integral in formulas (7) and (8) for such direction $h(t)$ can be calculated. It is equal to $(b-a)[F(A_1) - F(\sigma_{p-1})]$.

Consequently, the interpolation conditions for the polynomials L_{01} and L_{10} are executed.

Let $F(A)$ be of the form (6). Then the differential $\delta F[y; h]$ in the direction (9) at any point y equal to zero.

As $\sigma_k^{(p)}(t) = A_0^{(p)}(t)$ for $k = 0, 1, ..., p-1$, then for matrix functions of the form

$$F(A) = \int_T Q(t)A^{(p)}(t)G(t)\, dt \tag{10}$$

we have $F(\sigma_k) = F(\sigma_{k-1})$ or $l_{01}(A) = l_{10}(A) = F(A_0)$.

The integral term in the formulae (7) and (8) for the operator (10) is equal to the difference $F(A) - F(A_0)$. ¿From this follows a validity this theorem.

Linear interpolation of matrix-valued functions of two variables

Let $X = C^{(1,1)}(T^2)$ be the space of square matrices $A(t) = A(t_1\ t_2)$ such that $A^{(1,0)}(t_1, a_2) = \frac{\partial A(t_1, a_2)}{\partial t_1}$ and $A^{(0,1)}(a_1, t_2) = \frac{\partial A(a_1, t_2)}{\partial t_2}$ are continuous correspondingly on intervals $T_1 = [a_1, b_1]$ and $T_2 = [a_2, b_2]$, $A^{(1,1)}(t) = \frac{\partial^2 A(t_1, t_2)}{\partial t_1 \partial t_2}$ is continuous on the rectangle $T^2 = T_1 \times T_2$.

In this case we consider the operator polynomials of the first order $\mathcal{P}_1(A)$ of the following form

$$\mathcal{P}_1(A) = N + N_{00}A(a)M_{00} + \int_{T_1} N_{10}(\tau_1)A^{(1,0)}(\tau_1, a_2)M_{10}(\tau_1)\,d\tau_1 +$$

$$+ \int_{T_2} N_{01}(\tau_2)A^{(0,1)}(a_1, \tau_2)M_{01}(\tau_2)\,d\tau_2 + \int_{T^2} N_{11}(\tau)A^{(1,1)}(\tau)M_{11}(\tau)\,d\tau, \qquad (11)$$

where N, N_{ij} and M_{ij} $(i, j = 0, 1)$ are given matrices, $a = (a_1, a_2)$, $\tau = (\tau_1, \tau_2)$.

Theorem 3 *The operator polynomials of the first order*

$$L_{01}(A) = F(A_0) + [F(\sigma) - F(A_0)][A_1(a) - A_0(a)]^{-1}[A(a) - A_0(a)] +$$

$$+ \int_0^1 \delta F[\sigma + s(A_1 - \sigma); h]\,ds, \qquad (12)$$

$$L_{10}(A) = F(A_0) + [A(a) - A_0(a)][A_1(a) - A_0(a)]^{-1}[F(\sigma) - F(A_0)] +$$

$$+ \int_0^1 \delta F[\sigma + s(A_1 - \sigma); h]\,ds, \qquad (13)$$

where $\sigma(t) = A_0(t) + A_1(a) - A_0(a)$, $h(t) = A(t) - A_0(t) - A(a) + A_0(a)$, *are interpolating for* $F(A)$ *with respect to knots* $A_0(t)$ *and* $A_1(t)$.

If $F(A)$ *coincides with polynomial (11) for* $M_{00} = E$, *then* $L_{01}(A) \equiv F(A)$ *on* $C^{(1,1)}(T^2)$, *or if* $N_{00} = E$, *then the identity* $L_{10}(A) \equiv F(A)$ *is valid.*

Proof. For $F(A) = N_{00}A(a)M_{00}$ the integral in formulas (12) and (13) is equal to zero.

The first two summands coincide with this operator when $M_{00} = E$ for the formula (12) and $N_{00} = E$ for the formula (13).

Note further that $F(\sigma) = F(A_0)$ for all summands of the polynomial (11) exept of the second one.

If $F(x)$ has the form of one of the integrals in (11), then after calculation of the differential of $F(x)$ in the given direction $h(t)$, we obtain that the last summand in (12) and in (13) is equal to $F(x) - F(x_0)$, i.e. for the matrix polynomials of the form (11) we have $L_{01}(A) = L_{10}(A) \equiv \mathcal{P}_1(A)$. The execution of interpolation conditions in this case is verified simply.

Remark. Formulas (12)and (13) are invariant relatively operators of more general form than (11). For example, this fact holds true for the following integral operator

$$F(A) = \int_{T^2} N(t)\frac{\partial^\alpha A(t)}{\partial t_1^{\alpha_1}\partial t_2^{\alpha_2}}M(t)\,dt,$$

where $\alpha = \alpha_1 + \alpha_2 \geq 1$, α_1 and α_2 are not negative integers. The formula (12) is invariant also for the operator of the form

$$F(A) = \int_{T^2} K(t)A(t)\,dt.$$

Avariety of similar formulas of the linear interpolation in the spaces of continuously differentiable functions have been obtained in [2].

References

[1] Gantmacher F.R. Theory of Matrices (Chelsea, New York, (1959)).
[2] Yanovich L.A. Linear Interpolation of Operators in the Space of Conti nuously Differentiable Functions. Proc. Third Seminar "Nonlinear Phenomena in Comles Systems", Polatsk, February 14-16, 1994, p. 76-82.

"TG-GSS" MANAGER TO GENERATE APPLIED SOFTWARE FOR ANALYSIS OF DATA LIKE SET OF TIME SERIES AND ILL-POSED INVERSE PROBLEMS

P.N. Zaikin, T.B. Podosenova

Russia, MSU, Calculating Mathematics and Cybernetics Department, e-mail: pnz@cs.msu.su

Abstract. The principal problems of creating software manager for computer-aided procedure of simulating applied software for some typical kind of science investigations and industrial technologies, where images of observed data are like sets of digitised time series, are discussed.

1 Introduction.

Nowadays the situation around the using of computers and especially of personal ones has lead in many cases to necessity of essential reconstruction of methods, devices, simulating models, tools and numerical methods when one is dealing with applied systems of science investigations and industrial technologies. The red-lines of using computer-aided technologies are: the including of computer processors into detectors and devices, the using of data banks and of data knowledges banks, of dialogue-graphics tools of representation of subject domain objects and operating under their screen images, and at last providing the experts (the knowledge engineers) with the possibilities of planning the computer data processing and of choosing the final solution.

Notice that the enough reach software and experience of its using is accumulated in each subject domain. Therefore the last years the CASE (computer applied software environment) technology is widely using.

2 Methods. Mathematical models.

In our subject domain the observed data may be represented by unique digital one- or two-dimensional functions like time series. One get two-dimensional functions, for example, when the series of measurements are conducting according to the time factor. The mathematical model of phenomenon under investigation consists of two kinds of functions: one is enough smooth function, another part is a set of local resonances. The cause-consequence relations here are enough complicated and involve a non-linear calibration function, an efficiency function, a non-linear main dependence. Therefore the problems of data analysis are ill-posed ones.

It is often supposed that theoretically the radiation spectra may be represented by a set of single lines like delta-functions, whose locations are depending on energies of nuclei radiation spectra lines. But the registration devices change significantly the theoretical representation of radiation spectra. The process of registration may be described by integral equation of the first kind.

The different types of functions are using to describe the mathematical models of spectral curves. The forms of single lines in spectra are depending on registration devices and may be whether symmetric whether asymmetric ones to the left or to the right from the line

location. Sometime the single line is represented by a sum of components, which describe the complicated processes taking place in devices when the nuclei radiation spectrum is measuring.

The top idea of regular algorithms in our subject domain consists of computing, firstly, the most stable part of solution in that way then the model is nearly according to the experimental data. The computing of other part of solution is possible only by taking into account whether time dependence in model whether another strategy of conducting experiment, detecting data and data processing. The non-unique models of experimental data interpretation take place also. The user has a possibility to get information from estimated data bank and data expert knowledge base.

The different kinds of programming tools (packages of applied programs designed by different firms and authors) are using to process such experimental data. The most part of these packages use the fixed set of methods and non-changeable schemes of data processing. But in practice the necessity often arises to change the given (fixed) way of calculations. That is why one of the urgent problems nowadays is the design of such programming tools of data processing which provide the user with the possibility to change the schemes of calculations and to choose the sequence of computing stages in interactive mode.

The dialogue-graphics environment TG-GSS is designed to solve the wide range of problems of elements analysis, in particular, of alpha- and gamma- radiation analysis by using of methods of nuclear spectroscopy.

TG-GSS environment allows to reach two purposes:

- to create the semantic models of data and processes of user subject domain;
- to construct the semantic nets of complicated schemes of experimental data processing providing the user both with flexible setting up of system on concrete problems peculiarities and data frames and also with the possibility to get informational and expert support.

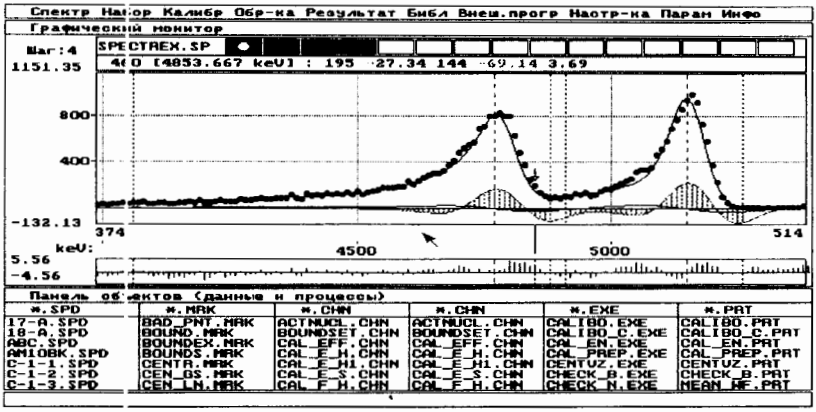
TG-GSS is designed in such a way that the applied mathematicians and/or the subject domain experts can construct themselves (without the help of system programmers) the applied programs packages to solve the concrete problems of elements analysis. There are special dialogue means in TG-GSS allowing to write the complicated schemes of data processing and to create the applied programs packages by using of sets of ready to run modules (utilities). There are formed some standards on data files formats and some agreements on running, organising and setting up utilities in TG-GSS.

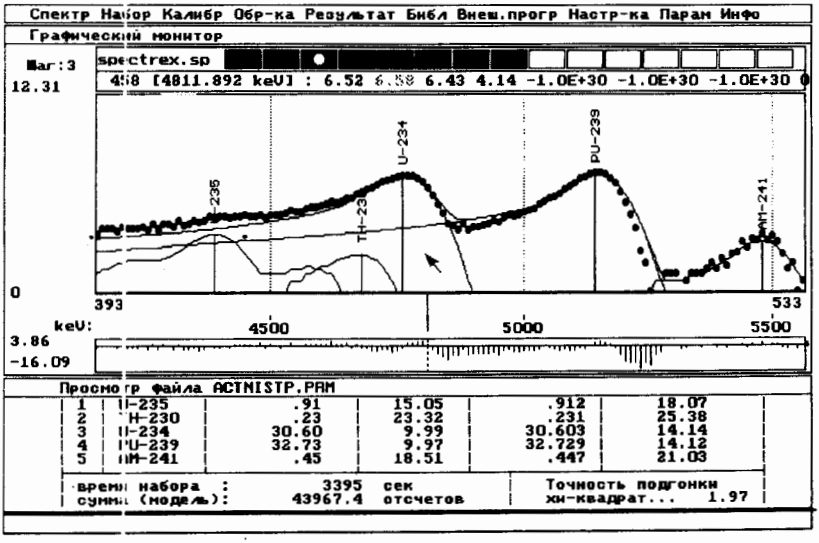
The main dialogue-graphics system TG-GSS components are:

- graphical monitor,
- manager of objects and processes,
- subsystem working with "sequences",
- subsystem working with estimated data banks,
- subsystem working with driver of registration instrument,
- informational support,
- menu subsystems.

At the pictures below one can see two main windows of TG-GSS: the graphical monitor window (where one can see graphs) and objects and processes Manager window (where one can see files names or messages).

The following information can be seen in graphical monitor window: the graphs of spectra, the images of nuclei spectral lines from current nuclide library, the diagrams


Спектр Набор Калибр Обр-ка Результат Библ Внеш.прогр Настр-ка Парам Инфо
Графический монитор
Шаг:4
SPECTREX.SP
1151.35
800
400
-132.13
374
514
keV:
4500
5000
5.56
-4.56
Панель объектов (данные и процессы)
*.SPD
*.MRK
*.CHN
*.EXE
*.PRT
17-A.SPD
18-A.SPD
ABC.SPD
AMIOBK.SPD
C-1-1.SPD
C-1-2.SPD
C-1-3.SPD
BAD_PNT.MRK
BOUND.MRK
BOUNDEX.MRK
BOUNDS.MRK
CENTR.MRK
CEN_GS.MRK
CEN_LN.MRK
ACTNUCL.CHN
BOUNDSET.CHN
CAL_EFF.CHN
CAL_E_H.CHN
CAL_E_H1.CHN
CAL_E_S.CHN
CAL_F_H.CHN
CALIBO.EXE
CALIBO_C.EXE
CAL_EN.EXE
CAL_PREP.EXE
CENTUZ.EXE
CHECK_B.EXE
CHECK_N.EXE
CALIBO.PRT
CALIBO_C.PRT
CAL_EN.PRT
CAL_PREP.PRT
CENTUZ.PRT
CHECK_B.PRT
MEAN_RF.PRT


Спектр Набор Калибр Обр-ка Результат Библ Внеш.прогр Настр-ка Парам Инфо
Графический монитор
Шаг:3
spectrex.sp
12.31
0
393
533
keV:
4500
5000
5500
3.86
-16.09
U-235
TH-230
U-234
PU-239
AM-241
Просмотр файла ACTNISTP.PRM
1 U-235 .91 15.05 .912 18.07
2 TH-230 .23 23.32 .231 25.38
3 U-234 30.60 9.99 30.603 14.14
4 PU-239 32.73 9.97 32.729 14.12
5 AM-241 .45 18.51 .447 21.03
время набора : 3395 сек
сумма (модель): 43967.4 отсчетов
Точность подгонки
хи-квадрат... 1.97

of busyness graphics (for example, the results of data processing can be represented as diagrams).

The graphical monitor is intended to show and to operate with the graphical images of subject domain objects (spectra, lines centres marks etc.). No more than 16 graphs of functions (function here is a one-dimensional vector of length up to 4096) can be loaded in two special windows of graphical monitor. Each graph may be showed in own style and colour. Both regular and irregular argument grid are allowed. There is mobile cursor in these two special windows by means of which it is possible to shift the pictures in horizontal direction, to get the values of loaded functions at current argument point etc.

Unlike the others data processing programs packages the dialogue-graphics environment TG-GSS includes the objects and processes Manager which is the key component of TG-GSS. This Manager window looks like a six-windows area. The list of files names of different groups of objects can be seen in different Manager windows. There are such files (objects) groups as: spectra files; the spectra marks files; auxiliary (service) system files; files of "chains"; running utilities (applied programs); listings of utilities; nuclide data libraries; "set up" files of utilities. A sets of available operations under files in each of six Manager windows are completely depending on and defining by semantics of corresponding objects groups.

One can see in objects and processes Manager window (at the same time with looking through the graphical information) depending on current mode of TG-GSS the following:

- list of files of subject domain objects (initial spectra, marks of peaks centres and of peaks bounds etc.),
- list of nuclei from current library,
- listings of utilities being run,
- comments, announcements, prompts, advices, questions,
- special editors window to view/to edit files or to set parameters of running utilities.

The looking of information in two main windows is organised in such a way that the user can work simultaneously with two mentioned above kinds of data presentation.

The subsystem working with "chains" is aimed to give to the user the possibilities to write, to edit and to run the data processing schemes. The special problem-oriented language (the "chains" language) was created to describe the semantic schemes of data processing. The "sequence" here is a sequence of commands of this special language. This language includes such structures as linear sequences of operators (commands), "enclosed sequences", loop, conditional and unconditional jump instructions. There are also the following groups of commands:

- utilities running commands,
- commands to load (in graphical window) graphical images of subject domain objects,
- commands to show some text information (comments, announcements, prompts, advices, questions) in objects and processes Manager window,
- special commands of data processing,
- commands to view and to edit files by two special editors,
- commands to copy or to erase files.

These mentioned above structures and commands are sufficient to write a very complicated computational block-charts. The processing of "chain" in TG-GSS involves the following steps: the reading of next command from "chain" file, the analysis of this com-

mand by interpreter and carrying out of indicated in the command actions.

It is worth to note that TG-GSS allows using of different algorithmical languages: the instrumental language here is Pascal but the running utilities (EXE-files) may be written in any language (Fortran, Pascal, Basic etc.).

3 Results.

The described methodology is realised in TG-GSS monitor to generate applied software to process, in particular, alpha- and gamma- spectra in elements analysis of samples. Other example of using TG-GSS is analysis of cosmic gamma-ray spectra in CORONAS-I space experiment.

References

[1] A. V. Dmitriev, S. N. Kuznetsov, S. P. Ryumin, P. N. Zaikin, T. B. Podosenova, I. I. Petsilyak, Yu. D. Kotov, A. S. Glyanenko. Using "TG-GSS" dialogue-graphics platform for processing the cosmic gamma-ray spectra in CORONAS-I space experiment, (this issue)

Методы решения некоторых многопараметрических спектральных задач математической физики [1]

А. А. Абрамов, А. Л. Дышко, Н. Б. Конюхова, Т. В. Левитина, В. И. Ульянова

Вычислительный центр РАН
Вавилова 40, 117967 Москва ГСП-1, Россия
e-mail: alalabr@ccas.ru, nadja@ccas.ru, lev@ccas.ru

Аннотация. Многопараметрическими спектральными задачами являются задачи на собственные значения, выделяющиеся тем, что возникает набор нескольких спектральных параметров, значение которых подлежит определению. Такие задачи очень трудны и для исследования, и для численного решения (см., например, [1]). В работе совсем кратко излагаются некоторые из результатов, содержащихся в [2–8], где для некоторых классов таких задач удалось разработать глобально сходящиеся методы поиска нужных собственных значений и вычисления соответствующих им собственных функций.

§1. Метод вычисления волновых эллипсоидальных функций (см.[2,3])

Мы начнем с задачи, послужившей основой нашей работы в этой области. Рассмотрим уравнение Гельмгольца

$$\Delta W + k^2 W = 0, \tag{1}$$

где k — заданное число, Δ — оператор Лапласа: $\Delta = \dfrac{\partial^2}{\partial x^2} + \dfrac{\partial^2}{\partial y^2} + \dfrac{\partial^2}{\partial z^2}$; x, y, z — декартовы координаты в пространстве. Всюду в работе все величины вещественны.

Перейдем в уравнении (1) к эллипсоидальным координатам. При использовании в этих координатах метода Фурье для (1) возникает система

$$\sqrt{f(\xi)}\frac{d}{d\xi}\left\{\sqrt{f(\xi)}\frac{d\Lambda}{d\xi}\right\} + \frac{1}{4}q(\xi,h,l)\Lambda = 0, \quad 0 < \xi < 1, \tag{2a}$$

$$\sqrt{-f(\xi)}\frac{d}{d\xi}\left\{\sqrt{-f(\xi)}\frac{d\Lambda}{d\xi}\right\} - \frac{1}{4}q(\xi,h,l)\Lambda = 0, \quad 1 < \xi < \rho^2, \tag{2b}$$

$$\sqrt{f(\xi)}\frac{d}{d\xi}\left\{\sqrt{f(\xi)}\frac{d\Lambda}{d\xi}\right\} + \frac{1}{4}q(\xi,h,l)\Lambda = 0, \quad \rho^2 < \xi < \infty. \tag{2c}$$

Здесь $f(\xi) = \xi(\xi-1)(\xi-\rho^2)$, $q(\xi,h,l) = h\rho^2 - l\rho^2\xi + \omega^2\xi^2$, $1 < \rho$ — число, определяемое значениями параметров выбранной эллипсоидальной координатной системы,

[1] Работа выполнена при поддержке Российского фонда фундаментальных исследований (проект N 96-01-00951).

ω зависит от этих параметров и значения k; h и l — неизвестные константы разделения. Это так называемая алгебраическая форма волнового уравнения Ламе. Мы рассматриваем искомую функцию $\Lambda(\xi)$ на $(0,1)\cup(1,\rho^2)\cup(\rho^2,\infty)$.

В дополнение к (2) мы получаем также граничные условия

$$\lim_{\xi\to+0}\Lambda(\xi)=0 \qquad \text{или} \qquad \lim_{\xi\to+0}\left(\sqrt{f(\xi)}\Lambda'(\xi)\right)=0, \tag{3a}$$

$$\lim_{\xi\to1\pm0}\Lambda(\xi)=0 \qquad \text{или} \qquad \lim_{\xi\to1\pm0}\left(\sqrt{\mp f(\xi)}\Lambda'(\xi)\right)=0, \tag{3b}$$

$$\lim_{\xi\to\rho^2\pm0}\Lambda(\xi)=0 \qquad \text{или} \qquad \lim_{\xi\to\rho^2\pm0}\left(\sqrt{\pm f(\xi)}\Lambda'(\xi)\right)=0. \tag{3c}$$

Функции $\Lambda(\xi)$, рассматриваемые на $(0,1)\bigcup(1,\rho^2)$, дают так называемые угловые эллипсоидальные волновые функции (у.э.в.ф.), на (ρ^2,∞) — радиальные. Уравнения (2) имеют особенности в точках 0, 1, и ρ^2. Несложной заменой независимого переменного легко избавиться от этих особенностей.

О трудностях, связанных с вычислением этих функций, см., например, [9]; за исключением небольшого числа простейших случаев, приведенных в этой работе, расчеты для у.э.в.ф. практически отсутствовали. В то же время волновые эллипсоидальные функции столь же важны при решении задач математической физики, рассматриваемых внутри или вне эллипсоида, как и функции Лежандра и Бесселя для задач, рассматриваемых внутри или вне сферы. Применение волновых эллипсоидальных функций задержано именно трудностями их вычисления (и исследования, конечно). Отметим важный вклад М. В. Федорюка в развитие теории этих функций.

Мы рассматриваем сейчас двухпараметрическую спектральную задачу для у.э.в.ф.: уравнения (2a) и (2b) с граничными условиями (3a), (3b), (3c) ((3c) только при $\xi\to\rho^2-0$). Изложим метод вычисления неизвестных нам спектральных параметров h и l. Возьмем для определенности левую колонку граничных условий (3). Обозначим: k_1 — число нулей $\Lambda(\xi)$ на $(0,1)$, k_2 — на $(1,\rho^2)$, $\sigma=h-l$, $\tau=l$. Зафиксировав k_1 и k_2, получаем в плоскости σ,τ две спектральные кривые: кривая (a) содержит точки, координаты которых дают нетривиальное решение (2a), кривая (b) содержит точки, координаты которых дают нетривиальное решение (2b). Можно показать, что на (a) имеет место $\dfrac{d\tau}{d\sigma}<0$ и на (b) имеет место $\dfrac{d\tau}{d\sigma}>0$. Далее мы излагаем несколько упрощенно рекомендуемые вычисления. Осуществив преобразование Прюфера $\sqrt{\pm f(\xi)}\dfrac{\Lambda'}{\Lambda}=\operatorname{ctg}\theta$, получаем для $\theta(\xi)$ уравнение первого порядка на $(0,1)$ и на $(1,\rho^2)$. Граничные условия (3) с учетом взятых k_1 и k_2 дают

$$\theta(+0)=0,\ \theta(1-0)=(k_1+1)\,\pi,\ \theta(1+0)=0,\ \theta(\rho^2-0)=(k_2+1)\,\pi.$$

Взяв какие–либо σ и τ, решим (численно) для $\theta(\xi)$ две задачи Коши: уравнение на $(0,1)$ с граничным условием в $+0$ и уравнение на $(1,\rho^2)$ с граничным условием в $1+0$, получаем два значения

$$\varphi_a(\sigma,\tau)=\theta(1-0)-(k_1+1)\,\pi,\quad \varphi_b(\sigma,\tau)=\theta(\rho^2-0)-(k_2+1)\,\pi.$$

Знаки φ_a и φ_b дают важную информацию о расположении взятой точки (σ, τ) относительно кривых (a) и (b). Можно доказать, что в одном из получившихся криволинейных квадрантов $\varphi_a < 0$ и $\varphi_b < 0$, в другом $\varphi_a < 0$ и $\varphi_b > 0$, в третьем $\varphi_a > 0$ и $\varphi_b > 0$, в четвертом $\varphi_a > 0$ и $\varphi_b < 0$.

Это дает возможность, отправляясь от любого начального приближения для (σ, τ), заключить искомую точку пересечения (a) и (b) в прямоугольник со сторонами, параллельными осям σ и τ, и далее вести процесс так: берем середину прямоугольника, решаем две задачи Коши, определяем по знакам φ_a и φ_b, в каком из криволинейных квадрантов лежит эта середина, разрезаем прямоугольник на две части прямой, проходящей через эту середину и параллельной оси σ или оси τ, отбрасываем пустую половину. В среднем мы за два шага уполовиним каждый из размеров прямоугольника. Цена одного шага — две задачи Коши для одного уравнения первого порядка. Тем самым этот процесс по числу операций только в четыре раза дороже часто употребляемого метода деления отрезка спектра пополам для однопараметрической спектральной задачи Штурма–Лиувилля. Отметим, что в процессе решения мы не вычисляем координат каких-либо точек, лежащих на (a) или (b), такое вычисление потребовало бы на каждом шаге решения однопараметрической спектральной задачи и стоило бы слишком дорого. Мы не останавливаемся на задаче вычисления у.э.в.ф., нормированных нужным образом, на вычислении $\Lambda(\xi)$ в пределах (ρ^2, ∞) и на многом другом. Отметим только одну трудность (на методах преодоления которой мы также не останавливаемся) — возможность появления в уравнениях (2) больших величин (из-за сильной сплюснутости или вытянутости эллипсоида, из-за больших значений ω^2, k_1 и/или k_2 и т. п.).

Разработанные методы вычисления эллипсоидальных волновых функций реализованы в виде пакета программ. Эти программы работоспособны для широкого диапазона используемых величин и проявили себя как очень эффективные даже при наличии указанных выше трудностей.

Работы [2,3] как содержащие алгоритмы вычисления эллипсоидальных волновых функций включены в международный обзор [10].

Упомянутые методы вычисления эллипсоидальных волновых функций были применены к решению известной задачи дифракции плоской звуковой волны на эллипсоиде (см. [11,12], а также статью *A. A. Abramov, A. L. Dyshko, N. B. Konyukhova, and T. V. Levitina* настоящего сборника). О некоторых полезных свойствах у.э.в.ф. смотри [13].

§2. Обобщение метода (см. [2,4,5,6])

1. Рассмотрим систему двух обыкновенных дифференциальных уравнений второго порядка

$$(p_i(x_i)y_i')' + q_i(x_i, \lambda, \mu)y_i = 0 \quad x_i \in [0,1],\ i = 1,2. \tag{1}$$

Здесь $p_i > 0$ на $[0,1]$, спектральные параметры λ, μ входят в (1), вообще говоря, нелинейно.

Дополним эту систему граничными условиями следующего (для простоты) вида

$$y_i(0) = y_i(1) = 0, \quad i = 1,2. \tag{2}$$

Пусть q_1 возрастает по λ и μ, q_2 возрастает по λ и убывает по μ. Тогда при соответствующих аккуратных оговорках тот метод решения двухпараметрической спектральной задачи, который был изложен в §1, без каких-либо существенных изменений применим и к решению задачи (1)&(2) [4]. В частном случае, когда

$$q_i(x_i) = Q_i(x_i) + \lambda P_{i1}(x_i) + \mu P_{i2}(x_i), \quad i = 1, 2,$$

достаточным условием применимости метода §1 при непрерывных $Q_i(x_i)$ и $P_{ij}(x_i)$ является требование $\det ||P_{ij}(x_i)|| \neq 0$ для всех $(x_1, x_2) \in [0,1] \times [0,1]$.

2. Метод §1 обобщается в [5] на некоторые системы трех уравнений с линейным вхождением спектральных параметров

$$y_i'' + \Big(q_i(x_i) + \sum_{j=1}^{3} \lambda_j p_{ij}(x_i)\Big) y_i = 0, \quad 0 \leq x_i \leq 1,$$

$$y_i(0) = y_i(1) = 0, \quad i = 1, 2, 3.$$

Предполагается, что $\det ||p_{ij}(x_i)|| \neq 0$ в $\Omega = [0,1] \times [0,1] \times [0,1]$. При выполнении последнего предположения, как известно, для спектральных параметров существует невырожденная линейная замена с постоянной матрицей такая, что знаки элементов p_{ij} новой матрицы в Ω такие, как в таблице

$$\left\| \begin{array}{ccc} + & - & - \\ + & + & - \\ + & + & + \end{array} \right\|.$$

Такие знаки коэффициентов при спектральных параметрах обеспечивают нужную монотонность уравнений относительно этих параметров.

Если дополнительно предположить, что все элементы матрицы обратной к новой матрице отличны от нуля в Ω и знаки их даются той же таблицей, то удается построить метод, обобщающий метод §1 и обладающий теми же свойствами: собственное значение (с.з.) заключается в параллелепипед со сторонами, параллельными координатным плоскостям, и затем уточняется отбрасыванием пустых частей параллелепипеда. Отметим существенное усложнение возникающей ситуации: среди восьми областей, выделяемых спектральными поверхностями, имеются две области такие, что результаты решения задач Коши для фиксированного набора спектральных параметров не дают основания для уменьшения указанного параллелепипеда.

3.В [2,6] рассмотрены n-параметрические задачи, возникающие при использовании метода Фурье в его простейшей форме. Мы имеем задачу на с.з.:

$$Lu = \lambda p u, \quad u \not\equiv 0,$$

где u — искомая собственная функция (с.ф.), функция N независимых переменных, λ — искомое с.з., L и p — заданные дифференциальные операторы. Рассмотрим случай $L = L_1 + \ldots + L_N$, $p = p_1 + \ldots + p_N$, где L_i и p_i — операторы по i-ому переменному. При использовании метода Фурье получим систему

$$L_i u_i = \lambda p_i u_i + \alpha_i u_i, \quad u_i \not\equiv 0, \quad i = 1, \ldots, N, \tag{3}$$

где u_i зависит только от i-го независимого переменного, $\alpha_1, \ldots, \alpha_N$ — неизвестные константы разделения, имеет место равенство

$$\alpha_1 + \ldots + \alpha_N = 0. \tag{4}$$

Мы предполагаем, что операторы L_i и p_i самосопряженные, p_i — положительно определенные.

Важным вопросом является вопрос идентификации отдельных решений этой задачи. Для простоты ограничимся частным случаем, когда L_i — дифференциальный оператор второго порядка без особенностей, а p_i — умножение на функцию (положительную). Пусть k_i — число нулей u_i на соответствующем интервале, $k_i \geq 0$. Тогда для каждого набора $\{k_1, \ldots, k_N\}$ существует, и притом единственный соответствующий ему набор $\{\lambda, \alpha_1, \ldots, \alpha_N\}$, дающий решение системы (3)&(4). Предложен и исследован глобально сходящийся (т.е. сходящийся при любых начальных приближениях) итерационный метод вычисления $\lambda, \alpha_1, \ldots, \alpha_N$ при заданных $k_1, \ldots, k_N$.

§3. Численное решение задачи о свободных акустических колебаниях внутри трехосного эллипсоида (см. [7])

Пусть трехосный эллипсоид S ограничивает область D. Нужно найти с.з. k^2 задачи

$$\Delta u + k^2 u = 0 \quad \text{в } D,$$

$$u|_S = 0 \quad \text{(идеально мягкий эллипсоид)},$$

или

$$\left.\frac{\partial u}{\partial n}\right|_S = 0 \quad \text{(идеально жесткий эллипсоид)}.$$

Задача решалась в эллипсоидальной системе координат. Здесь возникает следующая трехпараметрическая спектральная задача. Рассматриваются уравнения (2) из §1: одно на $(0,1)$, другое на $(1,\rho^2)$, третье на (ρ^2, ξ_*), где ξ_* определяется формой эллипсоида. Уравнения дополняются граничными условиями (3) из §1; в ξ_* ставится одно из двух условий

$$\Lambda(\xi_*) = 0 \quad \text{или} \quad \Lambda'(\xi_*) = 0.$$

Здесь неизвестны k^2 и две константы разделения.

Используя метод, сходный с упомянутым в п.2 §2, удалось провести детальные расчеты и вычислить собственные частоты колебаний акустической среды внутри эллипсоида при различных значениях параметров задачи.

§4. Метод решения n–параметрической спектральной задачи, основанный на введении параметра (см. [8])

Рассмотрим систему уравнений

$$(p_i(t_i)y_i')' + \Big(q_i(t_i) + \sum_{j=1}^{n} a_{ij}(t_i)\lambda_j\Big)y_i = 0, \quad \alpha_i \leq t_i \leq \beta_i, \quad i = 1, \ldots, n. \tag{1}$$

Предполагаем, что $p_i > 0$ на $[a_i, \beta_i]$ для всех i. Уравнения (1) дополняем граничными условиями вида (для простоты)

$$y_i(\alpha_i) = y_i(\beta_i) = 0. \tag{2}$$

Задаем целые неотрицательные числа $k_1, \ldots, k_n$. Ставится задача: найти числа $\lambda_1, \ldots, \lambda_n$ такие, что справедливо:

$$y_i \text{ имеет на } (\alpha_i, \beta_i) \text{ ровно } k_i \text{ нулей}, \quad i = 1, \ldots, n. \tag{2'}$$

Известно, что если $D(t_1, \ldots, t_n) = \det \|a_{ij}\| > 0$ при $\alpha_i \le t_i \le \beta_i$, $i = 1, \ldots, n$, то задача (1)&(2)&(2') имеет и притом единственное решение.

