INTERNATIONAL JOURNAL FOR NUMERICAL METHODS IN FLUIDS
Int. J. Numer. Meth. Fluids 2002; **40**:1083–1092 (DOI: 10.1002/fld.394)

Application of the FIDAP code to the 8:1 thermal cavity problem

P. M. Gresho[‡] and S. B. Sutton[*,†]

Lawrence Livermore National Laboratory, Livermore, CA 94550, U.S.A.

SUMMARY

We present results using FIDAP on 3 meshes with 3 different elements ($Q_1Q_0, Q_2P_{-1}, Q_2Q_{-1}$) employed on each. Whereas the bulk of the results were obtained via the 'classical' Galerkin finite element method (GFEM) and the trapezoidal (TR) time integrator, we also tested several alternative options: flux-conservative formulation, energy-conservative formulation, streamline upwinding, mass lumping, backward Euler time stepping, and one additional element (Q_2Q_1). The most accurate of these options was the Q_2Q_{-1} element in the advective formulation, no upwinding, trapezoidal time integration, and consistent mass; i.e. straight GFEM. Copyright © 2002 John Wiley & Sons, Ltd.

KEY WORDS: finite elements; FIDAP; thermal convection; CFD; Boussinesq equations; incompressible flow

1. INTRODUCTION

While performing three-dimensional, time-dependent, laminar thermal convection analyses for the National Ignition Facility (NIF) program at the Lawrence Livermore National Laboratory (LLNL) with the commercial code FIDAP [1], we were contacted by Prof. Bathe to organize a 'special session' for his planned 1st MIT Conference on Computational Fluid and Solid Mechanics, June 2001. After contacting Dr Mark Christon at Sandia National Laboratories to 'coerce' him into volunteering to lead the effort, we complied by 'lifting' a typical two-dimensional cross-section from one of our three-dimensional NIF simulations and choosing similar values of Rayleigh and Prandtl numbers, thus explaining the origin of $Ra = 3.4 \times 10^5$ and $Pr = 0.71$ (actual values used in most of our simulations were 3.407×10^5 and 0.7088, respectively, and we mistakenly rounded them in formulating the MIT conference problem, which will be commented on more fully later). See Reference [2] for the problem description.

*Correspondence to: S. B. Sutton, Lawrence Livermore National Laboratory, Livermore, CA 94550, U.S.A.
†E-mail: sutton4@llnl.gov
‡Retired.

Contract/grant sponsor: U.S. Department of Energy; contract/grant number: W-7405-Eng-48

Received 31 December 2001
Revised 1 July 2002

This contribution thus comes via the rather versatile and general commercial finite element code, FIDAP [1]. This code still offers the user a wide selection with respect to element choices, statement of governing equations, (e.g. advective form, divergence form) implicit time integrators (variable or fixed step, first order or second order), and solution techniques for both the non-linear and linear sets of equations. We have tested quite a number of these variations on this problem which are summarized herein.

2. METHODOLOGY

Most of the results were obtained using the classical 'plane vanilla' (and least expensive) Galerkin finite element method—no tricks, such as stability-enhancing upwind-related modifications to the advection terms—combined with an 'honest' non-dissipative second-order accurate time integrator: trapezoid rule [3]. However, to demonstrate the often-deleterious effects of 'stabilizing' modifications, we present some SUPG (streamline-upwind Petrov–Galerkin) results and one from a highly dissipative and only first-order accurate time integrator: backward Euler (BE). See References [1, 3] for algorithmic details.

In the results to be summarized herein, two types of solvers were employed on the linear systems resulting after the successive substitution (Picard) method was applied to each non-linear algebraic system: (1) After applying the penalty approximation method [3] to eliminate the pressure ($P = -\lambda \nabla \cdot \mathbf{u}$) the *fully-coupled* ($\mathbf{u}, \theta$) systems were solved using an efficient form of Gaussian elimination (skyline method [1]) for Meshes 1 and 2, and for all elements except Q_2Q_1—which must be solved 'more' fully coupled ($\mathbf{u}, P, \theta$) as the penalty method is then not so efficient. (2) For Mesh 3, the segregated solution method was employed to generate an iterative sequence of smaller (uncoupled) linear systems (for u, v, θ and P, as well as one for a Lagrange multiplier), each of which is solved by an iterative method. The symmetric systems (P and the Lagrange multiplier) were solved using the SSOR-preconditioned conjugate residual (CR) method and the unsymmetrical ones (u, v, θ) via conjugate gradient squared (CGS), preconditioned with diagonal (Jacobi) scaling combined with both explicit and implicit relaxation (see References [1, 3]). Convergence criteria employed were as follows: $\varepsilon_N = 10^{-7}$ for the outer (Picard) iterations and $\varepsilon_L = 10^{-4}$ for the linear subsystems. The outer iterations typically converged in 3–5 iterations and the linear subsystems required 2–6 via CGS and 20–80 via CR. Sufficient testing convinced us that the final convergence criterion selected was sufficiently small via both relative error and relative residual (Euclidean) norms; i.e. $\|\Delta x\|/\|x\| < \varepsilon$ and $\|R(x)\|/\|R(x_0)\| < \varepsilon$, where $R(x) \equiv Ax - b$.

