Mass Lumping for Constant Density Acoustics

William W. Symes

ABSTRACT

Mass lumping provides an avenue for efficient time-stepping of time-dependent problems with conforming finite element spatial discretization. Typical justifications for mass lumping use quadrature error estimates which do not hold for nonsmooth coefficients. In this paper, I show that the mass-lumped semidiscrete system for the constant-density acoustic wave equation with Q^1 elements exhibits optimal order convergence even when the coefficient (bulk modulus) is merely bounded and measurable, provided that the right-hand side possesses some smoothness in time.

INTRODUCTION

Mass lumping is the process of replacing the mass matrix, occurring in a finite element formulation of a time-dependent or eigenvalue problem, with a diagonal matrix. Such diagonal replacement is essential in time-domain finite element approximation of wave propagation: time steps must be fine enough to resolve the oscillations characteristic of wave problems, and solution of a nontrivial linear system at every time step is prohibitively expensive for large-scale 2D or 3D simulations (the reader can take this expense as a definition of "large-scale"). Cohen (2001) describes many instances of this construction for spectral C^0 element approximation of acoustic, elastic, and electromagnetic waves for example.

According to standard theory for conforming spectral FEM (Cohen, 2001), lumping is possible without loss of accuracy through use of a Gauss-Lobatto quadrature rule with nodes coinciding with those of the nodal conforming basis of the method. This approach to mass-lumping relies for its justification on quadrature error estimates such as that presented by **?**, problem 4.1.5: one obtains an error of order $O(h^k)$ in the mass integral over a regular famility of elements of diameter h for sufficiently regular trial solutions, provided that the quadrature rule is exact for polynomials of degree $2k-1$, *and* the multiplier (mass density or other coefficient) belongs to the space $W^{k,\infty}$. For lowest order $(k=1)$, energy errors of order h follow so long as the quadrature rule is exact for linear functions *and* the coefficient in the mass integral is of class $W^{1,\infty}$. This justification fails when the coefficients (factors in the mass matrix integrand) lack sufficient regularity, for instance exhibit discontinuities along piecewise smooth interfaces.

The purpose of this note is to sketch a different argument for lowest order mass-lumping, which holds for coefficients that are merely bounded and measurable, and which preserves convergence with optimal order $O(h)$ in energy. The essential ingredient is sufficient smoothness of the solution to yield optimal order convergence in energy of the Galerkin approximation. These two characteristics are somewhat in tension: lack of regularity in the coefficients limits the regularity of the solution. In this note I study the simplest special case, the constant-density acoustic wave equation, for which the required degree of regularity is easily established for solutions which are *smooth in time*. Smoothness in time is a natural feature of some problems, such as simulation of seismograms, in which wavefields have fixed temporal bandwidth. Time regularity induces just enough spatial regularity in the constant-density case to give both optimal order convergence in energy and *the same* order of convergence for the mass-lumped approximation.

A by-product of the argument is the observation that with additional coefficient regularity (a global LIpshitz bound, for example), the error in the lumped-matrix approximation is *the same*, to leading order, as in the consistent approximation. In this report, I demonstrate this stricter approximation property only for 1D problems.

The constant density acoustic wave equation is a very special case. Even for the variable density wave equation, coefficient discontinuities generally destroy the regularity underlying the optimal order error estimate, even for lowest-order elements, along with the corresponding convergence rate for the mass-lumped system. It is however possible to choose special elements for which optimal order convergence is restored. For isolated discontinuities along interfaces, this approach is known as *immersed finite elements* ("IFE") (Li and Ito, 2006). Owhadi has recently shown how to construct suitable elements for very general self-adjoint elliptic second order problems, vastly generalizing the IFE method (Ohwadi and Zhang, 2007; Owhadi and Zhang, 2008). It seems possible to base a mass-lumping strategy for more general wave problems on these concepts.

This report begins with a review of the properties of weak solutions of the acoustic wave equation. In the next two sections I establish general properties of the Galerkin approximation, and derive an error estimate for for mass lumping with Q^1 elements on a regular mesh. The more refined estimate for Lipshitz coefficients in 1D is the topic of the next section. I end with a brief discussion of prospects for generalization. An Appendix reviews the standard error estimate for Galerkin methods applied to the wave equation.