Мы накладываем более слабые ограничения:

$$D(t_1, \ldots, t_n) \ge 0, \tag{3}$$

существуют $t_1^0, \ldots, t_n^0$ такие, что

$$D(t_1^0, \ldots, t_n^0) > 0. \tag{4}$$

Возможность: D неотрицательно, но обращается кое-где (не везде) в 0 – реализуется в некоторых прикладных задачах.

При ограничениях (3)&(4) уже невозможно гарантировать существование решения задачи (1)&(2)&(2'). Мы предполагаем, что рассматриваемая задача, удовлетворяющая лишь (3)&(4), имеет решение, его единственность доказывается.

В случае, если $a_{ij}(t_i) = a_{ij}^0 = \text{const}$ для всех i, j, задача (1)&(2)&(2'), очевидно, решается легко. Для перехода к исходной общей задаче используем метод движения по параметру (см., например, [14]). При $0 \le \sigma \le 1$ определяем $\tilde{a}_{ij}(t_i, \sigma) = \sigma\, a_{ij}(t_i) + (1 - \sigma) a_{ij}(t_i^0)$, где $t_1^0, \ldots, t_n^0$ удовлетворяют (4). Используя преобразование Прюфера, получаем при $0 \le \sigma \le 1$ из (1)&(2)&(2') систему дифференциальных уравнений первого порядка с граничными условиями в α_i и β_i и параметрами $\tilde{\lambda}_1(\sigma), \ldots, \tilde{\lambda}_n(\sigma)$. Эта система решается приближенно шагами по σ на сетке $0 = \sigma_0 < \sigma_1 < \sigma_2 < \ldots < \sigma_N = 1$; для перехода к приближенным значениям $\tilde{\lambda}_1, \ldots, \tilde{\lambda}_n$ в следующей точке используется одна из формул, рекомендованных в [14].

Для выбранного способа введения параметра и выбранной формулы реализации одного шага по σ дается при произвольном n полное обоснование метода: доказано, что не возникает тупиковых ситуаций и что погрешность ответов получается $0(h^2)$, где $h = \max\limits_m (\sigma_{m+1} - \sigma_m)$.

Метод был опробован на двух- и трехпараметрических задачах, описанных в §1 и §3. Для получения высокой точности в рассматриваемых вариантах число N (главный показатель затрачиваемой работы) оказалось небольшим, как правило порядка десяти.

Литература

[1] Volkmer H., Multiparameter eigenvalue problems and expansion theorems (Springer–Verlag, Berlin–Heidelberg), (1988).

[2] А. А. Абрамов, А. Л. Дышко, Н. Б. Конюхова, Т. В. Левитина, Ж. вычисл. матем. и матем. физ. **29** (6), 813 (1989).
[3] А. А. Абрамов, А. Л. Дышко, Н. Б. Конюхова, Т. В. Левитина, Ж. вычисл. матем. и матем. физ. **31** (2), 212 (1991).
[4] Т. В. Левитина, Ж. вычисл. матем. и матем. физ. **31** (5), 689 (1991).
[5] T. Levitina, Numer. Anal. and Math. Modelling (Banach Center Publs , Warszawa) **29**, 275 (1994).
[6] А. А. Абрамов, Ж. вычисл. матем. и матем. физ. **34** (10), 1524 (1994).
[7] T.Levitina, Design Engineering Technical Conferences (ASME, New–York) **3**, Part B, 449 (1995).
[8] А. А. Абрамов, В. И. Ульянова, Вычисл. матем. и информатика. Труды научной сессии, посвященной 85-летию академика А.А.Дородницына (ВЦ РАН, Москва), 13 (1996).
[9] F. M. Arscott, P. J. Taylor, and R. V. M. Zahar, Math. Comput. **40** (161), 367 (1983).
[10] D. W. Lozier, and F. W. Olver, Proc. of Symposia in Appl. Math. (ed. W. Gautschi, AMS, Providence) **48**, 79 (1994).
[11] А. А. Абрамов, А. Л. Дышко, Н. Б. Конюхова, Т. В. Левитина, Ж. вычисл. матем. и матем. физ. **35** (9), 1374 (1995).
[12] A. A. Abramov, N. B. Konyukhova, and T. V. Levitina, Design Engineering Technical Conferences (ASME, New–York) **3**, Part B, 437 (1995).
[13] T. Levitina, and E. J. Brändas, Int. J. Quant. Chem. Symp. **30**, 5 (1996).
[14] Дж. Ортега, В. Рейнболдт, Итерационные методы решения нелинейных систем уравнений со многими неизвестными (Мир, Москва), (1975).

ПОВЕДЕНИЕ РЕШЕНИЙ СИНГУЛЯРНО ВОЗМУЩЕННОГО ДИФФЕРЕНЦИАЛЬНОГО УРАВНЕНИЯ БЕСКОНЕЧНОГО ПОРЯДКА

И.Е.Амирханов,С.А.Васильев,Е.П.Жидков,И.Е.Жидкова

141980 ЛВТА ОИЯИ

Дубна Московской области, Россия

Аннотация. В работе рассматривается уравнение Кадышевского (релятитистский аналог уравнения Шредингера):

$$[2c\sqrt{Q^2+m^2c^2} - H_0^{rad} - V(r)]\Psi_{ql}(r) = 0,$$

где

$$H_0^{rad} = 2mc^2 \mathrm{ch}\left(\frac{i\hbar}{mc}\frac{d}{dr}\right) + \frac{\hbar^2 l(l+1)}{mr\left(r+\frac{i\hbar}{mc}\right)} exp\left(\frac{i\hbar}{mc}\frac{d}{dr}\right).$$

Для этого уравнения исследуются способы постановки краевой задачи. Приводятся результаты численных расчетов по решению задачи на собственные функции и собственные значения для дифференциального уравнения восьмого порядка с малым параметром при старших производных.

1 ВВЕДЕНИЕ

Существует несколько подходов к проблеме описания релятивистских двухчастичных систем в квантовой теории поля. В 60-е годы Логуновым А.А. и Тавхелидзе А.Н. был развит квазипотенциальный формализм в рамках этой задачи, в котором сочетался полностью ковариантный подход, вероятностная интерпретация волновой функции в духе квантовой механики и трехмерность [1]. В этом подходе волновая функция высупает как непосредственное обобщение нерелятивистской волновой функции, поскольку зависит только от одного временного аргумента и подчиняется уравнению типа Шредингера.

В этом направлении Кадышевским В.Г. был получен релятивистский аналог уравнения Шредингера [2]. Для радиальной волновой функции $\Psi_{ql}(r)$, описывающей относительное движение двух скалярных частиц с импульсом q, угловым моментом l и равными массами, одномерное релятивистское уравнение будет иметь следующий вид:

$$[2c\sqrt{Q^2+m^2c^2} - H_0^{rad} - V(r)]\Psi_{ql}(r) = 0,$$

$$H_0^{rad} = 2mc^2 \mathrm{ch}\left(\frac{i\hbar}{mc}\frac{d}{dr}\right) + \frac{\hbar^2 l(l+1)}{mr\left(r+\frac{i\hbar}{mc}\right)} exp\left(\frac{i\hbar}{mc}\frac{d}{dr}\right),$$

где $V(r)$ - квазипотенциал, c - скорость света.

В работе [3] показано, что при формальном устремлении $c \to \infty$, решения уравнения Кадышевского регулярно сходятся к решениям уравнения Шредингера.

Если ограничится случаем S-волны ($l = 0$) и разложить оператор $ch\left(\frac{i\hbar}{mc}\frac{d}{dr}\right)$ в ряд, то уравнение Кадышевского будет иметь следующий вид:

$$\left[\frac{d^2}{dr^2} - \varepsilon\frac{2}{4!}\frac{d^4}{dr^4} + \varepsilon^2\frac{2}{6!}\frac{d^6}{dr^6} - \ldots - \frac{m}{\hbar^2}\left(V(r) - 2c\sqrt{Q^2 + m^2c^2}\right)\right]\Psi_{q0}(r) = 0,$$

где $\varepsilon = \left(\frac{\hbar}{mc}\right)^2$ – малый параметр.

Это уравнение представляет собой дифференциальное уравнение бесконечного порядка с малым параметром при старших производных, что позволяет отнести его к классу сингулярно возмущенных уравнений.

Трудности, возникающие при решении подобного рода уравнений, связаны с тем, что здесь невозможно применить обычную схему разложения в степенной ряд по малому параметру, ведь если формально устремить $\varepsilon \to 0$, то порядок уравнения изменится, а решения такого "упрощенного" уравнения не смогут удовлетворять всем условиям, которые налагались на решения исходного уравнения. Некоторые из этих условий окажутся лишними. Таким образом, здесь мы имеем дело с сингулярно возмущенным дифференциальным уравнением.

2 ИССЛЕДОВАНИЕ МОДЕЛЬНОЙ ЗАДАЧИ

Если ограничиться лишь 4–ой производной в уравнении Кадышевского и выбрать следующий потенциал:

$$V(r) = 0, \quad r \le \pi,$$
$$V(r) = \infty, \quad r > \pi,$$

тогда уравнение можно записать в таком виде:

$$\left(-\varepsilon\frac{d^4}{dr^4} + \frac{d^2}{dr^2} + q^2\right)\psi(r) = 0. \qquad (1)$$

Решение такого уравнения можно представить в виде $\psi = expikr$. После его подстановки в (1), получим следующее биквадратное уравнение:

$$\varepsilon k^4 + k^2 - q^2 = 0.$$

Корни этого уравнения имеют вид:

$$k_{1,2} = \pm i\sqrt{\frac{\sqrt{1 + \varepsilon q^2} - 1}{2\varepsilon}}; \qquad k_{3,4} = \pm\sqrt{\frac{\sqrt{1 + \varepsilon q^2} + 1}{2\varepsilon}}.$$

При $\varepsilon \to 0$ они ведут себя таким образом:

$$k_{1,2} \to \pm iq, \qquad k_{3,4} \to \frac{1}{\sqrt{\varepsilon}}.$$

Общее решение уравнения (1) можно записать как

$$\psi(r) = C_1 \sin k_1 r + C_2 \cos k_2 r + C_3 \mathrm{sh} k_3 r + C_4 \mathrm{ch} k_4 r,$$

причем коэффициенты C_1, C_2, C_3, C_4 определяются из конкретных краевых условий.

Если для этой задачи выбрать симметричные краевые условия вида:

$$\begin{aligned} \psi(0) &= 0, \qquad \psi(\pi) = 0, \\ \psi''(0) &= 0, \qquad \psi''(\pi) = 0, \end{aligned}$$

тогда уравнение для нахождения собственных значений будет:

$$f(q, \varepsilon) = \sin k_1 \pi = 0,$$

а его корни -

$$q_n^2 = n^2(1 + \varepsilon n^2).$$

Если положить в последнем выражении $\varepsilon = 0$, то получим набор собственных значений "вырожденной" задачи ($q_n^2 = n$). Надо заметить, что при несимметричных краевых условиях такой предельный переход не имеет места. Эти соображения дают повод выбрать именно симметричные краевые условия при исследовании нашей задачи.

3 РЕЗУЛЬТАТЫ ЧИСЛЕННЫХ ИССЛЕДОВАНИЙ

Для численных исследований были выбраны следующие дифференциальные уравнения со следующими краевыми условиями:

$$\begin{cases} -\dfrac{\varepsilon}{12}\dfrac{d^4\psi_{IV}}{dr^4} + \dfrac{d^2\psi_{IV}}{dr^2} + q_{IV}^2\psi_{IV} = 0, \\ \psi_{IV}(0) = 0, \qquad \psi_{IV}(\pi) = 0, \\ \psi''_{IV}(0) = 0, \qquad \psi''_{IV}(\pi) = 0, \end{cases}$$

$$\begin{cases} \dfrac{\varepsilon^2}{360}\dfrac{d^6\psi_{VI}}{dr^6} - \dfrac{\varepsilon}{12}\dfrac{d^4\psi_{VI}}{dr^4} + \dfrac{d^2\psi_{VI}}{dr^2} + q_{VI}^2\psi_{VI} = 0, \\ \psi_{VI}(0) = 0, \qquad \psi_{VI}(\pi) = 0, \\ \psi''_{VI}(0) = 0, \qquad \psi''_{VI}(\pi) = 0, \\ \psi^{(4)}_{VI}(0) = 0, \qquad \psi^{(4)}_{VI}(\pi) = 0, \end{cases}$$

$$\begin{cases} -\dfrac{\varepsilon^3}{20160}\dfrac{d^8\psi_{VIII}}{dr^8} + \dfrac{\varepsilon^2}{360}\dfrac{d^6\psi_{VIII}}{dr^6} - \dfrac{\varepsilon}{12}\dfrac{d^4\psi_{VIII}}{dr^4} + \dfrac{d^2\psi_{VIII}}{dr^2} + q_{VIII}^2\psi_{VIII} = 0, \\ \psi_{VIII}(0) = 0, \qquad \psi_{VIII}(\pi) = 0, \\ \psi''_{VIII}(0) = 0, \qquad \psi''_{VIII}(\pi) = 0, \\ \psi^{(4)}_{VIII}(0) = 0, \qquad \psi^{(4)}_{VIII}(\pi) = 0, \\ \psi^{(6)}_{VIII}(0) = 0, \qquad \psi^{(6)}_{VIII}(\pi) = 0. \end{cases}$$

Эти задачи решались численно на интервале $[0, \pi]$ с шагом $h = 0.0157(796$ методом матричной прогонки.

Были получены следующие результаты для $\Delta\psi_N$ при различных ε, где

$$\Delta\psi_N = \max ||\psi_{Nn}(\varepsilon, r) - \psi_n^{schr}||,$$

а ψ_n^{schr} - собственные функции вырожденной задачи.

ε	$\Delta\psi_{IV}$	$\Delta\psi_{VI}$	$\Delta\psi_{VIII}$
1.0	0.000358713	0.000395262	0.000404331
0.1	0.000071391	0.000077381	0.000079648
0.01	0.000010428	0.000012284	0.000016074
0.001	0.000005691	0.000003759	0.000001611

Для нескольких первых собственных значений имеем

ε	n	q^2_{IV}	q^2_{VI}	q^2_{VIII}
1.0	1	1.081520	1.084354	1.084276
	2	5.297297	5.471528	5.483795
	3	15.507309	17.438434	17.747194
0.1	1	1.008163	1.008188	1.008184
	2	4.130520	4.132255	4.132193
	3	9.659969	9.679848	9.679782
0.01	1	1.000814	1.000827	1.000547
	2	4.013033	4.013046	4.081444
	3	9.065955	9.065967	9.075564
0.001	1	1.000079	1.000089	1.000068
	2	4.001272	4.001292	4.033378
	3	9.006441	9.006839	9.054243

4 ВЫВОД

Из приведенных результатов можно сделать вывод о сходимости собственных функций (в выбранной норме) и собственных значений для сформулированных выше задач к решению вырожденной задачи. Это свидетельствует о том, что выбор симметричных краевых условий для дифференциальных уравнений высокого порядка - оправдан.

Литература

[1] Logunov A.A.,Tavkhelidze A.N. Nuovo Cimento, **v.29**, 380 (1963).

[2] Кадышевский В.Г., Мир-Касимов Р.М.,Скачков Н.Б. ЭЧАЯ, **т.2**,Вып.3,с.636-690,(1971).

[3] Жидков Е.П.,Кадышевский В.Г. ТМФ,**т.3**, N 2, (1970).

О некоторых нелинейных сингулярных задачах в моделях релятивистской космологии [1]

Воронов Н.А.*, Дышко А.Л., Конюхова Н.Б., Староверова И.Б.

Вычислительный центр РАН, Вавилова 40, 117967 Москва GSP-1, Россия
e-mail: nadja@ccas.ru
**Институт теоретической и экспериментальной физики, Б.Черемушкинская 25, 117259 Москва, Россия*

Аннотация. Дан краткий обзор [1]-[4]: представлены математические модели и основные результаты качественного анализа нелинейных обыкновенных дифференциальных уравнений второго порядка с вырождениями, описывающих эволюцию в пространствах Минковского, Фридмана и де Ситтера тонкого сферического "пузыря" – локализованного сферически симметричного решения уравнения скалярного поля Хиггса в приближении, когда можно пренебречь толщиной оболочки по сравнению с ее радиусом. Основное отличие в поведении объекта: расширения Вселенной по Фридману недостаточно для того, чтобы при любом начальном радиусе пузырь не схлопывался за конечное время, что аналогично его поведению в плоском пространстве-времени; в пространстве де Ситтера при величине начального радиуса, большем некоторого критического значения, определенного параметрами задачи, пузырь не схлопывается ни за конечное, ни за бесконечное время, т.е. имеет конечный предельный радиус при стремлении временной переменной к бесконечности.

1. *Поле Хиггса в пространстве Минковского* описывается лагранжианом [5]:

$$L = (\partial\varphi)^2/2 - U(\varphi), \qquad U(\varphi) = \lambda^2(\varphi^2 - \nu^2)^2, \tag{1}$$

где $(\partial\varphi)^2 = \eta^{ik}\frac{\partial\varphi}{\partial x^i}\frac{\partial\varphi}{\partial x^k}$, $\eta^{ik} = diag(1,-1,-1,-1)$ – метрический тензор плоского пространства-времени, $i,k = 0,\ldots,3$, $x^0 = t$ (скорость света в вакууме принята за единицу), λ, ν – параметры. Лагранжиану (1) соответствует волновое уравнение

$$\Box\varphi + 4\varphi\lambda^2(\varphi^2 - \nu^2) = 0, \tag{2}$$

где $\Box \equiv \partial^2/\partial t^2 - \Delta$, Δ – оператор Лапласа по переменным x^j, $j = 1,2,3$. Для (2) справедливо: сохраняется интеграл энергии, $E(t) = \int[\delta^{ik}\frac{\partial\varphi}{\partial x^i}\frac{\partial\varphi}{\partial x^k}/2 + U(\varphi)]dV \equiv const$, где δ^{ik} – символ Кронекера, а интеграл берется по всему R^3 (подынтегральное выражение – гамильтониан, или плотность энергии); постоянные $\varphi = \pm\nu$ – решения с нулевой энергией (значения $\varphi = \pm\nu : U(\pm\nu) = 0$ называют "вакуумами" одинаковой "глубины", а поле (1) – скалярным полем с двумя вырожденными вакуумами); 3) существует стационарное плоское решение $\varphi_c(\vec{r}) = \nu\tanh[\lambda\nu\sqrt{2}(\vec{n},\vec{r}-\vec{r}_0)]$, $\vec{r},\vec{r}_0 \in R^3$, порождающее бегущую уединенную волну $\varphi_c(\vec{r},t) = \nu\tanh\{\lambda\nu\sqrt{2}[(\vec{n},\vec{r}-\vec{r}_0) \pm v_0(t-t_0)]/\sqrt{1-v_0^2}\}$, $t \geq t_0$, если задана начальная скорость $v_0 : 0 < v_0 < 1$; здесь $\vec{n} = (n_1,n_2,n_3)$ – вектор единичной длины, характеризующий направление движения фронта волны. Решение $\varphi_c(\vec{r})$, названное в [6] "стенкой", обладает поверхностной плотностью энергии $\mu = 4\sqrt{2}\lambda\nu^3/3$,

[1] Работа поддержана Российским фондом фундаментальных исследований (проекты N95-02-04681 и N96-01-00951).

сосредоточенной в области толщиной $l = [\sqrt{2}\lambda\nu]^{-1}$ (величина μ получается вычислением для $\varphi_c(\vec{r})$ одномерного интеграла энергии от $-\infty$ до $+\infty$). В [6], [7] при изучении космологических следствий локализованных решений (2) было введено понятие пузыря как односвязной области, окруженной стенкой, т.е. объекта, вне которого поле φ экспоненциально стремится к одному из вакуумов, внутри – к другому, а почти вся энергия сосредоточена в оболочке пузыря. Для сферически симметричного пузыря поле в начальный момент состояния покоя задается в виде $\varphi(r,0) = \nu\tanh[\lambda\nu\sqrt{2}(r - R_0)]$, $R_0 > 0$, а эволюция поля описывается начально-краевой задачей:

$$\frac{\partial^2\varphi}{\partial t^2} - \frac{\partial^2\varphi}{\partial r^2} - \frac{2}{r}\frac{\partial\varphi}{\partial r} + 4\lambda^2\varphi(\varphi^2 - \nu^2) = 0, \quad t \geq 0, \quad 0 < r < \infty. \tag{3а}$$

$$\varphi(r,0) = \nu\tanh\left(\lambda\nu\sqrt{2}(r - R_0)\right), \quad \varphi_t(r,0) = 0 \quad \forall r \geq 0. \tag{3б}$$

$$|\lim_{r\to 0}\varphi(r,t)| < \infty, \quad \lim_{r\to\infty}\varphi(r,t) = \nu \quad \forall t \geq 0. \tag{3в}$$

В численно-аналитических исследованиях [8]-[10] задачи (3) обнаружень пульсация пузыря с излучением энергии на бесконечность и образование вблизи вакуума $\varphi = \nu$ долгоживущих локализованных осциллирующих состояний, устойчивость которых изучалась в [11].

2. *В расширяющейся Вселенной* используем равенство $S = S_M + S_H + \Re/(16\pi G)$, где S – общее действие для гравитирующих макроскопической материи и скалярного поля [12]. Здесь S_M – действие для материи, определяющей вид расширяющейся Вселенной; $\Re$ – скалярная кривизна пространства, G - ньютоновская гравитационная постоянная; S_H – действие для поля Хиггса (1) в гравитационном поле, $S_H = \int\sqrt{-g}\left[g^{ik}\varphi_{,i}\varphi_{,k}/2 - \lambda^2(\varphi^2-\nu^2)^2\right]d^4x$ (интеграл берется по всему R^3 и по временной переменной между двумя ее конечными значениями, причем предполагается минимальное взаимодействие поля φ с гравитационным полем g_{ik}. Вариация S по g^{ik} приводит к уравнению Эйнштейна:

$$\Re_{ik} - \Re g_{ik}/2 = 8\pi G\left[T_{Hik} + T_{Mik}\right], \tag{4}$$

где $\Re_{ik}$ - тензор Риччи, T_{Hik} и T_{Mik} – тензоры энергии-импульса поля Хиггса и макроскопической материи соответственно [12]. Варьируя S по φ, получаем:

$$\left(\sqrt{-g}\right)^{-1}\left(\sqrt{-g}g^{ik}\varphi_{,i}\right)_{,k} + 4\lambda^2\varphi(\varphi^2 - \nu^2) = 0. \tag{5}$$

Уравнения (4), (5) в совокупности с уравнением движения материи определяют полную систему уравнений для задачи трансцендентной сложности. В [1]-[4] делается физически оправданное предположение, что в (4) можно пренебречь компонентами T_{Hik} по сравнению с величинами T_{Mik}. Тогда исходная задача распадается на две: из (4) находим фоновую метрику; (5) определяет эволюцию скалярного поля Хиггса на фоне найденной метрики.

В пространствах Фридмана источник гравитационного поля – идеальный газ: $T_{Mik} = (p + \rho)u_iu_k - pg_{ik}$, где u^i, p, ρ – соответственно "четырехскорость", давление и плотность газа; p, ρ не зависят от x^j, $j = 1,2,3$ (изотропная модель). Используя по пространственным переменным сферические координаты и вводя новые "время" η и "радиус" χ по формулам $dt = a(\eta)d\eta$, $r = a(\eta)\sinh\chi$, где $a(\eta)$

– пока неизвестная неотрицательная функция, получаем из (4) метрику Фридмана: $ds^2 = a^2(\eta)[d\eta^2 - d\chi^2 - \sinh^2\chi(d\theta^2 + \sin^2\theta d\phi^2)]$. В двух случаях $a(\eta)$ получено явно [12]: а) $a(\eta) = a_0(\cosh\eta - 1)$, $0 < a_0 = const$, если $p = 0$ (газ без давления или пыль); б) $a(\eta) = a_0\sinh\eta$, $0 < a_0 = const$, если $p = \rho/3$ (ультрарелятивистский газ). В пространствах с метрикой Фридмана из (5) получаем задачу для поля сферического пузыря, образованного в начальный момент стенкой:

$$\varphi_{\eta\eta} + 2\varphi_\eta\dot{a}/a - \varphi_{\chi\chi} - 2\varphi_\chi\coth\chi + 4a^2\lambda^2\varphi(\varphi^2 - \nu^2) = 0,\ \eta \geq \eta_0,\ 0 < \chi < \infty. \quad (6а)$$

$$\varphi(\chi, \eta_0) = \nu\tanh\{a(\eta_0)\lambda\nu\sqrt{2}(\chi - X_0)\},\quad \varphi_\eta(\chi, \eta_0) = 0 \quad \forall\chi \geq 0, \quad (6б)$$

$$|\lim_{\chi\to 0}\varphi(\chi,\eta)| < \infty,\quad \lim_{\chi\to\infty}\varphi(\chi,\eta) = \nu \qquad \forall\eta \geq \eta_0 > 0, \quad (6в)$$

где $0 < X_0$ – фиксировано. В таком виде задача (6) не исследовалась. Дополнительные трудности для ее численного решения по сравнению с (3) связаны с невыполнением закона сохранения энергии (такой закон отсутствует в расширяющейся Вселенной), с вырождением задачи при $\eta_0 \to 0$, с экспоненциальным ростом $a(\eta)$ при $\eta \to \infty$ и др.

В пространстве де Ситтера справедливо $T_{Mik} = \Lambda g_{ik}/2$, где Λ – космологическая постоянная, и решением (4) является метрика $ds^2 = dt^2 - a^2(t)[dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2)]$, где $\Lambda = 3H^2$, $a(t) = a_0\exp(Ht)$, $0 < a_0 = const$, H – постоянная Хаббла [13]. В этом пространстве в сферически симметричном случае (5) имеет вид

$$\varphi_{tt} + 3\varphi_t\dot{a}/a - \varphi_{rr}/a^2 - 2\varphi_r/(a^2 r) + 4\lambda^2\varphi(\varphi^2 - \nu^2) = 0,\ t \geq 0, \qquad 0 < r < \infty. \quad (7а)$$

Для изучения эволюции поля сферического пузыря, который в начальный момент покоился, надо дополнить это уравнение условиями

$$\varphi(r,0) = \nu\tanh\{a_0\lambda\nu\sqrt{2}(r - R_0)\},\qquad \varphi_t(r,0) = 0 \quad \forall r \geq 0, \quad (7б)$$

$$|\lim_{r\to 0}\varphi(r,t)| < \infty,\quad \lim_{r\to\infty}\varphi(r,t) = \nu \quad \forall t \geq 0, \quad (7в)$$

где $0 < R_0$ – фиксировано. В таком виде задача (7) не исследовалась. (О поведении физических полей на фоне метрики де Ситтера см., например, в [14], [15].)

Радиусом $R(t)$ пузыря при фиксированном t называем нуль функции φ как решения задачи (3) (задачи (7)): $\varphi(R(t),t) = 0$, $t \geq 0$, $R(0) = R_0$, $\dot{R}(0) = 0$; для задачи (6): $\varphi(X(\eta),\eta) = 0$, $\eta \geq \eta_0$, $X(\eta_0) = X_0$, $\dot{X}(\eta_0) = 0$.

3. *В приближении тонкой стенки* задачи (3), (6), (7) рассмотрены в [1]-[4]: получены и исследованы нелинейные обыкновенные дифференциальные уравнения второго порядка с вырождениями, описывающие эволюцию $R(t)$ ($X(\eta)$) в приближении, когда можно пренебречь толщиной сферической оболочки по сравнению с ее радиусом. Соответствующие динамические задачи получены как "методом свернутого действия" [16], так и более формализованным способом: при достаточно большом начальном радиусе для решения (3) ((7)) полагаем $\varphi_\Pi(r,t) \approx \nu\tanh[\alpha(r,t)/l]$ (для (6) полагаем $\varphi_\Pi(\chi,\eta) \approx \nu\tanh[\alpha(\chi,\eta)/l]$), где $0 < l$ – мало, α – новая неизвестная функция, $R(t)$: $\alpha(R(t),t) = 0$ ($X(\eta)$: $\alpha(X(\eta),\eta) = 0$); при $l \to 0$ получаем уравнение для $R(t)$ ($X(\eta)$). Ранее в [8], [10] было показано, что такое приближение для задачи (3) очень хорошо описывает стадию сжатия пузыря вплоть до $R \sim l$, хотя само решение становится совсем непохожим на $\varphi_\Pi(r,t)$. Это дает дополнительное основание полагать, что

приближение тонкой стенки представляет интерес для описания поведения изучаемых объектов и в расширяющейся Вселенной.

3.1. *Динамическая задача для тонкого пузыря в пространстве Минковского:*

$$\ddot{R} = -(2/R)[1 - \dot{R}^2], \qquad t \geq 0, \tag{8а}$$

$$R(0) = R_0 > 0, \qquad \dot{R}(0) = 0. \tag{8б}$$

Интерес представляет положительное решение (8), его поведение с ростом t и интервал его существования при различных значениях R_0. Справедливы утверждения: (8а) обладает семейством частных решений $R_\pm(t,c) = \pm t + c$, где c -произвольная постоянная; для любого решения (8а), кроме $R_\pm(t,c)$, справедливо $R^2/\sqrt{|1-\dot{R}^2|} = const$; решение (8) убывает, не выходит из области $R_0 - t \leq R(t) \leq R_0$, $0 \leq t \leq t_{cx}$, где t_{cx} : $t_{cx} > R_0$, $\lim_{t\to t_{cx}} R(t) = 0$, $\lim_{t\to t_{cx}} \dot{R}(t) = -1$, и имеет вид [8]: $R(t) = R_0 \mathrm{cn}\{\sqrt{2}t/R_0, \sqrt{2}/2\}$, где $\mathrm{cn}(x,k)$ – эллиптический косинус с модулем k.

В [3] описано поведение решения (8) при его приближении к нулю (при схлопывании пузыря) и даны оценки $t_{cx} = t_{cx}(R_0)$. В окрестности любой фиксированной точки $t = t_{cx}$ уравнение (8а) имеет однопараметрическое семейство решений, представимое рядом Ляпунова: $R(a, t_{cx}, t) = (t - t_{cx})\{-1 + a(t - t_{cx})^4 + \sum_{k=2}^{\infty} c_k[a(t - t_{cx})^4]^k\}$, где a – параметр, $|a(t - t_{cx})^4|$ мало, $c_2 = -5/6, \ldots$ (c_k не зависят от a и t_{cx}); эти решения при $a > 0$ не выходят из области $0 \leq R(t) \leq t_{cx} - t$, $0 \leq t \leq t_{cx}$. Учитывая два слагаемых ряда и (8б), приближенно получаем $t_{cx} \approx 1.25R_0$. Из точного решения имеем $t_{cx} = R_0 K(\sqrt{2}/2)/\sqrt{2}$, где $K(k)$ – полный эллиптический интеграл первого рода с модулем k. Обозначив $m = (2 - \sqrt{2})/(2 + \sqrt{2})$, из приближений для $K(k)$ получаем более точно $t_{cx} \approx R_0(1 + m)(1 + m^2/4)/(2\sqrt{2}) \approx 1.310867R_0$.

3.2. *Динамическая задача для тонкого пузыря в пространствах Фридмана:*

$$\ddot{X} + [2\coth X + 3\dot{a}(\eta)\dot{X}/a(\eta)][1 - \dot{X}^2] = 0, \qquad \eta \geq \eta_0 > 0, \tag{9а}$$

$$X(\eta_0) = X_0 > 0, \quad \dot{X}(\eta_0) = \dot{X}_0, \quad |\dot{X}_0| < 1. \tag{9б}$$

Здесь рассматривается общая задача, когда в начальный момент пузырь необязательно покоится; вид $a(\eta)$ зафиксирован в п.2 при описании задачи (6) – случаи а) и б). Интерес представляют: положительное решение (9), его поведение с ростом η и интервал его существования для различных значений X_0, $\dot{X}_0$; поведение решения при его продолжении влево к особой точке $\eta = 0$; предельные сингулярные задачи Коши, отвечающие (9) при $\eta_0 = 0$ и $\dot{X}_0 = 0$ (или $\dot{X}_0 = 1$, или $\dot{X}_0 = -1$). Справедливы утверждения: (9а) обладает семейством частных решений $X_l(\eta, c) = l\eta + c$, где c – произвольная постоянная; для любого решения (9а), кроме $X_l(\eta, c)$, справедливо $\frac{d}{d\eta}\left\{\ln\left(\sinh^2 X/\sqrt{|1-\dot{X}^2|}\right)\right\} = -3\dot{X}^2\dot{a}(\eta)/a(\eta) \leq 0 \quad \forall\eta > 0$; $\eta = 0$ – особая точка (9а), так как $\dot{a}(\eta)/a(\eta) = 2/(j\eta) + O(\eta)$, $\eta \to 0$, где $j = 1$ для а) и $j = 2$ для б). Характер вырождения (9а) при $\eta \to 0$ позволяет поставить для него три сингулярные задачи Коши: (9а) при $\eta > 0$, условие $\lim_{\eta\to+0} X(\eta) = X_0$ и условие $\lim_{\eta\to+0} \dot{X}(\eta) = 0$ (*задача 1*), или условие $\lim_{\eta\to+0} \dot{X}(\eta) = 1$ (*задача 2*), или условие $\lim_{\eta\to+0} \dot{X}(\eta) = -1$ (*задача 3*). Следующие утверждения доказаны в [3]:

1) $\forall X_0 > 0$ и $\forall \dot{X}_0$: $|\dot{X}_0| < 1$, решение (9) не выходит из области, ограниченной прямыми $X_1(\eta) = \eta - \eta_0 + X_0$ и $X_2(\eta) = \eta_0 - \eta + X_0$, $\eta \geq 0$, т.е. при $\eta > \eta_0$ оно не выходит из "светового конуса в будущее", а при $\eta < \eta_0$ – из "светового конуса в прошлое".

