2001 SUR Nomination (partial)

5. PROJECT DESCRIPTION

Title for this project description:

Development of Computational Techniques with Application to Petroleum Exploration

5.1 Proposal Framework

The proposal focuses on the development and application of numerical computational fluid and solid dynamics.

In particular, we believe there are significant potential applications to oil and gas exploration. The approach to the problems is from a fundamental mechanics direction; consequently, the software will find applications in many exploration areas and in creeping flows in totally different areas (e.g. areas as diverse as plastic molding and food science).

Our research specifically focuses on the modelling of processes related to the development of sedimentary basins, and the assessment of their oil and gas resources. The modelling team comprises researchers from Dalhousie University, and the Geological Survey of Canada-Atlantic, at the Bedford Institute of Oceanography.

Among other projects, the numerical modelling part of the team's research will involve two main components; 1) the development of new numerical techniques to enable efficient computations with sufficient accuracy, and; 2) the applications of these techniques to specific problems.

5.2 Related Initiatives

THE ATLANTIC INNOVATION FUND (AIF)

The proposed research will comprise part of a much larger Oil and Gas Research and Development initiative in the Atlantic Provinces of Canada. The Federal Government of Canada has established a $300 million fund to support innovation and economic development in Atlantic Canada (the provinces of Newfoundland, New Brunswick, Nova Scotia and Prince Edward Island). Oil and Gas related research and development is one target area for this fund and a Pan-Atlantic consortium and network, the Petroleum Systems Network has been developed to undertake innovative research and to acquire infrastructure in support of economic development. Part of the research will focus on the development of sedimentary basins.

A proposal has been developed which will probably request up to $50 million from this fund. We are currently waiting for the request for proposals. It will include funds for a component of the Geodynamical modelling and model development described in this proposal. Specifically, we will request funding for up to 5yrs for students and postdoctoral fellows to work on the development and application of techniques described here. These people will become part of the pool of highly qualified personnel.

As part of the AIF proposal there will be additional opportunities to apply to the Canada Foundation for Innovation (CFI) for improved computational infrastructure.

CANADA RESEARCH CHAIR AND CANADA FOUNDATION FOR INNOVATION

We have already been awarded $C325,000 from CFI as part of the Canada Research Chair in Geodynamics held by Chris Beaumont. These funds may be used to upgrade the existing IBM-SP, or could be used in conjunction

with a SUR grant.

CANADIAN INSTITUTE FOR ADVANCED RESEARCH (CIAR)

We are members of the Earth System Evolution Programme of CIAR and other members of this group, for example, Drs Jean Braun (ANU) and Sean Willett (U. Washington) are also developing computational fluid and solid techniques for research in geodynamics. Access to their expertise is beneficial to our own research.

5.3 Development of Efficient Solvers for Complex Flows

BACKGROUND

Complex flows constitute a recurrent theme in geophysics - researchers in this field have made an increasing use of numerical model calculations for lithospheric deformation, mantle convection, and the Earth's core dynamo, which is responsible for the geomagnetic field. Codes like TERRA and DYNAMO (developed at Los Alamos national Laboratory) simulating respectively earth-scale thermal convection and magnetic field, were rated on the teraflops scale on massively parallel (>1000 processors) machines in 1998, to give but an end member example compared with other more modest codes and available resources.

We have developed a reasonable expertise in the modelling of complex incompressible flows with various finite-element codes using parallel structured and, so far, serial unstructured solvers, mostly in 2D. Significant progress in algorithms and hardware capabilities lead us to expect feasibility for similar problems in 3D over the next few years. Intensive algorithmic tests (FEATFLOW code : 10^8 cpu secs ~ 2.5 cpu years; sources: Stefan Turek, Universitat Heidelberg) have been performed to segregate solvers according to various measures of performance, and computational geometry studies are contributing to increasingly efficient meshing techniques.

We seek to develop along the line proposed by specialists of CFD solvers our own 3D (nearly) incompressible flow solver, and undertake numerical experiments on general (conceptual geophysical problems) and specific applications (see 5.4 for problems related to the Petroleum industry).

COMPLEX FLOWS

Flows to be considered are typically complex in the sense that few simplifications are allowed in the physical problem at hand. Modelled materials are characterized by strongly nonlinear, temperature/pressure/strain/ rotation dependent rheologies. They may dynamically localize deformation and are subjected to non-constant and non-uniform specific weight. Continua are heterogeneous, with free surface and moving or evolving (phase transition) internal boundaries. The representation of physical problems into a boundary value mathematical problem corresponds to somewhat arbitrary choices in terms of the boundary conditions and initial preparation of the fields. Extraction of useful information relies therefore on a large number of simulations spanning some set (chosen by educated guess) of input parameters. Average model running time has so far been of order a day and we contemplate a similar (wall clock time) scale for future 3D models.

We plan to approach our 3D solver with the design of a complexity graph. Each node of this graph represents a class of problems and each oriented edge corresponds to an incremental increase in complexity removing one previously assumed simplification. The merits of such an approach will be many :

- explicit simplifications will imply closer to optimal solver in each complexity class
- users will be able to use a solver well-suited to their problem
- solution of more complex problems could benefit from a (preconditioning-like) path through simpler problems.

The 'simplify' operation may involve (e.g.) :

- material assumptions (isoviscosity, linearity, incompressibility, no elasticity)
- boundary conditions and geometry (1 or 2 axis periodicity, fixed domain)
- level of deformation, presence or not of localized features
- ranges for dimensionless parameters (aspect ratios, Pe,Pr,Re,Da...in 0,1, infinity zones)
- topology and shape of domain
- assumption regarding the evolution of fields in time.

Each successfully solved class of problem will be added to the complexity graph as development progresses.