3. RESULTS

We present detailed results from 3 elements on 3 grids, all from the advective formulation, TR time integration, and conventional GFEM. The elements used (see Reference [3]) were: (1) Q_1Q_0 (bilinear velocity and temperature, piecewise-constant pressure on quadrilaterals), (2) Q_2P_{-1} (biquadratic velocity and temperature, piecewise-linear pressure) and (3) Q_2Q_{-1}, (same as Q_2P_{-1} except pressure is piecewise-bilinear). Even though the first and third have some (div-) stability problems [3], they produced excellent results and are still quite useful in general. The Q_2P_{-1} (9/3) element, while possibly the most popular (stable) higher-order

Table I. Point 1, wall, and mean data.

		Mesh 1 Grid resolution: 27×121 Steps/period: ~ 25			Mesh 2 Grid resolution: 53×241 Steps/period: ~ 25			Mesh 3 Grid resolution: 105×481 Steps/period: ~ 25		
		Avg.	Amp.	Period	Avg.	Amp.	Period	Avg.	Amp.	Period
u_1	Q_1Q_0	0.05605	0.05852	3.4583	0.05861	0.05644	3.4341	0.05693	0.05564	3.4279
	Q_2P_{-1}	0.05246	0.00542	3.4245	0.05703	0.0553	3.4265	0.05665	0.05548	3.4259
	Q_2Q_{-1}	0.05601	0.05194	3.4285	0.05688	0.05564	3.4261	0.05649	0.05541	3.4259
v_1	Q_1Q_0	0.46189	0.08246	3.4582	0.4651	0.07938	3.4342	0.46263	0.0782	3.4279
	Q_2P_{-1}	0.46409	0.0084	3.4428	0.4631	0.07754	3.4265	0.46251	0.078	3.4259
	Q_2Q_{-1}	0.46233	0.07316	3.4285	0.4627	0.07824	3.4261	0.46180	0.0778	3.4259
θ_1	Q_1Q_0	0.26385	0.04582	3.4582	0.2664	0.04394	3.4341	0.26515	0.04333	3.4279
	Q_2P_{-1}	0.26590	0.00442	3.4429	0.2658	0.04288	3.4265	0.26547	0.04312	3.4259
	Q_2Q_{-1}	0.26590	0.0405	3.4286	0.2651	0.04324	3.4261	0.26572	0.04314	3.4259
ε_{12}	Q_1Q_0	0	—	—	0	—	—	0	—	—
	Q_2P_{-1}	0	—	—	0	—	—	0	—	—
	Q_2Q_{-1}	0	—	—	0	—	—	0	—	—
ψ_1	Q_1Q_0	−0.07293	0.00738	3.4582	−0.07397	0.0072	3.4341	−0.07450	0.00712	3.4276
	Q_2P_{-1}	−0.07337	7E-4	3.4414	−0.07398	0.00706	3.4264	−0.07444	0.00708	3.4259
	Q_2Q_{-1}	−0.07218	0.00622	3.4286	−0.07409	0.0071	3.4261	−0.07439	0.0071	3.4259
ω_1	Q_1Q_0	−2.2379	1.1528	3.4581	−2.3428	1.0798	3.4341	−2.4144	1.0776	3.4279
	Q_2P_{-1}	−2.4106	0.1072	3.4414	−2.4240	1.0862	3.4266	−2.4498	1.0816	3.4259
	Q_2Q_{-1}	−2.2513	1.0264	3.4285	−2.4190	1.093	3.4257	−2.4455	1.081	3.4259
ΔP_{14}	Q_1Q_0	−0.00152	0.02086	3.4582	−0.00193	0.0209	3.4340	−0.00234	0.02057	3.4280
	Q_2P_{-1}	−0.00182	0.00226	3.4412	−0.00125	0.0208	3.4264	−0.00219	0.02068	3.4262