2) $\forall X_0 > 0$ *и* $\forall \dot{X}_0 : \ -1 < \dot{X}_0 \le 0$, *решение (9) с ростом* η *убывает и* $\exists \eta_{cx} : \ \eta_{cx} > \eta_0 + X_0$, $\lim_{\eta \to \eta_{cx}-0} X(\eta) = 0$, $\lim_{\eta \to \eta_{cx}-0} \dot{X}(\eta) = -1$; *при* $0 < \dot{X}_0 < 1$ *то же утверждение справедливо, начиная с некоторого* $\eta = \tilde{\eta}_0 : \ \dot{X}(\tilde{\eta}_0) = 0$.

3) $\forall X_0 \ne 0$ *решение сингулярной задачи 1 для достаточно малых* $|\eta|$ *существует и единственно, оно голоморфно при* $\eta = 0 : \ X(\eta) = X_0 + X^{(1)}\eta^2 + X^{(2)}\eta^4 + \ldots$, *где* $X^{(1)} = -\coth X_0/p$, $p = 7$ *для а) и* $p = 4$ *для б),... (коэффициенты ряда однозначно определяются из (9а)).* (В [3] для решения $X(\eta)$ *задачи 1* получена оценка выбора $\eta_0 : \ |X(\eta) - X_0 - X^{(1)}\eta^2| < \epsilon, \ 0 < \eta \le \eta_0$, где $0 < \epsilon$ – фиксировано; оценка важна для нахождения решения с заданной точностью вычислений.)

4) $\forall X_0 \ne 0$ *сингулярная задача 2 (задача 3) обладает однопараметрическим семейством решений* $X_+(\eta, c)$ $(X_-(\eta, c))$; *они голоморфны при* $\eta = 0$ *и включают в себя прямую* $X_1(\eta) = X_0 + \eta$ $(X_2(\eta) = X_0 - \eta)$ *при нулевом значении параметра:* $X_\pm(\eta, c, X_0) = X_0 \pm \eta + c\eta[\eta^\lambda + \beta_\pm \eta^{\lambda+1} + \ldots]$, *где* $\lambda = 12$, $\beta_\pm = \pm 13 \coth X_0/7$ *для а) и* $\lambda = 6$, $\beta_\pm = \pm 7 \coth X_0/4$ *для б); с – параметр,* $|c\eta^\lambda|$ *мало (коэффициенты ряда, вообще говоря, зависят от* c *и* X_0*); при* $c < 0$ $(c > 0)$ *справедливо* $X_0 - \eta \le X_\pm(\eta, c, X_0) \le X_0 + \eta$.

5) $\forall \eta_{cx} > 0$ *уравнение (9а) обладает однопараметрическим семейством решений, удовлетворяющих при* $\eta \to \eta_{cx}$ *условиям п.2); эти решения голоморфны при* $\eta = \eta_{cx}$ *и включают в себя прямую* $X_1(\eta) = -\eta + \eta_{cx}$ *при нулевом значении параметра:* $X(\eta, c, \eta_{cx}) = -(\eta - \eta_{cx}) + c\{(\eta - \eta_{cx})^5 + 5[\dot{a}(\eta_{cx})/a(\eta_{cx})](\eta - \eta_{cx})^6 + \ldots\}$, *где* c – *параметр,* $|c(\eta - \eta_{cx})^5|$ *мало (коэффициенты ряда, вообще говоря, зависят от* c *и* η_{cx}*); при* $c > 0$ *справедливо* $\eta - \eta_{cx} \le X(\eta, c, \eta_{cx}) \le \eta_{cx} - \eta$.

Из 1), 3), 4) следует, что решение (9), продолженное влево к особой точке $\eta = 0$, совпадает с решением *задачи 1*, или *задачи 2*, или *задачи 3* с некоторым значением $X(0) = \tilde{X}_0 : \ X_0 - \eta_0 < \tilde{X}_0 < X_0 + \eta_0$. Из 5), учитывая два слагаемых ряда и (9б), для решения (9) весьма приближенно получаем $\eta_{cx} \approx 5X_0/(4 - \dot{X}_0) + \eta_0$.

Напомним, что при $X \sim l$ истинная эволюция пузыря в пространствах Фридмана описывается задачей (6) для уравнения с частными производными.

3.3. *Динамическая задача для тонкого пузыря в пространстве де Ситтера:*

$$\ddot{R} + \left\{2/[a^2 R] + 3\dot{a}\dot{R}/a\right\}\left\{1 - a^2\dot{R}^2\right\} + \dot{a}\dot{R}/a = 0, \ t \ge 0, \ R(0) = R_0, \ \dot{R}(0) = 0, \qquad (10)$$

где $a(t) = a_0 \exp(Ht)$, R_0, a_0, $H > 0$. Интерес представляет положительное решение (10): его поведение с ростом t, интервал его существования для различных значений R_0, возможности его глобального существования на $[0, \infty)$ и асимптотическое поведение при $t \to \infty$. В [1] проведены предварительные расчеты, а в [2] анонсированы аналитические свойства (10), доказанные в [4], из которых следуют: 1) существование критического радиуса $R_C < (a_0 H)^{-1}$ и отвечающего ему критического решения (10) при $R_0 = R_C$, такого, что при $t \to \infty$ справедливо $\{R(t), \dot{R}(t)\} \to \{0, 0\}$ – пузырь схлопывается за бесконечное время, причем $\lim_{t\to\infty}[R(t)\exp(Ht)] = \sqrt{2/3}/(a_0 H)$, $\lim_{t\to\infty}[\dot{R}(t)\exp(Ht)] = -\sqrt{2/3}/a_0$; 2) $\forall R_0 > R_C$ глобальное существование положительного решения (10) на $[0, \infty)$, причем $\{R(t), \dot{R}(t)\} \to \{R_f, 0\}$ при $t \to \infty$, где $R_f = R_f(R_0, a_0 H) > 0$ и справедливо $\lim_{t\to\infty}\{[R(t) - R_f]\exp(2Ht)\} = [2R_f(a_0 H)^2]^{-1}$, $\lim_{t\to\infty}[\dot{R}(t)\exp(2Ht)] = [R_f a_0^2 H]^{-1}$; 3) $\forall R_0 < R_C$ схлопывание пузыря за свое конечное время: $\{R(t), \dot{R}(t)\} \to \{0, -\exp(-Ht_0)/a_0\}$ при $t \to t_0$, причем $\lim_{t\to t_0}\{R(t)/[1 - \exp[-H(t - t_0)]]\} = -\exp(-Ht_0)/(a_0 H)$; 4) приближенные оценки критического R_C и асимптотического R_f радиусов пузыря через значения R_0 и $a_0 H$.

Для строгой формулировки результатов введем в (10) новое "время" $\tau = -\exp(-Ht)/(a_0H)$ и рассмотрим более общую задачу:

$$R'' = -[2/R - 3R'/\tau](1 - R'^2), \quad \tau_0 \le \tau < 0. \tag{11a}$$

$$R(\tau_0) = R_0 > 0, \ R'(\tau_0) = \dot R_0, \ |\dot R_0| < 1, \ \tau_0 = -(a_0H)^{-1} < 0. \tag{11б}$$

где $R' = dR/d\tau$, а точка $t = \infty$ перешла в особую точку $\tau = 0$. Справедливы утверждения: (11а) обладает семейством частных решений $R^{(1)}_{\pm}(\tau, c) = \pm\tau + c$, где c -произвольная постоянная, и решениями $R^{(2)}_{\pm}(\tau) = \pm\sqrt{2/3}\tau$; для любого решения (11а), кроме $R^{(1)}_{\pm}(\tau, c)$, справедливо $\frac{d}{d\tau}\left\{\frac{R^2}{\sqrt{|1-R'^2|}}\right\} = \frac{3}{\tau}\frac{R^2R'^2}{\sqrt{|1-R'^2|}} \le 0$, $\tau_0 \le \tau < 0$. Следующие утверждения доказаны в [4]:

1) $\forall R_0 > 0$ и $\forall \dot R_0 : \ |\dot R_0| < 1$, решение (11) не выходит из "светового конуса в будущее": $\tau_0 - \tau + R_0 \le R(\tau) \le \tau - \tau_0 + R_0$, $\tau_0 \le \tau < 0$, т.е. оно существует глобально на $[\tau_0, 0)$, по крайней мере для $R_0 \ge -\tau_0$.

2) $\forall R_f$ не существует других решений (11а), удовлетворяющих в особой точке $\tau = 0$ условиям $\lim_{\tau\to-0} R(\tau) = R_f$, $\lim_{\tau\to-0} R'(\tau) = -1$ ($\lim_{\tau\to-0} R'(\tau) = 1$), кроме $R_-(\tau) = R_f - \tau$ ($R_+(\tau) = R_f + \tau$).

3) $\forall R_f \ne 0$ существует однопараметрическое семейство решений (11а), удовлетворяющих в особой точке $\tau = 0$ условиям $\lim_{\tau\to-0} R(\tau) = R_f$, $\lim_{\tau\to-0} R'(\tau) = 0$; они голоморфны при $\tau = 0$: $R(\tau, c, R_f) = R_f + \tau^2/(2R_f) + c\tau^4 - [(3/2)cR_f^{-2} + R_f^{-5}/8]\tau^6 + \ldots$, где c параметр, $|c\tau^4|$ мало (коэффициенты ряда, вообще говоря, зависят от c и R_f); эти решения не выходят из "светового конуса в прошлое": $R_f + \tau \le R(\tau, c\ R_f) \le R_f - \tau$.

4) Существует однопараметрическое семейство решений (11а), удовлетворяющих в особой точке $\tau = 0$ условиям $\lim_{\tau\to-0} R(\tau) = 0$, $\lim_{\tau\to-0} R'(\tau) = -\sqrt{2/3}$: почти все решения этого семейства не аналитичны в точке $\tau = 0$ за исключением одного $R^{(2)}_-(\tau) = -\sqrt{2/3}\tau$, которое отвечает нулевому значению параметра; справедливо разложение $R(\tau, a) = -\sqrt{2/3}\tau + \tau\{a(-\tau)^\lambda + \sum_{k=2}^{\infty} b_k[a(-\tau)^\lambda]^k\}$, сходящееся при малых значениях $|a(-\tau)^\lambda|$, $\tau < 0$, где $\lambda = \sqrt{2}$, a параметр, $b_2 = (17 + 12/\sqrt{2})/(2\sqrt{6}), \ldots$ (коэффициенты ряда не зависят от a); для неотрицательных решений справедливо $0 \le R(\tau, a) \le -\tau$.

5) $\exists R_C(R_0, \dot R_0, \tau_0) : \ R_C < -\sqrt{2/3}\tau_0$, $\forall R_0 > R_C$ решение (11) при $\tau \to 0$ удовлетворяет условиям п.3) с некоторым $R_f > 0$, при $R_0 = R_C$ оно является одним из решений семейства п.4), а $\forall R_0 < R_C \ \exists \tau_{cx} : \ \tau_{cx} > \tau_0 + R_0$, $\lim_{\tau\to\tau_{cx}-0} R(\tau) = 0$, $\lim_{\tau\to\tau_{cx}-0} R'(\tau) = -1$.

6) $\forall \tau_{cx} \in (\tau_0, 0)$ уравнение (11а) обладает однопараметрическим семейством решений, удовлетворяющих при $\tau \to \tau_{cx}$ условиям п.5); они голоморфны при $\tau = \tau_{cx}$ и включают в себя прямую $R_l(\tau) = \tau_{cx} - \tau$ при нулевом значении параметра. $R(\tau, a, \tau_{cx}) = -(\tau - \tau_{cx}) + a[(\tau - \tau_{cx})^5 + (15/14)(\tau - \tau_{cx})^6/\tau_{cx} + \ldots]$, где a - параметр, $|a(\tau - \tau_{cx})^5|$ мало (коэффициенты ряда, вообще говоря, зависят от a и τ_{cx}); при $a > 0$ для неотрицательных решений справедливо $0 \le R(\tau, a, \tau_{cx}) \le \tau_{cx} - \tau$.

Из 3), учитывая три слагаемых ряда и (11б), для решения (11) приближенно получаем $R_f \approx [R_0 - \dot R_0\tau_0/4 + \sqrt{(R_0 - \dot R_0\tau_0/4)^2 - \tau_0^2}]/2$; эта формула дает оценку снизу для R_f, при $R_0 > 3$ ее относительная погрешность не более 1% и с ростом R_0 быстро убывает. Из [2], [4] для решения (11) имеем $R_C > (2/\sqrt{3} - \dot R_0)/[a_0H(1 + \sqrt{2})]$, для случая

$\dot{R}_0 = 0$ имеем также $R_C < (6/7)\sqrt{2/3}/(a_0 H)$, а из расчетов для этого случая получили $R_C = C/(a_0 H)$, где $C \approx 0.6204032$.

В заключение отметим, что если явление спонтанного нарушения дискретной симметрии во взаимодействиях элементарных частиц будет окончательно установлено (с этой целью в CERN строится мощный ускоритель для поиска хиггсовского бозона), то неизбежными космологическими следствиями соответствующей теории будут явления и объекты во Вселенной, изучаемые в [1]-[4] и в данной работе. Если при этом считать справедливой изотропную модель Фридмана на ранней стадии существования Вселенной, то сингулярные задачи Коши, рассмотренные в п. 3.2, как бы отвечают образованию в момент ее зарождения (и их ранней эволюции) объектов типа тонкостенных сферических пузырей.

Авторы выражают благодарность Е.П.Жидкову за полезные замечания.

Литература

[1] Т. И. Белова, Н. А. Воронов, Письма в ЖЭТФ. **61**(5), 337 (1995).

[2] Н. А. Воронов, А. Л. Дышко, И. Ю. Кобзарев, Н. Б. Конюхова, ЯФ. **60**(1), 140 (1997).

[3] Н. А. Воронов, А. Л. Дышко, Н. Б. Конюхова, И. Б. Староверова, Ж. вычисл. матем. и матем. физ. (в печати).

[4] Н. А. Воронов, А. Л. Дышко, Н. Б. Конюхова, И. Б. Староверова, Ж. вычисл. матем. и матем. физ. (в печати).

[5] P. W. Higgs, Phys. Rev. Lett. **13**(16), 508 (1964).

[6] Я. Б. Зельдович, И. Ю. Кобзарев, Л. Б. Окунь, ЖЭТФ. **67**(1), 3 (1974).

[7] T. D. Lee, G. C. Wick, Phys. Rev. D9. **9**(8), 2291 (1974).

[8] Н. А. Воронов, И. Ю. Кобзарев, Н. Б. Конюхова, Письма в ЖЭТФ. **22**(11), 590 (1975).

[9] И. Л. Богочюбский, В. Г. Маханьков, Письма в ЖЭТФ. **24**(1), 15 (1976).

[10] Т. И. Белова, Н. А. Воронов, И. Ю. Кобзарев, Н. Б. Конюхова, ЖЭТФ. **73**(5), 1611 (1977).

[11] Т. И. Белова, Н. А. Воронов, Н. Б. Конюхова, Б. С. Парийский, Ж. вычисл. матем. и матем. физ. **21**(1), 89 (1981).

[12] Л. Д. Ландау, Е. М. Лифшиц. Теория поля. М.: Наука (1973).

[13] А. Д. Долгов, Я. Б. Зельдович, М. В. Сажин. Космология ранней Вселенной. М.: ИМГУ (1988).

[14] A. Starobinsky, J. Yokoyama, Phys. Rev. D. **50**(10), 6357 (1994).

[15] R. Basu, A. Vilenkin, Phys. Rev. D. **50**(12), 7150 (1994).

[16] М. Б. Волошин, И. Ю. Кобзарев, Л. Б. Окунь, ЯФ. **20**(6), 1229 (1974).

Модификация метода К.Ланцоша выявления скрытых периодичностей

Е.А.Гребеников, С.В.Миронов

(ИВВС РАН, Москва, 117312, ул.Вавилова, 37,
тел. (095) 129-04-74, e-mail:gea@ipk.msk.su)

1. Введение

В [1] К.Ланцошем изложен метод решения известной классической задачи теории аппроксимации, сформулированной Лагранжем и получившей название "задачи о выявлении скрытых периодичностей".

Задача формулируется следующим образом.
Пусть:

1. На отрезке $t \in [T_1, T_2]$ задан дискретный массив $\{f(t_s), s = 1, 2, \cdots, 2N+1, t_1 = T_1, T_{2N+1} = T_2\}$ значений некоторой функции $f(t)$;

2. Функция $f(t)$ является полигармонической, т.е. разлагается на некоторое число периодических гармоник с рационально несоизмеримыми частотами;

3. Число гармоник "m", их амплитуды a_k, b_k и частоты ω_k $(k = 1, \cdots, m)$ - неизвестны.

Требуется определить параметры $m, a_k, b_k, \omega_k (k = 1, \cdots, m)$.

Проведенный нами анализ метода Ланцоша показывает, что он может быть эффективно применен, если выполнены следующие условия:

1. Точки t_s равноотстоят друг от друга, т.е.

$$\Delta t_s = t_{s+1} - t_s = \frac{T_2 - T_1}{2N};$$

2. Отсутствует "интерференция" частот [2], т.е. частоты ω_k не только взаимно рационально несоизмеримы, но их численные значения "достаточно далеки" друг от друга. Наши численные эксперименты показывают, что пренебрежение этими условиями может привести к существенным ошибкам в определении параметров, поэтому представляется целесообразным решать представленную задачу без указанных выше ограничений. Такое решение

было впервые изложено в [2,3], и оно основано на так называемом итерационном методе с пошаговым уточнением значений всех определяемых до j - го шага итераций параметров $a_1, \cdots, a_j, \quad b_1, \cdots b_j, \quad \omega_1, \cdots \omega_j$.

Опираясь на математическую теорию оптимизации, изложим новую постановку задачи.

2. Необходимые условия минимума функционала Лежандра - Гаусса

Представим аппроксимирующий (интерполяционный) многочлен для полигармонической функции $f(t)$ в виде

$$P_m(t) = \sum_{k=1}^{m} (a_k \cos(\omega_k t) + b_k \sin(\omega_k t)) \tag{1}$$

и напишем хорошо известный квадратичный функционал Лежандра - Гаусса

$$L(a\ b, \omega, m) = \sum_{s=1}^{N} \left[f(t_s) - \sum_{k=1}^{m} (a_k \cos(\omega_k t_s) + b_k \sin(\omega_k t_s)) \right]^2, \tag{2}$$

где $a = (a_1, \cdots, a_m), \quad b = (b_1, \cdots, b_m), \quad \omega = (\omega_1, \cdots, \omega_m)$ - суть m - мерные векторы.

Сущность классического метода наименьших квадратов состоит в том, что неизвестные параметры аппроксимирующего полинома $P_m(t)$ определяются из необходимых условий минимума функционала $L(a, b, \omega, m)$.

Если предположить, что число гармоник m и частоты $\omega_1, \cdots, \omega_m$ известны, а параметры $a_1, \cdots, a_m, \quad b_1, \cdots b_m$ являются неизвестными, определение последних сводится к решению системы $2m$ линейных алгебраических уравнений

$$\frac{\partial L}{\partial a_i} = 0, \qquad \frac{\partial L}{\partial b_i} = 0 \qquad (i = 1, \cdots m). \tag{3}$$

Если допустить, что число гармоник m известно, а частоты $\omega_1, \cdots \omega_m$ и амплитуды $a_1, \cdots a_m, \quad b_1, \cdots b_m$ неизвестны, то необходимые условия минимума функционала $L(a, b, \omega, m)$ записываются в виде

$$\frac{\partial L}{\partial a_i} = 0, \qquad \frac{\partial L}{\partial b_i} = 0, \qquad \frac{\partial L}{\partial \omega_i} = 0 \qquad (i = 1, \cdots m). \tag{4}$$

Система уравнений (4) имеет размерность $3m$, является нелинейной (трансцендентной), и ее решение следует определять итерационными методами.

Предположим далее, что целочисленный параметр m также неизвестен. Тогда для "замыкания" необходимых условий экстремума функционала $L(a, b, \omega, m)$

следует ввести определение производной функционала по целому параметру m. Наиболее естественным является определение этой производной через отношение приращений, а именно, будем считать, что

$$\left.\frac{\partial L}{\partial m}\right|_{m-1} = \frac{\Delta_{m-1}L}{\Delta(m-1)} = \Delta_{m-1}L = L(a,b,\omega,m) - L(a^*,b^*,\omega^*,m-1), \quad (5)$$

где $\Delta(m-1) = m-(m-1) = 1$,
$a, b, \omega - m$ - мерные векторы $(a_1,\cdots a_m),(b_1,\cdots b_m),(\omega_1,\cdots \omega_m)$,
$a^*, b^*, \omega^* - (m-1)$- мерные векторы $(a_1,\cdots a_{m-1}),(b_1,\cdots b_{m-1}),(\omega_1,\cdots \omega_{m-1})$.

Поэтому в общем случае для решения сформулированной выше задачи необходимо решить следующую систему трансцедентных уравнений:

$$\frac{\partial L}{\partial a} = \frac{\partial L}{\partial b} = \frac{\partial L}{\partial \omega} = \Delta_{m-1}L = 0, \quad (6)$$

порядок которой равен $(3m+1)$ и число неизвестных тоже равно $(3m+1)$.

Несмотря на то, что уравнения (6) являются существенно нелинейными, "двумерным итерацнонным процессом" [2,4] они могут быть эффективно решены для достаточно больших значений $m(m \approx 100-500)$.

Литература

1. Ланцош К. Практические методы прикладного анализа.- М.: Физматгиз, 1961.

2. Миронов С.В. Редукционный метод выявления скрытых периодичностей. - В кн. Математические вопросы нелинейного анализа и управления динамическими системами - М.: Изд-во Моск. университета, 1985, с. 3-18.

3. Гребеников Е.А., Киоса М.Н., Миронов С.В. Численно - аналитические методы исследования регулярно возмущенных систем. - М.: Изд-во Моск. Университета, 1985.

4. Grebenicov E. and Mironov S. Reduction Method for the Determination of Unknown Periods in Processing a Sequence of Observations. - В кн. Annales de Physique. Collocue C1, supplément au $n^0$1, Vol. 21, février 1996, s. C1-175 - C1-179.

ЧИСЛЕННОЕ МОДЕЛИРОВАНИЕ НЕКОТОРЫХ МАГНИТНЫХ СИСТЕМ СОЛЕНОИДАЛЬНОГО ТИПА [1]

Е. П. Жидков, В. П. Овсянников, Р. В. Полякова, И. П. Юдин

Объединенный институт ядерных исследований. Дубна 141980, Московская обл., Россия

Процесс создания многих магнитных систем, включая тракты линейных ускорителей, современных источников частиц и др., наряду с этапом длительного и трудоемкого физического моделирования в значительной мере опирается на численное моделирование распределения поля соленоидальных магнитов в рабочей области установки. Значимость численного моделирования при конструировании рассматриваемых магнитных систем определяется не только известными достоинствами вычислительного эксперимента, но и техническими особенностями конкретной установки, например, наличием криогенных температур, затрудняющих проведение измерений магнитного поля. Благодаря достаточной точности выбранной математической модели, с помощью вычислительного эксперимента возможно решение проблемы выбора и оптимизации формы магнитных полюсов конкретной магнитной системы.

В данной работе рассматривается проблема моделирования поля магнитных систем соленоидального типа с железным экраном и без него; предлагается метод уточнения решения такой краевой задачи; изучается проблема получения путем численного моделирования конфигурации системы, дающей наиболее крутой спад поля на торцах магнита.

1. Для исследования распределения магнитного поля на оси z безжелезного соленоида использовалась формула Био–Савара–Лапласа:

$$H_z(r,z) = J\int_{R_1}^{R_2}\int_{z_1}^{z_2} G_z(r',z',r,z)dz'dr', \qquad (1)$$

где (r,z) - точка, принадлежащая области распределения магнитного поля, (r',z') - точка, принадлежащая токонесущей области, z_1, z_2 – продольные координаты катушки, R_1, R_2 – внутренний и внешний радиус катушки, $G_z(r',z',r,z)$ – функция Грина для продольной составляющей поля в аксиально-симметричной системе имеет вид [1]:

$$G_z(r',z'\ r,z) = \frac{k}{\sqrt{(r'+r)^2+(z'-z)^2}} \times \left[K + \frac{r'^2 - r^2 - (z'-z)^2}{(r'-r)^2+(z'-z)^2}E\right]$$

где K и E – полные эллиптические интегралы I-го и II-го рода от модуля $X^2 = \frac{4r'r}{(r'+r)^2+(z'-z)^2}$, k - коэффициент зависящий от системы единиц.

На основании метода решения некоторого класса обратных задач магнитостатики, описанного в работе [2], разработан численный алгоритм формирования поля магнитных систем соленоидального типа.

Для численного моделирования магнитного поля соленоида в присутствии железного экрана использовалась система программ КПММС-1.0 [3], в состав которой входят программы POISCR [4] и RELAX3D [5] .

[1] Работа выполнена при поддержке Российского Фонда Фундаментальных Исследований, грант No 95 – 01 – 00737а.

В частности изучалось распределение стационарного электромагнитного поля соленоида электронной пушки [см.6] (рис.1а) и криогенных электронно–лучевых ионизаторов [7] (рис. 2).

Подобное математическое моделирование совместно с измерениями магнитного поля фокусирующей системы являлось основополагающим шагом при создании аналогичных установок, например CRYEBIS-II, DIONE (Франция), CEBIS-II (США) и др. [8].

1а. Пусть B - вектор магнитной индукции, H - вектор напряженности поля, j - вектор объемной плотности тока. Тогда уравнения Максвелла для стационарного случая имеют вид [9]

$$\begin{gathered} \mathrm{rot} H = c_0 j, \\ \mathrm{div} B = 0, \\ B = \mu_0 \mu(\|B\|^2) \cdot H, \qquad \lim_{\|x\| \to \infty} H(x) = 0. \end{gathered} \tag{2}$$

где c_0 - константа, зависящая от системы единиц, μ_0 - магнитная проницаемость вакуума; $\mu = \mu(\|B\|^2)$ - магнитная проницаемость ферромагнетика. Уравнения (2), дополненные условиями непрерывности нормальной составляющей вектора B и касательной составляющей вектора H на границах раздела различных сред

$$(B_1 - B_2) \cdot \vec{n}_1 = 0, \quad (H_1 - H_2) \times \vec{n}_1 = 0. \tag{3}$$

где $\vec{n}_1$ - нормаль к поверхности раздела, однозначно определяют искомые физические характеристики.

Для осесимметричного поля численно решались уравнения для азимутальной компоненты $u = A_\Theta$ векторного потенциала A, определяемого равенством

$$B = \mathrm{rot} \mathrm{A}, \qquad \mathrm{A} = (\mathrm{A_r}, \mathrm{A_z}, \mathrm{A_\Theta})^{\mathrm{T}}, \tag{4}$$

где следует положить $A_r = A_z = 0$. Если обозначить

$$j = (0, 0, j_\Theta)^T, \qquad f = j_\Theta, \tag{5}$$

то уравнение для $u(r, z)$ примет вид

$$\begin{gathered} \frac{\partial}{\partial r}\left(\nu\|B\|^2 \cdot \left(\frac{\partial u}{\partial r} + \frac{u}{r}\right)\right) + \frac{\partial}{\partial z}\nu(\|B\|^2)\frac{\partial u}{\partial z} = c_0 f, \\ u(\infty) = 0, \qquad \nu = (\mu_0 \mu)^{-1}, \end{gathered} \tag{6}$$

где $\nu = (\mu_0\mu)^{-1}$, а компоненты B восстанавливаются по формулам

$$B_z = \frac{1}{r}\frac{\partial}{\partial r}(ru), \quad B_r = -\frac{\partial u}{\partial z}, \quad B_\Theta = 0. \tag{7}$$

На границах раздела сред используются условия (3). Соленоидальная магнитная система может являться "открытой",т.е. не замкнутой железным экраном. Поэтому при проведении численных расчетов уделялось внимание вопросу аппроксимации краевого условия на бесконечности $u(\infty) = 0$. Различные способы учета этого условия, требующие дополнительного программного обеспечения, обсуждаются в [9]. В комплексе

программ КПММС-1.0 [3] реализован способ, сводящийся к использованию условия Дирихле

$$u(\xi) = 0, \quad \xi \in \Gamma \tag{8}$$

где Γ - граница расчетной области Ω (в нашем случае это прямоугольник, содержащий токовые области и ферромагнетик (магнитные полюса) с координатами сторон l_z и l_r).

2а. В процессе проектирования модификаций магнита, который представляет совокупность взаимосвязанных различных технических и программных средств, необходим компромисс между такими факторами, как улучшение технических и эксплуатационных характеристик магнитной системы, сокращение ее стоимости и одновременно повышение требований к эксплуатационным качествам магнитной системы, необходимым для экспериментов на физической установке.

В комплекс программ КПММС-1.0 составной частью входит накопление банка данных исследуемых конфигураций магнитных систем для дальнейшего графического исследования полученных численных результатов, что значительно облегчает и ускоряет процесс отбора конфигураций магнитной системы для ее технической реализации.

Проведенные нами численные расчеты показали, что для безжелезного соленоида при включении дополнительных витков на торцах обмотки наблюдается более крутой спад поля на краях магнита.

Расчет поля для системы с "броневым" экраном (рис.1а) проведен при конечноэлементном разбиении расчетной области на $4 \cdot 10^4$ неизвестные. Результат, представленный на рис.1б, показывает, что железные наконечники на торцах обеспечивают здесь нужный крутой спад поля.

Расчеты поля фокусирующей установки КРИОН-С проводились раздельно для правого и левого магнитных полюсов. На общей границе расчетных областей ставилось условие Неймана. Расчеты магнитного поля для обоих полюсов проводились при конечноэлементном разбиении области Ω, содержащем $4 \cdot 10^4$ неизвестных.

При сравнении с экспериментальными значения $B(z)$ оказалось, что разность между экспериментальными и расчетными величинами находится в пределах точности измерений.

Полученные результаты показали, что в области магнитного полюса для электронного коллектора не обеспечивается необходимый спад магнитного поля. Последнее подтверждено и экспериментами с электронным пучком. Численные эксперименты при возможном варьировании размеров области ферромагнетика показали, что достигнуть необходимого спада магнитного поля таким способом не удается.

Расчеты для другого типа наконечников, соответствующих двухсвязанной области $G_1 = G \cup G_2$, показали, что добавление ферромагнитного кольца G_2 позволяет, сохраняя требования к конструкциям, добиться необходимого спада магнитного поля на оси системы. Серия выполненных численных экспериментов позволила найти оптимальную форму и расположение дополнительной области G_2 исходя из требований к данной конструкции (рис.3).

3. В процессе численных расчетов проводился дополнительный контроль точности аппроксимации условия $u(\infty) = 0$ на основе методики, использующей экстраполяцию по параметру R^{-1}, где

$$R = \max(l_z, l_r) \tag{9}$$

а l_z и l_r - длины сторон прямоугольника, ограничивающего расчетную область Ω. Предполагается, что погрешность аппроксимации краевого условия $u(\infty) = 0$ разлагается в ряд по степени параметра R^{-1}. Для рассматриваемой задачи использовался следующий алгоритм. Решаются последовательно краевые задачи (7), (8) для последовательности расширяющихся областей Ω_k, $k = 1, \cdots, M$, имеющих попарно различные значения R_k из (9). Далее сравниваются значения решений $u_k(p)$ (при различных K) в контрольных точках P расчетной области Ω. Если для данного выбранного $\epsilon > 0$ выполняется условие

$$\|u_k(p) - u_{k+1}(p)\| \le \epsilon, \quad p \in \Omega \tag{10}$$

то полагаем, что условие (8) удовлетворительно аппроксимирует краевое условие $u(\infty) = 0$. Если априорно предположить, что существует регулярное разложение погрешности аппроксимации условия на бесконечности по параметру R^{-1}

$$u_R(p) = u(p) + \frac{1}{R^\alpha} g(p) + O(R^{-\beta}), \quad \beta > \alpha > 0. \tag{11}$$

где $u(p)$ - точное решение уравнения (6), а функция $g(p)$ не зависит от R, то можно исключить второе слагаемое в (11) путем экстраполяции. Для этого определим числа γ_1, γ_2 из системы уравнений

$$\gamma_1 + \gamma_2 = 1, \quad \gamma_1 R_1^{-\alpha} + \gamma_2 R_2^{-\alpha} = 0, \quad R_1 \ne R_2 \tag{12}$$

где R_1, R_2 соответствуют двум различным расчетным областям. Тогда линейная комбинация

$$\bar{u}(p) = \gamma_1 u_{R_1}(p) + \gamma_2 u_{R_2}(p) \tag{13}$$

удовлетворяет оценке

$$\|u(p) - \bar{u}(p)\| \le O(R^{-\beta}) \tag{14}$$

т.е. приближает условие $u(\infty) = 0$ с большей точностью, чем каждое из u_{R_1} и u_{R_2}. Приведем пример расчетов по этой схеме. Полагаем $\alpha = 2$, $R_2 = 2.0 \cdot 10^{-1}$, $R_3 = 2.5 \cdot 10^{-1}$, $R_1 = 1.6 \cdot 10^{-1}$. На рис. 4 представлены графики функций $\|B - B(R_i)\|$, $i = 1, 3$; B соответствует $\bar{u}$. Наблюдаются уточнение результатов в области, близкой к магнитному полюсу, где предъявляются более высокие требования к точности вычислений.