WEAK SOLUTIONS

The constant density acoustic wave equation relates the fields $u(\mathbf{x},t)$, $\eta(\mathbf{x})$, and $f(\mathbf{x},t)$ through

$$\eta\frac{\partial^2 u}{\partial t^2} - \nabla^2 u = f. \tag{1}$$

In the acoustic setting, u is excess pressure; η is related to the sound velocity c, bulk modulus κ, and density ρ by $\eta = c^{-2} = \rho/\kappa$. The right-hand side f is either the divergence of a specific body force or the time rate of change of a constitutive law defect.

I assume that u and f are *causal*, that is, vanish for large negative times, and that the wave equation (1) holds in the domain $\Omega \times \mathbf{R}$, $\Omega \subset \mathbf{R}^d$. The field u vanishes on the boundary $\partial\Omega$.

The weak form of (1) is:

$$\int dt\,\left(\left\langle \eta u, \frac{d^2}{dt^2}\phi\right\rangle + a(u,\phi)\right) = \int dt\,\langle f,\phi\rangle,\ \phi\in C_0^2(\mathbf{R}, H_0^1(\Omega)), \tag{2}$$

in which $\langle\cdot,\cdot\rangle$ is the inner product in $L^2(\Omega)$ and $a(\cdot,\cdot)$ is the usual energy form:

$$a(v,w) = \int_\Omega \nabla v\cdot\nabla w \tag{3}$$

Weak (finite energy) causal solutions of (1) are known to exist and be uniquely determined by their data (in other words, f), under very weak hypotheses on η and f: $\log\eta$ should be bounded and measurable, and f should be locally square integrable as a function of time with values in $L^2(\Omega)$. A standard argument using energy estimates and duality (Lions, 1972; Stolk, 2000) establishes the existence of a causal weak solution $u\in C^1(\mathbf{R}, H_0^1(\Omega)$ of (2), satisfying for any $\alpha > 0$

$$e[u](t) \equiv \frac{1}{2}\left(\left\langle \eta\frac{\partial u}{\partial t}(t), \frac{\partial u}{\partial t}(t)\right\rangle + a(u(t),u(t))\right) \le \frac{1}{4\alpha}\int_{-\infty}^{t} d\tau\, e^{\alpha(t-\tau)}\|f(\tau)\|^2. \tag{4}$$

The symbol $\|\cdot\|$ on the RHS of the preceding equation denotes the norm in $L^2(\Omega)$: $\|g\|^2 = \langle g,g\rangle$.

In this generality, very little can be said about the convergence of finite element approximations, except that it takes place: the cited references establish existence of weak solutions precisely by demonstrating convergence of a generic Galerkin approximation. This report exploits the improved approximation possible when f is smoother *in* t. For instance, if $f\in H^k_{\rm loc}(\mathbf{R}, L^2(\Omega))$, then the same energy estimates show that the weak solution possesses additional regularity: $u\in H^{k+1}_{\rm loc}(\mathbf{R}, L^2(\Omega))$. For $k=1$, the strong form of the equation holds and it follows that $u\in C^0(\mathbf{R}, H^2(\Omega))$. The solution is thus smooth enough for to establish the usual optimal order error estimates for P^1 or Q^1 Galerkin approximations.

SEMIDISCRETE GALERKIN METHOD

Choose finite dimensional subspaces $V^h\subset H_0^1(\Omega)$, define Galerkin approximation $u^h\in C^2(\mathbf{R}, V^h)$ by

$$\frac{d^2}{dt^2}\langle\eta u^h,\phi^h\rangle + a(u^h,\phi^h) = \langle f,\phi^h\rangle,\ \phi^h\in V^h, \tag{5}$$

plus appropriate initial conditions.

Enumerate a basis $\{v_j : 1\le j\le N^h\}$ of V^h, write

$$u(\mathbf{x},t) = \sum_{j=1}^{N_h} U_j^h(t)v_j(\mathbf{x}). \tag{6}$$

Define mass, stiffness, and load matrices M^h, K^h, F^h by

$$M^h_{i,j} = \langle \eta v_i, v_j \rangle, \;\; K^h_{i,j} = a(v_i, v_j), \;\; F^h_j(t) = \langle f(\cdot, t), v_j \rangle. \tag{7}$$

Then the Galerkin system (5) is equivalent to the system of ODEs

$$M^h \frac{d^2 U^h}{dt^2} + K^h U^h = F^h \tag{8}$$

I presume that the family of subspaces $V^h \subset H^1_0(\Omega)$ has the standard approximation properties: there exists $C \geq 0, 0 < C_* \leq C^*$ (independent of h) so that