	Q_2Q_{-1}	−0.00135	0.01974	3.4285	−0.00200	0.02096	3.4259	−0.00203	0.02068	3.4259
ΔP_{51}	Q_1Q_0	−0.5337	0.02238	3.4582	−0.5332	0.023	3.4343	−0.5323	0.02292	3.4280
	Q_2P_{-1}	−0.5342	0.00262	3.4414	−0.5348	0.02298	3.4266	−0.5348	0.02292	3.4262
	Q_2Q_{-1}	−0.5360	0.02172	3.4286	−0.5338	0.02314	3.4260	−0.5349	0.02292	3.4260
ΔP_{35}	Q_1Q_0	0.5362	0.0101	3.4581	0.5354	0.01026	3.4341	0.5357	0.0102	3.4283
	Q_2P_{-1}	0.5360	0.0012	3.4422	0.5360	0.01026	3.4266	0.5370	0.01022	3.4261
	Q_2Q_{-1}	0.5373	0.00982	3.4286	0.5358	0.01034	3.4261	0.5370	0.0102	3.4257
$-Nu_{x=0}$	Q_1Q_0	4.5661	0.00776	3.4582	4.5796	0.0074	3.4343	4.5821	0.00726	3.4279
	Q_2P_{-1}	4.6318	8.6E-4	3.4410	4.5893	0.00724	3.4265	4.5825	0.00722	3.4259
	Q_2Q_{-1}	4.6328	0.0074	3.4286	4.5893	0.0073	3.4260	4.5821	0.00722	3.4258
$-Nu_{x=W}$	Q_1Q_0	4.5661	0.00776	3.4582	4.5796	0.0074	3.4343	4.5821	0.00726	3.4263
	Q_2P_{-1}	4.6318	8.78E-4	3.4421	4.5888	0.00724	3.4265	4.5825	0.00722	3.4259
	Q_2Q_{-1}	4.6328	0.0074	3.4286	4.5893	0.0073	3.4261	4.5821	0.00722	3.4258
u	Q_1Q_0	0.2396	4.42E-5	3.4582	0.2396	3.56E-5	3.4335	0.2397	3.46E-5	3.4271
	Q_2P_{-1}	0.2393	5.98E-6	3.4336	0.2397	3.38E-5	3.4265	0.2397	3.42E-5	3.4254
	Q_2Q_{-1}	0.2396	3.64E-5	3.4286	0.2397	3.38E-5	3.4250	0.2397	3.4E-5	3.4271
ω	Q_1Q_0	2.8728	0.0031	3.4582	2.9769	0.0032	3.4342	3.0075	0.00322	3.4280
	Q_2P_{-1}	3.0188	3.8E-4	3.4419	3.0180	0.00322	3.4275	3.0179	0.00322	3.4258
	Q_2Q_{-1}	3.0171	0.00306	3.4285	3.0180	0.00324	3.4260	3.0179	0.00322	3.4259

Table II. Problem timing information.

	Mesh 1	Mesh 2	Mesh 3
Q_1Q_0	Machine: C360 Solver: Gaussian elimination Timing: 1230 μs/node/step Memory: 13.75 Mbytes	Machine: C360 Solver: Gaussian elimination Timing: 2375 μs/node/step Memory: 102.1 Mbytes	Machine: J5000 Solver: Segregated/iterative Timing: 3085 μs/node/step Memory: 26.78 Mbytes
Q_2P_{-1}	Machine: J5000 Solver: Gaussian elimination Timing: 786 μs/node/step Memory: 16.99 Mbytes	Machine: J5000 Solver: Gaussian elimination Timing: 2370 μs/node/step Memory: 118.9 Mbytes	Machine: J5000 Solver: Segregated/iterative Timing: 3747 μs/node/step Memory: 33.53 Mbytes
Q_2P_{-1}	Machine: J5000 Solver: Gaussian elimination Timing: 1000 μs/node/step Memory: 16.99 Mbytes	Machine: J5000 Solver: Gaussian elimination Timing: 2469 μs/node/step Memory: 118.9 Mbytes	Machine: J5000 Solver: Segregated/iterative Timing: 6102 μs/node/step Memory: 33.73 Mbytes