4. В качестве заключения можно сказать следующее:

а) в результате численного моделирования выше приведенных магнитных систем соленоидального типа создан банк данных, включающих информацию для широкого набора расчетных систем и соответствующей им карты поля, а также других параметров (в частности относительно геометрии магнитных систем), что позволяет расширить "физическую" интуицию для дальнейшего развития магнитных систем такого класса;

б) продемонстрирована эффективность предложенной методики уточнения решения поставленной краевой задачи магнитостатики;

в) полученные распределения магнитного поля для широкого набора расчетных систем предлагаются для использования в качестве "заготовок" при конструировании электрофизических установок, содержащих соленоиды.

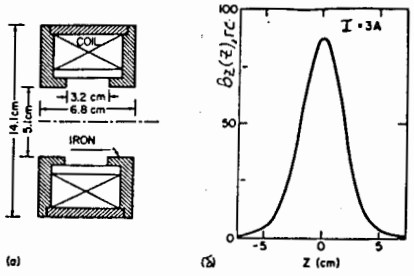


Рис.1 Соленоид с железным экраном закрытого типа: а) схематический вид и б) магнитное поле B(z) на оси z при рабочем токе в обмотке I = 3 А.

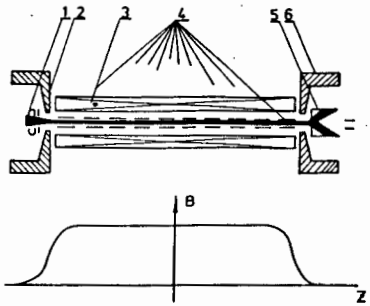


Рис.2 Соленоид с полюсными наконечниками для электронно-лучевого источника КРИОН-С: схематический вид и магнитное поле B(z) на оси z. 1 - электронная пушка, 2,6 - магнитные полюса, 3 - сверхпроводящий соленоид, 4 - трубка дрейфа, 5 - электронный коллектор.

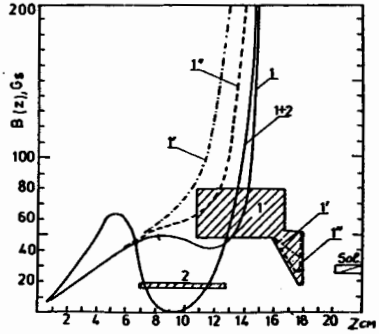


Рис.3 Спад магнитного поля на торце соленоида при различных конфигурациях торцевой области.

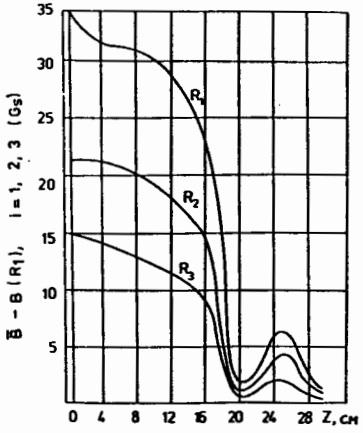


Рис.4 Графики функций $|B - B(R_i)|$, i=1,2,3 в зависимости от координаты z.

ЛИТЕРАТУРА

1. Ландау Л.Д., Лифшиц Е.М. Электродинамика сплошных сред. Физматгиз, М., 1959.

2. Жидков Е.П., Куц И.В., Полякова Р.В., Шелаев И.А. Юдин И.П. Решение одной нелинейной обратной задачи магнитостатики методом регуляризации. ЖВМ и МФ, т.29, вып.4, с.617-621, Москва, 1989.

3. Жидков Е.П., Лима С., Полякова Р.В., Фернандес Нодарсе Ф., Юдин И.П. Комплекс Программ для Моделирования Магнитных Систем, Препринт ОИЯИ, Р11-93-256. Дубна, 1993, 12с.

4. Holsinger R.F., Iselin C., POISCR-r604, User Guide, CERN, SPS/EMA, 1982.

5. RELAX3D. User's Guide and Reference Manual, TRIUMF 4004 Wesbrook Mall, Vancouver, B.C. Canada V6T 2A3, 1992.

6. Andreev V.V., Yudin I.P. Three-Order Optics of the Real Solenoid Lens, In: Proc. Intern. Conf. on HEACC'92, July 1992, Humburg, Germany.

7. Донец Е.Д. А.с. N 248860 (СССР). ОИПОТЗ, 1969, N 24, с. 65.
Донец Е.Д., Овсянников В.П. ОИЯИ, Р7-9799. Дубна, 1976.

8. Proc. Third International EBIS Workshop, Cornell University, May, 1985.

9. Айрян Э.А. и др. Численные алгоритмы расчета магнитных систем ускорителей заряженных частиц. ЭЧАЯ, М., "Энергоатомиздат", 1990. т.1, вып.1, с.251-307.

ПОВЫШЕНИЕ ТОЧНОСТИ ПРИБЛИЖЕННЫХ РЕШЕНИЙ КРАЕВОЙ ЗАДАЧИ НА ПОЛУПРЯМОЙ [1]

Е.П.Жидков, А.Г.Соловьев

141980 ЛВТА ОИЯИ
Дубна Московской области, Россия
e-mail: solovjev@main1.jinr.dubna.su

Аннотация. Предложен метод повышения точности приближенных решений краевой задачи на полупрямой для обыкновенного дифференциального уравнения второго порядка. При численном решении такой задачи интервал $[x_0, \infty)$ заменяется отрезком $[x_0, R]$, условие в точке R выбирается исходя из асимптотики решения на бесконечности, и краевая задача решается на этом конечном отрезке. В результате такой замены возникает погрешность. Предложенный метод основан на многократном решении задачи на конечных отрезках разной длины. Последующая комбинация полученных решений ("экстраполяция по длине отрезка") дает уточненное решение.

1 Постановка задачи

Рассматривается следующая краевая задача на полупрямой:

$$\begin{cases} u'' - \left(1 + \dfrac{c_1}{x} + \psi(x)\right) u = 0, \\ u(x_0) = A, \\ u(+\infty) = 0, \end{cases} \tag{1}$$

где $\psi(x)$ – непрерывная на интервале $[x_0, \infty)$ функция, такая что $\psi(x) = o\left(\dfrac{1}{x}\right)$ при $x \to \infty$, а заданная постоянная $A \neq 0$.

При численном решении таких задач интервал $[x_0, \infty)$, как правило, заменяется отрезком $[x_0, R]$, условие из бесконечности переносится в точку R (см. [1 – 3]), и решается следующая краевая задача на конечном отрезке (см. [4 – 5]):

$$\begin{cases} y'' - \left(1 + \dfrac{c_1}{x} + \psi(x)\right) y = 0, \\ y(x_0) = A, \\ y'(R) + \left(1 + \dfrac{c_1}{2R}\right) y(R) = 0. \end{cases} \tag{2}$$

В результате такой замены решение $y(x)$ задачи (2) отличается от решения $u(x)$ исходной задачи (1):

$$y(x) = u(x) + \lambda \cdot v(x),$$

[1] Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований (проект 95 – 01 – 01467).

где $v(x)$ – линейно независимое от $u(x)$ решение уравнения, а поведение коэффициента $\lambda = \lambda(R)$ при больших значениях R описывается следующей теоремой, доказанной в работе [5]:

Теорема 1 Пусть функция $\psi(x)$ имеет асимптотику

$$\psi(x) = O\left(\frac{1}{x^2}\right), \quad x \to \infty.$$

Тогда $\lambda(R) \to 0$ при $R \to \infty$, причем

$$\lambda(R) = \frac{1}{4}\mathrm{e}^{-2R}R^{-c_1}\left[\psi(R) - \frac{c_1(c_1+2)}{4R^2}\right]\left(1 + O\left(\frac{1}{R}\right)\right), \quad R \to \infty.$$

Таким образом, $\lambda(R) \to 0$ при $R \to \infty$. Это означает, что решение задачи (2) на отрезке $[x_0, R]$ будет тем ближе к решению задачи (1), чем больше R. Изложим теперь метод, позволяющий путем многократного решения задачи (2) на различных отрезках получить решение с погрешностью, существенно меньшей, чем погрешности каждого из решений по-отдельности.

2 Идея метода

Пусть $y_1(x)$ и $y_2(x)$ – решения задачи (2) на отрезках $[x_0, R_1]$ и $[x_0, R_2]$ соответственно $(R_2 > R_1)$. Запишем

$$y_1(x) = u(x) + \lambda_1 \cdot v(x),$$
$$y_2(x) = u(x) + \lambda_2 \cdot v(x).$$

Отношение погрешностей решений $y_1(x)$ и $y_2(x)$ есть величина, не зависящая от x:

$$\frac{y_2(x) - u(x)}{y_1(x) - u(x)} = \frac{\lambda_2}{\lambda_1}.$$

Эта величина может быть получена приближенно, исходя из асимптотики решения $u(x)$ при $x \to \infty$ (см. [4 – 5]):

$$\frac{\lambda_2}{\lambda_1} = \alpha \cdot (1 + o(1)), \quad R_1 \to \infty, \quad R_2 \to \infty.$$

Определим уточненное решение следующим образом:

$$\Psi(x) = \frac{y_2(x) - \alpha \cdot y_1(x)}{1 - \alpha}. \tag{3}$$

Если R_1 и R_2 достаточно велики, то погрешность $\Psi(x)$ меньше, чем погрешности $y_1(x)$ и $y_2(x)$ по-отдельности. Этот факт установлен в работе [5], где доказана следующая

Теорема 2. Пусть $y_1(x)$ и $y_2(x)$ – решения задачи (2) на отрезках $[x_0, R_1]$ и $[x_0, R_2]$ соответственно, где $R_2 > R_1$, и выполнены условия теоремы 1. Пусть, кроме того, R_1 и R_2 выбраны таким образом, чтобы параметр

$$\alpha = \frac{\mathrm{e}^{-2R_2} R_2^{-c_1} \left[\psi(R_2) - \frac{c_1(c_1+2)}{4R_2^2}\right]}{\mathrm{e}^{-2R_1} R_1^{-c_1} \left[\psi(R_1) - \frac{c_1(c_1+2)}{4R_1^2}\right]}$$

удовлетворял условию $0 < |\alpha| < 1 - \delta$, где $0 < \delta < 1$ и не зависит от R_1 и R_2.

Тогда линейная комбинация

$$\Psi(x) = \frac{y_2(x) - \alpha \cdot y_1(x)}{1 - \alpha}, \quad x \in [x_0, R_1]$$

приближает искомое решение $u(x)$ с погрешностью

$$||\Psi - u|| = O\left(\frac{||y_2 - u||}{R_1}\right),$$

где $||\cdot|| = \max\limits_{x \in [x_0, R_1]} |\cdot(x)|$.

Процесс уточнения можно продолжить, если имеются три решения $y_1(x)$, $y_2(x)$ и $y_3(x)$ задачи (2) на отрезках $[x_0, R_1]$, $[x_0, R_2]$ и $[x_0, R_3]$ соответственно ($R_3 > R_2 > R_1$). В этом случае уточнение производится в два этапа. Сначала по формуле (3) находятся два уточненных решения: одно – линейная комбинация $y_1(x)$ и $y_2(x)$, и другое – линейная комбинация $y_2(x)$ и $y_3(x)$. Затем из этих двух полученных решений снова составляется линейная комбинация по формуле (3). В результате погрешность уточненного решения будет меньше по сравнению со случаем, когда задача (2) решалась дважды. Уточнение по большему количеству решений позволяет получить, соответственно, и бо́льшую точность. В работе [4] предложен основанный на этой идее алгоритм получения приближенного решения с заданной точностью.

3 Пример

В качестве примера, иллюстрирующего применение предложенного метода, рассмотрим следующую задачу:

$$\begin{cases} u'' - \left(1 - \frac{8}{x} + \frac{2}{x^2}\right) u = 0, \\ u(1) = \frac{1}{5\mathrm{e}}, \\ u(\infty) = 0, \end{cases} \tag{4}$$

решение которой известно:

$$u(x) = \mathrm{e}^{-x} x^4 \left(\frac{1}{5} - \frac{1}{x} + \frac{1}{x^2}\right). \tag{5}$$

Заменим исходную задачу (4) задачей на конечном отрезке $[1, R]$:

$$\begin{cases} y'' - \left(1 - \dfrac{8}{x} + \dfrac{2}{x^2}\right) y = 0, \\ y(1) = u(1), \\ y'(R) + \left(1 - \dfrac{4}{R}\right) y(R) = 0. \end{cases} \tag{6}$$

Задачу (6) будем решать численно, получая ее решения с погрешностью порядка 10^{-4}.

Выберем $R_1 = 10.0$ и $R_2 = 10.1$. В таблице 1 приведены точное решение (5) исходной задачи (4), решения задачи (6) на отрезках $[1, R_1]$ и $[1, R_2]$ (с погрешностью порядка 10^{-4}), а также уточненное решение, построенное по формуле (3).

Таблица 1. **Уточнение по двум решениям**

x	$u(x)$	$y_1(x)$	$y_2(x)$	$\Psi(x)$
1.00	0.0735759	0.0735759	0.0735759	0.0735759
2.00	-0.1082682	-0.1094405	-0.1092989	-0.1082872
3.00	-0.0896167	-0.0894744	-0.0894915	-0.0896141
4.00	0.0586100	0.0600683	0.0598922	0.0586340
5.00	0.1684487	0.1700714	0.1698754	0.1684748
6.00	0.1963172	0.1972682	0.1971533	0.1963319
7.00	0.1697924	0.1697796	0.1697810	0.1697912
8.00	0.1245237	0.1235234	0.1236440	0.1245062
9.00	0.0819688	0.0799324	0.0801781	0.0819343
10.00	0.0499399	0.0465732	0.0469797	0.0498834

Выберем $R_3 = 10.2$ и добавим в рассмотрение третье решение задачи (6) на отрезке $[1, R_3]$. Результат уточнения по трем решениям приведен в таблице 2.

Таблица 2. **Уточнение по трем решениям**

x	$u(x)$	$y_1(x)$	$y_2(x)$	$y_3(x)$	$\Psi(x)$
1.00	0.0735759	0.0735759	0.0735759	0.0735759	0.0735759
2.00	-0.1082682	-0.1094405	-0.1092989	-0.1091736	-0.1082648
3.00	-0.0896167	-0.0894744	-0.0894915	-0.0895067	-0.0896168
4.00	0.0586100	0.0600683	0.0598922	0.0597364	0.0586062
5.00	0.1684487	0.1700714	0.1698754	0.1697019	0.1684438
6.00	0.1963172	0.1972682	0.1971533	0.1970516	0.1963137
7.00	0.1697924	0.1697796	0.1697810	0.1697823	0.1697914
8.00	0.1245237	0.1235234	0.1236440	0.1237508	0.1245253
9.00	0.0819688	0.0799324	0.0801781	0.0803956	0.0819731
10.00	0.0499399	0.0465732	0.0469797	0.0473392	0.0499476

При получении решения с заданной точностью по предложенному в работе [4] алгоритму выберем $\epsilon = 10^{-5}$ (допустимая погрешность), $R_1 = 10.0$ и $R_{n+1} = R_n + h$, $n = 1, 2, \ldots$ В этом случае задача решается 4 раза. В таблице 3 приведены точное решение и конечный результат.

Таблица 3. **Получение решения с заданной точностью**

x	$u(x)$	$\Psi(x)$
1.00	0.0735759	0.0735759
2.00	-0.1082682	-0.1082683
3.00	-0.0896167	-0.0896163
4.00	0.0586100	0.0586105
5.00	0.1684487	0.1684486
6.00	0.1963172	0.1963165
7.00	0.1697924	0.1697914
8.00	0.1245237	0.1245224
9.00	0.0819688	0.0819671
10.00	0.0499399	0.0499377

Как видно из таблиц 1 – 3, каждое дополнительное решение задачи (6) приносит один верный знак в уточненное решение. Этот факт представлен в таблице 4, где $\Psi_n(x)$ – уточненное решение, построенное по n решениям задачи (6), $n = 2, 3, 4$.

Таблица 4. **Иллюстрация процесса уточнения при получении решения с заданной точностью**

	$x = 5$	$x = 10$
$u(x)$	0.1684487	0.0499399
$y(x)$	0.1700714	0.0465732
$\Psi_2(x)$	0.1684748	0.0498834
$\Psi_3(x)$	0.1684438	0.0499476
$\Psi_4(x)$	0.1684486	0.0499377

4 Выводы

Предложен метод повышения точности приближенных решений краевой задачи на полупрямой для обыкновенного дифференциального уравнения второго порядка. Разработан алгоритм нахождения приближенного решения на данном отрезке с заданной точностью. Причем можно получить любую наперед заданную точность, но для этого необходимо обеспечить получение достаточно точных решений задачи на конечном отрезке. Численные эксперименты, тестирование программы, основанной на предложенном алгоритме, дали хорошие результаты.

Литература

[1] *А.А.Абрамов.* Сравнение решений прогоночных уравнений при переносе граничных условий из бесконечности для гамильтоновых линейных систем. – ЖВМ и МФ, 1995, **35**, N12, 1808 - 1818.

[2] *А.А.Абрамов.* О приближенных решениях, основанных на теоремах сравнения, скалярных и матричных уравнений Риккати на бесконечном интервале. – ЖВМ и МФ, 1993, **33**, N1, 35 - 51.

[3] *Е.С.Биргер.* Об оценке погрешности замены условия ограниченности решения линейного дифференциального уравнения на бесконечном интервале. – ЖВМ и МФ, 1968, **8**, N3, 674 - 678.

[4] *Е.П.Жидков, А.Г.Соловьев.* Повышение точности приближенных решений краевой задачи на полупрямой. – 1996, Дубна, ОИЯИ, Р11-96-153.

[5] *Е.П.Жидков, А.Г.Соловьев.* Уточнение приближенных решений краевой задачи на полупрямой. – 1996, Дубна, ОИЯИ, Р11-96-480.

Применение тёплицевых матриц для решения задач квантовой механики

И.В. Ибрагимов, Е.Е. Тыртышников

Институт Вычислительной Математики, РАН,
117333, Москва, Губкина 8
e-mail: ilghiz@inm.ras.ru, tee@inm.ras.ru

Abstract. В данной работе предложен новый подход к решению уравнений Хартри – Фока для многоэлектронных систем, который использует идею регулярных базисных функций. В этом случае матрица интегралов обладает свойством тёплицевости. Это значительно сокращает число операций с плавающей точкой при решении этих уравнений. А для двухэлектронной задачи в некотором приближении из нелинейной задачи на собственные значения трансформируется в задачу решения системы линейных уравнений с тёплицевой или циркулянтной матрицей той же размерности.

Данный подход был проверен на тестовых примерах определения энергии взаимодействия атомов в кристаллической решетке и определения вида волновой функции для свободного атома. Были получены хорошие результаты.

1 Введение

В настоящее время одной из важных задач квантовой механики является нахождение энергии взаимодействия двух или нескольких атомов. Для решения такой задачи разработано много различных методов, позволяющих учитывать различные свойства рассматриваемых систем.

Наиболее общие (и, в то же время, значительно более точные методы на основе уравнений Хартри–Фока [1] или теории функционала плотности [2]) сводятся к решению вариационной задачи для нелинейного оператора, что требует решения нелинейной полиномиальной задачи на собственные значения.

Решается стационарное уравнение Шредингера в приближении фиксированных ядер [3] $\hat{H}\Psi = E\Psi$, где $\hat{H} = \hat{T}_e + C_{nn} + \hat{V}_{ne} + \hat{V}_{ee}$, а

$\hat{T}_e = -\frac{\hbar^2}{2}\sum_{i=1}^{M}\frac{1}{m_e}\left(\frac{d^2}{dx_i^2} + \frac{d^2}{dy_i^2} + \frac{d^2}{dz_i^2}\right)$ – оператор кинетической энергии всех M электронов;

$C_{nn} = \sum_{i<j}^{N}\frac{Q_iQ_j}{|R_i - R_j|}$ – энергия взаимодействия всех ядер (в данном приближении является константой, в дальнейших выкладках фигурировать не будет, так как можно от E перейти к $E_C = E - C_{nn}$);

$\hat{V}_{ne} = \sum_{i=1}^{N}\sum_{j=1}^{M}\frac{Q_iq_e}{|R_i - r_j|}$ – оператор энергии взаимодействия ядер с электронами;

$\hat{V}_{ee} = \sum_{i<j}^{M}\frac{q_e^2}{|r_i - r_j|}$ – оператор энергии взаимодействия всех электронов,

Ψ является, таким образом, собственной функцией оператора $\hat{H}$, зависящей от координат всех частиц в трехмерном пространстве,

E – собственное значение оператора $\hat{H}$, минимальное из них соответствует наблюдаемой энергии системы.

Минимальное E и соответствующую функцию Ψ можно найти вариационным методом в виде детерминанта Слейтера [1,3]:

$$\tilde{\Psi} = \frac{1}{\sqrt{M!}} \det \begin{vmatrix} \psi_1(\mathbf{r}_1)\alpha_1 & \psi_1(\mathbf{r}_1)\beta_1 & \dots & \psi_{M/2}(\mathbf{r}_1)\alpha_1 & \psi_{M/2}(\mathbf{r}_1)\beta_1 \\ \psi_1(\mathbf{r}_2)\alpha_2 & \psi_1(\mathbf{r}_2)\beta_2 & \dots & \psi_{M/2}(\mathbf{r}_2)\alpha_2 & \psi_{M/2}(\mathbf{r}_2)\beta_2 \\ \vdots & \vdots & & \vdots & \vdots \\ \psi_1(\mathbf{r}_M)\alpha_M & \psi_1(\mathbf{r}_M)\beta_M & \dots & \psi_{M/2}(\mathbf{r}_M)\alpha_M & \psi_{M/2}(\mathbf{r}_M)\beta_M \end{vmatrix}$$

где

α_i и $\beta_i, i = 1, \dots, M$ – спиновые компоненты волновых функций (заданы в данном приближении),

$\mathbf{r}_i, i = 1, \dots, M$, – радиус-векторы всех электронов;

$\psi_i, i = 1 \dots, M$, – некоторые неизвестные ортонормированные функции из L_2 с областью определения аргументов R^3.

Для нахождения $\psi_i, i = 1, \dots, M$ дискретизуют задачу. Для некоторого базиса функций $\phi_j, j = 1, \dots, K$ представляя $\psi_i = \sum_{j=1}^{K} C_{ji}\phi_j$, ищут такие C_{ji}, для которых $\int_\Omega \Psi^* H \Psi d\Omega$ минимально при $\int_\Omega \Psi^* \Psi d\Omega = 1$. Это приводит к уравнениям Хартри–Фока:

$$\sum_{p=1}^{K} (F_{pq} - E_k S_{pq}) C_{pk} = 0; \forall q = 1, \dots, K; \forall k = 1, \dots, M/2; \qquad (1)$$

где

$$F_{pq} = < p|H|q > + \sum_{i,j=1; i \neq j}^{K} P_{ij}(< ij|pq > - \frac{1}{2} < ip|jq >)$$

$$< p|H'|q > = \int \phi_p^*(\mathbf{r}) \left(-\frac{1}{2} \left(\frac{d^2}{dx^2} + \frac{d^2}{dy^2} + \frac{d^2}{dz^2} \right) + \sum_{i=1}^{N} \frac{\mathbf{Q}_i}{|\mathbf{R}_i - \mathbf{r}|} \right) \phi_q(\mathbf{r}) d\mathbf{r}$$

$$< ij|kl > = \int \frac{\phi_i^*(\mathbf{r}_1)\phi_j^*(\mathbf{r}_2)\phi_k(\mathbf{r}_1)\phi_l(\mathbf{r}_2)}{\mathbf{r}_{12}} d\mathbf{r}_1 d\mathbf{r}_2$$

$$S_{ij} = < i|j > = \int \phi_i^*(\mathbf{r})\phi_j(\mathbf{r}) d\mathbf{r}, \qquad P_{ij} = \sum_{x=1}^{M/2} C_{ix}^* C_{jx}$$

В настоящее время предложено большое количество различных базисных наборов функций [4]: это и функции, являющиеся собственными функциями для частного случая решения этой задачи, и функции, являющиеся собственными функциями для

различных операторов, близких к $\hat{H}$, и позволяющих аналитически получить вид собственных функций, а также некоторые, не имеющие физического смысла функции типа B-сплайнов.

Рассмотрим вычислительные затраты при решении этой задачи для различных базисов. Во-первых, для этого необходимо вычислить порядка $\frac{K^4}{8}$ интегралов перекрывания $< ij|kl >$ и порядка $\frac{K^2}{2}$ интегралов перекрывания $< i|\hat{T}|j >$, если все базисные функции имеют ненулевые значения на всей области определения. В то же время известно из численных экспериментов [1,3], что функции φ_i обычно экспоненциально стремятся к нулю на всей области определения этой функции за исключением одного или нескольких локальных участков; таким образом оправдывается применение в качестве базисного набора функций B-сплайнов. Так в работе [5] сравнивались два таких базисных набора и было получено, что 512 специальным функциям соответствует базис из 10000 B-сплайнов (в обоих таких базисах достигалась одинаковая точность вычислений). Но базисные сплайны требуют гораздо меньше вычислительных затрат во время нахождения интегралов перекрывания: соответственно порядка K^2 и K. Кроме этого, такие интегралы проще вычисляются [6].

На стадии решения нелинейной системы уравнений набор функций, состоящий из B-сплайнов значительно проигрывает: так как обычный метод решения системы (1) – метод простой итерации [4], на каждом шаге которого надо решать проблему собственных значений для плотной эрмитовой матрицы. В случае B-сплайнов размер задачи значительно больше, чем в случае специальных функций. Нередко это полностью компенсирует быстроту вычисления интегралов перекрывания.

2 Свойства базиса на основе регулярных базисных сплайнов

Оригинальным методом решения поставленной выше задачи является выбор базиса на основе регулярных базисных сплайнов. Тогда система (1) переписывается в матричном виде:

$$\left(-\hat{L} + \hat{S}E_j - \frac{1}{2}\sum_{i=1}^{M/2}\hat{C}_i^*\hat{T}\hat{C}_i + diag\left(\hat{T}\sum_{i=1}^{M/2}\hat{C}_i^*\bar{C}_i\right)\right)\bar{C}_j = 0, \forall j = 1, \ldots, M/2, \qquad (2)$$

где $\hat{L}$ и $\hat{S}$ – ленточные трехуровневые [9] положительно определенные симметричные матрицы, причем $\hat{S}$ может быть единичной; $\hat{T}$ – теплицева трехуровневая положительно определенная симметричная матрица; $\hat{C}_i$ – ленточные трехуровневые положительно определенные симметричные матрицы, составленные из элементов i-ого собственного вектора; $\bar{C}_j$ и E_j – j-ый собственный вектор и соответствующее ему собственное значение.

Соответствие этих матриц аналогичным в формуле (1) следующее:

(1)	(2)
$\hat{L}$	$< p\|H\|q >$
$\hat{S}$	S_{pq}
$diag\left(\hat{T}\sum_{i=1}^{M/2}\hat{C}_i\bar{C}_i\right)$	$\sum_{i,j=1;i\neq j}^{K} P_{ij} < ij\|pq >$
$-\frac{1}{2}\sum_{i=1}^{M/2}\hat{C}_i\hat{T}\hat{C}_i$	$-\frac{1}{2}\sum_{i,j=1;i\neq j}^{K} P_{ij} < ip\|jq >$

Для того, чтобы посчитать все элементы матриц $\hat{L}$, $\hat{S}$ и $\hat{T}$, необходимо выполнить всего K операций — это затраты на предварительную работу, причем если не изменять объем пространства, в котором находится интересующий нас объект, то $\hat{T}$ также остается неизменным.

В качестве итерационной схемы можно предложить предобусловленный метод сопряженных градиентов, который использовался для аналогичной задачи для не регулярных B-сплайнов в [7,8]. Затраты на умножение матрицы на вектор сократятся до $K\log_2 K$ против K^2.

Рассмотрим случай только действительных коэффициентов C_{ij} для приближения, в котором базисные функции представляют собой ступеньку, то есть B-сплайн нулевой степени. Очевидно, что $\hat{L}\bar{C} = \hat{C}\bar{L}$, а так как вектор $\bar{C}$ не должен содержать нулевых элементов, то уравнение (2) при $M = 2$ запишется в виде

$$\hat{T}\hat{C}\bar{C} = 2\bar{L} + 2\bar{S}E. \tag{3}$$

Поскольку на вектор $\bar{C}$ наложено условие нормировки, то решение системы (3) находится решением теплицевой трехуровневой положительно определенной симметричной системы линейных уравнений с двумя правыми частями.

Кроме этого обнаружено, что существует большой набор физических объектов, таких, как кристаллические решетки и регулярные поверхности, для которых такая теплицева матрица является также и симметричным циркулянтом по двум или трем уровням. Это означает, что можно применять эффективные методы решения линейных систем с такими матрицами.

В этом случае будет необходимо решить систему уравнений с единственной правой частью (так как решение системы уравнений с циркулянтной матрицей и правой частью, состоящей из одинаковых элементов очевидно и требует только K вычислений для суммирования вектора). Это позволит решать такую систему за порядка $K\log_2 K$ операций с плавающей точкой.

3 Численная проверка

Для подтверждения вышеизложенных результатов были проведены численные эксперименты. Рассматривались задача о распределении в пространстве около ядра водорода двух спаренных электронов. Была взята самая простая аппроксимация степени 0, B-сплайн представлял собой ступеньку. Целью данного эксперимента являлось:

1) получить значения для электронной плотности и сравнить ее с известными ранее;

2) проследить зависимость точности получаемого решения от объема пространства, в котором рассматривается данное разбиение и значения K.

Были проведены численные эксперименты. Изменялись следующие параметры: $K = 4^3, 8^3, 16^3, 32^3, 64^3, 128^3$ (бо́льшие значения K было сложно использовать, так как потребовалась бы оперативная память объемом более 32 МБайт).

Рассматриваемый объем варьировался в пределах от $2^3\mathring{A}$ до $32^3\mathring{A}$.

Время расчета любого эксперимента на персональном компьютере *Pentium 120* не превышало минуты.

Надо отметить, что, представляя функции ϕ_i в виде базисных сплайнов нулевой степени, мы заранее отказались правильно учитывать вклад кинетической энергии электронов, что не могло не сказаться на полученных результатах.

Обсудим полученные результаты.

Наибольшая точность достигалась при максимальном разбиении $K = 128^3$ и выборе области больше $4^3\mathring{A}$. Результат хорошо кореллирует с экспериментом: главный максимум в $0.75\mathring{A}$ от ядра (экспериментально $0.59\mathring{A}$). Вид $|\psi|^2$ также соответствует эксперименту: при $r \to \infty$ и при $r \to 0$: $|\psi|^2 \to 0$. При попытке уменьшить объем пространства, в котором рассматривалось данное разбиение решение становилось значительно неустойчивым.

При понижении значения K понижалась и точность аппроксимации: так на 32 отрезках $|\psi|^2 \gg 0$, что не соответствует действительности. Это можно объяснить тем, что $|\psi|^2$ в большой окрестности точки $r = 0$ может значительно отличаться от нуля.

Список литературы

[1] W. J. Herbe, L. Radom, P. v.R. Schleyer, J.A.Pople. Ab initio Molecular orbital theory. Springer-Verlag. (1986).

[2] J. K. Labanowski, J. W. Andzelm. Density Functional Methods in Chemistry. Springer-Verlag. (1986).

[3] В. М. Татевский. Строение и физико-химические свойства молекул и веществ. М.: МГУ. 463с (1993).

[4] W. Jones, Int. J. Quant. Chem., **51**, 417, (1994).

[5] A. S. Umar, Phys. Rev. C: Nucl. Phys., **44**, 2512, (1991).

[6] J. Rao, W. Lui, Phys. Rev. A:, **50**(2), 1916, (1994).

[7] A. P. Seitsonen, M. J. Puska, and R. M. Nieminen, Phys. Rev. B:, **51**, 14057, (1995).

[8] M. C. Payne, M. P. Teter, et. al, Rev. Mod. Phys., **64**, n-4, 1045, (1992).

[9] В. В. Воеводин. Е. Е. Тыртышников. Вычислительные процессы с тёплицевыми матрицами. М.: Наука. 316с (1987).

МЕТОДИКА ПРЕОБРАЗОВАНИЯ КООРДИНАТ ЛОКАЛЬНЫХ ДЕТЕКТОРОВ В ЕДИНУЮ СИСТЕМУ [1]

З.М.Иванченко, И.М.Иванченко, Н.Н.Карпенко

ЛВТА ОИЯИ, 141980 Дубна Московской обл., Россия
e-mail: iva@cv.jinr.dubna.su

Аннотация. Анализируется математическая модель (переопределенная система нелинейных уравнений) и методы определения параметров локальных систем координат позиционно-чувствительных детекторов, образующих единый комплекс регистрации траекторной информации. Программа CORPAR, реализующая рассмотренный метод, используется для геометрической калибровки системы пропорциональных камер установки ЭКСЧАРМ на ускорителе У-70 (Протвино).

Современные позиционно-чувствительные детектирующие модули позволяют регистрировать координату точки траектории частицы в собственной локальной системе координат, так что для перехода в единую систему необходимо определить параметры соответствующего преобразования. Юстировка системы и непосредственные геодезические измерения всех требуемых величин для многокомпонентной установки сопряжены со значительными трудностями и расходованием времени. Нередки ситуации, когда происходят перемещения отдельных измерительных модулей в процессе проведения эксперимента. Погрешности передаточной функции, принципиально неустранимые только за счет геодезических измерений, не позволяют адекватно реализовать пространственное разрешение используемых траекторных детекторов. Все это обуславливает необходимость оперативного и надежного способа получения параметров, характеризующих эффективную геометрию установки.