EFFICIENT SOLVERS

We take the view that low order finite element methods for parallel Stokes flow solvers will be part of the high complexity nodes in this hierarchy. Theoreticians do not agree on which of the stabilized or non-conformal (e.g. solenoidal trilinear-rotated q1-q0) methods are best for element discretization, and we advocate to implement both types. Most Stokes flow solvers may be regrouped generically into multilevel pressure Schur complement methods [22] and their efficiency is controlled by a delicate and somewhat unpredictable balance between machine architecture, operation count per iteration and number of iterations to convergence. We are of the view that scalability and robustness of convergence have priority over speed, at least within certain limits, and that special attention should be paid to preconditioners adapted to the prevailing physics. Each solver itself corresponds to a tree involving e.g. geometry discretizations (meshing, remeshing), nonlinear solvers (Newton-Krylov-Schwarz, defect correction), linear solvers (fully coupled multigrid, Krylov), boundary operators (Poincaré-Steklov functionals), error estimators and analysis tools. We are not yet in a position to prefer one approach over another and contemplate the building in our code of all basic algorithmic blocks.

As explained earlier, we don't exclude mixed methods (e.g. finite-element/spectral) as a way to increase efficiency.

5.4 Applications Related To Oil-Gas Research and Development

These applications include the following.

Sedimentary Basin/Tectonics Interactions

Sedimentary basins are depressions in the Earth's surface where sediments accumulate. They form as a consequence of deformation and subsidence of the Earth's outer strong rind, the lithosphere. Deformation may also cause uplift of the surface and many sedimentary basins undergo superimposed phases of subsidence and uplift. The primary mechanisms for the formation of sedimentary basins are lithospheric extension, which creates rift basins and passive continental margins, and lithospheric contraction, which creates forearc and foreland basins. Lithospheric extension and contraction that are normal to plate boundaries can be addressed to first order with plane-strain vertical cross-section models. However, if the motion includes a transcurrent component parallel to the plate boundary, leading to transtension or transpression, the problem needs to be addressed in 2.5 or 3 dimensions. Even when motion is contractional/extensional lithospheric deformation varies along the length of the mountain belt/rift system again creating a fundamentally 3 dimensional system.

A crucial problem at the heart of rift and passive margin geometries concerns partitioning of strain and deformation during rifting. Even the controls of symmetry and asymmetry require a solution of this problem and there is merit in both 2D and 3D approaches. Deformation is both focussed (on faults and within shear zones) and distributed (where there is no localization) and, in our opinion, the key to rift geometries is an understanding of the reason for this partitioning of distributed and focussed strain. Progress will be made by

combining model and observational studies. Rifting processes are also one target of the NSF Margins program and our work is intended to complement the observational component of this initiative. Our proposed research is also intended to provide insight into the development of sedimentary basins that are superimposed on the lithospheric rifting fabric.

Two end-member hypotheses to be investigated are that reactivation of existing faults/shears, or strain softening/hardening of newly developed ones, are the primary controls on rifting fabric. Our initial approach to these problems has been to solve Stokes flow problems at the scale of the lithosphere and sublithospheric mantle. The 2D cross section coupled thermal-mechanical ALE FE models (e.g. Huismans and Beaumont, [1]; Pysklywec et al., [2]; Pysklywec et al. [3]) have 200x100 Eulerian meshes, 700 x 300 Lagrangian meshes, a domain of approx. 2000x600km, and are driven by specified velocity boundary conditions and internal density variations. The surface is a deformable stress-free boundary which provides an accurate calculation of the gravitational forcing by topography. Rheologies include cohesive (von Mises), and pressure sensitive (Drucker-Prager) frictional plasticity, thermally activated creep (grain-size sensitive, power-law, and Peierl's law) and may include strain rate- or strain-softening/hardening. Inclusion of strain-dependent material behaviour allows us to investigate strain partitioning during deformation in addition to that predicted by the basic rheologies. Our provisional conclusion like that of others, is that strain softening is required for plate tectonics to operate. Plate boundaries with Byerlee's law and laboratory based creep rheologies are too strong and do not localize deformation. Softening/hardening is also necessary for reactivation to be important, without them there would be no weak faults/shears to reactivate. Examples of the model results (Huismans and Beaumont, [1]) demonstrate the sensitivity of rift geometries to initial conditions, rifting velocity, and strain softening in the frictional regime. They show symmetric (McKenzie type) and asymmetric (Wernicke type) styles can be produced when strain softening is included.

The next steps are of three types :1) to broaden the range of model experiments to determine the sensitivity to other material (see also Frederiksen and Braun, [4] for possible mechanisms of ductile strain softening) and thermal initial conditions, and to other processes e.g. decompression melting and sedimentation during rifting; 2) to investigate the post-rift evolution and the sensitivity of the evolving passive margin to extensional and contractional reactivation as it cools, and; 3) to parameterize possible natural fault and shear zone softening and hardening mechanisms (e.g observational research on 'weak faults', e.g. Rutter et. al. [5], the 'damage' approach to fault zone weakening e.g. Tackley [6]), and to determine the model sensitivity to these, as opposed to empirical, mechanisms. This approach will be undertaken in parallel with the development of improved numerical techniques, first in 2D (for example using high resolution meshes in parallel codes, and adaptive mesh refinement to follow local shear intensification), and later in 3D as the proposed developments are achieved.

Motivation for the 3D models stems from the need to understand strike variations in rift and passive margin architecture. Are there characteristic segmentation patterns and if so how is strain partitioned within the segments and across their boundaries? Some indication of possible patterns exists from studies of rifts, for example, the fundamental or overlapping half-graben units (Rosendahl [7]). This interpretation is also complemented by sandbox experiments (e.g. Schreurs and Coletta [8]). A first-order problem concerns the reason for strike polarity changes, the half-graben architecture, and the particular geometries where segments overlap. Can these be understood by applying a minimum work principle coupled with strain softening? The hypothesis is that initial frictional-plastic failure of the upper crust when viewed in cross section occurs on cross-conjugate structures that define the early rift-bounding faults. At any location either of the conjugates may preferentially strain soften, thereby creating the characteristic asymmetric half-grabens, seen both in nature and in the cross-section models (Huismans and Beaumont, [1]). The polarity chosen will, however vary along strike and the question to be addressed in the 3D models is the way in which the system organizes into finite length segments of alternating polarity. Here the conjecture is that strain softening on both the half-graben faults and in the transfer regions, that link the opposite polarity segments, allows the system to minimize work and yet retain

the segmentation. By analyzing the system as a whole, it is potentially possible to determine the amount of strain softening required for a given geometry to be a minimum work configuration.