- $h^d C_* I \leq M^h \leq h^d C^* I$, and a similar inequality holds for the matrix M_1 in which η is replaces by 1;
- a similar elementwise bound holds: for any i, j, $0 \leq M^h_{i,j} \leq C^* h^d$;
- for all $g \in H^2(\Omega) \cap H^1_0(\Omega)$, there is $v \in V^h$ for which

  $$\|g - v\|_{H^1_0(\Omega)} \leq Ch \|g\|_{H^2(\Omega)};$$

- for all $v \in V^h$,

  $$\|v\|_{H^2(\Omega)} \leq Ch^{-1} \|v\|_{H^1_0(\Omega)},$$

  and

  $$\|v\|_{H^1_0(\Omega)} \leq Ch^{-1} \|v\|_{L^2(\Omega)}.$$

Q^1 elements on a regular rectangular grid, for example, have these properties REF.

Note that the first of these properties implies that

$$u^h = \sum_i U^h_i v^h_i \;\Rightarrow\; C_* h^d (U^h)^T U^h \leq \|u^h\|_{L^2(\Omega)} \leq C^* h^d (U^h)^T U^h. \tag{9}$$

Also

$$e[u^h] = \frac{1}{2} \left(\left(\frac{dU^h}{dt} \right)^T M \frac{dU^h}{dt} + (U^h)^T K^h U^h \right) \equiv E[U^h]. \tag{10}$$

An argument similar to that used to establish (4) yields

$$E[u^h](t) \leq \frac{1}{4\alpha} \int_{-\infty}^t d\tau \, e^{\alpha(t-\tau)} F(\tau)^T (M^h)^{-1} F(\tau). \tag{11}$$

As remarked at the end of the last section, if $f \in C^1(\mathbf{R}, L^2(\Omega))$ and $f = 0$ for $t << 0$, then the solution u is smooth enough that the usual optimal order error estimate in energy holds:

$$e[u - u^h](t) \leq Ch^2 \int_{-\infty}^t \left\| \frac{\partial f}{\partial t} \right\|^2. \tag{12}$$

See the appendix for a derivation. If f has yet more square-integrable derivatives, corresponding estimates hold for higher time derivatives of $u - u^h$.

MASS LUMPING FOR RECTANGULAR Q^1 ELEMENTS

In this section I suppose that Ω is a rectangle, which after translation could be taken to be

$$\Omega = [0,l_1]\times[0,l_2]\times ... \times[0,l_d].$$

The usual reflection construction makes $H^1_0(\Omega)$ isomorphic to the odd subspace of $H^1_{\rm per}(\mathbf{R}^d)$, where subscript per signifies multiple periodicity with multiperiod $(2l_1,...,2l_d)$. I will use this identification without comment in what follows. Also note that the odd multiperiodic extenstion of $H^1_0(\Omega)\cap H^2(\Omega)$ is a subspace of $H^2_{\rm per}(\mathbf{R}^d)$.

Let $\mathbf{h}=(h_1,...,h_d)$ be the vector of cell side lengths, and $V^{\mathbf{h}}\subset H^1_0(\Omega)$ the corresponding space of Q^1 elements. For a multiindex $\alpha=(\alpha_1,...\alpha_d)$, let v_α be the element of $V^{\mathbf{h}}$ satisfying

$$v_\alpha(\alpha_1h_1,...,\alpha_dh_d)=1;\ v_\alpha(\beta_1h_1,...,\beta_dh_d)=0,\ \alpha\ne\beta\in\mathbf{N}^d.$$

Set $\Delta=\{\beta\in\mathbf{Z}^d:|\beta_i|\le 1, i=1,...,d\}$. Then

$$(M^{\mathbf{h}}U^{\mathbf{h}})_\alpha=\sum_{\beta\in\Delta}M^{\mathbf{h}}_{\alpha,\alpha+\beta}U^{\mathbf{h}}_{\alpha+\beta}=\sum_{\beta\in\Delta}M^{\mathbf{h}}_{\alpha,\alpha+\beta}(U^{\mathbf{h}}_{\alpha+\beta}-U^{\mathbf{h}}_\alpha)+\tilde{M}^{\mathbf{h}}_\alpha U^{\mathbf{h}}_\alpha,$$

in which

$$\tilde{M}^{\mathbf{h}}_\alpha=\sum_{\beta\in\Delta}M^{\mathbf{h}}_{\alpha,\alpha+\beta}$$

is the diagonal (α,α) entry of the usual row-sum lumped mass matrix.