Рассмотрим решение задачи определения параметров преобразования информации, регистрируемой координатными детекторами, в единую прямоугольную декартову систему координат XYZ. Каждому детектору соответствует локальная система координат X_iY_i $(i = 1, 2, ..., m)$ таким образом, что связь между координатой в локальной системе и соответствующими ей координатами в единой системе выражается функцией $X_i = X\cos\alpha_i + Y\sin\alpha_i + h_i$. Параметры α_i, h_i, z_i ищем из условия минимума функционала:

$$M = \frac{1}{\zeta}\sum_{j=1}^{N}\sum_{i=1}^{m} w_{ij}\Big((a_{xj}z_i + b_{xj})\cos\alpha_i + (a_{yj}z_i + b_{yj})\sin\alpha_i + h_i - x_{ij}\Big)^2. \qquad (1)$$

Здесь

N – количество траекторий, прошедших через m детекторов;

$a_{xj}, a_{yj}, b_{xj}, b_{yj}$ – параметры траектории j в единой системе координат XYZ (полный набор параметров пространственной траектории, аппроксимируемой линейной функцией);

x_{ij}, w_{ij} – измеренная координата и ее вес.

[1] Работа выполнена при поддержке Российского фонда фундаментальных исследований, гранты N 95-01-00736, N 96-07-89190

Для уменьшения влияния вычислительных погрешностей [1] представим α_i, h_i, z_i в виде:

$$\alpha_i = \alpha_i^0 + \Delta\alpha_i\ , \qquad h_i = h_i^0 + \Delta h_i\ , \qquad z_i = z_i^0 + \Delta z_i\ , \tag{2}$$

где α_i^0, h_i^0, z_i^0 – результаты геодезических измерений. Подставляя (2) в (1), получим выражение функционала в виде

$$M = \frac{1}{2}\sum_{j=1}^{N}\sum_{i=1}^{m} w_{ij}\Big((a_{xj}(z_i^0 + \Delta z_i) + b_{xj})\cos(\alpha_i^0 + \Delta\alpha_i) +$$

$$+(a_{yj}(z_i^0 + \Delta z_i) + b_{yj})\sin(\alpha_i^0 + \Delta\alpha_i) + h_i^0 + \Delta h_i - x_{ij}\Big)^2. \tag{3}$$

Используя измерения траекторий частиц опорными детекторами, находим оценки параметров пространственных траекторий $a_{xj}, a_{yj}, b_{xj}, b_{yj}$. Для решения этой задачи разработана программа фитирования точек, регистрируемых в косоугольной системе координат.

Введя новые обозначения, функционал (3) можно записать в следующем виде:

$$M = \frac{1}{2}\sum_{j=1}^{N}\sum_{i=1}^{m} w_{ij}\Big(a_{xj}^i\Delta z_i\cos\Delta\alpha_i + b_{xj}^i(\cos\Delta\alpha_i - 1) +$$

$$+(a_{yj}^i\Delta z_i + b_{yj}^i)\sin\Delta\alpha_i + c_{ij} + \Delta h_i\Big)^2, \tag{4}$$

где

$$a_{xj}^i = a_{xj}\cos\alpha_i^0 + a_{yj}\sin\alpha_i^0,$$
$$a_{yj}^i = a_{yj}\cos\alpha_i^0 - a_{xj}\sin\alpha_i^0,$$
$$b_{xj}^i = (a_{xj}z_i^0 + b_{xj})\cos\alpha_i^0 + (a_{yj}z_i^0 + b_{yj})\sin\alpha_i^0,$$
$$b_{yj}^i = (a_{yj}z_i^0 + b_{yj})\cos\alpha_i^0 - (a_{xj}z_i^0 + b_{xj})\sin\alpha_i^0,$$
$$c_{ij} = b_{xj}^i - x_{ij} + h_i^0.$$

Из условия минимума функционала (4) получим систему нормальных уравнений

$$\sum_{j=1}^{N} w_{ij}(a_{xj}^i\Delta z_i\cos\Delta\alpha_i + b_{xj}^i(\cos\Delta\alpha_i - 1) + (a_{yj}^i\Delta z_i + b_{yj}^i)\sin\Delta\alpha_i + c_{ij} + \Delta h_i)*$$

$$*(-a_{xj}^i\Delta z_i\sin\Delta\alpha_i - b_{xj}^i\sin\Delta\alpha_i + (a_{yj}^i\Delta z_i + b_{yj}^i)\cos\Delta\alpha_i) = 0,$$

$$\sum_{j=1}^{N} w_{ij}(a_{xj}^i\Delta z_i\cos\Delta\alpha_i + b_{xj}^i(\cos\Delta\alpha_i - 1) + (a_{yj}^i\Delta z_i + b_{yj}^i)\sin\Delta\alpha_i + c_{ij} + \Delta h_i)*$$

$$*(a_{xj}^i\cos\Delta\alpha_i + a_{yj}^i\sin\Delta\alpha_i) = 0, \tag{5}$$

$$\sum_{j=1}^{N} w_{ij}(a_{xj}^i\Delta z_i\cos\Delta\alpha_i + b_{xj}^i(\cos\Delta\alpha_i - 1) + (a_{yj}^i\Delta z_i + b_{yj}^i)\sin\Delta\alpha_i +$$

$$+c_{ij} + \Delta h_i) = 0,$$

где $i = 1, 2, \ldots, m$.

Полученная система трансцендентных уравнений решается итерационным методом.

Для определения множества параметров, включающего аппликату и угол поворота, необходимо, чтобы траектории были непараллельны, а дисперсии распределения координат траекторий в каждом детекторе - ненулевые. В качестве исходных данных используются результаты геодезических измерений и результаты регистрации кусочно-прямолинейных траекторий частиц (специальная тест-калибровочная и рабочая информация).

Рассмотренный алгоритм положен в основу программы CORPAR[2], используемой для геометрической калибровки координатных детекторов установки ЭКСЧАРМ[3]. Главной целью научно-физических исследований на этой установке является получение новых данных о рождении и свойствах очарованных частиц и узких резонансов, содержащих странные кварки. Установка функционирует на канале 5Н ускорителя У-70 (Протвино) в пучке нейтронов с энергией 50-70 ГэВ.

В программе CORPAR реализовано несколько динамически выбираемых режимов (методов), каждый из которых допускает параметрическую настройку. Режимы определяются множеством искомых параметров: (h_i), (h_i, z_i), (α_i, h_i, z_i). Следует отметить, что выбор множества параметров определяется не только точностью геодезических измерений и качеством юстировки (одним из главных требований к которой является устранение до минимума угловых смещений детекторов), но и уровнем шумов в исходных данных и множественностью траекторий в событии.

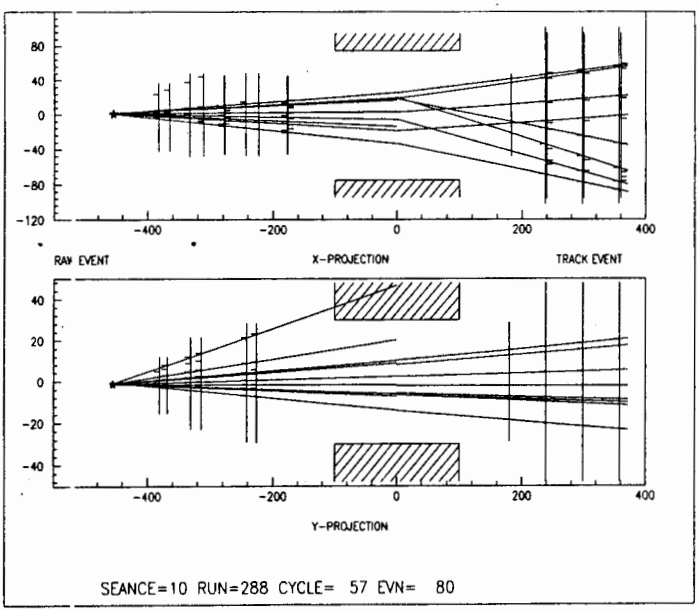


Рис.1. Типичное событие, зарегистрированное спектрометром ЭКСЧАРМ

В условиях реальных экспериментов процедура геометрической калибровки реализуется методом последовательных приближений. На первых итерациях используется (h_i) - метод с пошаговым уменьшением ширины дорожки при распознавании траекторий. На последующих итерациях используется (h_i, z_i) или (α_i, h_i, z_i) - метод в зависимости от качества юстировки и точности геодезических измерений. При таком подходе учитывается сравнительно малая чувствительность (h_i) - метода к статистическим ошибкам второго рода благодаря линейной зависимости функционального аргумента от искомого параметра.

Актуальность вычисления геометрических поправок обусловлена, в частности, сложностью изучаемых событий. Функция магнитного поля и типичное событие, зарегистрированное спектрометром ЭКСЧАРМ, соответствующее рождению очарованного бариона Σ_c^0 , показаны на рис.2.,1.

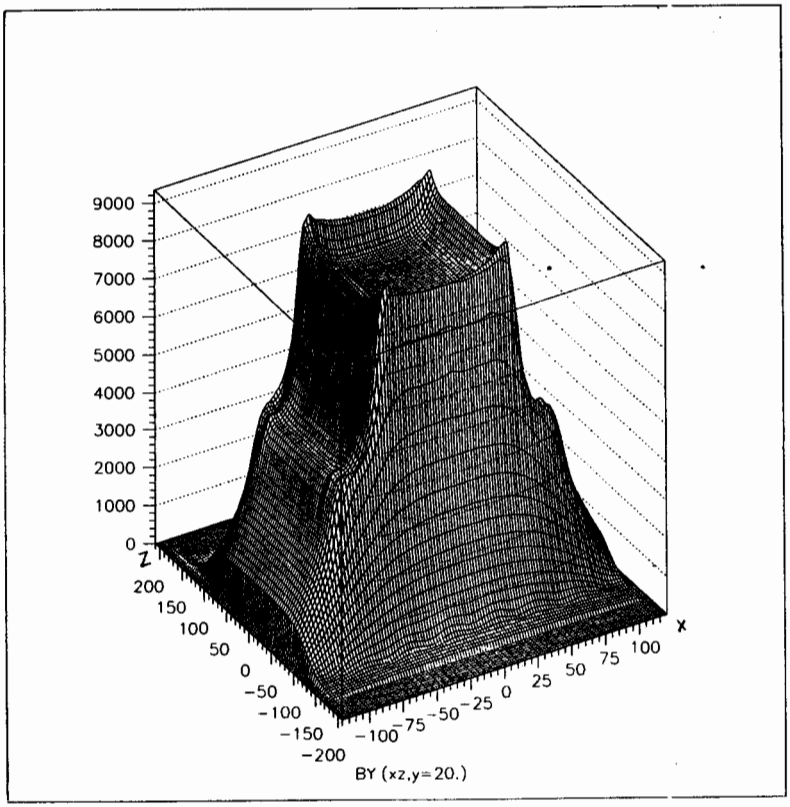


Рис.2. Функция магнитного поля

Применение геометрических поправок улучшает точность вычисления параметров найденных траекторий и событий и радикально уменьшает статистические погрешности первого и второго рода на всех этапах распознавания и анализа траекторной информации. Например, даже для двухчастичной системы (таблица) использование полученных описанной методикой величин позволяет устранить значимую систематическую погрешность, более, чем в два раза улучшить отношение сигнал/фон, повысить эффективность идентификации полезных событий, улучшить массовое разрешение и т.д.

Таблица. Статистические характеристики спектра эффективной массы системы $p\pi^-$ в области массы Λ^0 (МэВ/c^2)

Номер сеанса экспозиции ЭКСЧАРМ	10	
Номера файлов (RUNS)	268-270	
Количество триггеров	782752	
F - окрестность	$\pm$ 12	
B - окрестность	$\pm$ 5	
	CORPAR	**ГЕОДЕЗИЯ**
Смещение оценки массы	0.17 $\pm$ 0.02	0.67 $\pm$ 0.08
Разрешение	1.50 $\pm$ 0.03	2.66 $\pm$ 0.14
Количество Λ^0	8348	4276
Отношение сигнал/фон	1.5	0.67
Отношение сигнал/триггер	1.1 %	0.55 %

Программа CORPAR является составной частью математического обеспечения экспериментов, проводимых на установке ЭКСЧАРМ. Результаты определения геометрических параметров спектрометра используются при моделировании, математической обработке и физическом анализе[3,4] экспериментальной информации.

Литература.

1. Вентцель Е.С. Теория вероятностей.Изд-во "Наука", М., 1964.

2. Зинченко А.И. и др. Методика геометрической калибровки координатных детекторов установки ЭКСЧАРМ. ОИЯИ, Р10-95-541, Дубна, 1995.

3. Алеев А.Н. и др. Наблюдение очарованного бариона Σ_c^0 в эксперименте ЭКСЧАРМ. Краткие сообщения ОИЯИ 3[77]-96, Дубна,1996,с.32.

4. Алеев А.Н. и др. Исследование инклюзивного рождения φ -мезонов на серпуховском ускорителе. ОИЯИ, Р1-96-437, Дубна, 1996.

ПАРАЛЛЕЛЬНАЯ СИСТЕМА ОБРАБОТКИ ЭКСПЕРИМЕНТАЛЬНЫХ ДАННЫХ НА БАЗЕ МОДЕЛИ КЛИЕНТ-СЕРВЕР [1]

И.М.Иванченко, В.В.Кореньков, Ю.А.Кретов

ЛВТА ОИЯИ, 141980 Дубна Московской обл., Россия
e-mail: ivanko@cv.jinr.dubna.su
korenkov@cv.jinr.dubna.su
kretovyu@dct119.jinr.dubna.su

Аннотация. Разработка эффективных методов и средств обработки данных экспериментов в физике высоких энергий связана с решением двуединой задачи обеспечения скоростных вычислительных процессов и обслуживания высокоинтенсивных информационных потоков при высоком уровне автоматизации процессов массовой обработки.В работе анализируется разработанная и развиваемая в ЛВТА ОИЯИ система параллельной обработки данных в рамках архитектуры клиент-сервер на базе кластера ЭВМ, работающих под ОС типа UNIX и соединённых сетью ETHERNET.

Разработка эффективных методов и средств обработки данных экспериментов в физике высоких энергий должна прямо учитывать циклопические объёмы экспериментальных данных, требующих огромного количества вычислений. Например, в цикле отнюдь не самых ресурсоёмких экспериментов EXCHARM на ускорителе У-70 (протонный синхротрон ИФВЭ, Протвино) для одного файла данных, регистрируемых в течение 18-20 минут реального времени, только для первичной обработки требуется около 10 часов процессорного времени на самой мощной UNIX-ЭВМ общеинститутского пользования. Одно из направлений повышения производительности и эффективности систем массовой обработки связано с распараллеливанием процесса обработки.Именно этот подход положен в основу создания ПАраллельной Расширяемой Интегрированной Системы (ПАРИС-96). Основой параллелизма в рассматриваемой системе является распараллеливание данных. Компьютерный кластер включает ЭВМ под ОС типа UNIX, связанные через сеть ETHERNET. Кластер ПАРИС создается по схеме клиент-сервер. Сервером является выделенный компьютер под ОС Linux, предоставляющий клиентам (рабочим макромодулям) такие ресурсы, как общее дисковое пространство и программу-монитор для доступа к очередям ввода/вывода данных. Управление всем кластером локализовано на сервере. Клиенты используют собственный ЦП для обработки информации, а данные берут с сервера. Запуск удалённых заданий на компьютерах-клиентах производится автоматически специальной резидентной программой, которая выполняется на сервере. Все устройства внешней памяти (стримеры типа EXABYTE) подключены к серверу, поэтому гигабайтные информационные потоки локализованы. Макромодули системы благодаря их мультипрограммности могут использоваться в интерактивном режиме для мониторинга, анализа данных , подготовки программ и т.п. на фоне прохождения основного процесса - автоматической массовой обработки данных. Основные свойства системы: высокая производительность; масштабируемость; невысокие запросы

[1] Работа выполнена при поддержке Российского фонда фундаментальных исследований, грант N 95-01-00736

ресурсов рабочих макромодулей (интенсивно используется только ЦП); возможность использования удаленных ЭВМ, так как интенсивный обмен с внешними носителями локализован внутри сервера; высокая надежность; высокий уровень автоматизации.

Реализация идеи автоматизированной обработки данных экспериментов ЭКСЧАРМ одновременно на нескольких ЭВМ началась в ЛВТА ОИЯИ в 1992 г. До этого обработка велась независимо на нескольких ЕС ЭВМ. При этом не использовалось то, что эти машины имеют общую дисковую память. Компьютерную платформу параллельной системы обработки ПАРИС-ВМ [2] представлял кластер Виртуальных Машин (ВМ) с общим пространством минидисков, где хранились входные и выходные очереди данных, а также набор функциональных программ обработки данных и доступа к очередям.Архитектуру ПАРИС-ВМ можно охарактеризовать как симметричную (без выделенной управляющей ВМ), где функция доступа к общим ресурсам кластера возлагалась на штатную универсальную операционную систему VM-CMS.

Быстрое развитие и внедрение так называемых персональных ЭВМ на базе процессора Intel показало перспективность использования таких машин для решения задач в области физики высоких энергий [3]. ПЭВМ довольно быстро достигли приемлемого уровня интеграции со сложившимся компьютерным окружением ядерно-физических исследований. Тогда было решено реализовать принципы параллельной обработки данных кластера ПАРИС-ВМ на новой компьютерной платформе.В ПАРИС-94 использовались ПЭВМ различных конфигураций под MS DOS. В качестве коммуникационной среды использовалась общеинститутская сеть ETHERNET. В качестве файлового сервера кластера сначала использовалась PC-486/66 под MS DOS с сетевой надстройкой NWLite, а затем выделенный файл-сервер Novell NetWare 3.12. Архитектура ПАРИС-94 также является симметричной. Контроль за использованием общих дисковых ресурсов осуществляла сетевая ОС (NetWare или NWLite). Использование выделенного сервера NetWare породило следующую проблему. Объем данных, обрабатываемых за сутки, составлял порядка 0.5-1 Гб. Так как файл-сервер являлся выделенным (никаких дополнительных функций он не выполнял), а коммуникации были общими, то информационный поток по сети ETHERNET, создаваемый системой ПАРИС-94, приводил к ощутимым перегрузкам в общеинститутской коммуникационной среде. Было введено искусственное ограничение скорости передачи данных в сеть (не более 100 Кб/с), но такое разравнивание информационного потока значимо снижало производительность кластера ПАРИС-94. Чтобы не строить специальный локальный сегмент сети, функции файл-сервера и сервера ввода-вывода (имеющего доступ к стримеру EXABYTE) были совмещены в рамках одного компьютера. Эффективное решение этой частной , но немаловажной задачи, было одним из условий при выборе операционной системы на очередном этапе развития ситемы ПАРИС

Как и в кластере ПАРИС-94, коммуникационной средой кластера ПАРИС-96 осталась общеинститутская сеть ETHERNET. Сервер системы (Pentium-166), защищенный от колебаний напряжения в сети источником бесперебойного питания BackUPS, функционирует под ОС Linux. Гигабайтные потоки данных локализованы внутри сервера: общее дисковое пространство кластера и стримерные накопители подключены к серверу напрямую. Таким образом, поток данных в общей сети ETHERNET, порождаемый деятельностью ПАРИС-96, определяется только количеством и производительностью ра-

бочих узлов кластера. Следовательно, нет необходимости создавать отдельный сегмент сети, можно строить распределенную систему обработки данных. В качестве функциональных звеньев кластера ПАРИС-96 (клиентов) можно использовать любые ЭВМ под управлением ОС UNIX (ПЭВМ-486, Pentium, рабочие станции, машины коллективного пользования), в то время как в ПАРИС-92 это могли быть только ЕС-ЭВМ, а в кластере ПАРИС-94 - ПЭВМ под управлением MS DOS. Со стороны клиентов требуется лишь поддержка сетевых протоколов TCP/IP, rsh/rcp.

В отличие от предыдущих версий ПАРИС, система ПАРИС-96 построена по технологии клиент-сервер, где существует не только выделенный файл-сервер, но и специальная задача (процесс), управляющая всей работой кластера. Такой подход позволяет оптимально синхронизировать взаимодействие узлов кластера. Наличие в кластере двух устройств внешней памяти типа EXABYTE (для хранения входных и результирующих данных) позволяет поддерживать рабочую фазу комплекса без вмешательства оператора на время обработки одной кассеты (емкость кассеты EXB-8500 составляет около 4.5 Гб).

Все данные, хранящиеся в общем дисковом хранилище кластера, можно разделить на следующие логические очереди:
- входная очередь (считанные с ленты файлы данных, ожидающие начала обработки);
- рабочая очередь (данные в процессе обработки);
- очередь отложенных заданий (сюда данные попадают из рабочей очереди после аварийной остановки счета);
- выходная очередь (содержит обработанные данные, готовые для сброса на результирующую ленту).

Сервер системы обработки предназначен для выполнения следующих задач:
- автоматизированное пополнение входной очереди;
- освобождение выходной очереди;
- обеспечение целостности данных кластера при переносе элементов из одной очереди в другую (также восстановление целостности после сбоя сервера);
- автоматический запуск удаленных заданий на рабочих узлах кластера (протокол rsh, сервер выступает в качестве инициатора).

Клиент (рабочий узел) кластера предназначен только для ведения обработки того фрагмента данных из общего хранилища, который укажет сервер при запуске счета. Единственный ресурс клиента, который будет использоваться интенсивно - это ЦП. Скорость сетевого интерфейса и емкость локального диска не являются критическими параметрами клиента. Например, для обработки данных экспериментов EXCHARM достаточной пропускной способностью сетевого адаптера является 40 Кб/с, а свободный объём локального диска должен составлять менее 4 Мб (включая размер исполняемого модуля обработки данных).

Исходя из архитектуры кластера можно выделить несколько последовательных участков, через которые проходят данные в процессе их обработки:
1) Копирование входных данных со стримера на общий диск.
2) Обработка данных звеньями кластера (чтение и запись данных производится с сетевого диска).
3) Копирование полученных результатов на стример.

Что касается участков 1 и 3, то здесь ограничение скорости передачи данных (не более 500 Кб/с) определяется характеристиками накопителя EXB-8500. На участке 2 эффективность следующих узлов сервера может влиять на производительность кластера ПАРИС-96:

- дисковой подсистемы (диски и дисковый контроллер);
- сетевого интерфейса (ETHERNET-адаптер);
- реализации используемой ОС UNIX (ядро UNIX и сетевые протоколы).

Для ОС Linux 2.0.0 на Pentium в общем случае можно добиться функционирования всех узлов, описанных выше, в темпе 700-1000 Кб/с.

Подведя итог, можно видеть, что критическим звеном в кластере ПАРИС-96 выступает стример EXB-8500 с ограничением 500 Кб/с. Из этого ограничения, зная производительность программы обработки данных, можно вычислить максимальное число рабочих узлов (клиентов) кластера. Например, для рассматриваемых экспериментов пиковая производительность (предел масштабируемости) достигается в следующей конфигурации: 20 компьютеров типа Pentium/166 или HP-9000/712, либо 45 компьютеров типа Pentium/66 или SUN-20, либо 111 компьютеров типа IBM PC 486/66.

Для повышения пиковой производительности (уровня плато) системы хотя бы в два раза нужно радикально менять звенья кластера (переходить на FastEthernet, использовать более скоростные стримеры и т.д.), что приведет к снижению такого важного параметра , как отношение производительность/цена. Повышение уровня автоматизации без снижения других характеристик достигается за счёт использования робота или, по крайней мере, последовательного блока для нескольких кассет.

Потенциал описанного комплекса ПАРИС-96 для обработки данных достаточно высок (до 700 событий в секунду, что почти в 3 раза превышает скорость их измерения), но на практике он еще не был полностью использован из-за недостатка финансирования. Практическое использование системы [4] подтвердило её основные свойства: высокая производительность, экономичность, масштабируемость, возможность использования удаленных ЭВМ , высокая надежность.

Литература

1. Nash T. et al. The ACP multiprocessor system at Fermilab. Proc. of 23-rd int. conf. on HEP, Berkly, CA, USA, 1986, p.1459.

2. Иванченко И.М. и др. Архитектура системы массовой обработки экспериментальных данных с установки EXCHARM на базе кластера виртуальных машин.//XVI Межд. Симпозиум по Ядерной Электронике и VI Межд. Школа по Автоматизации ...,Варна,Болгария,Сент.,1994. ОИЯИ,D13-94-491,Дубна,1994,с.85.

3. Иванченко И.М.,Седых Ю.В. Рабочая станция на базе ПЭВМ для разработки программного обеспечения экспериментов в области физики высоких энергий. Программирование,1988,N4,с.102.

4. Алеев А.Н. и др. Исследование инклюзивного образования φ-мезонов нейтронами на серпуховском ускорителе. ОИЯИ,Р1-96-437,Дубна,1996.

РАСПРОСТРАНЕНИЕ РАДИОВОЛН В НЕОДНОРОДНОМ ПРОСТРАНСТВЕ С НЕРЕГУЛЯРНОЙ ГРАНИЦЕЙ РАЗДЕЛА МЕЖДУ ДВУМЯ СРЕДАМИ

А.С.Ильинский, Т.Н.Галишникова

Факультет вычислительной математики и кибернетики
Московского государственного университета им. М.В.Ломоносова.
Россия, 119899, Москва, Воробьевы горы
e-mail: celd@cs.msu.su

Abstract. The problem of electromagnetic field diffraction by a cylindrical nonregular interface between two media is discussed. The boundary value problem for the Maxwell set of equations in the half-space is reduced to sets of integral equations. The numerical algorithms are based on the summability theory of series with logarithmic singularity which define the integral equation kernels. The numerical algorithms are used to evaluate the mathematical models domain of application.

Одним из распространенных методов исследования окружающей среды является метод обзора земной поверхности с помощью различных технических средств, установленных на летательных аппаратах. Среди таких средств важное место занимают радиолокационные станции обзора земной поверхности, позволяющие получить изображение с большого расстояния. Интерпретация результатов электромагнитных зондирований позволяет обнаружить малоразмерные объекты, получить детальные изображения местности, дать заключение о состоянии почв, загрязнении водной поверхности, а также решить ряд других актуальных задач.

Для интерпретации результатов зондирования большой интерес представляет изучение зависимости характеристик отраженного поля от формы и проницаемости участка отражающей поверхности. Для таких исследований требуется разработка эффективных методов решения задач дифракции электромагнитного поля, падающего на поверхность раздела сред с достаточно произвольными характеристиками и формой. Важны как постановка прямой задачи - определение зависимости отраженного сигнала от формы и физических характеристик отражающей поверхности, так и обратная задача восстановления характеристик отражающей поверхности по величине принятого сигнала. Для решения обратных задач необходимо иметь эффективный алгоритм решения прямых задач.

В настоящей работе обсуждаются численные методы анализа задач дифракции плоской электромагнитной волны, падающей на нерегулярную поверхность раздела двух различных сред. Рассматриваются как строго периодическая цилиндрическая поверхность, так и цилиндрическая поверхность, отличная от плоскости на конечном участке.

1. Интегральные уравнения в задаче дифракции на периодической границе раздела двух сред. Будем рассматривать периодическую границу раздела двух сред. Выберем декартову систему координат xyz таким образом,

чтобы ось z была параллельна цилиндрической образующей поверхности раздела, периодической в направлении оси x. Будем считать, что что каждая точка на границе раздела однозначно проектируется на ось x, т.е. уравнение границы есть $y = f(x)$. Пусть среда над границей, в которой распространяется падающее поле (область D_1), и под границей (область D_2) определяется диэлектрическими $\varepsilon_{1,2}$ и магнитными $\mu_{1,2}$ проницаемостями соответственно. Волновые числа есть $k_{1,2} = \omega\sqrt{\varepsilon_{1,2}\mu_{1,2}}$, ω - круговая частота. Будем предполагать, что ε_1 - действительное число. Зависимость от времени $\exp(-i\omega t)$.

Обозначим через $\mathbf{E}_0(x,y,z)$, $\mathbf{H}_0(x,y,z)$ падающее поле плоской волны вида

$$\mathbf{E}_0(x,y,z) = \mathbf{E}_0 \exp\left(-i\alpha_0 x - i\beta_0 y + i\gamma_0 z\right),$$

$$\mathbf{H}_0(x,y,z) = \mathbf{E}_0\sqrt{\frac{\varepsilon_0}{\mu_0}} \exp\left(-i\alpha_0 x - i\beta_0 y + i\gamma_0 z\right), \qquad (1)$$

где

$$\alpha_0 = k_1 \sin\varphi_0 \sin\theta_0, \qquad \beta_0 = k_1 \sin\varphi_0 \cos\theta_0, \qquad \gamma_0 = k_1 \cos\varphi_0; \qquad (2)$$

здесь φ_0 - угол между осью z и направлением распространения падающей волны, θ_0 - угол между отрицательным направлением оси y и проекцией на плоскость $z = 0$ направления распространения падающего поля, имеющего амплитуду $\mathbf{E}_0$.

Требуется найти решение однородной системы уравнений Максвелла в областях D_1, D_2

$$\operatorname{rot}\mathbf{H}^{(1,2)}(x,y,z) + i\omega\varepsilon_{1,2}\mathbf{E}^{(1,2)}(x,y,z) = \mathbf{0},$$

$$\operatorname{rot}\mathbf{E}^{(1,2)}(x,y,z) - i\omega\mu_{1,2}\mathbf{H}^{(1,2)}(x,y,z) = \mathbf{0}, \qquad (3)$$

с граничными условиями непрерывности тангенциальных компонент электромагнитного поля при переходе через границу раздела сред

$$\left[\mathbf{n}, \mathbf{E}^{(1)}\right] = \left[\mathbf{n}, \mathbf{E}^{(2)}\right], \qquad \left[\mathbf{n}, \mathbf{H}^{(1)}\right] = \left[\mathbf{n}, \mathbf{H}^{(2)}\right], \qquad (4)$$

отраженное и прошедшее поля удовлетворяют условиям излучения.

С учетом геометрии структуры и характера падающего поля задача дифракции во всем пространстве сводится к решению внутри одного периода, границу раздела в котором обозначим S_0.

Будем искать полное электромагнитное поле в виде

$$\mathbf{E}(x,y,z) = \mathbf{E}(x,y)\exp\left(i\gamma_0 z\right), \qquad \mathbf{H}(x,y,z) = \mathbf{H}(x,y)\exp\left(i\gamma_0 z\right). \qquad (5)$$

Можно показать, что проекции на ось z координат полей $\mathbf{E}^{(1,2)}(x,y)$, $\mathbf{H}^{(1,2)}(x,y)$ удовлетворяют двумерному уравнению Гельмгольца

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + k_{1,2}^2 - \gamma_0^2\right)\begin{pmatrix} E_z^{(1,2)}(x,y) \\ H_z^{(1,2)}(x,y)\end{pmatrix} = 0, \qquad (6)$$

граничные условия (4) примут вид

$$E_z^{(1)} = E_z^{(2)}, \qquad H_z^{(1)} = H_z^{(2)},$$

$$-\frac{\omega\mu_1}{\beta_1^2}\frac{\partial H_z^{(1)}}{\partial n}+\frac{\gamma_0}{\beta_1^2}\frac{\partial E_z^{(1)}}{\partial \tau}=-\frac{\omega\mu_2}{\beta_2^2}\frac{\partial H_z^{(2)}}{\partial n}+\frac{\gamma_0}{\beta_2^2}\frac{\partial E_z^{(2)}}{\partial \tau},$$

$$\frac{\omega\varepsilon_1}{\beta_1^2}\frac{\partial E_z^{(1)}}{\partial n}+\frac{\gamma_0}{\beta_1^2}\frac{\partial H_z^{(1)}}{\partial \tau}=\frac{\omega\varepsilon_2}{\beta_2^2}\frac{\partial E_z^{(2)}}{\partial n}+\frac{\gamma_0}{\beta_2^2}\frac{\partial H_z^{(2)}}{\partial \tau}, \qquad (7)$$

где

$$\frac{\partial}{\partial n}=n_x\frac{\partial}{\partial x}+n_y\frac{\partial}{\partial y}, \qquad \frac{\partial}{\partial \tau}=n_x\frac{\partial}{\partial y}-n_y\frac{\partial}{\partial x},$$

n_x, n_y - направляющие косинусы нормали $\mathbf{n}$, внешней к области D_1, $\beta_{1,2}^2=k_{1,2}^2-\gamma_0^2$. Далее будем предполагать, что $\mu_1=\mu_2=\mu_0$.

Для получения интегральных уравнений используем формулы Грина и двумерные квазипериодические фундаментальные решения $g_{1,2}(M,P)$ для уравнения Гельмгольца:

$$\Delta g_{1,2}(M,P)+\left(k_{1,2}^2-\gamma_0^2\right)g_{1,2}(M,P)=-2\pi\delta(M,P). \qquad (8)$$

Функции $g_{1,2}(M,P)$ определены во всем пространстве R^2, удовлетворяют условиям излучения и имеют вид

$$g_{1,2}(M,P)=\frac{i\pi}{2}\sum_{n=-\infty}^{\infty}\frac{\exp\left(i\Lambda_n\Delta x\right)\exp\left(i\Gamma_n^{(1,2)}|\Delta y|\right)}{\Gamma_n^{(1,2)}},$$

$$\Lambda_n=(t+2\pi n)/b, \qquad \Gamma_n^{(1,2)}=\sqrt{k_{1,2}^2-\gamma_0^2-\Lambda_n^2}, \qquad \operatorname{Im}\Gamma_n^{(1,2)}>0,$$

$$\operatorname{Im}\Gamma_n^{(1,2)}=0, \qquad \operatorname{Re}\Gamma_n^{(1,2)}>0, \qquad (9)$$

где b - период решетки, $k_2=k_1\sqrt{\varepsilon_2}$, $t=k_1 b\sin\theta_0$, $\Delta x=x_M-x_P$, $\Delta y=y_M-y_P$.

Используя граничные условия (7) и свойства поверхностных потенциалов, получим систему четырех интегральных уравнений второго рода относительно неизвестных $u(P)=E_z^{(1)}(P)$, $v(P)=H_z^{(1)}(P)$, $\partial u(P)/\partial n$, $\partial v(P)/\partial n$ на границе S_0 довольно громоздкого вида, содержащую, кроме того, и производные искомого решения по касательной [1].