Computer resources are given below:

- Machines: Hewlett-Packard J5000, Hewlett-Packard C360.
- Clock rate: 440 MHz (J5000), 340 MHz (C360).
- Total memory: 2 GBytes (J5000), 1 GByte (C360).
- SPECfp95 rate: 52.3 (J5000), 28.1 (C360).
- Number of processors: On J5000, all problems run in single processor mode.

element (at least when using quadrilaterals), was often slightly less accurate than Q_2Q_{-1}. This may be more important in 3D simulations, where neither of these higher-order elements has been adequately tested/evaluated. Some [3] suspect that Q_2Q_{-1}, even though somewhat unstable, may be the winner in this race—at least if and when the so-called 'pesky modes' are properly dealt with, so as not to slow down the iterative solvers (see Reference [3])—as occurred herein with FIDAP.

The results presented in Tables I and II are self-explanatory, with the possible need to explain one 'outlier': The Q_2P_{-1} element performed poorly (low amplitudes) on Mesh 1, but recovered strongly on Mesh 2. Also, the seemingly large extra cost for Q_2Q_{-1} over Q_2P_{-1} on Mesh 3 is probably related to the slight instability (pesky modes) for the former, combined with the use of iterative solvers; Q_2Q_{-1}required about twice as many iterations for each pressure solve. The calculations addressed in Table I results employed approximately 25 timesteps per oscillation period. When the number of timesteps was doubled, the oscillation amplitude changed by less than 1%. Finally, we restate that parameter values used herein are very slightly different than those requested (for 'historical' reasons); viz., $Ra = 3.407 \times 10^5$ ($\sim 0.2\%$ high) and $Pr = 0.7088$ ($\sim 0.2\%$ low). We have since determined that this inconsistency has less than a 0.1% effect on the amplitude.

Figure 1 gives the time history of the temperature at Point 1 for Q_2Q_{-1} on Mesh 2. Figure 1(a) focuses on the developing time regime, showing the frequency beating during the early stages that gives way to a single frequency. Figure 1(b) shows the single frequency behaviour at later times in the solution. Figure 2 shows the skew-symmetric pattern of temperature variations with respect to the local time average. The dark regions have an instantaneous temperature less than the local mean while in the grey regions it is greater. The

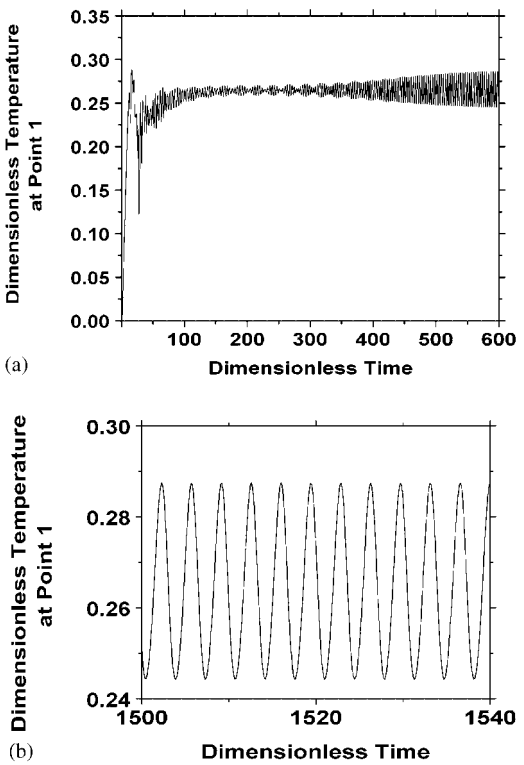

Figure 1. The temperature at Point 1. (a) The early stages of flow development showing the frequency beating. (b) The latter stages of the solution, after stabilization, showing single frequency behaviour.

arrows track a single disturbance 'bubble' over one oscillation period as it propagates up the hot wall.