Let $\tilde{U}^{\mathbf{h}}$ be the solution of the mass-lumped Galerkin system

$$\tilde{M}^h\frac{d^2\tilde{U}^{\mathbf{h}}}{dt^2}+K^{\mathbf{h}}\tilde{U}^{\mathbf{h}}=F^{\mathbf{h}}. \tag{13}$$

For any solution of such a system, an energy inequality similar to (11) holds: for any $\alpha>0$,

$$\tilde{E}[u^{\mathbf{h}}](t)\equiv\frac{1}{2}\left(\left(\frac{dU^{\mathbf{h}}}{dt}\right)^T\tilde{M}\frac{dU^{\mathbf{h}}}{dt}+(U^{\mathbf{h}})^TK^{\mathbf{h}}U^{\mathbf{h}}\right)(t)$$

$$\le\frac{1}{4\alpha}\int_{-\infty}^t d\tau\, e^{\alpha(t-\tau)}(F^{\mathbf{h}}(\tau))^T(M^{\mathbf{h}})^{-1}F^{\mathbf{h}}(\tau). \tag{14}$$

A little algebra shows that the difference between $\tilde{U}^{\mathbf{h}}$ and the Galerkin trajectory $U^{\mathbf{h}}$ satisfies

$$\left[\tilde{M}^{\mathbf{h}}\frac{d^2}{dt^2}(U^{\mathbf{h}}-\tilde{U}^{\mathbf{h}})+K^h(U^{\mathbf{h}}-\tilde{U}^{\mathbf{h}})\right]_\alpha=\sum_{\beta\in\Delta}M^{\mathbf{h}}_{\alpha,\alpha+\beta}\left(\frac{d^2U^{\mathbf{h}}_{\alpha+\beta}}{dt^2}-\frac{d^2U^{\mathbf{h}}_\alpha}{dt^2}\right). \tag{15}$$

According to (14),

$$\tilde{E}[U^{\mathbf{h}}-\tilde{U}^{\mathbf{h}}]\le\frac{1}{4\alpha}\int_{-\infty}^t d\tau\, e^{\alpha(t-\tau)}$$

$$\times \left[\sum_{\beta\in\Delta} M^{\mathbf{h}}_{\cdot,\cdot+\beta} \left(\frac{d^2U^{\mathbf{h}}_{\cdot+\beta}}{dt^2} - \frac{d^2U^{\mathbf{h}}_{\cdot}}{dt^2} \right)(\tau) \right]^T (M^{\mathbf{h}})^{-1} \left[\sum_{\beta\in\Delta} M^{\mathbf{h}}_{\cdot,\cdot+\beta} \left(\frac{d^2U^{\mathbf{h}}_{\cdot+\beta}}{dt^2} - \frac{d^2U^{\mathbf{h}}_{\cdot}}{dt^2} \right)(\tau) \right].$$

$$\leq Ch^d \sum_{\beta\in\Delta} \int_{-\infty}^t d\tau \left| \frac{d^2U^{\mathbf{h}}_{\cdot+\beta}}{dt^2} - \frac{d^2U^{\mathbf{h}}_{\cdot}}{dt^2} \right|^2, \tag{16}$$

in which $h = \max(h_1, ..., h_d)$ and C from here on will stand for a constant depending on the coefficients and the geometry of the problem but not on h.

Define the unitary translation operator $T^{\mathbf{h}}_\beta : L^2_{\mathbf{per}}(\mathbf{R}^d) \to L^2_{\mathbf{per}}(\mathbf{R}^d)$ by

$$T^{\mathbf{h}}_\beta u(\mathbf{x}) = u(\mathbf{x} + \mathrm{diag}(\beta)\mathbf{h}).$$