The lack of strong self-organized patterns in rifts also suggests that other effects compete with the mechanism proposed above. By introducing initial weak structures into the FE models it will be possible to determine the threshold level of weakening and density of these structures required to interfere with self-organization. This approach offers a possible way to address the relative roles of reactivation of pre-existing fabric versus the creation of new rifting fabric.

5.5 Salt Tectonics

A second application concerns the role of 'salt' layers deposited as part of the sedimentary fill of basins. Salt is much weaker and is less dense that most other sediments and commonly becomes mobile, deforming under the weight of sediments accumulating above. Salt tectonics has a long and rich history, particularly because salt mobility can influence several aspects of the petroleum system e.g. Jackson et al. [9], Alsop et al.[10], Vendeville et al.[11], Cobbold [12].

In the last decade the petroleum industry has become increasingly interested in the oil and gas potential of reservoirs in the deep water, slope and rise regions of passive continental margins e.g. Mohriak and Talwani [13]. These regions are also commonly salt provinces, e.g. parts of west Africa and the Gulf of Mexico, and there is a need to understanding the deep water petroleum system taking into account salt deformation and its interaction with the overall sedimentary prism, e.g. Jackson and Vendeville [14], and Talbot [15]. Parts of the margin of Atlantic Canada are also salt provinces and the application of the more general problem of salt tectonics to this region will be a focus in the Oil and Gas Proposal to the Atlantic Innovation Fund.

The passive margin salt tectonics problem has been addressed through direct, mainly seismic, observations to determine the current geometries of salt-related structures, mathematical analysis of simplified salt systems, for example diapiric instabilities, laboratory analogue modelling, and numerical models. To date, the analogue models may have been the most successful in providing insight into salt tectonics. However, as numerical techniques have improved and computational power has increased, numerical approaches are poised to address questions that complement those addressed by analogue models and mathematical analysis.

The minimum requirements for numerical experiments concerning salt tectonics is an ability to calculate large gravitationally (buoyancy) driven deformation of sediments and salt using appropriate frictional-plastic and thermally-activated ductile rheologies. Given that salt commonly becomes mobile during accumulation of overlying sediment, the numerical experiments should also include an upper free surface (subject to hydrostatic pressure if the section is submarine) which both deforms and is modified by sedimentation and erosion. The role of evolving fluid pressure within the sedimentary wedge is certainly important under some circumstances, as is the isostatic compensation of the system. Our current numerical capabilities satisfy all of the above requirements, except the fluid pressure component, in 2D. Although we have not focussed on salt tectonics, preliminary experiments indicate that we can address the combined salt/sediment deformation and surface process problem. In analogous problems we have considerable experience in calculating flows at the crustal scale which involve weak materials flowing in channels under gravity forcing and interacting with surface erosion and sediment deposition (Beaumont et al. [16], Beaumont et al. [17]). We have also developed and used surface process models e.g. Kooi and Beaumont [18], Johnson and Beaumont [19] and coupled them to crustal scale tectonic models, e.g. Beaumont et al. [20].

Our proposed research focus will initially be at the larger scale, with the intent of understanding the way sedimentation triggers extensional and compressional deformation of passive continental margins underlain by salt (e.g. deformation like that described by Marton et al. [21] and centrifuge modelled by Talbot [15]). We will use the existing 2D structured ALE FE code for coarse experiments. In the same manner as in the rifted margin

tectonic research the experiments will be progressively refined using higher resolution parallel techniques and adaptive meshing. The problem can be addressed in 2D at the large, i.e, margin scale. However, when considering deformation of individual salt bodies and the development of diapirs versus salt sheets it will be necessary to use 3D techniques that we propose to develop.

These techniques and software we are developing, while applicable to the Petroleum industry, may also find applications elsewhere because they will solve general problems of visco-plastic creeping flows with free surfaces under a range of boundary conditions.

Project Web Site – http://adder.ocean.dal.ca

References for Section 5

[1] Huismans, R., and Beaumont, C., sub. Asymmetric lithospheric extension: the role of frictional-plastic strain softening inferred from numerical experiments, Geology.

[2] Pysklywec, R.N., Beaumont, C., and Fullsack, P., 2000. Modeling the behavior of the continental mantle lithosphere during plate convergence, Geology, v. 28, p. 655-658.

[3] Pysklywec, R.N., Beaumont, C., and Fullsack, P., sub. Lithospheric deformation during the early stages of continental collision: numerical experiments and comparison with South Island, New Zealand, Journal of Geophysical Research.

[4] Frederiksen, S., and Braun, J., 2001. Numerical modelling of strain localization during extension of the continental lithosphere, Earth and Planetary Science Letters, v. 188, p. 241-251.

[5] Rutter, E.H., Holdsworth, R.E., and Knipe, R.J., in press. The nature and tectonic significance of fault zone weakening: an introduction, in: R.E. Holdsworth et al. eds., The Nature and Significance of Fault Zone Weakening, Geological Society Spec. Publ. 186.

[6] Tackley, P.J., 2000. The quest for self-consistent generation of plate tectonics in mantle convection models, in: M.A. Richards et al. eds., Geophysical Monograph 121, American Geophysical Union, p. 47-72.

[7] Rosendahl, B.C., 1987. Architecture of continental rifts with special reference to east Africa, Annual Reviews Earth and Planetary Science, v. 15, p. 445-503.