В случае, когда падающее поле не зависит от координаты z, т.е. $\gamma_0=0$, граничные условия принимают вид

$$E_z^{(1)}=E_z^{(2)}, \qquad H_z^{(1)}=H_z^{(2)}, \qquad \frac{\partial E_z^{(1)}}{\partial n}=\frac{\partial E_z^{(2)}}{\partial n}, \qquad \frac{1}{\varepsilon_1}\frac{\partial H_z^{(1)}}{\partial n}=\frac{1}{\varepsilon_2}\frac{\partial H_z^{(2)}}{\partial n}, \qquad (10)$$

а полученная ранее система четырех уравнений распадается на две независимые системы из двух интегральных уравнений относительно $u(P)$, $v(P)$, $\partial u(P)/\partial n$, $\partial v(P)/\partial n$:

$$\hat{u}(M)=\frac{1}{2\pi}\int_{S_0}\left[\hat{u}(p)\frac{\partial}{\partial n_P}(g_2-g_1)+\frac{\partial\hat{u}(P)}{\partial n}(g_1-\alpha g_2)\right]ds_P+\hat{u}_0(M), \qquad (11)$$

$$\frac{1}{2}(1+\alpha)\frac{\partial\hat{u}(M)}{\partial n}=\frac{1}{2\pi}\int_{S_0}\left[\hat{u}(P)\frac{\partial^2}{\partial n_M\partial n_P}(g_2-g_1)-\right.$$

$$- \frac{\partial \hat{u}(P)}{\partial n_P} \frac{\partial}{\partial n_M} (\alpha g_2 - g_1) \Big] ds_P + \frac{\partial \hat{u}_0(M)}{\partial n_M}, \quad M \in S_0, \tag{12}$$

где

$$\alpha = 1, \qquad \hat{u}(P) = u(P) = E_z^{(1)}(P), \qquad E - p;$$

$$\alpha = \varepsilon_2/\varepsilon_1, \qquad \hat{u}(P) = v(P) = H_z^{(1)}(P), \qquad H - p;$$

$$u_0(M) = E_{0,z}(x, y), \qquad v_0(M) = H_{0,z}(x, y).$$

В уравнениях (11), (12) ядра интегральных уравнений имеют слабую особенность при совпадении аргументов, поскольку выражаются через разность фундаментальных решений и их вторых нормальных производных. Построены алгоритмы для их вычисления, основанные на суммировании в явном виде рядов с логарифмической особенностью и улучшении сходимости входящих в первую производную рядов. При построении этих алгоритмов учитывалась асимптотика исследуемых функций.

В частном случае, когда область D_2 является хорошо проводящей, на границе раздела сред электромагнитное поле удовлетворяет условиям Леонтовича

$$[\mathbf{n}, \mathbf{E_1}] = -W [\mathbf{n}, [\mathbf{n}, \mathbf{H_1}]], \tag{13}$$

где $W = \sqrt{\mu_0/\varepsilon_2}$ - импеданс среды D_2. Дифракционная задача сводится к решению в области D_1 двух независимых уравнений Гельмгольца вида ($\gamma_0 = 0$)

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + k_1^2 \right) \begin{pmatrix} u(P) \\ v(P) \end{pmatrix} = 0, \tag{14}$$

с импедансным граничным условием

$$u(P) + \alpha_E \frac{\partial u(P)}{\partial n} = 0, \qquad \alpha_E = \frac{iW}{k_1}, \qquad E - p,$$

$$\alpha_H v(P) + \frac{\partial v(P)}{\partial n} = 0, \qquad \alpha_H = ik_1 W, \qquad H - p.$$

Рассеянное поле в области D_1 должно удовлетворять условиям излучения. Соответствующие интегральные уравнения записываются следующим образом:

$$\frac{1}{2}\alpha_E \frac{\partial u(P)}{\partial n} + \frac{1}{2\pi} \int_{S_0} \left[g_1(M, P) + \alpha_E \frac{\partial g_1(M, P)}{\partial n_P} \right] \frac{\partial u(P)}{\partial n} ds_P = -u_0(M),$$

$$M \in S_0, \qquad E - p; \tag{15}$$

$$\frac{1}{2} v(M) + \frac{1}{2\pi} \int_{S_0} \left[\frac{\partial g_1(M, P)}{\partial n_P} - \alpha_H g_1(M, P) \right] v(p) ds_P = v_0(P),$$

$$M \in S_0, \qquad H - p. \tag{16}$$

Здесь $u_0(P)$, $v_0(P)$ - падающее поле.

Так как $\varepsilon_2 = \varepsilon_2' + i\varepsilon_2''$, $\varepsilon_2'' = i\sigma/\omega$, то для хорошо проводящих сред сравнение численных результатов решения систем интегральных уравнений (11), (12)

с решением уравнений (15), (16) позволяет оценить границы применимости более простой математической модели.

2. Интегральные уравнения в задаче дифракции на конечном импедансном участке взволнованной поверхности.

Рассмотрим дифракционную задачу в предположении, что отражающая поверхность представляет собой отражающую плоскость, деформированную в конечной области, причем деформированная поверхность однозначно проектируется на плоскость. Будем рассматривать такие поверхности S, которые задаются уравнением $y = f(x)$, где $\operatorname{supp} f(x)$ ограничен, и которые целиком содержатся в полупространстве $y \geq 0$, что позволяет свести дифракционную задачу к решению одного (а не системы) интегрального уравнения.

Обозначим через D_1 полупространство $y \geq f(x)$ с нерегулярной границей, на которую падает плоская волна

$$u_0(x, y) = \exp\left(ik_1\left(-x\sin\theta_0 - y\cos\theta_0\right)\right). \tag{17}$$

Требуется найти в области D_1 решение двумерного уравнения Гельмгольца (14) относительно полного поля $u(P)$ в случае E-поляризации или $v(P)$ в случае H-поляризации, удовлетворяющее на S импедансным граничным условиям, а на полуплоскости условиям идеальной проводимости

$$u(P) = 0, \qquad E - p, \tag{18}$$

$$\frac{\partial v(P)}{\partial n} = 0, \qquad H - p. \tag{19}$$

На бесконечности рассеянное поле должно удовлетворять условиям излучения.

Введем функции Грина, определенные в полупространстве $y \geq 0$:

$$g^E(M, P) = \frac{i\pi}{2}\left[H_0^{(1)}\left(k_1\sqrt{(\triangle x)^2 + (\triangle y)^2}\right) - H_0^{(1)}\left(k_1\sqrt{(\triangle x)^2 + (sy)^2}\right)\right], \tag{20}$$

$$g^H(M, P) = \frac{i\pi}{2}\left[H_0^{(1)}\left(k_1\sqrt{(\triangle x)^2 + (\triangle y)^2}\right) + H_0^{(1)}\left(k_1\sqrt{(\triangle x)^2 + (sy)^2}\right)\right], \tag{21}$$

удовлетворяющие на границе $y = f(x)$ соответственно краевым условиям первого или второго рода, где $\triangle x = x_M - x_P$, $\triangle y = y_M - y_P$, $sy = y_M + y_P$. Соответствующие интегральные уравнения имеют вид, аналогичный уравнениям (15), (16), только вместо $g_1(M, P)$ берется либо функция $g^E(M, P)$ (E-поляризация), либо $g^H(M, P)$ (H-поляризация). Интегрирование производится по границе S, а правая часть в уравнениях задается следующим образом:

$$u_0(x, y) = \exp\left(ik_1\left(-x\sin\theta_0 - y\cos\theta_0\right)\right) - \exp\left(ik_1\left(-x\sin\theta_0 + y\cos\theta_0\right)\right). \tag{22}$$

$$v_0(x, y) = \exp\left(ik_1\left(-x\sin\theta_0 - y\cos\theta_0\right)\right) + \exp\left(ik_1\left(-x\sin\theta_0 + y\cos\theta_0\right)\right). \tag{23}$$

В соответствии с разработанной методикой построены вычислительные алгоритмы и проведен анализ основных особенностей поведения характеристик отраженного поля. Для периодической поверхности импедансное приближение хорошо описывает характеристики рассеянного поля при наличии проводящей среды [1].

Список литературы

[1] Т.Н. Галишникова, А.С.Ильинский. Рассеяние плоской волны на волнистой поверхности Математические модели естествознания. М.: Изд-во МГУ. С. 86-111. (1995).

МОНОТОННЫЕ РАЗНОСТНЫЕ СХЕМЫ БЕГУЩЕГО СЧЕТА НА НЕРАВНОМЕРНЫХ СЕТКАХ [1]

П. П. Матус

Институт математики АНБ,
Республика Беларусь, 220072, г. Минск, ул. Сурганова, 11
e-mail: imanb%imanb.belpak.minsk.by@demos.su, dns@im.basnet.minsk.by

Abstract. В работе для одномерного и многомерного параболического уравнения с постоянными коэффициентами строятся монотонные разностные схемы бегущего счета второго порядка аппроксимации на произвольной неравномерной сетке. Доказана сходимость предложенных схем в норме C со вторым порядком.

При численном моделировании современных задач науки и техники с особенностями (ударные волны, пограничные слои и др.) часто используют неравномерные сетки. Как правило, они имеют лишь первый порядок аппроксимации по пространственной переменной.

В [1] для одномерных нестационарных задач построены и исследованы различные классы разностных схем повышенного порядка аппроксимации на неравномерных прямоугольных сетках. Даже для неявных схем получены несколько нетипичные условия устойчивости: $\tau \geq ch^2$, $h = \max\limits_i h_i$; $\tau \geq c \max\limits_i |h_{i+1} - h_i|^2$, $c = \mathrm{const} > 0$. В [2, 3] полученные результаты обобщаются на многомерные эллиптические уравнения с краевыми условиями Дирихле. Основная идея метода заключается в аппроксимации уравнения в некоторой неузловой точке (в случае прямоугольных сеток — это центр масс системы материальных точек единичной массы, входящих в шаблон схемы). Поскольку для нестационарных уравнений в этом случае производную по времени приходится интерполировать по соседним узлам, то проблема построения и исследования явных разностных схем в этом случае представляет самостоятельный интерес.

1 Одномерное уравнение

Для простоты рассмотрим параболическое уравнение

$$\frac{\partial u}{\partial t} = k\frac{\partial^2 u}{\partial x^2} + f(x,t), \qquad 0 < x < l, \quad 0 < t \leq T \tag{1}$$

с однородными граничными условиями и заданными начальными данными, $k = \mathrm{const} > 0$. Введем в рассмотрение неравномерную сетку узлов по пространству

$$\hat{\bar{\omega}}_h = \{x_i = x_{i-1} + h_i,\ i = 1, 2, \ldots, N,\ x_0 = 0,\ x_N = l\} \tag{2}$$

и равномерную по времени

$$\bar{\omega}_\tau = \{t_j = j\tau,\ j = 0, 1, \ldots, j_0;\ \tau = T/j_0\}.$$

[1] Работа выполнена при финансовой поддержке Белорусского Фонда Фундаментальных Исследований

Для определенности будем предполагать, что $h_{i+1} \geq h_i$ (сетка сгущается к началу координат). Для аппроксимации задачи (1) применим двухслойную схему вида

$$y_t + \frac{h_+ - h}{3} y_{t\bar{x}} = k y_{\bar{x}\hat{x}} + f(\bar{x}, t), \tag{3}$$

$$\hat{y}_0 = \hat{y}_N = 0, \quad y^0(x) = u_0(x), \quad x \in \hat{\omega}_h. \tag{4}$$

Здесь $h_\pm = h_{i\pm1}$, $y_{\bar{x}} = (y_i - y_{i-1})/h_i$, $\hbar = 0{,}5(h + h_+)$,

$$y_{\bar{x}\hat{x}} = \frac{1}{\hbar_i}\left(\frac{y_{i+1} - y_i}{h_{i+1}} - \frac{y_i - y_{i-1}}{h_i}\right), \qquad \bar{x} = x + \frac{h_+ - h}{3}, \qquad x \in \hat{\omega}_h.$$

В работе [1] показано, что схемы вида (3), (4) аппроксимируют исходную дифференциальную задачу на произвольной неравномерной пространственной сетке с порядком $O(\hbar^2 + \tau)$ относительно узла $(\bar{x}, t)$.

Теорема 1. *Пусть шаг τ подчинен условиям:*

$$\frac{h_{i+1}^2 - h_i^2}{6k} \leq \tau \leq \frac{h_{i+1}^2 + 2h_i h_{i+1}}{6k}, \quad h_{i+1} \geq h_i, \quad i = \overline{1, N-1}. \tag{5}$$

Тогда разностная схема бегущего счета (3), (4) монотонна по начальным данным, правой части и имеет место оценка

$$\max_{t \in \omega_\tau} \|y(t+\tau)\|_C \leq \|y(0)\|_C + \sum_{t'=\tau}^{t} \tau \|\bar{f}(t')\|_C, \qquad t \in \omega_\tau, \tag{6}$$

где $y(t) = y(t_j) = (y_1^j, y_2^j, \ldots, y_{N-1}^j)$, $\|\cdot\|_C = \max\limits_{x \in \hat{\omega}_h} |\cdot|$, $\bar{v} = v(\bar{x})$.

Доказательство. Разностное уравнение (3) запишем в индексной форме

$$\begin{aligned}(1 + \tilde{h}_i) y_i^{j+1} = &\left(\frac{\tau k}{h_i \hbar_i} - \tilde{h}_i\right) y_{i-1}^j + \left(1 + \tilde{h}_i - \frac{2\tau k}{h_i h_{i+1}}\right) + \\ &y_i^j + \frac{\tau k}{\hbar_i h_{i+1}} y_{i+1}^j + \tilde{h}_i y_{i-1}^{j+1} + \tau \bar{f}_i, \qquad \tilde{h} = \frac{h_+ - h}{3h}.\end{aligned} \tag{7}$$

Так как при выполнении системы неравенств (5) все коэффициенты в правой части положительны (выполняется свойство монотонности), то из (7) находим

$$(1 + \tilde{h}_i)|y_i^{j+1}| \leq \|y\|_C + \tilde{h}_i \|y^{j+1}\|_C + \tau \|f(t_j)\|_C. \tag{8}$$

Так как последнее неравенство выполнено для всех $i = 1, 2, .., N-1$, то из (8) следует (6). Теорема доказана.

Замечание 1. Так как y_0^{j+1} задано, то решение схемы находится последовательно слева направо по формулам бегущего счета.

Замечание 2. Если неравномерная сетка сгущается по другому закону ($h_{i+1} \leq h_i$), то в уравнении (3) следует взять другую направленную разность

$$y_t + \frac{h_+ - h}{3} y_{t\underline{x}} = y_{\bar{x}\hat{x}} + f(\bar{x}, t). \tag{9}$$

Замечание 3. При наличии в уравнении производных первого порядка следует использовать следующую аппроксимацию [4]

$$\partial u/\partial x \sim u_{\tilde{x}}, \quad u_{\tilde{x}} = (u_x + u_{\overset{\circ}{x}} + u_{\bar{x}})/3,$$

$$u_x = (u_+ - u)/h_+, \quad u_{\overset{\circ}{x}} = (u_+ - u_-)/(h_+ + h), \quad u_{\bar{x}} = (u - u_-)/h.$$

Причем, как показано в этой же работе $\dfrac{\partial u}{\partial x}(\bar{x}, t_j) - u_{\tilde{x}} = O(h^2)$ на произвольной неравномерной пространственной сетке.

Замечание 4. Для уравнения (1) зададим неоднородные граничные условия первого рода

$$u(0,t) = \mu_1(t), \quad u(l,t) = \mu_2(t), \quad t \geq 0, \qquad u(x,0) = u_0(x),$$

которое на сетке $\hat{\bar{\omega}}_h \times \bar{\omega}_\tau$ аппроксимируются точно

$$\hat{y}_0 = \mu_1(\hat{t}), \quad \hat{y}_N = \mu_2(\hat{t}), \qquad y^0(x) = u_0(x), \quad x \in \hat{\bar{\omega}}_h.$$

Тогда теорема 1 остается справедливой при $\mu_1(t) = 0$, $\mu_2(t) \neq 0$. В общем случае оценка (6) имеет место, если первые два шага равны $h_1 = h_2$.

Замечание 5. На основе априорной оценки (6) можно получить и соответствующие оценки скорости сходимости. Подставляя $y = z + u$ в уравнения (3), (10), получим задачу для погрешности z:

$$z_t + \frac{h_+ - h}{3} z_{t\bar{x}} = k z_{\bar{x}\tilde{x}} + \psi(\bar{x}, t), \tag{12}$$

$$\hat{z}_0 = \hat{z}_N = 0; \qquad z^0(x) = 0, \quad x \in \hat{\bar{\omega}}_h. \tag{13}$$

Как показано в [1], для $\psi = \psi_1 + \psi_2$,

$$\psi_1 = \frac{\partial u}{\partial t}(\bar{x}, t) - \left(u_t + \frac{h_+ - h}{3} u_{t\bar{x}}\right), \qquad \psi_2 = k\left(u_{\bar{x}\tilde{x}} - \frac{\partial^2 u}{\partial x^2}\right)$$

имеют место оценки

$$\|\psi_k\|_C \leq c(h^2 + \tau), \quad k = 1, 2, \ h = \max_i h_i, \ c = \text{const} > 0. \tag{14}$$

Следовательно, на основании теоремы 1 для решения задачи (12), (13) имеет место оценка

$$\max_{t \in \omega_\tau} \|z(t)\|_C \leq T\|\psi\|_C \leq c_1(h^2 + \tau), \quad c_1 = \text{const} > 0. \tag{15}$$

2 Многомерное уравнение

Используя идеи, изложенные в работах [2, 3], нетрудно построить аналогичные алгоритмы и для p-мерного параболического уравнения

$$\frac{\partial u}{\partial t} = \sum_{\alpha=1}^{p} \frac{\partial^2 u}{\partial x_\alpha^2} + f(x,t), \quad x = (x_1, x_2, \ldots, x_p) \in G, \ 0 < t \leq T, \tag{16}$$

$$u(x,t)=0,\ x\in\Gamma;\quad u(x,0)=u_0(x),\ x\in\bar G. \tag{17}$$

где $\bar G=\{0\le x_\alpha\le l_\alpha,\ \alpha=1,2,...,p\}$. В $\bar G$ обычным образом вводится неравномерная сетка

$$\hat\omega_h=\{\,x=x_{i_1i_2}=(x_1^{i_1},x_2^{i_2},\dots,x_p^{i_p});\ x_\alpha^{i_\alpha}=x_\alpha^{i_\alpha-1}+h_\alpha^{i_\alpha},\\ i_\alpha=1,2,...,N_\alpha,\ x_\alpha^0=0,\ x_\alpha^{N_\alpha}=l_\alpha;\ \alpha=1,2\}.$$

Соответствующая разностная схема второго порядка аппроксимации на неравномерной сетке $\omega=\hat\omega_{i_1}\times\omega_\tau$ имеет вид

$$y_t+\sum_{k=1}^{p}(h_k^+y_{t\bar x_k}+h_k^-y_{tx_k})=\sum_{\alpha=1}^{p}\sum_{\substack{k=1\\k\ne\alpha}}^{p}(y+h_k^+y_{x_k}+h_k^-y_{\bar x_k})^{(\sigma)}_{\bar x_\alpha\hat x_\alpha}+f(\bar x,t), \tag{18}$$

$$y(x,\hat t)=0,\ x\in\gamma_h,\ t\in\omega_\tau;\quad y(x,0)=u_0(x),\ x\in\hat\omega_h. \tag{19}$$

Здесь $h_k^\pm=0,5(\bar h_k\pm|\bar h_k|)$, $\bar h_k=\dfrac{h_{k+}-h_k}{3}$, $v^{(\sigma)}=\sigma\hat v+(1-\sigma)v$, $\bar x=(\bar x_1,\bar x_2,...,\bar x_p)$, $\bar x_\alpha=x_\alpha+\bar h_\alpha$.

В соответствии с работами [1 — 3] нетрудно показать, что погрешность аппроксимации

$$\psi(\bar x,t)=O(|h|^2+\tau^{m_\sigma}),$$

где

$$|h|^2=h_1^2+h_2^2+..+h_p^2,\ h_\alpha=\max_{1\le i_\alpha\le N_\alpha}h_\alpha^{i_\alpha},\ m_\sigma=\begin{cases}2, & \sigma=0,5,\\ 1, & \sigma\ne0,5.\end{cases}$$

Интересно отметить, что в случае равномерной сетки ω_h разностное уравнение (18) вырождается в обычную схему "крест" и в этом случае $\bar x=x\in\omega_h$ — является расчетной точкой. В настоящей работе ограничимся исследованием лишь двумерного уравнения при $\sigma=0$ и $h_{k+}\ge h_k$, $k=1,2$. Тогда из (18) будем иметь

$$y_t+\bar h_1y_{t\bar x_1}+\bar h_2y_{t\bar x_2}=(y+\bar h_2y_{x_1})_{\bar x_2\hat x_2}+(y+\bar h_1y_{x_2})_{\bar x_1\hat x_1}+f(\bar x,t). \tag{20}$$

Для простоты рассуждений также предположим, что сетка по направлению x_2 равномерная ($h_{2+}=h_2$). Тогда имеет место

Теорема 2. *Пусть выполнены следующие условия:*

$$\frac{h_{1+}^2-h_1^2}{6}\le\tau\le\frac{h^2}{2c^2},\quad c^{-1}\le\frac{h}{h_+}\le c,\quad h_2^2\ge\frac{h_{1+}^2-h_1^2}{3}$$

для всех $i_1=1,2,...,N_1-1$. *Тогда разностная схема (19), (20) монотонна и имеет место оценка*

$$\max_{t\in\omega_\tau}\|y(t+\tau)\|_C\le\|y(0)\|_C+\sum_{t'=\tau}^{t}\tau\|\bar f(t')\|_C. \tag{21}$$

Доказательство. Уравнение (20) преобразуем к каноническому виду

$$\begin{aligned}(1+\bar h_1)\hat y=a_1y+a_2y_{+1}+a_3y_{-1}+a_4(y_{+2}+y_{-2})+\\ a_5(y_{+1,+2}+y_{1+,2-})+\bar h_1\hat y_{-1}+\tau f(\bar x,t),\quad t\in\omega_\tau,\end{aligned} \tag{22}$$

где

$$a_1 = 1 + h_1 - 2\tau\Big(\frac{1}{h_1 h_{1+}} - \frac{\tilde{h}_{1+}}{h_2^2}\Big),\ a_2 = \frac{\tau}{h_{1+}}\Big(\frac{1}{h_1} - \frac{2\tilde{h}_{1+}}{3h_2^2}\Big),\ h = \max\{h_1, h_2\},$$

$$a_3 = \frac{\tau}{h_1 h_1} - \tilde{h}_1,\ a_4 = \frac{\tau(1-\tilde{h}_{1+})}{h_2^2},\ a_5 = \frac{\tau\tilde{h}_{1+}}{h_2^2},\ \tilde{h}_1 = \frac{h_{1+}-h}{3h_1},\ \tilde{h}_{1+} = \frac{h_{1+}-h_1}{3h_{1+}},$$

$$y_{\pm k} = y_{i_k \pm 1},\ y_{+1,+2} = y_{i_1+1, i_2+1},\ y_{1+,2-} = y_{i_1+1, i_2-1}.$$

При выполнении условий теоремы коэффициенты уравнения (22) a — неотрицательны. Беря по модулю обе части уравнения (22), приходим к неравенству

$$(1 + \tilde{h}_{1i_1})|\hat{y}_{i_1 i_2}| \le \|y\|_C + h_{1i_1}\|\hat{y}\|_C + \tau\|\bar{f}(\tau)\|_C.$$

В силу произвольности $i_1 i_2$ из последней оценки следует (21). Теорема доказана.

Замечание 6. Пусть $p = 2$ и для уравнения (16) заданы неоднородные гравичные условия $u(x,t) = \mu(x,t)$, $x \in \gamma$. Подставляя теперь в (20) $y = u + z$ и полагая $h_{2+} = h_2$ (равномерная сетка по направлению x_2) получим задачу

$$z_t + \frac{h_{1+} - h_1}{3} z_{t\bar{x}_1} = z_{\bar{x}_2 \hat{x}_2} + \Big(z + \frac{h_{1+} - h_1}{3} y_{x_2}\Big)_{\bar{x}_2 \hat{x}_2} + \psi. \tag{23}$$

Согласно [2, 3] $\|\psi\|_C \le (h^2 + \tau)$. Следовательно, при выполнении условий теоремы 2

$$\max_{t \in \omega_\tau} \|z(t)\| \le c_2(h^2 + \tau), \qquad c_2 = \mathrm{const} > 0.$$

References

[1] А.А.Самарский, П.Н.Вабищевич, П.П.Матус. Разностные схемы повышенного порядка точности на неравномерных сетках // Дифференц. уравнения. 1996. Т. 32, № 2. С. 313 — 322.

[2] А.А.Самарский, П.Н.Вабищевич, П.П.Матус. Разностные схемы повышенного порядка аппроксимации на неравномерных сетках для эллиптических уравнений // Докл. АН Беларуси. 1996. Т. 40, № 5. С. 9 — 14.

[3] А.А.Самарский, П.Н.Вабищевич, П.П.Матус. Разностные схемы второго порядка точности на неравномерных сетках // ЖВМ и МФ. 1997. Т. 37, № 11.

[4] С.Э. Ананич. Монотонные консервативные разностные схемы второго порядка аппроксимации на неравномерных сетках для уравнения Фоккера — Планка. Мн., 1996. — 14 с. (Препринт / АН Беларуси. Ин-т математики; № 9(512)).

ГИДРОДИНАМИЧЕСКИЙ ПОДХОД К МОДЕЛИРОВАНИЮ ДИНАМИКИ МОРСКИХ ВОЛН, СОЗДАВАЕМЫХ УРАГАНОМ [1]

В.Н. Масленникова, М.Е. Боговский

Российский Университет дружбы народов
Кафедра дифференциальных уравнений и функционального анализа
117198 Москва, ул. Миклухо-Маклая, д. 6
phone: +7 095 955 0936
fax: +7 095 952 0745
e-mail: vmaslennikova@mx.pfu.edu.ru

В математическом моделировании ураганов, тайфунов и создаваемых ими волн до недавнего времени по вполне понятным причинам предпочитали избегать непосредственного использования системы Навье–Стокса, лежащей, как известно, в основе всей гидродинамики. Этот пробел отчасти восполняется работами [1–5], в которых сопутствующее ураганам эмпирически выявленное наличие разрывов [6–8] моделируется зависящим от времени разрывом в типе граничных условий. При этом непосредственное использование системы Навье–Стокса повышает уровень математической модели, а соответствующее значительное усложнение алгоритма решения компенсируется значительным ростом вычислительных ресурсов.

В настоящей работе изучается построенная на основе системы Навье–Стокса математическая модель динамики волн, создаваемых ураганом на поверхности океана. Предполагается, что уравнения Навье–Стокса описывают одновременно динамику атмосферы и океана, а неизвестная поверхность раздела двух сред (воды и воздуха) представляет собой внутреннюю поверхность разрыва. Наша главная цель - определить форму этой поверхности.

Специфика ураганной силы ветра учитывается путем замены традиционных условий прилипания на поверхности раздела сред условиями Навье [9], которые состоят в условии непротекания и в условиях пропорциональности касательных к поверхности раздела сред напряжений разностям касательных скоростей, что соответствует эмпирически выявленному трению урагана о поверхность океана [6]. Рассматриваемая математическая модель сводится к начально-краевой задаче для нелинейной системы Навье–Стокса в цилиндре $Q_T = \Omega \times (0,T)$, где область $\Omega \subset \mathbb{R}^3$ также представляет собой цилиндр $\Omega = \omega \times (-h,h)$ с ограниченным основанием $\omega \subset \mathbb{R}^2$, $h > 0$. Предполагается, что поверхность раздела сред $\sigma(t)$ делит цилиндр Ω на две части: верхнюю $\Omega_+(t)$ и нижнюю $\Omega_-(t)$ так, что $\Omega = \Omega_+(t) \cup \Omega_-(t) \cup \sigma(t)$. При этом верхняя часть занята воздухом с плотностью ρ_+ и вязкостью μ_+, а нижняя - водой с плотностью ρ_- и вязкостью μ_-.

Для простоты ограничимся случаем, когда условия Навье заданы на всей поверхности раздела $\sigma(t)$. Разумеется, при этом теряется такая характерная особенность урагана, как разрыв в типе граничных условий, но достигаемое за счет этого упрощение позволяет сосредоточить все внимание на решении и без того сложной задачи

[1] Работа выполнена при поддержке Российского Фонда Фундаментальных Исследований, Грант No. 96-01-01130

о нахождении неизвестной поверхности раздела сред $\sigma(t)$.

Для поля скоростей $v = v(x,t) = (v_1, v_2, v_3)$ и давления $p = p(x,t)$ на зависящем от времени открытом подмножестве $\Omega_+(t) \cup \Omega_-(t)$ рассматривается система Навье–Стокса

$$\begin{cases} \rho\left[\dfrac{\partial v}{\partial t} + (v, \nabla)v\right] - \mu\Delta v + \nabla p = f(x,t), \\ \operatorname{div} v = 0, \quad x \in \Omega_+(t) \cup \Omega_-(t), \ t \in (0,T). \end{cases} \tag{1}$$

где плотность $\rho = \rho(x,t)$ и вязкость $\mu = \mu(x,t)$ имеют вид:

$$\rho(x,t) = \begin{cases} \rho_+, \ x \in \Omega_+(t), \\ \rho_-, \ x \in \Omega_-(t), \end{cases} \qquad \mu(x,t) = \begin{cases} \mu_+, \ x \in \Omega_+(t), \\ \mu_-, \ x \in \Omega_-(t), \end{cases}$$

при $t \in (0,T)$ с заданными положительными постоянными ρ_+, ρ_-, μ_+, μ_-. Для системы (1) решается начально-краевая задача с начальными условиями

$$v|_{t=0} = 0, \quad x \in \Omega_+(0) \cup \Omega_-(0), \tag{2}$$

и с граничными условиями на поверхности цилиндра Ω:

$$v|_{\partial\Omega} = 0, \quad t \in (0,T), \tag{3}$$

которые берутся здесь однородными только для того, чтобы не отвлекаться на не относящиеся к существу дела технические подробности. На поверхности раздела сред $\sigma(t)$ задаются шесть скалярных условий. Первые три условия состоят в равенстве напряжений по обе стороны от $\sigma(t)$:

$$\mathcal{T}\nu\Big|_{\sigma_+(t)} = \mathcal{T}\nu\Big|_{\sigma_-(t)}, \quad t \in (0,T), \tag{4}$$

где $\mathcal{T}$ – тензор напряжений в несжимаемой вязкой сплошной среде, ν – нормаль к $\sigma(t)$, а $\sigma_+(t)$ и $\sigma_-(t)$ – стороны поверхности раздела $\sigma(t)$, прилегающие к $\Omega_+(t)$ и $\Omega_-(t)$, соответственно.

Оставшиеся три условия на $\sigma(t)$ являются условиями Навье, состоящими в условии непротекания

$$(v,\nu)|_{\sigma_+(t)} = (v,\nu)|_{\sigma_-(t)}, \quad t \in (0,T), \tag{5}$$

через движущуюся поверхность раздела сред $\sigma(t)$ и в трении

$$(\mathcal{T}\nu,\tau_j)|_{\sigma(t)} = E\left[(v,\tau_j)|_{\sigma_+(t)} - (v,\tau_j)|_{\sigma_-(t)}\right], \quad t \in (0,T), j = 1,2, \tag{6}$$

с коэффициентом $E > 0$, где τ_1, τ_2 – касательные к $\sigma(t)$ направления, $(\tau_1,\tau_2) = 0$, а круглыми скобками $(\cdot,\cdot)$ обозначается скалярное произведение в $\mathbb{R}^3$. Из физических соображений нормаль ν в левой части (6) при условии $E > 0$ должна быть внешней по отношению к $\Omega_-(t)$.

Поверхность $\sigma(t)$ задается уравнением

$$x_3 = \psi(x',t), \quad x' = (x_1,x_2) \in \omega, \ t \in (0,T). \tag{7}$$

Обозначив $\phi(x,t) = x_3 - \psi(x',t)$, можно задать $\sigma(t)$ неявно уравнением

$$\phi(x,t) = 0, \quad t \in (0,T) \tag{8}$$

В рассматриваемую математическую модель входит также и дифференциальное уравнение свободной поверхности $\sigma(t)$:

$$\phi_t + (v, \nabla_x \phi) = 0, \quad x \in \sigma(t), \quad t \in (0,T), \tag{9}$$

которое для $\psi(x',t)$ принимает вид

$$\psi_t + \sum_{j=1}^{2} v_j(x',\psi,t)\frac{\partial \psi}{\partial x_j} = v_3(x',\psi,t), \quad x' \in \omega, \quad t \in (0,T), \tag{10}$$

с начальным условием

$$\psi|_{t=0} = \psi_0(x'), \quad x' \in \omega. \tag{11}$$

Рассматриваемая математическая модель сводится, таким образом, к решению начально-краевой задачи (1)–(6), (9)–(10) с неизвестной поверхностью раздела сред $\sigma(t)$. Решение этой задачи аппроксимируется при $\varepsilon \to 0$ некоторой последовательностью решений вспомогательной задачи с малым параметром $\varepsilon > 0$.