Since $v_{\alpha+\beta} = T^{\mathbf{h}}_\beta v_\alpha$,

$$\sum_\alpha \left(\frac{d^2U^{\mathbf{h}}_{\alpha+\beta}}{dt^2} - \frac{d^2U^{\mathbf{h}}_{\alpha}}{dt^2} \right) v^{\mathbf{h}}_\alpha = \sum_\alpha \frac{d^2U^{\mathbf{h}}_{\alpha}}{dt^2} v^{\mathbf{h}}_{\alpha-\beta} - \frac{d^2U^{\mathbf{h}}_{\alpha}}{dt^2} v^{\mathbf{h}}_\alpha = (T^{\mathbf{h}}_{-\beta} - I)\frac{\partial^2 u^{\mathbf{h}}}{\partial t^2}.$$

Using the bounds on mass matrices described in the last section, obtain

$$\left| \frac{d^2U^{\mathbf{h}}_{\cdot+\beta}}{dt^2} - \frac{d^2U^{\mathbf{h}}_{\cdot}}{dt^2} \right|^2 \leq Ch^{-d} \left\| \sum_\alpha \left(\frac{d^2U^{\mathbf{h}}_{\alpha+\beta}}{dt^2} - \frac{d^2U^{\mathbf{h}}_{\alpha}}{dt^2} \right) v^{\mathbf{h}}_\alpha \right\|^2_{L^2(\Omega)}$$

$$= Ch^{-d} \left\| (T^{\mathbf{h}}_{-\beta} - I)\frac{\partial^2 u^{\mathbf{h}}}{\partial t^2} \right\|^2_{L^2(\Omega)}. \tag{17}$$

Combining (16) and (17) gives

$$\tilde{E}[U^{\mathbf{h}} - \tilde{U}^{\mathbf{h}}] \leq C \sum_{\beta\in\Delta} \int_{-\infty}^t d\tau \left\| (T^{\mathbf{h}}_{-\beta} - I)\frac{\partial^2 u^{\mathbf{h}}}{\partial t^2}(\tau) \right\|^2_{L^2(\Omega)}. \tag{18}$$

Suppose that in addition to the previous hypotheses, $f \in H^3_{\mathrm{loc}}(\mathbf{R}, L^2(\Omega))$. Then (4) applied to the $\partial u/\partial t$ gives

$$\left\| \frac{\partial^2 u^{\mathbf{h}}}{\partial t^2} - \frac{\partial^2 u}{\partial t^2} \right\|^2_{L^2(\Omega)} \leq Ch^2 \int_{-\infty}^t d\tau \left\| \frac{\partial^2 f}{\partial t^2}(\tau) \right\|^2_{L^2(\Omega)},$$

whence

$$\tilde{E}[U^{\mathbf{h}} - \tilde{U}^{\mathbf{h}}] \leq C \sum_{\beta\in\Delta} \int_{-\infty}^t d\tau \left\| (T^{\mathbf{h}}_{-\beta} - I)\frac{\partial^2 u}{\partial t^2}(\tau) \right\|^2_{L^2(\Omega)} + Ch^2 \int_{-\infty}^t d\tau \left\| \frac{\partial^2 f}{\partial t^2}(\tau) \right\|^2_{L^2(\Omega)}$$

As noted earlier, under these hypotheses on f, $\partial^2 u/\partial t^2 \in C^0(\mathbf{R}, H^2(\Omega))$, and even $\in C^0(\mathbf{R}, H^2_{\mathrm{per}}(\mathbf{R}^d)$, with

$$\left\| \frac{\partial^2 u}{\partial t^2}(t) \right\|^2_{H^2(\Omega)} \leq C \int_{-\infty}^t d\tau \left\| \frac{\partial^3 f}{\partial t^3}(\tau) \right\|^2_{L^2(\Omega)}.$$

A straightforward argument based on the trace theorem shows that

$$\left\|(T_{-\beta}-I)\frac{\partial^2 u}{\partial t^2}(t)\right\|^2_{L^2(\Omega)} \le Ch^2\int_{-\infty}^{t} d\tau\,\left\|\frac{\partial^3 f}{\partial t^3}(\tau)\right\|^2_{L^2(\Omega)}.$$

Combining the last few inequalities, obtain

$$e[\tilde{u}^{\mathbf{h}}-u] \le C(\tilde{E}[U^{\mathbf{h}}-\tilde{U}^{\mathbf{h}}]+e[u^{\mathbf{h}}-u]) \le Ch^2\int_{-\infty}^{t} d\tau\,\left\|\frac{\partial^3 f}{\partial t^3}(\tau)\right\|^2_{L^2(\Omega)}. \quad (19)$$

That is, the error in the lumped mass Galerkin approximation is the same order as the error in the consistent mass approximation.