[8] Schreurs, G., and Coletta, B., 1998. Analogue modelling of faulting in zones of continental transpression and transtension, in: R.E. Holdsworth et al., eds., Continental Transpressional and Transtensional Tectonics, Geological Society Special Publications, v. 135, p. 59-79.

[9] Jackson, M.P.A., Roberts, D.G., and Snelson, S., eds., 1995. Salt Tectonics a Global Perspective, American Association of Petroleum Geologists Memoir 65, 454 pp.

[10] Alsop, G.I., Blundell, D.J., and Davison, I., eds., 1996. Salt Tectonics, Geological Society Special Publication 100, 310 pp.

[11] Vendeville, B.C., Mart, Y., and Vigneresse, J-L., eds., 2000. Salt, Shale and Igneous Diapirs in and around Europe. Geological Society Special Publication 174, 207 pp.

[12] Cobbold, P.R., ed., 1993. New Insights into Salt Tectonics, Tectonophysics, v.228, p.141-450.

[13] Mohriak, W., and Talwani, M., eds., 2000. Atlantic Rifts and Continental Margins, Geophysical Monograph 115, American Geophysical Union. 354 pp.

[14] Jackson, M.P.A., and Vendeville, B.C., 1996. Regional extension as a geologic trigger for diapirism, Geological Society of America Bulletin, v.106, p.57-73.

[15] Talbot, C.J., 1992. Centrifuged models of Gulf of Mexico profiles, Marine and Petroleum Geology, v.9, p.412-432.

[16] Beaumont, C., Jamieson, R.A., Nguyen, M.H, and Lee., B., 2001. Mid-crustal channel flow in large hot orogens: results from coupled thermal-mechanical models, Lithoprobe SNORCLE Transect Report.

[17] Beaumont, C., Jamieson, R.A., Nguyen, M.H., and Lee, B., sub. Himalayan tectonics explained by extrusion of a low-viscosity crustal channel coupled to focused surface denudation, Nature.

[18] Kooi, H., and Beaumont, C., 1994. Escarpment evolution on high-elevation rifted margins: insights derived from a surface processes model that combines diffusion, advection and reaction. Journal of Geophysical Research, v.99, p.12191-12209.

[19] Johnson D.D., and Beaumont, C., 1995. Preliminary results from a planform kinematic model of orogen evolution, surface processes, and the development of clastic foreland basin stratigraphy, in: S. Dorobek and G. Ross eds., Stratigraphic Evolution of Foreland Basins, SEPM Special Publication v.52, p.3-24.

[20] Beaumont, C., Fullsack, P., and Hamilton, J., 1992. Erosional control of active compressional orogens, in: K.R. McClay ed., Thrust Tectonics, p. 1-18. Chapman and Hall.

[21] Marton, L.G., Tari, G.C., and Lehmann, C.T., 2000. Evolution of the Angolan passive margin, west Africa, with emphasis on the post-salt structural styles, in: W. Mohriak and M. Talwani eds., Atlantic Rifts and Continental Margins, Geophysical Monograph 115, American Geophysical Union, p.129- 149.

APPENDICES: DEVELOPMENT OF NUMERICAL TECHNIQUES

APPENDIX 1. Physical Problem Model

A1.1 Problem Statement (P)

Given a Lagrangian medium $D_0 \subset R^d$, and a set of (scalar or tensorial) fields, F_0, we want to compute the evolution in time of these fields subject to a series of constraints such as:

- moving stress-free boundaries
- motion or forces imposed, or controlled, on moving boundaries
- incompressibility of motion
- mechanical balance of continuum stresses
- rheological laws relating stresses (stress rates) to deformation and other control fields (temperature, pressure, strain, anisotropy, etc)
- evolution models (energy conservation, damage diffusion/production/advection, etc) linking various control fields.

A1.2 Comments

Heterogeneous materials involve a material labelling field $m \in F_0$ which may be represented as discrete or continuous. The system under consideration may be open in term of Lagrangian fluxes through the boundaries as a result of coupling with other processes (e.g. erosion/deposition, apparent fluxes coming from Eulerian formulations). Constraints of both linear and nonlinear, unilateral or bilateral nature coexist. The dynamics of the system may encompass a wide range of spatial scales and may include phase transitions or localization of gradients. Time scales vary from some lower limit preset by the coarse description, given in (P) of materials and initial/boundary conditions, to upper limits corresponding e.g. to a full lithospheric rifting episode, or inversion of a sedimentary basin which imply extremely large deformations.

A1.3 Choice and Definition of Model Parameters

We bear in mind that implementing models that duplicate the behaviour of real materials may be extremely difficult owing to the many contributing mechanisms and their complex interactions, as the following examples demonstrate.

i) Porous rocks. The rheological behaviour of porous rocks is a complex function of the confining pressure, p, among other factors. At low p fracture occurs at a low level of elastic deformation by microscale cracking, initiated by stress heterogeneity and microcrack dynamics/coalescence, with changes in bulk volume. At higher confinement, dilatation is inhibited, narrow shear bands appear at an angle with the most compressive stress that depends on p. At still higher p shear banding is replaced by homogeneous microcracking of microsopic brittle nature but macroscopic nonlinear ductile expression. Even in its most basic form, phenomenological modelling of a hardening Drucker Prager yield cone with a non-associative flow rule may require fitting 15 parameters [1], to which should be added those describing the damage mechanisms altering elasticity. In addition, sediment strength increases with decreasing porosity and increasing lithification. Water is involved in this evolution and may contribute through fluid overpressure to slippage along dynamically created decollement surfaces [2].

ii) Salt flow. Deformation of salt involves dislocation creep (cross slip of screw dislocations and climb dislocations) and fluid assisted grain boundary diffusional creep. Observations of domal/antidomal salt microstructures do not, however, demonstrate the dominant deformation mechanism (e.g. as grain size varies by a factor of 10, the viscosity may vary by a factor of 100) [3].

iii) Grain size sensitive creep. In another example of grain size sensitive creep, at stress levels of approximately 300 MPa, deformation microstructures of peridotites show evidence of shear and melting, indicative of dynamic recrystallization and shear localization controlled by grain growth and temperature [4].

iv) Polymineralic assemblages. Rocks may also exhibit complex rheological behaviour when they are polymineralic assemblages of contrasting rheological components, such as frictional-viscous clastomylonites (e.g. feldspar clasts within dynamically recrystallized interconnected layers of quartz) which lubricate fault zones at the brittle-ductile transition [5].