Вспомогательную задачу с параметром ε удобнее сформулировать сразу в обобщенной постановке. Для этого вводятся следующие обозначения. Пусть

$$Q_T^+ = \{(x,t) \in Q_T : x \in \Omega_+(t), t \in (0,T)\}, \quad Q_T^- = \{(x,t) \in Q_T : x \in \Omega_-(t), t \in (0,T)\}.$$

Через S_T обозначим поверхность

$$S_T = \{(x,t) \in Q_T : x \in \sigma(t), \quad t \in (0,T)\}.$$

И считая l достаточно большим наперед заданным натуральным числом, введем пространство

$$H^l(Q_T) = \Big\{ v(x,t) : v|_{Q_T^+} \in W_2^l(Q_T^+;\mathbb{R}^3),\ v|_{Q_T^-} \in W_2^l(Q_T^-;\mathbb{R}^3);\ v|_{t=0} = 0;\ v|_{\partial\Omega} = 0,$$
$$(v,\nu)|_{\sigma_+(t)} = (v,\nu)|_{\sigma_-(t)},\ t \in (0,T);\ \mathrm{div}_x v(x,t) = 0,\ (x,t) \in Q_T \Big\},$$

на котором при условии липшицевости S_T определены линейные непрерывные операторы продолжения

$$L^{(+)}:\ W_2^l(Q_T^+;\mathbb{R}^3) \to W_2^l(Q_T;\mathbb{R}^3),$$
$$L^{(-)}:\ W_2^l(Q_T^-;\mathbb{R}^3) \to W_2^l(Q_T;\mathbb{R}^3),$$

с подмножеств Q_T^+ и Q_T^- на весь цилиндр Q_T с сохранением класса. При этом, очевидно, $L^{(+)}u|_{Q_T^+} = u$ и $L^{(-)}u|_{Q_T^-} = u$. В сохранении же соленоидальности продолжения никакой необходимости нет.

Используя скалярное произведение $[u,v]$ гильбертова пространства $W_2^l(Q_T;\mathbb{R}^3)$, зададим на $H^l(Q_T)$ билинейную форму

$$B(u,v) = [L^{(+)}u, L^{(+)}v] + [L^{(-)}u, L^{(-)}v].$$

Символом $\mathcal{T}^j$ обозначим j-ю строку тензора напряжений $\mathcal{T}$. Очевидно, $\mathcal{T}^j$ – вектор-функция со значениями в $\mathbb{R}^3$. Обобщенное решение $v^\varepsilon(x,t)$ вспомогательной задачи

с параметром $\varepsilon > 0$ будет определяется с помощью интегрального тождества

$$\int\limits_{Q_T}(v_t^\varepsilon,u)\rho\,dxdt+\int\limits_{Q_T}((v^\varepsilon,\nabla)v^\varepsilon,u)\rho\,dxdt+\sum_{j=1}^{3}\int\limits_{Q_T}(T^j,\nabla u_j)\,dxdt+$$
$$+\sum_{j=1}^{2}\int\limits_{S_T}E[(v^\varepsilon,\tau_j)|_{\sigma_+(t)}-(v^\varepsilon,\tau_j)|_{\sigma_-(t)}][(u,\tau_j)|_{\sigma_+(t)}-(u,\tau_j)|_{\sigma_-(t)}]\,dsdt+ \qquad (12)$$
$$+\varepsilon B(v^\varepsilon,u)=\int\limits_{Q_T}(f,u)\,dxdt\quad\forall\,u\in H^1(Q_T)$$

Теперь наконец, можно сформулировать определение обобщенного решения вспомогательной задачи (10)–(12) с неизвестной поверхностью $\sigma(t)$.

Определение. *Вектор-функцию $v^\varepsilon(x,t)\in H^1(Q_T)$ и функцию $\psi(x',t)$, удовлетворяющую в $\omega\times(0,T)$ условию Липшица по всем переменным, будем называть обобщенным решением задачи (10)–(12), если $v^\varepsilon(x,t)$ удовлетворяет тождеству (12) с функцией $\psi(x',t)$, удовлетворяющей (10)–(11). При этом $\psi(x',t)$ тоже зависит от параметра ε.*

Задача с неизвестной границей (10)–(12) решается итерационным методом. Пусть $m\geq 1$ – номер итерации, и пусть уже известна m–я итерация $\sigma^m(t)$ для поверхности $\sigma(t)$. Решение $v^{\varepsilon,m}(x,t)$ интегрального тождества (12) будем искать методом Галеркина в классе $H^1(Q_T^m)$, соответствующем поверхности $\sigma^m(t)$.

Выберем в $H^1(Q_T^m)$ базисные вектор–функции $\{w^k(x,t)\}_{k=1}^\infty$. Галеркинские приближения $v^{\varepsilon,m,N}(x,t)$ определим равенством

$$v^{\varepsilon,m,N}(x,t)=\sum_{k=1}^{N}C_k^N w^k(x,t)$$

с числовыми коэффициентами C_k^N. Для определения коэффициентов C_k^N имеем нелинейную систему N алгебраических уравнений. Эта система получается из тождества (12) после подстановки $v^{\varepsilon,m,N}(x,t)$ вместо $v^\varepsilon(x,t)$ и $w^k(x,t)$ – вместо $u(x,t)$ для $k=1,\ldots,N$. Разрешимость подобной алгебраической системы обычно гарантируется априорной ограниченностью галеркинских приближений, которая получается умножением каждого из N уравнений на соответствующий коэффициент C_k^N с последующим суммированием по всем k. В данном случае доказательство априорной ограниченности наталкивается на препятствие, заключающееся в несовпадении на $\sigma^m(t)$ нормальной проекции скорости сплошной среды $(v^{\varepsilon,m,N},\nu)$ с нормальной проекцией скорости движения самой поверхности $\sigma^m(t)$. Это связано с формулой дифференцирования

$$\frac{\partial}{\partial t}\int\limits_\Omega|v^{\varepsilon,m,N}(x,t)|^2\,dx=\frac{\partial}{\partial t}\int\limits_{\Omega_+^m(t)}|v^{\varepsilon,m,N}(x,t)|^2\,dx+$$
$$+\frac{\partial}{\partial t}\int\limits_{\Omega_-^m(t)}|v^{\varepsilon,m,N}(x,t)|^2\,dx=2\int\limits_\Omega\left(v_t^{\varepsilon,m,N},v^{\varepsilon,m,N}\right)dx+$$
$$+2\int\limits_\Omega\left((v^{\varepsilon,m,N},\nabla)v^{\varepsilon,m,N},v^{\varepsilon,m,N}\right)dx+ \qquad (13)$$
$$+\sum_{j=1}^{2}\int\limits_{\sigma^m(t)}[(v,\tau_j)^2|_{\sigma_+^m(t)}-(v,\tau_j)^2|_{\sigma_-^m(t)}][(v^{\varepsilon,m,N},\nu)-(U^m,\nu)]\,ds,$$

где ν – нормаль, внешняя к $\Omega_-^m(t)$, а (U^m,ν) – нормальная проекция скорости движения поверхности $\sigma^m(t)$.

Это препятствие легко преодолеть, добавив в коэффициент трения E слагаемое, которое обратится в нуль при $m\to\infty$. Так, если нам требуется решить задачу (10)–(12) с постоянным коэффициентом трения $E=E_0>0$, то в определении обобщенного решения $v^{\varepsilon,m,N}(x,t)$ на каждой итерации следует исходить из тождества (12) с коэффициентом трения E, имеющим, например, вид

$$E=E_0+M|(v^{\varepsilon,m,N},\nu)-(U^m,\nu)|^2\sum_{j=1}^{2}(v^{\varepsilon,m,N},\tau_j)^2 \tag{14}$$

с некоторой наперед заданной достаточно большой постоянной $M>0$, величина которой, в конечном счете, окажется несущественной, так как добавка к E_0 в (14) обратится в нуль при $m\to\infty$. Подчеркнем однако, что M не зависит от ε, m и N.

С коэффициентом трения E вида (14) легко устанавливается оценка вида

$$\begin{aligned}&\sum_{i,j=1}^{3}\int\limits_{Q_T}\left(\frac{\partial v_i^{\varepsilon,m,N}}{\partial x_j}+\frac{\partial v_j^{\varepsilon,m,N}}{\partial x_i}\right)^2dxdt+\\&+\sum_{j=1}^{2}\int\limits_{S_T^m}[(v^{\varepsilon,m,N},\tau_j)|_{\sigma+(t)}-(v^{\varepsilon,m,N},\tau_j)|_{\sigma-(t)}]^2dsdt+\\&+\varepsilon\parallel L_m^{(+)}v^{\varepsilon,m,N}\parallel^2_{W_2^1(Q_T;\mathbb{R}^3)}+\varepsilon\parallel L_m^{(-)}v^{\varepsilon,m,N}\parallel^2_{W_2^1(Q_T;\mathbb{R}^3)}\le C\parallel f\parallel^2_{L_2(Q_T;\mathbb{R}^3)}\end{aligned} \tag{15}$$

с постоянной $C>0$, не зависящей от ε, m и N, где $L_m^{(+)}$ и $L_m^{(-)}$ -операторы продолжения через поверхность S_T^m, соответствующую $\sigma^m(t)$. Из (15), в частности, следует

$$\varepsilon\parallel L_m^{(+)}v^{\varepsilon,m,N}\parallel^2_{W_2^1(Q_T;\mathbb{R}^3)}+\varepsilon\parallel L_m^{(-)}v^{\varepsilon,m,N}\parallel^2_{W_2^1(Q_T;\mathbb{R}^3)}\le C\parallel f\parallel^2_{L_2(Q_T;\mathbb{R}^3)} \tag{16}$$

А оценки (16) достаточно для разрешимости нелинейной алгебраической системы, определяющей коэффициенты галеркинских приближений C_k^N. Выделяя из $\{v^{\varepsilon,m,N}(x,t)\}_{N=1}^{\infty}$ подпоследовательность, для которой $L_m^{(+)}v^{\varepsilon,m,N_n}$ и $L_m^{(-)}v^{\varepsilon,m,N_n}$ сходятся слабо в $W_2^1(Q_T;\mathbb{R}^3)$, получим m-ю итерацию $v^{\varepsilon,m}(x,t)$ обобщенного решения интегрального тождества (12)в классе $H^1(Q_T^m)$, так как $L_m^{(+)}v^{\varepsilon,m,N}=v^{\varepsilon,m,N}$ при $x_3>\psi^m(x',t)$ и $L_m^{(-)}v^{\varepsilon,m,N}=v^{\varepsilon,m,N}$ при $x_3<\psi^m(x',t)$. При этом в силу (14) тождество (12) для $v^{\varepsilon,m}(x,t)$ будет выполняться с коэффициентом трения

$$E=E_0+M|(v^{\varepsilon,m},\nu)-(U^m,\nu)|^2\sum_{j=1}^{2}(v^{\varepsilon,m},\tau_j)^2, \tag{17}$$

а в силу (15) имеет место оценка

$$\begin{aligned}&\sum_{i,j=1}^{3}\int\limits_{Q_T}\left(\frac{\partial v_i^{\varepsilon,m}}{\partial x_j}+\frac{\partial v_j^{\varepsilon,m}}{\partial x_i}\right)^2dxdt+\\&+\sum_{j=1}^{2}\int\limits_{S_T^m}[(v^{\varepsilon,m},\tau_j)|_{\sigma+(t)}-(v^{\varepsilon,m},\tau_j)|_{\sigma-(t)}]^2dsdt+\\&+\varepsilon\parallel L_n^{+)}v^{\varepsilon,m}\parallel^2_{W_2^1(Q_T;\mathbb{R}^3)}+\varepsilon\parallel L_m^{(-)}v^{\varepsilon,m}\parallel^2_{W_2^1(Q_T;\mathbb{R}^3)}\le C\parallel f\parallel^2_{L_2(Q_T;\mathbb{R}^3)}\end{aligned} \tag{18}$$

с постоянной C, не зависящей от ε и m. Из (18) следует, в частности,

$$\varepsilon \parallel L_m^{(+)} v^{\varepsilon,m} \parallel^2_{W_2^l(Q_T;\mathbb{R}^3)} + \varepsilon \parallel L_m^{(-)} v^{\varepsilon,m} \parallel^2_{W_2^l(Q_T;\mathbb{R}^3)} \le C \parallel f \parallel^2_{L_2(Q_T;\mathbb{R}^3)} \qquad (19)$$

Поскольку $v^{\varepsilon,m}(x,t) \in H^l(Q_T^m)$ с достаточно большим l и $v^{\varepsilon,m}|_{S\Omega} = 0$, то задача

$$\begin{cases} \psi_t^{m+1} + \sum\limits_{j=1}^{2} v_j^{\varepsilon,m}(x', \psi^{m+1}, t) \dfrac{\partial \psi^{m+1}}{\partial x_j} = v_3^{\varepsilon,m}(x', \psi^{m+1}, t), \quad x' \in \omega, \quad t \in (0,T), \\ \psi^{m+1}|_{t=0} = \psi_0(x'), \quad x' \in \omega. \end{cases}$$

однозначно разрешима, причем гладкость решения $\psi^{m+1}(x,t)$ определяется величиной заранее выбранного порядка l и гладкостью начальных данных $\psi_0(x')$. Уравнение $x_3 = \psi^{m+1}(x',t)$ будет теперь задавать $(m+1)$-ю итерацию $\sigma^{m+1}(t)$ для поверхности $\sigma(t)$. В качестве первой итерации $\sigma^1(t)$ можно взять, например, поверхность $x_3 = \psi_0(x')$.

Итак, построена последовательность решений $\{v^{\varepsilon,m}(x,t)\}_{m=1}^{\infty}$ и $\{\psi^{\varepsilon,m}(x',t)\}_{m=1}^{\infty}$. Выделяя из нее подпоследовательность, для которой $L_m^{(+)} v^{\varepsilon,m_n}$ и $L_m^{(-)} v^{\varepsilon,m_n}$ сходятся слабо в $W_2^l(Q_T;\mathbb{R}^3)$, получим обобщенное решение задачи (10)-(12) в классе $H^l(Q_T)$, так как $L_m^{(+)} v^{\varepsilon,m} = v^{\varepsilon,m}$ при $x_3 > \psi^m(x',t)$ и $L_m^{(-)} v^{\varepsilon,m} = v^{\varepsilon,m}$ при $x_3 < \psi^m(x',t)$. Отметим, что в доказательстве разрешимости задачи (10)-(12) наряду с оценкой (19) существенную роль играет компактное вложение $W_2^l(Q_T) \to C^{l-3}(Q_T)$, которое применяется к подпоследовательностям $L_m^{(+)} v^{\varepsilon,m_n}$ и $L_m^{(-)} v^{\varepsilon,m_n}$. Благодаря этому вложению удается оценить производные по x,t от $v^{\varepsilon,m}$ вплоть до $\sigma_+(t)$ и $\sigma_-(t)$ постоянными, не зависящими от m. Порядок оцениваемых производных зависит от величины заранее выбранного числа l и от гладкости начальных данных $\psi_0(x')$. Очень важно, что уравнение (10) представляет собой равенство на $\sigma(t)$ нормальной проекции (v^ε, ν) и нормальной проекции (U,ν) скорости движения самой поверхности $\sigma(t)$. В то же время

$$\lim_{m_n \to \infty} (v^{\varepsilon,m_n}, \nu) = (v^\varepsilon, \nu), \quad \lim_{m_n \to \infty} (U^{m_n}(x,t), \nu) = (U, \nu)$$

равномерно в $\overline{Q}_T$. В силу (17) это означает, что тождество (12) выполняется для $v^\varepsilon(x,t)$ с коэффициентом трения $E = E_0$. Таким образом, справедлива

Теорема. *Пусть $\rho_+ > 0$, $\rho_- > 0$, $\mu_+ > 0$, $\mu_+ > 0$, $T > 0$, $\varepsilon > 0$. И пусть $l \ge 3$ целое. Для всех $f(x,t) \in L_2(Q_T;\mathbb{R}^3)$ и $\psi_0(x') \in C^{l-2}(\overline{\omega})$ существует обобщенное решение задачи (10)-(12) в смысле данного выше определения.*

Для функции $\Phi_t^\varepsilon(x,t)$ рассмотрим задачу Коши

$$\begin{cases} \Phi_t^\varepsilon + (v^\varepsilon(x,t), \nabla_x \Phi^\varepsilon) = 0, \quad x \in \Omega, \quad t \in (0,T) \\ \Phi_t^\varepsilon|_{t=0} = x_3 - \psi_0(x'), \quad x \in \Omega. \end{cases} \qquad (20)$$

Здесь, в отличие от (9), уравнение выполняется в Q_T, а не только на S_T. Ввиду гладкости $v^\varepsilon(x,t)$ задача Коши (20) для каждого $\varepsilon > 0$ имеет единственное гладкое решение $\Phi^\varepsilon(x,t)$. Нетрудно убедиться, что поверхность

$$\Phi^\varepsilon(x,t) = 0 \qquad (21)$$

совпадает с поверхностью $x_3 = \psi(x', t)$, где функция $\psi(x', t)$ определяется из задачи (10)–(11) с v^ε вместо v.

Поскольку на зависящей от ε поверхности $x_3 = \psi(x', t)$ выполняется условие непротекания, то обобщенная (в смысле Соболева) дивергенция $\operatorname{div}_x v^\varepsilon(x, t) = 0$ в Q_T. Отсюда и из принципа максимума [10] (см. также [11]) следует, что

$$m_0 \le \min_{(x,t)\in\overline{Q}_T} \Phi^\varepsilon(x, t) \le \max_{(x,t)\in\overline{Q}_T} \Phi^\varepsilon(x, t) \le M_0 \tag{22}$$

для всех $\varepsilon > 0$, где числа

$$m_0 = \min_{x\in\Omega}(x_3 - \psi_0(x')), \quad M_0 = \max_{x\in\Omega}(x_3 - \psi_0(x'))$$

не зависят от ε.

Из оценки (22) следует ограниченность последовательности $\Phi^\varepsilon(x, t)$ в $L_p(Q_T)$ для любого $p \in (1, \infty)$ при $\varepsilon \to 0$. Выделяя подпоследовательность $\Phi^{\varepsilon_n}(x, t)$, слабо сходящуюся в $L_p(Q_T)$, получим (см. [10]) некоторую существенно ограниченную в Q_T функцию $\Phi(x, t)$. Искомая свободная поверхность раздела сред задается тогда уравнением

$$\Phi(x, t) = 0. \tag{23}$$

Отметим, что задача о нахождении свободной поверхности (23) решена здесь "в целом", т.е. без каких-либо ограничений на величину T и нормы функций $f(x, t), \psi_0(x')$. Если найденная при этом $\Phi(x, t)$ из $L_\infty(Q_T)$ не имеет соответствующего следа, то свободная поверхность раздела сред может стать негладкой или вообще разрушиться. Тем не менее, построенная выше аппроксимация с параметром ε в любом случае позволяет найти функцию $\Phi(x, t)$, задающую поверхность (23). Что же касается обобщенного решения $v(x, t)$ для (1), то его существование обусловлено гладкостью поверхности (23) и характером сходимости поверхности (21) к поверхности (23) при $\varepsilon = \varepsilon_n \to 0$. Вопрос о существовании обобщенного решения $v(x, t)$ для (1) в случае, когда на свободной поверхности раздела сред задаются условия Навье, значительно сложнее, чем в случае, когда заданы традиционные условия прилипания. Физически это вполне очевидно, так как условия прилипания соответствуют условиям Навье с коэффициентами трения $E_0 = +\infty$.

ЛИТЕРАТУРА

[1] В.Н. Масленникова. Начально-краевая задача для системы Навье–Стокса с теплопереносом с граничными условиями, меняющими тип на многообразиях, зависящих от времени. // Вестник РУДН, сер. Математика, **3**, вып. 1, 92–102 (1996).

[2] В.Н. Масленникова, М.Е. Боговский. Динамика урагана над поверхностью суши. // Вестник РУДН, сер. Математика. **1**, 37–47 (1994).

[3] В.Н. Масленникова, М.Е. Боговский. Некоторые математические модели динамики урагана. // Математическое моделирование. **6**, No 9, 26–32 (1994).

[4] В.Н. Масленникова, М.Е. Боговский, Ж.А. Уэрдиан. Об одной краевой задаче с разрывными краевыми условиями. // ДАН СССР. **318**, No 5, 1037-1040 (1991).

[5] V.N. Maslennikova, M.E. Bogovskiĭ. On a new boundary value problem for the Navier–Stokes system with discontinuity in the type of boundary conditions depending in time. // Acta Applicandae Mathematicae (Kluwer Acad. Publ., Dordrecht). **37**, 99–107 (1994).
[6] В.В. Шулейкин. Расчет развития, движения и затухания тропических ураганов и главных волн, создаваемых ураганами. Л., Гидрометеоиздат, (1978).
[7] G.E. Dunn, B.I. Miller. Atlantic hurricanes, Luisiana St. Univ. Press (1960).
[8] I.R. Tannehill. Hurricanes, Princeton Univ. Press (1956).
[9] L.M.H. Navier. Mémoire sur les lois du mouvement des fluides. Mé noires de l'Acad. Royale des Sci. de l'Inst. de France, 1823. Paris. **6**, 389–440 (1827).
[10] В.Н. Масленникова, М.Е. Боговский. Принцип максимума для уравнения неразрывности сжимаемой среды. // ДАН. **339**, No 4, 446–450 (1994)
[11] R.J. Di Perna, P.L. Lions. Ordinary differential equations, transport theory and Sobolev spaces. // Invent. Math. **98**, No 3, 511–547 (1989).

Maple-пакет для построения решений линейных обыкновенных дифференциальных уравнений в виде степенных рядов [1]

А. А. Рябенко

МГУ ВМиК,
Москва 1199899, Россия.
e-mail:ryabenko@cs.msu.su

1 Введение

В этой статье описывается пакет процедур для системы компьютерной алгебры Maple, предлагающий средства для решения следующей задачи. Дано однородное линейное обыкновенное дифференциальное уравнение с полиномиальными коэффициентами. Для него строится, когда это возможно, решение в виде формального степенного ряда $\sum_{n=0}^{\infty} c(n)(x-a)^n$ в такой точке a, что его коэффициенты $c(n)$ являются "хорошими" функциями от n. Это может быть полином, рациональная функция от n или гипергеометрическая последовательность (т.е. $c(n+1)/c(n)$ – рациональная функция). Последний случай, очевидно, включает в себя первые два.

Для решения поставленной задачи выполняются следующие действия. Определяем подходящие точки a, так как это описано в статье [1]. После этого находим неотрицательное целое M и строим линейную рекурренцию, с помощью которой можно вычислить $c(n)$ для всех $n > M$, определяем начальные элементы $c(0), ..., c(M)$. Начальные элементы задают линейное пространство всех решений данного уравнения в виде формальных рядов. Поэтому значения $c(0), ..., c(M)$ содержат произвольные постоянные вида $_B_0, _B_1, ...$ Далее, используя стандартный Maple пакет LREtools, находим решение построенной рекурренции. Оно будет содержать произвольные постоянные вида $_C_0, _C_1, ...$ Вычисляем значения произвольных постоянных, при которых полученное решение рекурренции согласуется с начальными значениями последовательности коэффициентов.

В нижеследующих разделах продемонстрируем работу основных процедур пакета.

2 Полиномиальные коэффициенты

Как было указано выше, полином является частным случаем гипергеоментрической последовательности. Но построение полиномиального решения для рекурренции осуществляется на много быстрее. Кроме того, в статье [1] указано в каких именно точках можно ожидать, что решение дифференциального уравнения в виде степенного ряда будет иметь полиномиальные коэффициенты. Это такие точки a, что $a+1$ – особая точка данного дифференциального уравнения. В следующем примере особые точки – $\{0, -1, 2\}$:

```
> L := (-2*x^4+2*x^3+4*x^2)*diff(y(x),x$2) +
```

[1]Эта работа поддерживается грантом INTAS-RFBR No. 95-412

```
> (-2*x^3-16*x^2-8*x) * diff(y(x),x) +
> (2*x^2 + 16*x + 8)*y(x):
```

Процедура 'Slode/polycoeffs' пытается построить решение с полиномиальными коэффициентами в точках $\{-1,-2,1\}$, но сделать это удается только в точке $a=1$:

```
>'Slode/polycoeffs'(L,y(x));
```

$$\{\frac{1}{2}_C_1\,(\sum_{n=0}^{\infty}(2\,n^2+2\,n+1)\,(x-1)^n)\}$$

3 Рациональные коэффициенты

Аналогично предыдущей процедуре работает процедура построения решений в виде рядов с рациональными коэффициентами: 'Slode/ratcoeffs'. Здесь в качестве значения точки a выбираются особые точки уравнения, в приведенном примере – это $\{0,1\}$:

```
> L1 := 2*x*(x-1)*diff(y(x),x$2) + (7*x-3)*diff(y(x),x) + 2*y(x):
> 'Slode/ratcoeffs'(L1,y(x));}{%
```

$$\left\{_C_1\,\left(\sum_{n=0}^{\infty}\frac{(n+1)\,x^n}{2\,n+1}\right)\right\}$$

Если мы исследуем особые точки уравнения, то $c(n)$ может выражаться рациональной функцией для всех $n \geq 0$. Но рассмотрение некоторых дополнительных точек позволяет построить решения, где коэффициенты ряда будут рациональными, начиная только с некоторого места ($n > N$):

```
> L2 := (3-x)*diff(y(x),x$2)-diff(y(x),x):
> 'Slode/ratcoeffs'(L2,y(x));
```

$$\left\{_B_0+_C_1\,\left(\sum_{n=1}^{\infty}\frac{(x-2)^n}{n}\right)\right\}$$

4 Гипергеометрические коэффициенты

В [1] показано, что линейное уравнение с полиномиальными коэффициентами может в особых точках иметь решения в виде степенного ряда, коэффициенты которого начиная с некоторого места образуют гипергеометрическую последовательность. Но если решение такого вида существует в какой-либо неособой точке, то это верно и для всех неособых. Поэтому процедура 'Slode/hypercoeffs' проверяет все особые и одну неособую (а именно такую, где может быть решение в виде ряда с полиномиальными коэффициентами). Например, следующее уравнение имеет одну особую точку 1. Но решение с "хорошими" коэффициентами обнаружилось в неособой точке $a=0$:

```
> L3 := (3*x^2-6*x+3)*diff(y(x),x$2) + (12*x-12)*diff(y(x),x) + 6*y(x):
> 'Slode/hypercoeffs'(L3,y(x),_CC);
```

$$\{\sum_{n=0}^{\infty} (_C_1 + _C_2\, n)\, x^n\}$$

С помощью следующей процедуры можно построить решение в любой точке:

```
> 'Slode/hyper_in_point'(L3,y(x),_CC,2);
```

$$-(\sum_{n=0}^{\infty} (-1)^n\, (-_C_1 + _C_1\, n + n\, _CC_1)\, (x-2)^n)$$

Эта процедура также может быть полезна в случае, если исследование всех возможных точек занимает очень много ресурсов:

```
> L4 := (z^3-2*z^2+z)*diff(y(z),z$3) +
> (7*z^2-15*z+8)*diff(y(z),z$2) +
> (10*z-13)*diff(y(z),z) - 2*y(z):
>'Slode/hyper_in_point'(L4,y(z),_CC,1);
```

$$_CC_4\, (z-1)^4 + _CC_4 \left(\sum_{n=5}^{\infty} \left(\prod_{k=4}^{n-1} (-\frac{k^3+4\,k^2+5\,k-2}{(k-3)\,(k+1)^2})\right)\, (z-1)^n\right)$$

На следующем примере видно, что процедура 'Slode/hypercoeffs' может построить полиномиальное решение дифференциального уравнения. Это означает, что начиная с некоторого места $c(n) = 0$:

```
> L5 := -z*(-z-1)*diff(w(z),z$2)+diff(w(z),z)-2*w(z):
> 'Slode/hypercoeffs'(L5,w(z));
```

$$\left\{_B_0 + (-3\,_C_2 - 2\,_B_0)\, (z+2) + (_B_0 + \frac{3}{4}\,_C_2)\, (z+2)^2 \right.$$

$$+ 6\,_C_2 \left(\sum_{n=1}^{\infty} \frac{\Gamma(n-2)\, (-2 + (\frac{1}{2})^n\, n^2 + (\frac{1}{2})^n\, n + 2\,(\frac{1}{2})^n)\, (z+2)^n}{\Gamma(n+1)}\right),\ _B_0 + 2\,_B_0\, z + _B_0\, z^2,$$

$$\left._B_2\, (z+1)^2\right\}$$

Список литературы

[1] S. A. Abramov, M. Petkovšek (1996): Special Power Series Solutions of Linear Differential Equations, *Proc. of FPSAC'96*, 1-10.

ОПЕРАТОРНО-РАЗНОСТНЫЕ СХЕМЫ ДЛЯ ЗАДАЧ МАТЕМАТИЧЕСКОЙ ФИЗИКИ [1]

А.А.Самарский, П.Н.Вабищевич

Институт математического моделирования РАН, Москва,
E-mail: samarskii@imamod.msk.su, vab@imamod.msk.su

Abstract. Обсуждаются общие проблемы приближенного решения нестационарных задач математической физики: теория устойчивости (корректности) операторно-разностных схем и аддитивные разностные схемы. Формулируются общие критерии устойчивости двухслойных операторно-разностных схем по начальным данным и правой части. Приведены новые априорные оценки устойчивости в интегральных по времени нормах. В настоящее время теория устойчивости применяется для исследования разностных схем для некорректных эволюционных задач, сформулированы критерии устойчивости проекционно-разностных схем. При решении многомерных задач, построении безитерационных алгоритмов для параллельных компьютеров необходимо ориентироваться на использование аддитивных разностных схем (схем расщепления). Приведены основные результаты теории аддитивных разностных схем суммарной аппроксимации, исследуются многокомпонентные аддитивные схемы, строятся регуляризованные аддитивные схемы полной аппроксимации.

1 Введение

После разностной или конечно-элементной дискретизации по пространству нестационарной задачи математической физики получаем задачу Коши для системы обыкновенных дифференциальных уравнений, которая рассматривается в сеточном гильбертовом пространстве. Дискретизация по времени дает операторно-разностную схему. Здесь мы отметим некоторые важнейшие направления развития теории операторно-разностных схем для нестационарных задач математической физики.

При исследовании разностных схем для нестационарных задач математической физики широко используется общая теория устойчивости (корректности) операторно-разностных схем [1,2]. В настоящее время получены точные (совпадающие необходимые и достаточные) условия широкого класса двух- и трехслойных разностных схем в конечномерных гильбертовых пространствах. Основные теоретические результаты теории устойчивости суммированы в книгах [3,4].

Необходимо особо подчеркнуть конструктивность общей теории устойчивости операторно-разностных схем, в которой критерии устойчивости формулируются в виде легко проверяемых неравенств для операторов. Среди наиболее важных обобщений отметим использование общей теории устойчивости для некорректных эволюционных задач [5,6] и для исследования проекционно-разностных схем (схем конечных элементов) [7,8]. Приведены также новые априорные оценки устойчивости в интегральных по времени нормах. На их основе может исследоваться, в частности, сходимость разностных схем для задач с обобщенными решениями.

[1] Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований, проект 96-01-00657.

При построении экономичных разностных схем для многомерных задач [3,4], безитерационных алгоритмов декомпозиции области для параллельных компьютеров [9,10] необходимо ориентироваться на использование аддитивных разностных схем (схем расщепления). Приведены основные результаты теории аддитивных разностных схем суммарной аппроксимации, строятся регуляризованные аддитивные схемы полной аппроксимации. Отметим также векторные аддитивные схемы [11,12].

2 Устойчивость операторно-разностных схем

Напомним основные результаты теории устойчивости двухслойных операторно-разностных схем [3,4] и отметим некоторые наиболее важные направления исследований.

2.1 Основные понятия

Пусть H — конечномерное гильбертово пространство, D, A — линейные операторы в H. Скалярное произведение и норма в H есть $(\cdot,\cdot)$ и $\|\cdot\|$ соответственно. Через H_D, где $D = D^* > 0$ обозначим пространство H со скалярным произведением и нормой

$$(y,v)_D = (Dy,v), \quad \|y\| = \sqrt{(Dy,y)}.$$

Пусть $\tau > 0$ — шаг сетки по времени и $y_n = y(t_n)$, $t_n = n\tau$. Рассмотрим однородную (для простоты) двухслойную операторно-разностную схему, записанную в канонической форме

$$B\frac{y_{n+1} - y_n}{\tau} + Ay_n = 0, \quad n = 0,1,\ldots \tag{1}$$

при заданном y_0. Будем считать, что в (1) операторы A и B — постоянны (не зависят от n), оператор A самосопряжен и положителен ($A = A^* > 0$).

Разностная схема (1) называется ρ-устойчивой (равномерно устойчивой) по начальным данным в H_D, если существуют постоянная $\rho > 0$ и постоянная M, не зависящая от τ, n, такие, что при любых n и при всех $y_n \in H$ для решения y_{n+1} разностного уравнения (1) справедлива оценка

$$\|y_{n+1}\|_D \le \rho\|y_n\|_D, \quad n = 0,1,\ldots,$$

причем $\rho^n \le M$. В качестве константы ρ выбирается обычно одна из величин

$$\begin{aligned} \rho &= 1, \\ \rho &= 1 + c\tau, \quad c > 0, \\ \rho &= \exp(c\tau), \end{aligned}$$

где постоянная c не зависит от τ, n. В случае $\rho = 1$ говорят об устойчивости разностной схемы

2.2 Условия устойчивости

Сформулируем стандартные условия устойчивости разностных схем [1,2].

Т е о р е м а 1. *Для разностной схемы (1) с оператором* $A = A^* > 0$ *условие*

$$B \geq \frac{\tau}{2} A \tag{2}$$

является необходимым и достаточным для устойчивости в H_A, *т.е. для выполнения оценки*

$$\|y_{n+1}\|_A \leq \rho \|y_n\|_A, \quad n = 0, 1, \ldots. \tag{3}$$

Если оператор $B = B^* > 0$, *то условие (2) является необходимым и достаточным для устойчивости и в* H_B.

При рассмотрении многих нестационарных задач, примером которых выступают задачи конвекции-диффузии-реакции, необходимо ориентироваться на условия ρ-устойчивости.

Т е о р е м а 2. *Для разностной схемы (1) с операторами* $A = A^*$, $B = B^* > 0$ *условия*

$$\frac{1-\rho}{\tau} B \leq A \leq \frac{1+\rho}{\tau} B \tag{4}$$

необходимы и достаточны для устойчивости в H_B.