MASS LUMPING FOR P^1 ELEMENTS FOR DIMENSION 1

If the coefficient η has a bit more regularity, then a more precise estimate than (19) is possible. In this section I describe the case $d=1$; it seems likely that similar estimates hold for $d>1$.

In this section, $d=1$, $\Omega=[0,1]$. Define the element mesh

$$x_i = ih, i=0,...N^h+1=\frac{1}{h}. \quad (20)$$

Choose for V^h the piecewise linear functions with nodes $\{x_i\}$, that is, the P1 (or equally well Q1) element space. The usual nodal basis is given by

$$v_j(x) = \begin{cases} \frac{(x-x_{j-1})}{h},\ x_{j-1}\le x\le x_j \\ \frac{(x_{j+1}-x)}{h},\ x_j\le x\le x_{j+1} \\ 0,\ \text{else} \end{cases} \quad (21)$$

Then $K^h_{i,i}=2/h, K^h_{i,i\pm1}=-1/h$, and $K^h_{i,j}=0$ for $|i-j|>1$, with appropriate modifications at the endpoints.

The usual prescription for mass lumping is to replace M^h by the diagonal matrix $\tilde{M}^h$ of its row sums. Note that M^h is also tridiagonal. Thus it is always possible to write

$$M^h_{i,\cdot} = \tilde{M}^h_{i,\cdot} + aK^h_{i,\cdot} + bD^h_{i,\cdot}. \quad (22)$$

D^h is the centered difference matrix, the tridiagonal matrix with

$$D^h_{i,i\pm1}=\frac{\pm1}{2h},\ D^h_{i,i}=0,\ i=2,...N^h-1.$$

A little algebra shows that

$$a=\frac{h}{2}(M^h_{i,i+1}+M^h_{i,i-1})\equiv hP_+,\ b=h(M^h_{i,i+1}-M^h_{i,i-1})\equiv hP_-.$$

The lumped version of the Galerkin ODE system is

$$\tilde{M}^h \frac{d^2\tilde{U}^h}{dt^2} + K^h\tilde{U}^h = F^h. \tag{23}$$

A little more algebra shows that the difference between its solution $\tilde{U}^h$ and the Galerkin trajectory U^h satisfies

$$\tilde{M}^h \frac{d^2}{dt^2}(U^h - \tilde{U}^h) + K^h(U^h - \tilde{U}^h) = hP_+K^h\frac{d^2U^h}{dt^2} + hP_-D^h\frac{d^2U^h}{dt^2}. \tag{24}$$

Since

$$\tilde{M}^h_{i,i} = \int_0^1 \eta v_i,$$

the first of the following two inequalities follows:

$$h\eta_{\min} \le \tilde{M}^h_{i,i} \le h\eta_{\max}\, i = 1, ..., N^h, \tag{25}$$

$$h\eta_{\min} I \le M^h \le h\eta_{\max} I \tag{26}$$

where

$$\eta_{\min} = \inf_{0\le x\le 1}\eta(x),\;\; \eta_{\max} = \sup_{0\le x\le 1}\eta(x).$$

So the *lumped energy form*

$$\tilde{E}^h(U) = \frac{1}{2}\left(\frac{dU}{dt}^T \tilde{M}^h \frac{dU}{dt} + U^T KU\right)$$

is positive definite, as is the consistent energy

$$E^h(U) = \frac{1}{2}\left(\frac{dU}{dt}^T M^h \frac{dU}{dt} + U^T K^h U\right).$$

Note that

$$u = \sum_i U_i v_i^h \;\Rightarrow\; E^h(U) = e[u], \tag{27}$$

and

$$C_* E^h[U] \le \tilde{E}^h[U] \le C^* E^h[U] \tag{28}$$

for suitable positive constants C_* and C^*.