Returning to the question of how to model such complex materials, it may be necessary to view any proposed rheological model in the context of a particular operating environment. That is, the problem should be selected to isolate the behaviour of specific components of the rheological behaviour. Alternatively, we may view systems modelled by (P) as having a synthetic value, similar to the way in which macroscopic fluid properties are recovered through non-equilibrium molecular dynamics simulations from basic nonlinear local binary interactions [6].

In summary, we recognize that the representation of real materials in the form of quantitative physical models is ambiguous and difficult. A vast array of models exist between discrete and continuous end members (e.g. models carried by fractal supports), with bridges going both ways (integration/homogenization). For example, in materials characterized by compact granular texture, the equivalent continuum description may require additional state variables such as the couple stress as in Cosserat models. Molecular dynamics or numerical discretization on the other hand effectively represent continua as a set of interacting particles. Ambiguity in the physical representation of materials often leads to associated ambiguity in the mathematical or numerical representation. Are effects such as the loss of connectivity (e.g during the elastic-viscoplastic transition or melting), and anisotropy development (e.g. during localization) best described by continuum dynamics, granular models, or homogenization by variational formulation of locally nonconvex models into continuum models [7][8]?

APPENDIX 2. Mathematics of the Problem Model

A2.0 Introduction

In this section we define the class of mathematical problems to be solved, outline their well posedness, review items needed to reduce the problem complexity, and outline the practical mapping of the problem on computers. The techniques we refer to have not been developed by our group, unless this is specifically mentioned.

In the past ten years we have used, implemented or tested a number of now classical methods which have resulted in successful nonlinear 2D thermomechanical serial (both structured and unstructured) and parallel (structured) finite element Lagrangian-Eulerian codes. Below we provide some background in modern CFD as a basis for future (particularly 3d) developments.

A2.1 Problem Statement

Most of the mathematical subproblems involved in the solution of problem class (P) are nonlinear constrained optimization problems :

$$(M): \min_{x \in X} F(x)$$

F is typically a functional (Frechet differentiable or not) over some Banach or Hilbert vector space, a subset X of which constitutes the set of admissible states. Spaces may be finite or infinite dimensional. The definition of admissibility may be very complex, although typically local, and corresponds to simultaneous systems of

(possibly nonlinear) equations /inequalities. The solution of (M) imposes conflicting partial goals (e.g. how to split the solution of nonlinear flow motion into components addressing boundary conditions, incompressibility, nonlinear rheology, etc).

The minimization form of the problem may come from natural principles (minimum work) or more ad hoc mathematical manipulation (minimization of residuals). Very importantly, the representation of (M) is far from unique. For example, strong forms of PDE's may lead to various minimization/variational principles depending on the choices of the norm to be minimized and the choice of independent field variables (e.g. minimium work/Hu-Washizu, conformal/mixed, Galerkin/first order least square residual, primal/dual formulations). Each choice of F,X and the meaning given to the minimization operation leads to different mappings of the original subproblem.

2.2 Comments on Well-Posedness of the Problem.

We comment here first on the solid, then the fluid description.
For Prandtl-Reuss and Hencky elastoplastic rheological models, convex dissipation functions are amenable to existence, unicity and stability results using (M) in the form of elliptical variational inequalities of the second kind and convex duality. The extension of these results to large deformation problems is fraught with difficulties regarding, for example non-convex dissipation, and definition of objective rates and strain measures [9].

At another level, the whole concept of a continuum probably needs to be modified using a model in which the elastic fabric/structure of a material is progressively disconnected/broken (and possibly reconnected during healing phases). Such an approach would differ significantly from present continuum models in which elastic stresses are heuristically regridded onto a new Eulerian configuration without consideration of the material limits of elasticity.

On the fluid side, we routinely solve very large deformation viscoplastic flow problems while having no proof of well posedness for this category of problems. Non-associative (incompressible) flow rules are used (implying a deviation from the classical Drucker-Prager postulate) as well as a Coulomb yield function that tends to a singularity at the free surface. We use various heuristic regularization methods (penalty of incompressibility, surface cohesion, L^2 pressure smoothing) which seem to indicate that these very nonlinear problems are indeed well posed.

2.3 Problem Reduction.

In this section we first look at ways to reduce the problem size, then at ways to reduce the theoretical complexity of the problem.

2.3.1. Reducing the problem size N.

Starting with N infinite, various methods can be used to transform the original problem to finite size, or optimum size problems.

- 'Finitization' may come through spatial discretization (using grids or clouds), but also from other truncation methods such as selection of a finite subset of an infinite basis (e.g. first N smallest frequencies or largest N sorted wavelet amplitudes). There are numerous types of cloud or meshless methods (Element Free Galerkin, Partitions of Unity, hp-clouds (h refers to mesh adaptivity, p refers to interpolation order adaptivity), Reproducing Kernel Particle Method, Meshless Local Petrov Galerkin, adaptive wavelet methods)[10].
- The user may be asked which functional, J(x0), of the solution he is interested in computing accurately.

This partial goal is incorporated into (M) through control theory to provide a reduced problem (M-onlyJ). Multiple trade offs are of course possible here but J may be viewed also as a flexible tool to explicitly focus, on parts or all the solution; in particular, exactly the same generic method is used to analyze the a posteriori error as will be explained later [11].