Having complied with the 'compulsory results' stated in Reference [2], we now move on to some further comparisons, semi-quantitatively, that we believe should be of interest to both FIDAP users and others—in which we briefly address some seemingly ever-present questions like:

(1) Should I use the advective form (e.g. $u \cdot \nabla\theta$) or the flux-conservative divergence form (e.g. $\nabla \cdot (u\theta)$) or even the quadratically conserving (energy) form (e.g. $1/2\,[u \cdot \nabla\theta + \nabla \cdot (u\theta]$)?
(2) When using Q2 (9-node) for velocity, should I use a continuous or discontinuous pressure approximation? If the latter, should it be linear or bilinear?
(3) Should I use *any* 9-node elements or should I stay with the simpler 4-node element? (In 3D, an important 'extrapolation' not considered herein—this question applies to the 8-node (trilinear) brick vs. the various 27-node (triquadratic) bricks).
(4) Should I ever use first-order, implicit Euler (BE) for time integration of time-dependent simulations?

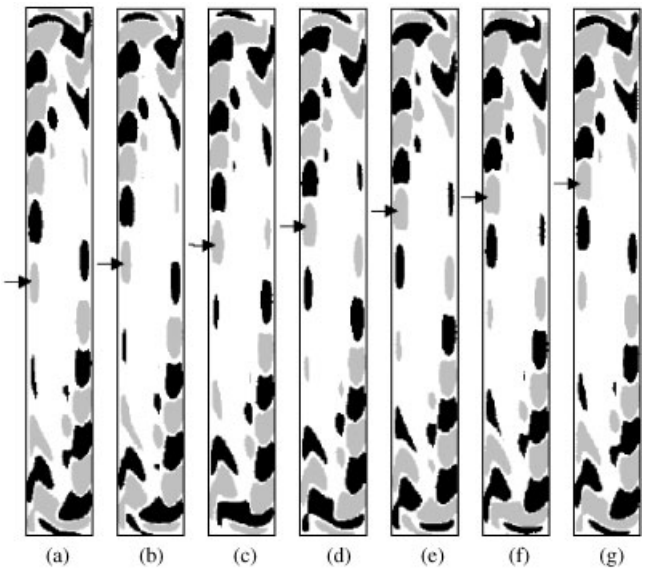

Figure 2. Patterns of the instantaneous temperature variation from the local time averaged mean; time interval between plots is (approximately) 1/6 of one period. In dark regions the temperature is less than the local mean, in grey regions it is greater.

We supply next a small amount of information that might help to answer some of these questions—but will only state our opinion when we are fairly confident. Figures 3–9 summarize our solution errors relative to the 'truth' solution provided in Reference [4]. In these figures the labels 'coarse', 'medium', and 'fine', which are adjacent to the data points, refer to Meshes 1, 2, and 3, respectively (see Table I). Where two labels are seemingly identifying the same point, this merely indicates that the points are nearly coincident and the positioning of the label is used to denote the relative error.

Figure 3 shows, for the Q_1Q_0 element at least, that the simple advective form is more accurate than either 'conservative' form. In the sequel, we discard the energy conservative form on the assumption that the results would carry over, i.e. it will lie between the other two in accuracy.

Two 9-node elements are similarly compared in Figures 4 and 5—with similar results: advective form generally wins. Also suggested is that Q_2Q_{-1} is somewhat more accurate than Q_2P_{-1}—especially on a coarse mesh.

In Figures 6 and 7 we compare all three 9-node elements in advective form, from which we see: (1) Q_2Q_1 (continuous P) is very bad on a coarse mesh (it gives a steady-state solution), but recovers strongly on better meshes; (2) Q_2Q_{-1} is more accurate than the other two, with Q_2P_{-1} not far behind (except on the coarse mesh); (3) the global (x, y) pressure approximation is slightly better than the local (ξ, η) for Q_2P_{-1} (see for example Reference [3, p. 554, 894] for further arguments in favour of global pressure approximation).

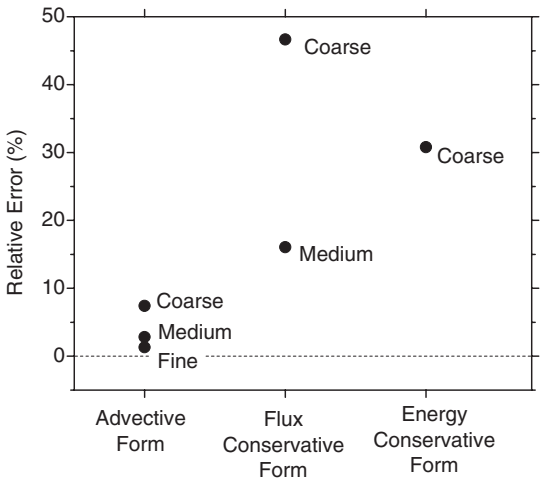

Figure 3. 4-node element (Q_1Q_0) results with different advection formulations.