The standard energy estimate for causal solutions of system (8) (see Appendix) is

$$E^h(U^h)(t) \le \frac{1}{4\alpha}\int_{-\infty}^t e^{\alpha(t-\tau)} F^h(\tau)^T (M^h)^{-1} F^h(\tau) \tag{29}$$

for any $\alpha > 0$. A similar estimate holds for solutions of the lumped system; in particular,

$$\tilde{E}^h(U^h - \tilde{U}^h) \le \frac{1}{2\alpha}\int_{-\infty}^t e^{\alpha(t-\tau)} \left[\left(hP_+K^h\frac{d^2U^h}{dt^2}\right)^T (\tilde{M}^h)^{-1}\left(hP_+K^h\frac{d^2U^h}{dt^2}\right)\right.$$

$$+\left(hP_{-}D^h\frac{d^2U^h}{dt^2}\right)^T(\tilde{M}^h)^{-1}\left(hP_{-}D^h\frac{d^2U^h}{dt^2}\right)\Bigg] \tag{30}$$

To estimate the first term on the RHS of (30), note first that the equation satisfied by the 3rd time derivative of U^h,

$$M^h\frac{d^5U^h}{dt^5}+K^h\frac{d^3U^h}{dt^3} = \frac{d^3F^h}{dt^3}, \tag{31}$$

implies via (29) that

$$\left\langle\frac{d^4U^h}{dt^4},M^h\frac{d^4U^h}{dt^4}\right\rangle \le \frac{1}{4\alpha}\int_{-\infty}^t d\tau e^{\alpha(t-\tau)}\left\langle\frac{d^3F^h}{dt^3},(M^h)^{-1}\frac{d^3F^h}{dt^3}\right\rangle. \tag{32}$$

On the other hand, the second time derivative satisfies

$$M^h\frac{d^4U^h}{dt^4}+K^h\frac{d^2U^h}{dt^2} = \frac{d^2F^h}{dt^2}, \tag{33}$$

whence

$$\left\langle\frac{d^4U^h}{dt^4},M^h\frac{d^4U^h}{dt^4}\right\rangle = \left\langle\left(K^h\frac{d^2U^h}{dt^2}-\frac{d^2F^h}{dt^2}\right),(M^h)^{-1}\left(K^h\frac{d^2U^h}{dt^2}-\frac{d^2F^h}{dt^2}\right)\right\rangle$$
$$\ge\frac{1}{2}\left\langle K^h\frac{d^2U^h}{dt^2},(M^h)^{-1}K^h\frac{d^2U^h}{dt^2}\right\rangle-\left\langle\frac{d^2F^h}{dt^2},(M^h)^{-1}\frac{d^2F^h}{dt^2}\right\rangle \tag{34}$$

From this point on, I will use C to denote a constant depending on upper and lower bounds for η and on an upper bound for the time t, but independent of h.

It follows from the matrix bounds (25) that $hP_\pm$ is bounded by Ch^2 in any weighted l^2 norm with diagonal weight matrix. Thus

$$\left\langle hP_+K^h\frac{d^2U^h}{dt^2}(t),(\tilde{M}^h)^{-1}hP_+K^h\frac{d^2U^h}{dt^2}(t)\right\rangle$$
$$\le Ch^4\left\langle K^h\frac{d^2U^h}{dt^2}(t),(\tilde{M}^h)^{-1}K^h\frac{d^2U^h}{dt^2}(t)\right\rangle$$
$$\le Ch^4\left\langle K^h\frac{d^2U^h}{dt^2}(t),(M^h)^{-1}K^h\frac{d^2U^h}{dt^2}(t)\right\rangle$$
$$\le Ch^4\left(\left\langle\frac{d^2F^h}{dt^2}(t)(t),(M^h)^{-1}\frac{d^2F^h}{dt^2}(t)\right\rangle+\int_{-\infty}^t\left\langle\frac{d^3F^h}{dt^3},(M^h)^{-1}\frac{d^3F^h}{dt^3}\right\rangle\right). \tag{35}$$

To bound the second term in (30), note that except in the first and last rows, K satisfies

$$K=-hD^+D^-=-hD^-D^+,\ (D^\pm U)_k=\pm\frac{1}{h}(U_{k\pm1}-U_k),$$

whereas

$$D = \frac{D^+ + D^-}{2}.$$

Also note that

$$(D^\pm)^T = -D^\mp.$$

Denote by Γ^h the operator norm of hP_- in the weighted l^2 norm with weight matrix $(\tilde{M}^h)^{-1}$. According to (25), $\Gamma^h = O(h^2)$ in general. Then