- Subsequent problem size reduction is attained by trying to optimize the number of equations (nodes) needed to solve the problem up to some accuracy using h-adaptivity. (Our problem, with Lagrangian material contrasts makes p-adaptivity more delicate). In heavy (e.g. 3d) computations, sharp (reliable and efficient) error estimators will contribute significantly to decrease N. Superconvergent projectors like the Zienkiewicz-Zhu [12] or Fortin projectors [13] for conformal or mixed elements have been used to build a posteriori estimators [14], and their efficiency of asymptotic scaling computed (under some simplifying hypothesis e.g. regarding the grid shape)[13]. In general cases (e.g. nonlinear plastic rheologies) the dual weighted residual method advocated by Rannacher [11] seems to offer powerful prospects: a tangent Hessian is computed from nonlinear Galerkin truncature equation, then injected into a linear dual problem, à la Aubin-Nitsche, to get an estimator of the global error. This error may be split into local (cell-integrated) components which then, by sorting or error balancing, drives h-refinement. The cost added by this control is amortized over the outer fixed point or Newton nonlinear loop. Here we may ask whether (M) while often being transformed into a sequential series of linearized problems :

  (M1) while (error is too big)
  { create linear problem : Ai. xi = bi;
  solve it}

  may be approached otherwise (nonlinear gradient methods, nonlinear multilevel, defect correction).

Of course h-adaptivity implies a way to locally increase the density of equations. This may be done with grid refinement tools, propagating refinement or keeping and treating specially hanging nodes. Mesh methods [15] are appealing, especially for problems involving material (free or internal) which would be fitted by mesh edges/faces. They also provide computable data access, and may be used to produce spatial partitions. Yet their generation/refinement/coarsening in 3D are very complex and all careful topological and geometrical interface cleaning would seem to be very demanding (e.g. before entering a constrained tetrahedralization). It is therefore interesting to consider alternatives. Some form of Lagrangian or semi-Lagrangian wavelet approach may be superior (especially in 3D) to an ALE-FEM approach. Reproducing Kernel Particle methods [16] offer a practical way to generate pseudo wavelets (which do not satisfy biorthogonality relations) in terms of difference of adjacent scales window functions. This approach gives access to a multiresolution analysis that replaces heuristically the previously mentioned error estimators. In strong contrast there are infinite adaptive wavelets methods. Here wavelet transformation is used to map the original problem into a infinite set of equations, A(x)=b, which system is then solved adaptively by considering larger and larger sections of the infinite vectors. This approach is very elegant but needs as other meshless methods to deal with boundaries and numerical integration. Using Lagrange penalization, fictitious domain methods [17], background grids such as the simplicial algebraic mesher [18] we have developed in 2D, we may keep track of space in some semi-structured manner. We notice (the importance of this feature will be explained below) that the wavelet methods have been used in all sorts of flavors : divergence free wavelets, LBB stable mixed u-p wavelets, infinite adaptive u-p wavelets [19], but the infinite adaptive wavelet approach automatically generates stability, as would be expected from the interpretation of stabilization as dissipating mechanism at the Nyquist frequency (infinite in this case).

2.3.2 Reducing the problem solving complexity.

So far we have not even started to solve the problem, but just considered the aspect of problem size. What complexity do we now expect from the solver handling this problem? This complexity is itself complex to estimate. We want to classify basic operations (e.g. linear algebra operations: mv = matrix-vector product; dp = dotproducts, axpy = a*x+y (a is a scalar, x and y are vectors)) and estimate how many of which will be

performed, and the complexity of each operation. A crude but important measure of complexity is the way the number of floating point operation scales with N : $N_o < C\ N^{a_o}$ (1). The second measure of complexity is the memory space needed; $N_m < C\ N^{a_m}$ (2). Banded Crout direct methods would lead to $a_o=5/3$ and $a_m=7/3$ [20]. Multifrontal methods improve these exponents by providing fill-in reduction equation reordering and architecture-fitting (see section A3) computing technologies. Yet, as proven in [21] Krylov methods like PCR are proven to scale much better ($a_m=1$ $a_o= 3/2$ in 3D), and the operation count per equation even improves from 2D to 3D.

Currently, a large number of authors claim asymptotic scaling $a_o \sim 1$ and $a_m \sim 1$. We postpone to section 3 other aspects of complexity not dealt with by these 2 measures. Constants in 1 and 2 are significantly affected by the choice of discretization, which offers a first way to discriminate between different methods. In table 1, compiled from [22], three mixed elements are compared in terms of their 2D/3D mv floating point count, stencil volumes for both divergence and elliptical operators, taking as unit a proxy for N (number of cells). The table speaks for itself. Unfortunately, low order elements typically need stabilization, which while being seriously studied [21][23] is not always easy to control, and may be detrimental to accuracy and outer Krylov iteration count. In this respect Turek advocates the use of the rotated trilinear element Qt/Q0 because it offers a unique combination of a LBB stable element of minimum complexity (small stencils) that is robust to grid stretching (case of boundary layers) [22][24].

To summarize, asymptotic complexity depends strongly on the choice of outer Krylov method (e.g. MINRES, PCR, PCG, GMRES), the choice of preconditioners, and the choice of element/discretization.