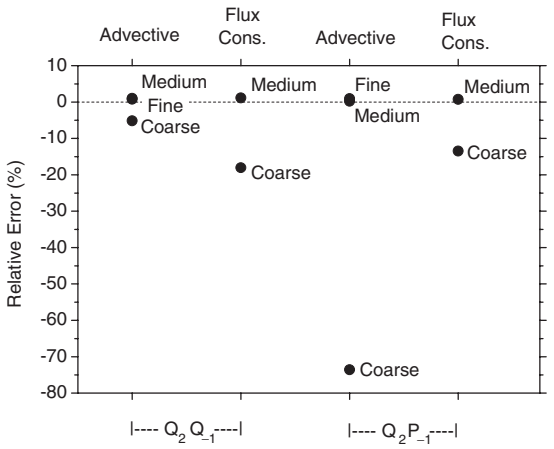

Figure 4. Some 9-node element results with different advection formulations.

Finally, Figures 8 and 9 compare the 4-node element against the three 9-node elements. In general we see improved accuracy with mesh refinement, with the exception of Q_2P_{-1} which, surprisingly, exhibits a larger error on the fine mesh (see also Figure 7). Interestingly, the 'cheaper' Q_1Q_0 provides comparable accuracy to the more expensive quadratic elements on the fine mesh (note that machine differences must be accounted for in comparing some Table II values).

One key result in Figure 8 is this: streamline upwinding (SUPG) via Q_1Q_0 on both coarse and medium meshes failed—went to steady-state. This is hardly surprising when it is realized that the fluid dynamic instability in this problem is in the boundary layer *in the flow direction*—just where streamline upwinding adds the extra damping.

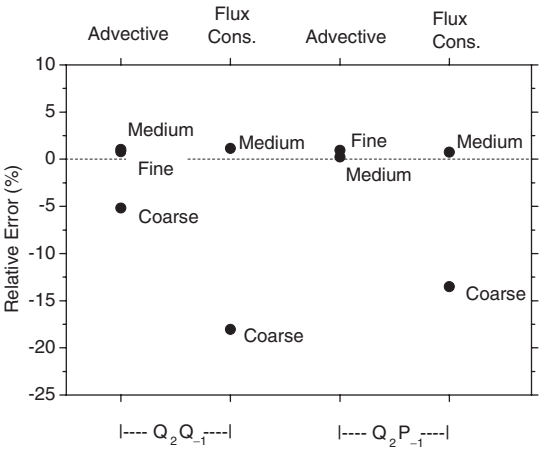

Figure 5. Zoom-in of Figure 4.

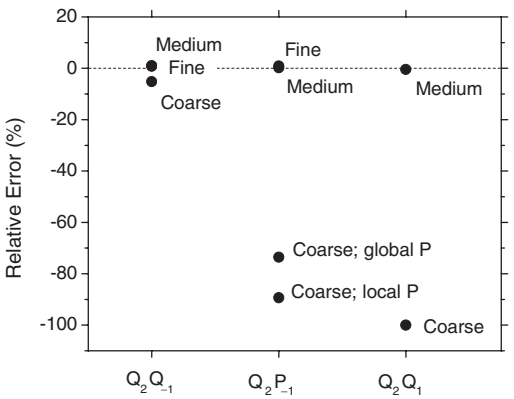

Figure 6. 9-node element results for 3 different pressures (advective form).

Two other runs that failed (went to steady-state) are these: (1) Q_1Q_0 on the coarse mesh when the mass was lumped—caused by poor phase and group velocity accuracy (see Reference [3]), (2) Backward Euler (BE) time integration with both 25 and 50 steps per cycle. At least 100 steps per cycle would be required for this inaccurate and highly dissipative integrator to succeed.

4. SUMMARY AND CONCLUSIONS

Having applied many variations in CFD-methodology via the FIDAP code, we believe that the following conclusions are valid:

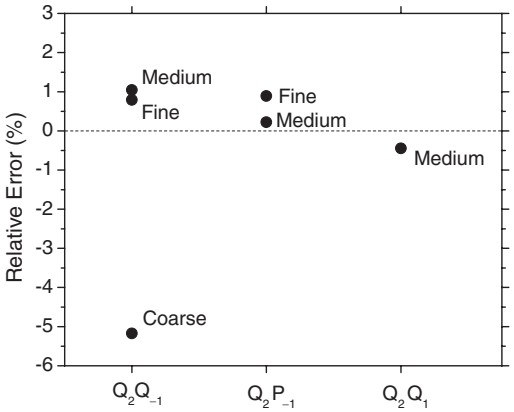

Figure 7. Zoom-in of Figure 6.