Сходимость разностных схем устанавливается в различных нормах, которые должны согласовываться с классом гладкости решений дифференциальной задачи. В силу этого необходимо иметь спектр оценок для разностного решения. При рассмотрении разностных схем для нестационарных краевых задач с обобщенными решениями особого внимания заслуживают оценки разностного решения в интегральных по времени нормах.

В работе [13] отмечается возможность получения условий устойчивости в интегральных по времени нормах. Получены априорные оценки для двухслойных разностных схем, записанных в канонической форме. Принципиальный момент связан с получением оценки для разностного решения в полуцелых узлах по времени, которое определяется линейной интерполяцией по сеточным функциям в узлах.

Т е о р е м а 3. *Для разностной схемы (1) с операторами* $A = A^* > 0$, $B = B^* > 0$ *при выполнении условия (2) имеет место априорная оценка*

$$\sum_{k=0}^{n} \tau \left\| \frac{1}{2}(y_{k+1} + y_k) \right\|_A^2 \leq \frac{1}{2}((B - \frac{\tau}{2}A) y_0, y_0). \tag{5}$$

Оценка (5) получена для разностного решения в полуцелых узлах, где решение определяется выражением $y_{k+1/2} = (y_{k+1} + y_k)/2$. Отметим, что устойчивость в такой интегральной по времени норме установлена при условиях (2), которые являются необходимыми и достаточными для устойчивости в равномерных по времени нормах (см. оценку (3)).

При загрублении неравенства (2) можно получить априорную оценку устойчивости по правой части в интегральных по времени нормах для разностного решения на целых временных шагах.

2.3 Специальные классы операторно-разностных схем

При решении нестационарных задач математической физики особое внимание уделяется схемам с весовыми множителями, когда, например, в (1)

$$B = E + \tau\sigma A, \quad \sigma = \text{const} \geq 0.$$

Построение адаптивных вычислительных алгоритмов на локально сгущающейся сетке по времени или пространству [14,15], регионально-аддитивных разностных схем [9,10] (безытерационных схем декомпозиции области), которые ориентированы на параллельные компьютеры, разностных схем для уравнений смешанного типа базируется на использовании схем с переменными весовыми множителями. Схемы отмеченного класса рассматриваются в работах [16,17]. Приведем характерный результат [18] об устойчивости по начальным данным в несколько более общей операторной формулировке.

Т е о р е м а 4. *Пусть в (1) A — постоянный и положительный оператор, а*

$$B = E + \tau GA, \quad GA \neq A^*G^*. \tag{6}$$

Тогда условие

$$G \geq \frac{1}{2}E \tag{7}$$

достаточно для устойчивости схемы в H_{A^*A}. Схемам с переменными весовыми множителями соответствует выбор диагонального сеточного оператора G в (6).

При ограничениях (7) устанавливается [18] устойчивость разностной схемы (1), в которой

$$B = E + \tau AG, \quad GA \neq A^*G^*.$$

Такой выбор оператора B связывается, например, с новым типом взвешивания при использовании схем с переменными весовыми множителями.

Для самосопряженного оператора A рассмотрены [19] также схемы типа

$$B = E + \tau T^*GT, \quad A = T^*T.$$

Исследование трехслойных разностных схем с переменными весовыми множителями проведено в работе [20].

3 Аддитивные разностные схемы

При рассмотрении многомерных задач большое внимание уделяется построению схем расщепления. Эти схемы связываются с аддитивным представлением оператора задачи и поэтому их называют аддитивными разностными схемами. Дается краткая характеристика наиболее важных классов аддитивных схем.

3.1 Постановка задачи

Рассматривается задача Коши для эволюционного уравнения первого порядка, когда ищется функция $y(t) \in H$, удовлетворяющая уравнению

$$\frac{dy}{dt} + \Lambda y = f(t), \quad 0 < t \le T \tag{8}$$

и начальному условию

$$y(0) = u_0. \tag{9}$$

Будем считать, что оператор Λ положительный, несамосопряженный и нестационарный, т.е. $\Lambda(t) \neq \Lambda^*(t) > 0$. При таких условиях для решения задачи (8),(9) верна априорная оценка

$$\|y(0)\| \le \|u_0\| + \int_0^t \|f(s)\| ds, \tag{10}$$

выражающая устойчивость по начальным данным и правой части.

При расщеплении оператора Λ на сумму отдельных (более простых) операторов можно строить неявные схемы, в которых переход на новый временной слой связан с решением задачи не с оператором Λ, а с решением ряда задач для отдельных операторных слагаемых. Среди таких аддитивных схем мы отметим здесь классические безусловно устойчивые схемы покомпонентного расщепления (локально-одномерные разностные схемы), ориентированные на общее многокомпонентное расщепление.

Будем считать, что для оператора Λ справедливо следующее аддитивное представление

$$\Lambda = \sum_{\alpha=1}^{p} \Lambda_\alpha, \quad \Lambda_\alpha \ge 0, \quad \alpha = 1, 2, ..., p. \tag{11}$$

Аддитивные разностные схемы строятся на основе представления (11), причем переход с одного временного слоя t_n на другой временной слой $t_{n+1} = t_n + \tau$ связан с решением задач для отдельных операторов $\Lambda_\alpha, \alpha = 1, 2, ..., p$ в аддитивном разложении (11), т.е. задача распадается на p подзадач.

Примером аддитивных схем являются классические схемы покомпонентного расщепления (локально-одномерные схемы) [3,4]. Необходимо отдельно упомянуть аддитивно-усредненные схемы покомпонентного расщепления [21,22]. Такие схемы построены не только для эволюционных уравнений первого порядка, но и для уравнений второго порядка [23].

3.2 Схемы покомпонентного расщепления

Аддитивные разностные схемы для задач с расщеплением на три и более попарно некоммутируемых операторов традиционно строятся на основе понятия суммарной аппроксимации – схемы покомпонентного расщепления (локально-одномерные схемы). Для задачи (8),(9),(11) имеем

$$\frac{y^{n+\alpha/p} - y^{n+(\alpha-1)/p}}{\tau} + \Lambda_\alpha(\sigma_\alpha y^{n+\alpha/p} + (1-\sigma_\alpha) y^{n+(\alpha-1)/p}) = \varphi_\alpha^n, \tag{12}$$

$$\alpha = 1, 2, ..., p, \quad n = 0, 1, ...,$$

где

$$f^n = \sum_{\alpha=1}^{p} f_\alpha^n.$$

Хорошо известно, что при $\sigma_\alpha \geq 0.5$ схема покомпонентного расщепления (12) безусловно устойчива. Приведем соответствующую (см. [3,4,23]) априорную оценку устойчивости по начальным данным и правой части.

Для правых частей $f_\alpha^n, \alpha = 1, 2, ..., p$ используется специальное представление

$$f_\alpha^n = \overset{\circ}{f}{}_\alpha^n + \overset{*}{f}{}_\alpha^n, \quad \alpha = 1, 2, ..., p, \quad \sum_{\alpha=1}^{p} \overset{\circ}{f}{}_\alpha^n = 0. \tag{13}$$

Для схемы покомпонентного расщепления верно следующее утверждение.

Т е о р е м а 5. *При* $0.5 \leq \sigma_\alpha \leq 2$, $\alpha = 1, 2, ..., p$ *и* $\tau > 0$ *для (12),(13) выполняется априорная оценка*

$$\|y^{n+1}\| \leq \|u_0\| + \sum_{k=0}^{n} \tau \sum_{\alpha=1}^{p} \left(\| \overset{*}{f}{}_\alpha^k \| + \tau \|\Lambda_\alpha \sum_{\beta=\alpha}^{p} \overset{\circ}{f}{}_\beta^k \| \right). \tag{14}$$

При ориентации на современные параллельные компьютеры можно использовать аддитивно-усредненные схемы покомпонентного расщепления. В этом случае [21,22] переход на новый временной слой осуществляется следующим образом:

$$\frac{y_\alpha^{n+1} - y^n}{p\tau} + \Lambda_\alpha(\sigma_\alpha y_\alpha^{n+1} + (1 - \sigma_\alpha) y^n) = f_\alpha^n, \tag{15}$$

$$\alpha = 1, 2, ..., p, \quad n = 0, 1, ...,$$

$$y^{n+1} = \frac{1}{p} \sum_{\alpha=1}^{p} y_\alpha^{n+1}.$$

Условия устойчивости таких схем такие же как и стандартных схем покомпонентного расщепления. Подобно теореме 5 доказывается следующее утверждение.

Т е о р е м а 6. *При* $\sigma_\alpha \geq 0.5$, $\alpha = 1, 2, ..., p$ *и любых* $\tau > 0$ *для (13),(15) выполняется априорная оценка*

$$\|y^{n+1}\| \leq \|u_0\| + \sum_{k=0}^{n} \tau \sum_{\alpha=1}^{p} \left(\| \overset{*}{f}{}_\alpha^k \| + p\tau\sigma_\alpha \|\Lambda_\alpha \overset{\circ}{f}{}_\alpha^k \| \right). \tag{16}$$

Потенциальное преимущество аддитивно-усредненных схем (15) связано с тем, что допускается параллельная организация вычислений сеточных функций $y_\alpha^{n+1}, \alpha = 1, 2, ..., p$.

Приведенные выше оценки устойчивости (14) и (16) служат основой при исследовании точности рассматриваемых схем расщепления. В виде (12), (15) формулируется задача для погрешности разностного решения. Принципиальный момент связан с тем, что в исследуемых схемах покомпонентного расщепления и аддитивно-усредненных схемах оценки устойчивости существенно зависят от расщепления (13),

от того как выделяются отдельные слагаемые. Фактически это означает то, что точность таких аддитивных схем зависит от того, как мы формулируем промежуточные задачи, как их аппроксимируем и т.д.

С другой стороны промежуточные задачи (вспомогательные сеточные величины $y^{n+\alpha/p}, y^n_\alpha, \alpha = 1, 2, ..., p$) какого-то самостоятельного значения не имеют. В идеале хотелось бы обойтись без них, т.е. строить схемы не привлекая понятия суммарной аппроксимации. Ниже отмечаются некоторые возможности в этом, как нам представляется, очень перспективном направлении.

3.3 Регуляризованные аддитивные схемы

Построение разностных схем заданного качества можно провести на основе общего методологического принципа – принципа регуляризации разностных схем [3,24]. Здесь мы используем этот принцип для получения новых безусловно устойчивых аддитивных схем.

Для построения безусловно устойчивых разностных схем на основе принципа регуляризации используется следующая технологическая цепочка:

- Выбирается какая-либо простейшая разностная схема (производящая разностная схема), которая не относится к необходимому классу безусловно устойчивых.
- Данная схема записывается в канонической форме, для которой из общей теории устойчивости операторно-разностных схем уже известны критерии устойчивости.
- Для введения в класс безусловно устойчивых разностных схем операторы порождающей разностной схемы возмущаются в соответствиями с критериями устойчивости.

В качестве производящей при построении безусловно устойчивых аддитивных схем рассмотрим простейшую явную схему

$$\frac{y^{n+1} - y^n}{\tau} + \sum_{\alpha=1}^{p} \Lambda_\alpha y^n = f^n, \quad n = 0, 1, ... \tag{17}$$

Аддитивные схемы построим на основе возмущения каждого отдельного операторного слагаемого в аддитивном представлении (11):

$$\frac{y^{n+1} - y^n}{\tau} + \sum_{\alpha=1}^{p} (E + \sigma\tau\Lambda_\alpha)^{-1} \Lambda_\alpha y^n = f^n, \quad n = 0, 1, ... \tag{18}$$

Т е о р е м а 7. *При $\sigma_\alpha \geq p/2$, $\alpha = 1, 2, ..., p$ и любых $\tau > 0$ для (18) выполняется априорная оценка*

$$\|y^{n+1}\| \leq \|u_0\| + \sum_{k=0}^{n} \tau \|f^k\|. \tag{19}$$

Рассматриваемая регуляризованная схема (18) имеет тесную связь с аддитивно-усредненной схемой суммарной аппроксимации. Для того чтобы продемонстрировать

это, введем фиктивные сеточные величины $y_\alpha^{n+1}, \alpha = 1, 2, ..., p$, которым теперь никакого самостоятельного значения не придается. Будем реализовывать схему (18) в следующей форме:

$$\frac{y_\alpha^{n+1} - y^n}{p\tau} + (E + \sigma\tau\Lambda_\alpha)^{-1}\Lambda_\alpha y^n = f_\alpha^n,$$

$$\alpha = 1, 2, ..., p, \quad n = 0, 1, ...,$$

$$y^{n+1} = \frac{1}{p}\sum_{\alpha=1}^{p} y_\alpha^{n+1}.$$

Тем самым мы снова приходим к аддитивно-усредненной схеме, построение которой проводится теперь без привлечения понятия суммарной аппроксимации. Отличие от ранее приведенной схемы (15) связано с различными аппроксимациями правых частей.

Сравнение оценок устойчивости схем (16) и (19) существенно в пользу регуляризованной схемы. А именно, для схемы (18) оценка (19) никак не связывается с расщеплением правой части, с выбором функций f_α^n, $\alpha = 1, 2, ..., p$. Именно это обстоятельство часто, как отмечалось выше, порождает трудности исследования аддитивных разностных схем суммарной аппроксимации.

Список литературы

[1] А.А.Самарский, ЖВМиМФ. **7**, 1096 (1967).
[2] А.А.Самарский, ДАН. **181**, 808 (1968).
[3] А.А.Самарский. Теория разностных схем. – М.: Наука, 1983.
[4] А.А.Самарский, А.В.Гулин. Устойчивость разностных схем. – М.: 1973.
[5] А.А.Самарский, П.Н.Вабищевич, Математическое моделирование. **2**(11), 89 (1990).
[6] A.A.Samarskii, P.N.Vabishchevich, Math. Models and Methods in Applied Schiences. **2**, 295 (1992).
[7] П.Н.Вабищевич, А.А.Самарский, ЖВМиМФ. **35**, 1011 (1995).
[8] А.А.Самарский, П.Н.Вабищевич, Математическое моделирование. **8**(9), 74 (1996).
[9] А.А.Самарский, П.Н.Вабищевич, Дифференциальные уравнения. **31**, 1563 (1995).
[10] А.А.Самарский, П.Н.Вабищевич, Докл. АН. **346**, 742 (1996).
[11] В.Н.Абрашин, Дифференциальные уравнения. **26**, 313 (1990).
[12] П.Н.Вабищевич, ЖВМиМФ. **36**(3), 44 (1996).
[13] А.А.Самарский, П.Н.Вабищевич, П.П.Матус, Докл. АН. (1997).
[14] П.Н.Вабищевич, П.П.Матус, В.Г.Рычагов, Дифференциальные уравнения. **31**, 849 (1995).
[15] П.Н.Вабищевич, Известия вузов. Математика. (4), 22 (1995).
[16] А.А.Самарский, А.В.Гулин, Докл. АН. **330**, 694 (1993).
[17] А.В.Гулин, А.А.Самарский, Дифференциальные уравнения. **29**, 1163 (1993).
[18] П.Н.Вабищевич, П.П.Матус, Докл. АН Беларуси. **37** (6), 15 (1993).
[19] П.Н.Вабищевич, П.П.Матус, В.С.Щеглик, Дифференциальные уравнения. **30**, 1175 (1994).
[20] П.Н.Вабищевич, П.П.Матус, В.С.Щеглик, Докл. АН Беларуси. **38** (3), 13 (1994).

[21] Д.Г.Гордезиани, Г.В.Меладзе, ЖВМиМФ. **14**, 246 (1974).
[22] A.A.Samarskii, P.N.Vabishchevich. Computational Heat Transfer. – Chichester: Wiley, 1995.
[23] Д.Г.Гордезиани, А.А.Самарский, Комплексный анализ и его приложения. – М.: Наука, 1978, 173.
[24] А.А.Самарский, ЖВМиМФ, **7**, 62 (1967).

ИТЕРАЦИОННОЕ РЕШЕНИЕ РЕТРОСПЕКТИВНОЙ ОБРАТНОЙ ЗАДАЧИ ТЕПЛОПРОВОДНОСТИ [1]

А.А.Самарский, П.Н.Вабищевич, В.И.Васильев

Институт математического моделирования РАН, Москва,
E-mail: samarskii@imamod.msk.su, vab@imamod.msk.su,
Якутский государственный университет, Якутск,
E-mail: niip@niipmi.yacc.yakutia.su

Abstract. В работе для ретроспективной обратной задачи теплопроводности (задачи с обратным временем для параболического уравнения) разработан итерационный вычислительный алгоритм, в котором уточняется начальное условие, т.е. на каждой итерации решается обычная краевая задача для уравнения теплопроводности. На основе общей теории итерационных методов решения операторных уравнений устанавливаются достаточные условия сходимости итерационного процесса, проводится выбор итерационных параметров.

1 Введение

Среди обратных задач теплообмена [1,2] можно выделить ретроспективную обратную задачу (задачу с обратным временем) по определению начального температурного состояния по температурным измерениям на конечный момент времени. Для ее приближенного устойчивого решения используются различные подходы. Среди основных можно отметить методы с возмущением исходного уравнения — методы квазиобращения в различных вариантах [3]. Второй класс методов связан с возмущением начальных условий. Такой подход реализуется при использовании формулировки задачи как задачи оптимального управления [4]. Аналогичная процедура фактически порождается при использовании процедур нелокального возмущения начальных условий [5].

В настоящее время в теории приближенных методов решения некорректных задач [6] сформировалось направление исследований, которое связано с использованием итерационных методов [7]. При таком подходе в качестве параметра регуляризации выступает число итераций, которое согласуется с погрешностью входных данных. При решении обратных задач для уравнений математической физики широко [8] используются градиентные итерационные методы при вариационной формулировке обратной задачи. В данной работе рассматривается наиболее простой итерационный метод при приближенном решении ретроспективной обратной задачи теплопроводности. Для поставленной обратной задачи итерационно уточняется начальное условие, т.е. на каждой итерации решается обычная краевая задача для уравнения теплопроводности. На основе общей теории итерационных методов решения операторных уравнений [9,10] устанавливаются достаточные условия сходимости итерационного процесса, проводится выбор итерационных параметров. В таких задачах оператор

[1] Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований, проект 96-01-00657.

перехода на новое итерационное приближение позволяет выделить приближенное решение искомого класса гладкости.

2 Постановка задачи

Рассмотрим в качестве модельной двумерную задачу в прямоугольнике

$$\Omega = \{\mathbf{x} \mid \mathbf{x} = (x_1, x_2), \quad 0 < x_\alpha < l_\alpha, \ \alpha = 1, 2\}.$$

В Ω ищется решение параболического уравнения

$$\frac{\partial u}{\partial t} - \sum_{\alpha=1}^{2} \frac{\partial}{\partial x_\alpha} k(\mathbf{x}) \frac{\partial u}{\partial x_\alpha} = 0, \quad \mathbf{x} \in \Omega, \quad 0 < t < T, \tag{1}$$

дополненного простейшими однородными граничными условия первого рода:

$$u(\mathbf{x}, t) = 0, \quad \mathbf{x} \in \partial\Omega, \quad 0 < t < T. \tag{2}$$

В рассматриваемой обратной задаче вместо задания решения на начальный момент времени $t = 0$ задается решение на конечный момент времени:

$$u(\mathbf{x}, T) = g(\mathbf{x}), \quad \mathbf{x} \in \Omega. \tag{3}$$

Такая обратная задача корректна, например, в классах ограниченных решений (положительных) решений.

3 Дифференциально-разностная задача

Поставим в соответствие дифференциальной задаче (1)-(3) дифференциально-разностную задачу, проведя дискретизацию по пространству. Будем считать, для простоты, что в области Ω введена равномерная по каждому направлению сетка с шагами h_α, $\alpha = 1, 2$. Пусть ω — множество внутренних узлов.

На множестве сеточных функций $y(\mathbf{x})$ таких, что $y(\mathbf{x}) = 0$, $\mathbf{x} \notin \omega$ определим сеточный оператор Λ соотношением

$$\Lambda y = -\sum_{\alpha=1}^{2} (a_\alpha y_{\bar{x}_\alpha})_{x_\alpha}, \tag{4}$$

положив, например,

$$a_1(\mathbf{x}) = k(x_1 - 0.5h_1, x_2), \quad a_2(\mathbf{x}) = k(x_1, x_2 - 0.5h_2).$$

В сеточном гильбертовом пространстве H скалярное произведение и норму введем соотношениями

$$(y, w) = \sum_{\mathbf{x} \in \omega} y(\mathbf{x}) w(\mathbf{x}) h_1 h_2, \quad \|y\| = \sqrt{(y, y)}.$$

В H имеем [9,10] $\Lambda = \Lambda^* > 0$. От (1)-(3) перейдем к дифференциально-операторному уравнению

$$\frac{dy}{dt} + \Lambda y = 0, \quad \mathbf{x} \in \omega, \quad 0 < t < T \tag{5}$$

при заданном

$$y(T) = g, \quad \mathbf{x} \in \omega. \tag{6}$$

4 Итерационное уточнение начального условия

Для приближенного решения обратной задачи (5),(6) будем использовать итерационные методы, связанные с уточнением начального условия. На каждой итерации соответствующие корректные задачи решаются на основе использования стандартных двухслойных разностных схем [9,11].

Пусть вместо обратной задачи (5),(6) рассматривается прямая задача для уравнения (5), когда вместо (6) используется начальное условие

$$y(0) = v, \quad \mathbf{x} \in \omega. \tag{7}$$

Обозначим через y_n разностное решение на момент времени $t_n = n\tau$, где $\tau > 0$ — шаг по времени, причем $N\tau = T$. При использовании обычной двухслойной схемы с весами [9,11] переход на новый временной слой в задаче (5),(7) осуществляется в соответствии с

$$\frac{y_{n+1} - y_n}{\tau} + \Lambda(\sigma y_{n+1} + (1-\sigma)y_n) = 0, \quad n = 0, 1, \ldots, N-1, \tag{8}$$

$$y_0 = v, \quad \mathbf{x} \in \omega. \tag{9}$$

Как хорошо известно [9], схема с весами (8),(9) безусловно устойчива при $\sigma \geq 0.5$ и справедлива следующая оценка устойчивости

$$\|y_{n+1}\| \leq \|y_n\| \leq \ldots \leq \|y_0\| = \|v\|, \quad n = 0, 1, \ldots, N-1. \tag{10}$$

Тем самым норма решения со временем убывает.

Для приближенного решения обратной задачи (5),(6) будем использовать простейший итерационный процесс, основанный на последовательном уточнении начального условия и решении на каждой итерации прямой задачи. Придадим этой задаче соответствующую операторную формулировку.

Из (8),(9) для заданного y_0 на конечный момент времени получим

$$y_N = S^N v, \tag{11}$$

где S — оператор перехода с одного временного слоя на другой:

$$S = (E + \sigma\tau\Lambda)^{-1}(E + (\sigma - 1)\tau\Lambda). \tag{12}$$

С учетом (5),(6) и (11) приближенному решению обратной задачи естественно сопоставить решение следующего сеточного операторного уравнения

$$Av = g, \quad \mathbf{x} \in \omega, \quad A = S^N. \tag{13}$$

В силу самосопряженности оператора Λ самосопряженным является оператор перехода S и оператор A в (13). Однозначная разрешимость сеточного уравнения (13) будет имет место, например, при положительности оператора A. Это условие будет выполнено для положительного оператора перехода S. Принимая во внимание представление (12), получим $S > 0$ при

$$\sigma \geq 1. \tag{14}$$

Условие (14) на вес схемы (8),(9) является более жестким, чем обычное условие устойчивости. В нашем случае в (13) при ограничениях (14)

$$0 < A = A^* < E. \tag{15}$$

Для решения уравнения (13),(15) можно использовать явный двухслойный итерационный метод, который записывается [10] в виде

$$\frac{v_{k+1} - v_k}{s_{k+1}} + Av_k = g, \tag{16}$$

где s_{k+1} — итерационные параметры.

Рассматриваемый итерационный метод соответствует следующей организации вычислений при приближенном решении ретроспективной обратной задачи (1),(2).

- При заданном v_k решается прямая задача с использованием разностной схемы

$$\frac{y_{n+1}^k - y_n^k}{\tau} + \Lambda(\sigma y_{n+1}^k + (1-\sigma)y_n^k) = 0, \quad n = 0, 1, \ldots, N-1, \tag{17}$$

$$y_0^k = v_k, \quad x \in \omega \tag{18}$$

для определения y_N^k.

- В соответствии с (16) уточняется начальное условие:

$$v_{k+1} = v_k - s_{k+1}(y_N^k - g). \tag{19}$$

Как следует из общей теории итерационных методов [10] скорость сходимости метода (13) определяется постоянными энергетической эквивалентности γ_α, $\alpha = 1, 2$:

$$\gamma_1 E \le A \le \gamma_2 E, \quad \gamma_1 > 0. \tag{20}$$

Принимая во внимание (15), можно положить $\gamma_2 = 1$. Положительная постоянная γ_1 зависит от сетки и лежит вблизи нуля.

Для стационарного итерационного метода ($s_k = s_0 = \text{const}$ в (16)) условия сходимости имеют [10] в наших условиях вид $s_0 \le 2$. Для оптимального постоянного значения итерационного параметра имеем $s_0 \approx 1$. Для ускорения сходимости необходимо ориентироваться на применение итерационных методов вариационного типа. При использовании итерационного метода минимальных невязок для итерационных параметров имеем

$$s_{k+1} = \frac{(Ar_k, r_k)}{(Ar_k, Ar_k)}, \quad r_k = Av_k - g.$$

В этом случае на каждой итерации минимизируется норма невязки, т.е. верна оценка

$$\|r_{k+1}\| < \|r_k\| < \ldots < \|r_0\|.$$

При использовании более общего неявного итерационного метода вместо (16) имеем

$$B\frac{v_{k+1} - v_k}{s_{k+1}} + Av_k = g, \tag{21}$$

где $B = B^* > 0$ В этом случае (метод минимальных поправок) для итерационных параметров используются расчетные формулы

$$s_{k+1} = \frac{(Aw_k, w_k)}{(B^{-1}Aw_k, Aw_k)}, \quad w_k = B^{-1}r_k.$$

Для этого случая минимизируется поправка w_{k+1} на каждой итерации, т.е. верна оценка

$$\|w_{k+1}\| < \|w_k\| < \ldots < \|w_0\|.$$

Аналогично рассматриваются и более быстрые трехслойные вариационные итерационные методы.

Отметим некоторые особенности выбора оператора B при решении некорректных задач. В обычных итерационных методах выбор оператора B подчинен исключительно ускорению скорости сходимости итерационного метода. При решении некорректных задач итерационный процесс обрывается до достижения невязки, величина которой определяется погрешностью входных данных. Для нас важно не только с какой скоростью сходится итерационный процесс на этом участке убывания, но и в каком классе гладкости этот итерационный процесс сходится, в какой норме необходимый уровень невязки достигается. Принципиальная особенность приближенного решения некорректных задач итерационными методами состоит в том, что выделение приближенного решения из необходимого класса гладкости достигается выбором оператора B.

Список литературы

[1] О.М.Алифанов. Обратные задачи теплообмена. – М.: Машиностроение, 1988.

[2] Дж.Бек, Б.Блакуэлл, Ч.Сент-Клэр. Некорректные обратные задачи теплопроводности. – М.: Мир, 1989.

[3] Р.Латтес, Ж.Л.Лионс. Метод квазиобращения и его приложения. – М.: Мир, 1970.

[4] Ж.Л.Лионс Оптимальное управление системами, описываемыми уравнениями с частными производными. – М.: Мир, 1972.

[5] П.Н.Вабищевич, Математическое моделирование. 4 (5),109 (1992).

[6] А.Н.Тихонов, В.Я.Арсенин. Методы решения некорректных задач. – М.: Наука, 1986.

[7] Г.М.Вайникко, А.Ю.Веретенников. Итерационные процедуры в некорректных задачах. – М.: Наука, 1986.

[8] О.М.Алифанов, Е.А.Артюхин, С.В.Румянцев. Экстремальные методы решения некорректных задач. – М.: Наука, 1988.

[9] А.А.Самарский. Теория разностных схем. – М.: Наука, 1983.

[10] А.А.Самарский, Е.С.Николаев. Методы решения сеточных уравнений. – М.: Наука, 1978.

[11] А.А.Самарский, А.В.Гулин. Устойчивость разностных схем. – М.: 1973.

Author Index

A

Abramov A.A. • 9, 293
Abramov S.A. • 16
Airapetyan R.G. • 27
Akishin P.G. • 32
Amirkhanov I.V. • 40, 48, 300
Andrianov S. • 55, 60
Anikeeva V.A. • 65
Anykeyev V.B. • 70
Aristova E.N. • 77
Avdeev M.V. • 82

B

Barashenkov V.S. • 91
Batgerel B. • 95
Beliaev A. • 100
Belyaev V.B. • 105
Berezhnoj V.A. • 110
Bogoljubskaya A.A. • 115
Bogoljubsky I.L. • 115
Bogovskij M.E. • 352
Borisovskij V.E. • 110
Brushlinsky K.V. • 121
Budnik A.D. • 126
Busha J. • 27

C

Casolino M. • 130
Cox M. • 186

D

Dautov R. • 181
Davidzhan E.A. • 137
Dikoussar N.D. • 82
Dmitriev A.V. • 100, 142
Dyachin V.V. • 146
Dyshko A.L. • 9, 293, 304

E

Egorov A.A. • 153
Egorov A.D. • 156

F

Filipova T.V. • 126

G

Galishnikova T.N. • 341
Gilyanenko A.S. • 142
Gol'din V.Ya. • 77
Grebenikov E.A. • 311

H

Hohlov Yu.E. • 159
Howison S. • 159
Hu C.-K. • 165

I

Ibragimov I.V. • 327
Ilinskij A.S. • 341
Isaev Yu.N. • 110
Ivanchenko I.M. • 332, 337
Ivanchenko Z.M. • 332
Ivanova T.M. • 137

K

Karpenko N.N. • 332
Kartavtsev O.I. • 105
Khoromskij B.N. • 172
Klimenko S.V. • 146
Kochkin V.I. • 105
Kolganova E.A. • 105, 177
Konyukhova N.B. • 9, 293, 304
Kononov A.V. • 234, 237
Korenkov V.V. • 110, 337
Kornev K. • 159, 181
Kostomarov D.P. • 186
Kotina E.D. • 193
Kotov Yu.D. • 142
Kovalenko V. • 198
Kretov Yu.A. • 337
Kuznetsov S.N. • 100, 142
Kuznetsov Yu.I. • 203

L

Lakhno V.D. • 40
Levitina V.V. • 9, 293
Loukianova A. • 210

M

Manyuk I.M. • 282
Maslennikova V.N. • 352
Matus P.P. • 347
Mironov S.V. • 311
Motovilov A.K. • 177
Mourzenko V. • 181
Muresan L. • 214
Muresan R. • 214
Musulmanbekov G. • 110

N

Nechaev O. • 100
Nefedov V.V. • 86
Nikitin I.N. • 145
Nikonov E.G. • 110, 172
Niukkanen A.W • 219

O

O'Brien M.R. • 186
Ososkov G.A. • 214
Ovsyannikov A.D. • 224, 314

P

Panasyuk M. • 100
Panebratsev Yu. • 214
Petsilyak I.I. • 42
Picozza P. • 13)
Plotnikov A.A. • 65
Plyashkevich V.S. • 110
Podosenova T.B. • 142, 288
Polyakova R.V. • 314
Popov A.V. • 70
Popov B.A. • 229
Poritsky P.V. • 234, 237
Puzynin I.V. • 32, 40, 48, 240,
Puzynina T.P. • 40, 48, 240

R

Ratnikova T.A. • 121
Ribalkin A.V. • 110
Robinson D.C. • 186
Ryabenko A.A. • 360
Ryumin S.P. • 142

S

Samarskij A.A. • 363, 372
Sapogin L.G. • 91
Savyolova T.I. • 137
Serdyukova S. • 246
Shirikov V.P. • 110
Shishkin A.G. • 186
Shishkin G.I. • 253
Smirnov Yu.S. • 32
Sokolov A.P. • 110
Solovjov A.G. • 321
Starchyk P.D. • 234, 237
Staroverova I.B. • 304
Strizh T.A. • 40, 48

T

Thuné M. • 246
Tyrtyshnikov E.E. • 327
Tyukhtyaev A.Yu. • 240

U

Ul'janova V.I. • 293

V

Vabishchevich P.N. • 363, 372
Vasil'eva A.B. • 65
Vasil'ev S.A. • 300
Vasil'ev V.I. • 372
Vinitsky S.I. • 240
Voitenko L.M. • 234, 237
Volkov B.S. • 110
Voronov N.A. • 304

W

Wong T.M.Y. • 259

Y

Yang T. • 265
Yanovich L.A. • 282
Yudin I.P. • 314

Z

Zaikin P.N. • 142, 288
Zaitsev F.S. • 186
Zemlyanaya E.V. • 40, 48
Zhang J.F. • 259
Zhanlav T. • 95
Zherelo A.V. • 156
Zhidkov E.P. • 27, 300, 314, 321
Zhidkova I.V. • 300
Zhigunov V.P. • 70
Zima E.V. • 16

Computational Modelling and Computing in Physics

September 16 – 21 1996, Dubna, Russia

Математическое моделирование
и компьютерные методы в физике

16 – 21 сентября 1996, Дубна, Россия

D5,11-97-112

Ответственный за подготовку сборника к печати А.В.Федоров

Макет Т.Е.Попеко

Рукопись сборника поступила 02.04.97. Подписано в печать 20.06.97
Формат 60 × 90/16. Офсетная печать. Уч.-изд. листов 35,10
Тираж 200. Заказ 50049. Цена 42800 р.

Издательский отдел Объединенного института ядерных исследований
Дубна Московской области