Table 1: Role of Discretization on Storage and Operation Count (compiled from [22])

D	MVF	SA	SB	MVS1	MVS2	R
Q_1/Q_1	54/243	9/27	9/27	38/165	25/125	2/2
Q_1/Q_0	34/129	9/27	4/8	18/51	9/27	4/5
$\tilde{Q}_1/Q_0$	44/135	7/11	4/6	20/45	5/7	9/13

Table legend:

D: discretization choice: Q_1/Q_1: bilinear u-p
Q_1/Q_0: bilinear u, constant p
$\tilde{Q}_1/Q_0$: Turek element

MVF: number of operations involved in a full Stokes matrix vector multiplication
SA: size of u-stencil
SB: size of divergence-stencil
MSV1: number of operations in preconditioned Schur matrix vector multiplication
MSV2: storage demand in assembled forms of preconditioned Schur operators
R: improvement ratios over column 1

[all results in 2D/3D, unit is the number of elements in the mesh]

Krylov methods differ in the type of problem they allow to be solved, their robustness, the need for internal tuning (e.g. subcycles) and (mv, dp, axpy) count per iteration ((1,2,6) e.g. for MINRES). Their efficiency is also (see A3.) strongly datastructure dependent. Minmax properties of Krylov iterations (which we recall are nonlinear methods of solution of linear systems) allow the error reduction per iteration to be related to the spectrum of the solved operator. Preconditioners are used to keep the condition number (of the preconditioned system) uniformly bounded from above, by a constant, as close to 1 as possible. This is achieved in essence by identification of all 'destabilizing mechanisms' and adequate treatment. For example, preconditioners like

multigrid stabilize the effect of grid refinement on the asymptotically singular highest frequencies, and Poincaré-Steklov overlap operators precondition the destabilizing effect of increased domain tearing in domain decomposition methods. Stable (LBB-stable or stabilized) approximations for the mixed u-p Stokes primal formulation may take special forms on stretched or distorted grids if we want steep Krylov iterations. Becker [24] gives examples of different stabilization methods for the Q1/Q1 mixed element, which over- under- or rightly stabilize mesh anisotropy. Preconditioners in the form of diagonal spectrally equivalent (and easier to invert) operators for the Stokes flow solver are all derived today in the same generic form of multilevel pressure Schur complement methods [22]. Global examples are the use of velocity-multigrid combined with discrete laplace ($B.D^{-1}.B^t$ type) approximate pressure Schur complement. Local examples involve e.g. multigrid acceleration combined with 'nodal patch exact local exact solves' such as the Vanka smoother or element by element cholesky multiplicative method.

We speculate that considerable potential exist for a next generation of solvers involving some form of dynamic domain decomposition (DDD). A first conceptual example of DDD would be the optimization of domain partitioning (e.g. in terms of nodal separators) such that the initial solution state is likely to maximize error-reduction say through a FETI[25]-preconditioned Krylov method solving for interface unknowns. A guide here is the recognition of surfaces across which several simultaneous constraints are satisfied (e.g. minimization of normal stress gradients, avoidance of material jumps, etc). Other (weaker) versions of DDD have been used, e.g. to use anisotropy with some profit. Mesh anisotropy can be used to improve accuracy, as mesh anisotropy may best reflect anisotropy in fields in boundary layers. This clearly impacts on the need for similar stabilization in physical situations, whereby for example, anisotropy would accompany strain localization in a narrow boundary layer type zone. Mesh anisotropy could best reflect physical anisotropy in such regions. This remark also sets some limits on the power of triangular and tetrahedral meshing in the sense that they do not provide a natural local cartesian frame. Becker [24] offers to 'stretch' the Vanka local direct solves to nodal patches spanning boundary layers . We think that graph theoretical methods will, at some point (like they have done with elimination graph and use of DAGS in direct solvers), appear to solve (with reasonable expectation) this kind of problem. How data should be assigned and migrate for example in a data parallel implementation would also need to be addressed.

A2.3. Mapping the Mathematical Problem on a Computer.

We now address aspects concerning problem-solver efficiency not accounted for by the type of analysis mentioned so far, namely those which relate to the particular computing architecture and programming model which undertakes the actual computation.

Considering the asymptotic estimation of their complexity, direct methods should not be considered as global solvers in 3D owing to their prohibitive storage and run time cost, even though their modern implementations (PSPASES, PWSSMP [26]) have extremely high (and scalable) throughput. By contrast multigrid based solvers, while having optimum operation count and low storage demand, have high data access complexity on unstructured grids (the only case we consider in 3D) that significantly affect their throughput. Codes like Terra using MG as their core solver for mantle convection problems have been running at 100 Gflops on 1024-node CRAY-T3E with an efficiency of order 10%, very far from the close to optimum throughput of some carefully crafted direct solvers (like PWSSMP on IBM SPs)

Work in progress [27] replacing tree data structures by hash tables is expected to lower the expected waited time to access data. Generally speaking, the art of smart datastructures rules here. They should allow parallel balancing (to this effect e.g. we have implemented the Teng-Miller Thurston [28] randomized algorithm for the partitioning of an unstructured mesh on an array of processors), equilibrating data while minimizing communication volumes. We have started developing code involving a mix of multithreading and message passing for use on clusters of SMPs (such as Dalhousie's SP). Preconditioners that are theoretically less efficient may turn out to be more, or much more, efficient due to the twist of machine architecture and problem

mapping. Global multilevel pressure Schur complement approaches may e.g. be better suited to long vectors/pipelines architectures while local (e.g. Vanka like) versions would make better use of caches, available registers (e.g. 4 daxpy on Power3 SP nodes), and translation lookaside buffer/virtual memory paging, achieving high throughput in a RISC type environment.

REFERENCES

[1] V. A. Lubarda, S. Mastilovic, and J. Knap, 1996. Brittle-ductile transition in porous rocks by cap model, Journal of Engineering Mechanics, v. 122, pp 633-642.

[2] J. Zhang, D. M. Davis, and T-F Wong, sub. The brittle-ductile transition in porous sedimentary rocks: geological implications for accretionary wedge aseismicity, Journal of Structural Geology.

[3] P.E. van Keken, C.J. Spiers, A.P. van den Berg, and E.J. Muyzert, 1993. The effective viscosity of rock salt: implementation of steady-state creep laws in numerical models of salt diapirism, Tectonophysics, v. 225, p. 457-476.