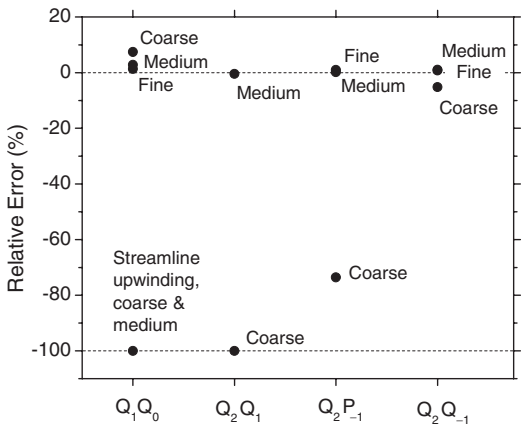

Figure 8. 4-node element vs 9-node elements (advective form).

(1) The 'plane-vanilla' GFEM with 'advective form' and trapezoidal rule for time integration is the best way to solve the time-dependent Boussinesq (and Navier–Stokes) equations.
(2) If the pesky-mode instability could be efficiently dealt with, then the Q_2Q_{-1} element should be employed over the Q_2P_{-1}—especially in 3D (we believe).
(3) The slightly unstable (pesky modes again) Q_1Q_0 element remains surprisingly competitive with the higher-order elements—another result that will (may) be more important in 3D.
(4) Neither the first-order backward Euler time integrator nor mass lumping should be employed if accurate results are desired in truly time-dependent flows.

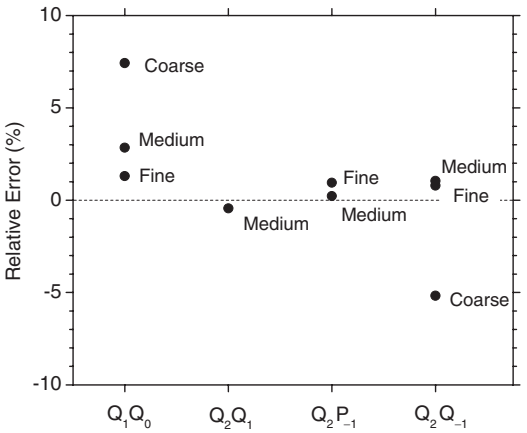

Figure 9. Zoom-in of Figure 8.

ACKNOWLEDGEMENTS

This work was performed under the auspices of the U.S. Department of Energy by the University of California, Lawrence Livermore National Laboratory under Contract No. W-7405-Eng-48.

Note added in proof (PMG): It is interesting to recall that the Q_2Q_{-1} element performed EXTREMELY WELL in an earlier thermal convection 'test problem' (see Reference [5]). Using only 745 nodes in a unit cavity, our results (see Reference [6] for details) were within 1% of the benchmark. That (steady-state) code was written at LLNL.

REFERENCES

1. FIDAP 8.52. *Theory Manual*. Fluent, Inc.: Lebanon, NH, December 1999.
2. Christon MA, Gresho PM, Sutton SB. Computational predictability of natural convection flows in enclosures. *International Journal for Numerical Methods in Fluids*, This Issue.
3. Gresho P, Sani R. *Incompressible Flow and the Finite Element Method*. Wiley: Chichester, 2000.
4. Xin S, Le Quéré P. An extended Chebyshev pseudo-spectral contribution to the CPNCFE benchmark. In *Computational Fluid and Solid Mechanics: Proceedings First M.I.T. Conference on Computational Fluid and Solid Mechanics*, 12–15 June 2001. Elsevier: Amsterdam, 2001; 1509–1513. (See also this issue for the final version of this paper.)
5. de Vahl Davis G, Jones IP. Natural convection in a square cavity: a comparison exercise. *International Journal for Numerical Methods in Fluids* 1983; **3**:227–249.
6. Upson CD, Gresho PM, Lee RL. *Finite Element Simulations of Thermally Induced Convection in an Enclosed Cavity*, Lawrence Livermore National Laboratory *Report UCID-18602*, March 1980.