$$\begin{aligned}
&\left(hP_- D^h \frac{d^2U^h}{dt^2}\right)^T (\tilde{M}^h)^{-1} \left(hP_- D^h \frac{d^2U^h}{dt^2}\right) \\
&\le (\Gamma^h)^2 \left((D^+ + D^-)\frac{d^2U^h}{dt^2}\right)^T (\tilde{M}^h)^{-1} \left((D^+ + D^-)\frac{d^2U^h}{dt^2}\right) \\
&\le \frac{2\Gamma^2}{h\eta_{\min}} \left[\left(D^+\frac{d^2U^h}{dt^2}\right)^T \left(D^+\frac{d^2U^h}{dt^2}\right) + \left(D^-\frac{d^2U^h}{dt^2}\right)^T \left(D^-\frac{d^2U^h}{dt^2}\right)\right] \\
&\le \frac{2\Gamma^2}{h\eta_{\min}} \left[\left(\frac{d^2U^h}{dt^2}\right)^T \left(-D^-D^+\frac{d^2U^h}{dt^2}\right) + \left(\frac{d^2U^h}{dt^2}\right)^T \left(-D^+D^-\frac{d^2U^h}{dt^2}\right)\right] \\
&\le \frac{4\Gamma^2}{h^2\eta_{\min}} \left(\frac{d^2U^h}{dt^2}\right)^T K \left(\frac{d^2U^h}{dt^2}\right) \le \frac{8\Gamma^2}{h^2\eta_{\min}} E\left(\frac{d^2U^h}{dt^2}\right). \\
&\le Ch^{-2}\Gamma^2 \int_{-\infty}^t \left(\frac{d^2F}{dt^2}\right)^T (M^h)^{-1} \left(\frac{d^2F}{dt^2}\right). \qquad (36)
\end{aligned}$$

Combine (30), (35), and (36) to yield

$$\tilde{E}^h[U^h - \tilde{U}^h] \le C(h^4 + h^{-2}\Gamma^2) \sum_{i=0}^3 \int_{-\infty}^t \left(\frac{d^iF}{dt^i}\right)^T (M^h)^{-1} \left(\frac{d^iF}{dt^i}\right). \qquad (37)$$

Define $\tilde{u}^h = \sum_i \tilde{U}_i^h v_i^h$; then

$$e[u^h - \tilde{u}^h] = E^h[U^h - \tilde{U}^h] \le C\tilde{E}^h[U^h - \tilde{U}^h] \le C(h^4 + h^{-2}\Gamma^2) \sum_{i=0}^3 \int_{-\infty}^t \left(\frac{d^iF}{dt^i}\right)^T (M^h)^{-1} \left(\frac{d^iF}{dt^i}\right). \qquad (38)$$

On the other hand, the right hand side is

$$\le C(h^4 + h^{-2}\Gamma^2) \sum_{i=0}^3 \int \left\|\frac{\partial^i f}{\partial t^i}\right\|^2 \qquad (39)$$

We distinguish two cases. First, suppose that we know only that $\log\eta$ is bounded and measurable. Then $\Gamma = O(h^2)$, and this is sharp if η has discontinuities. Accordingly, the Galerkin error bound (12) combines with (38) and (39) to yield

$$e[u - \tilde{u}^h] \le Ch^2 \sum_{i=0}^3 \int_{-\infty}^t \left\|\frac{\partial^i f}{\partial t^i}\right\|^2. \qquad (40)$$

Second, suppose that η is Lipshitz continuous. Then $\Gamma = O(h^3)$, and we get instead

$$e[u - \tilde{u}^h] = e[u - u^h] + O(h^4), \tag{41}$$

that is, lumping the mass matrix does not change the leading behaviour of the error at all.

REFERENCES

Ciarlet, P., 2002, The finite element method for elliptic problems: SIAM.

Cohen, G. C., 2001, Higher order numerical methods for transient wave equations: Springer.

Li, Z. and K. Ito, 2006, The immersed interface method: Numerical solutions of pdes involving interfaces and irregular domains: SIAM.

Lions, J., 1972, Nonhomogeneous boundary value problems and applications: Springer Verlag.

Ohwadi, H. and L. Zhang, 2007, Metric based upscaling: Communications in Pure and Applied Mathematics, **60**, 675–723.

Owhadi, H. and L. Zhang, 2008, Homogenization of the acoustic wave equation with a continuum of scales: Computer Methods in Applied Mechanics and Engineering, **198 (3-4)**, 397–406.

Stolk, C., 2000, On the modeling and inversion of seismic data: PhD thesis, Universiteit Utrecht.

.