[4] D.H. Jin, S. Kerato, and M. Obata, 1998. Mechanisms of localization in the continental lithosphere: Inference from the deformation microstructure of periodotites from the Ivrea zone Northwestern Italy, Journal of Structural Geology, v. 20, p.195-209.

[5] S. Wissing, and M.R. Handy, 1997. The strength and structure of bimineralic rocks with combined frictional and viscous ("semi-brittle") behaviour. International Conference on Deformation Mechanisms in Nature and Experiments, 1-7-19. March 1997, Switzerland, Abstract and Poster

[6] http://www.asgard.ethz.ch/day/ws0001/pfp_pj6.phtml

[7] A. Braides and M.S. Gelli
SISSA, via Beirut. 4, 34013 Trieste Italy
braides@sissa.it, gelli@sissa.it

Limits of discrete systems without convexity hypothesis

[8] A. Braides, A. Defranceschi and E. Vitali. Relaxation of elastic energies with free discontinuties and constraint on the strain. SISSA functional Analysis Preprints 23/2000/M.

[9] W. Han, and B.D. Reddy, 1995. Computational plasticity, the variational basis and numerical analysis. Reports on Computational Mathematics, v. 68, University of Iowa, Dept. of Mathematics.

[10] S.N. Atluri, and T. Zhu, 1998. A new meshless local Petrov-Galerkin (MLPG) approach in computation mechanics, Computational Mechanics v. 22, p.117-127.

[11] R. Rannacher
rannacher@iwr.uni-heidelberg.de
Error control in finite element computations: an introduction to error estimations and mesh size adaptation. A posteriori error estimation in least square stabilized finite element schemes, Special Issue Advances in Stabilized Methods in Computations Mechanics, Comp. Methods Appl. Mech. Engnrg., 1998

[12] O.C. Zienkiewicz, and J.Z. Zhu, 1987. A simple error estimator and adaptive procedure for practical engineering analysis, Int. J. Numer. Methods Engnrg, v. 24, p. 337-357.

[13] J.H. Brandts, 1994. Superconvergence and a posteriori error estimation for triangular mixed finite elements, Numerische Mathematik Electronic Edition v. 68, p. 311-324.

[14] C. Johnson, and P. Hansbo, 1992. Adaptive finite element methods in computational mechanics, Computer Methods in Applied Mechanics and Engineering v. 101, North-Holland, Amsterdam.

[15] M. Bern, and P. Plassman. Mesh Generation, supported by the Mathematical, Information, and Computational Sciences Division subprogram of the Office of Computational and Technology Research, U.S. Department of Energy, under contract W-31-109-Eng-38.

[16] W.K. Liu, S. Jun, Y. Chen, D.T. Sihling, and W. Hao, 1997. Multiresolution reproducing kernel particle method for computational fluid dynamics. International Journal of Numerical Methods in Fluids, v. 24, p. 1392-1415.

[17] R. Glowinski, A. Rieder, R.O. Wells, and X. Zhou, 1996. A wavelet multigrid preconditioner for Dirichlet boundary-value problems in general domains, Generd Domain, Math. Modelling and Num. Anal. (M^2AN), v. 30, p. 711-729.

[18] J.M. Maubach
jmaubach@vms.cis.pitt.edu
The efficient location of neighbours for locally defined n-simplicial grids.

[19] W. Dahmen, A. Kunoth, and K. Urban, 1996. A wavelet-Galerkin method for the Stokes equations, Computing v. 56, 259-301.

[20] T. Coupez, and S. Marie, 1997. From a direct solver to a parallel iterative solver in 3d forming simulation, International Journal of Supercomputer Applications and High Performance Computing, v. 11, p. 277-285.

[21] A. Wathen, and D. Silvester, 1993. Fast iterative solution of stabilized Stokes systems. Part I: Using simple diagonal preconditioners. SIAM J. Numer. Anal v. 30, p. 630-649.
Fast iterative solution of stabilized Stokes systems. Part II: Using general block preconditioners. Technical Report AM-93-08, University of Bristol, Department of Mathematics, 1993.

[22] Stefan Turek
Institut fur Angewandte Mathematik, Universtat Heidelberg, Im Neuenheimer Feld 294, 69120 Heidelberg, Germany.
ture@gria.iwr.uni-heidelberg.de
http://gria.iwr.uni-heidelberg.de/~ture

Multilevel Pressure Schur Complement techniques for the numerical solution of the incompressible Navier-Stokes equations

Multigrid techniques for simple discretely divergence-free finite element spaces

[23] L.P. Franca
Department of Mathematics, University of Colorado at Denver, P.O. Box 173364, Campus Box 170, Denver, CO 80217-3364 USA
A.Russo
Istituto di Analisi Numerica del CNR, via Abbiategrasso 209, 27100 Pavia Italy

Approximation of the Stokes Problem by Residual-Free Macro Bubbles. East West Journal of Applied Mathematics, v. 4, p. 265-278, 1996.

[24] R. Becker, 1995. An adaptive finite element method for the incompressible Navier Stokes equations on time-dependent domains. http://citeseer.nj.nec.com/becker95adaptive.html

[25] L.F. Pavarino, and O.B. Widlund. Balancing Neuman-Neuman methods for incompressible Stokes equations, February 27, 2001.
http://www.cs.nyu.edu/csweb/Research/TechReports/TR2001-813/TR2001-813.ps.gz

[26] http://www-users.cs.umn.edu/~mjoshi/pspases

[27] M. Griebel, and G. Zumbusch, 1997. Parallel multigrid in an adaptive PDE solver based on hashing and space filling curves, Parallel Computing Fundamentals, Applications and New Directions, Proceedings of the conference ParCo '97, 19-22, September 1997, Bonn, Germany

[28] G.L. Miller, S.H.Teng, William P. Thurston, and S.A.Vavasis, 1998. Geometric separators for finite element meshes, SIAM J. of Scientific Computing v. 19, p. 364-396.