ADAPTIVE TIMESTEPPING FOR ELASTOHYDRODYNAMIC LUBRICATION SOLVERS[*]

C. E. GOODYER[†] AND M. BERZINS[‡]

Abstract. The application of variable-step time-integration techniques in a method of lines approach for the numerical solution of transient elastohydrodynamic lubrication calculations is described. Spatial discretization of the coupled integro-differential equations leads to a system of differential-algebraic equations in time. At each timestep a large system of nonlinear equations of 10^4–10^7 variables is solved by using multigrid methods and multilevel multi-integration. The application of temporal error control methods is considered with reference to both the differential-algebraic formulation of the system to be solved and the multigrid-based nonlinear equation solver to be used. The error control method used takes account of the magnitude of the spatial error present by using a local time error control based on the magnitude of this error. This approach is shown to be beneficial in reducing the computational work required. Experimental results on a variety of lubrication problems, including a challenging rough surface problem, are used to display the effectiveness of the methods.

Key words. elastohydrodynamic lubrication, method of lines, differential-algebraic equations, variable timestepping, error balancing

AMS subject classifications. 35Q35, 65M20, 65M15, 65M55

DOI. 10.1137/050622092

1. Introduction. Elastohydrodynamic lubrication (EHL) problems arise in the mathematical modeling of highly loaded components undergoing deformation when lubricated by fluids such as oil. In particular, EHL may occur in journal bearings and gears where a load is applied on a very small contact area. This load causes pressures of up to 3 GPa and hence elastic deformation of the solid metal contact. Under these extreme conditions lubricating oil may have a glass-like consistency. EHL problems are most often modeled by nonlinear coupled integro-differential equations which require the use of numerical methods. Such EHL problems are of practical industrial interest but as the demand is for fast and robust solvers for increasingly more complicated problems, the ability to reduce the work needed for individual problems is thus paramount. A thorough background of EHL models and computations is provided by [38].

Numerical solutions have been computed since Petrusevich in 1951 [28], and finite difference, and to a lesser extent finite elements (see, e.g., [41]), have been used since then. The biggest advances have come from Lubrecht et al. with the application of multigrid in 1986 [24, 25, 36, 38] and multilevel multi-integration in 1990 [2, 38]. Since then the emphasis has largely been on solving more realistic and hence more physically challenging problems rather than further algorithmic developments. Most of the results are still computed by using a fixed regular grid and a fixed timestep with multilevel methods.

The need to move EHL simulations closer to the design process for both oils and components requires ever more complex problems to be formulated and solved;
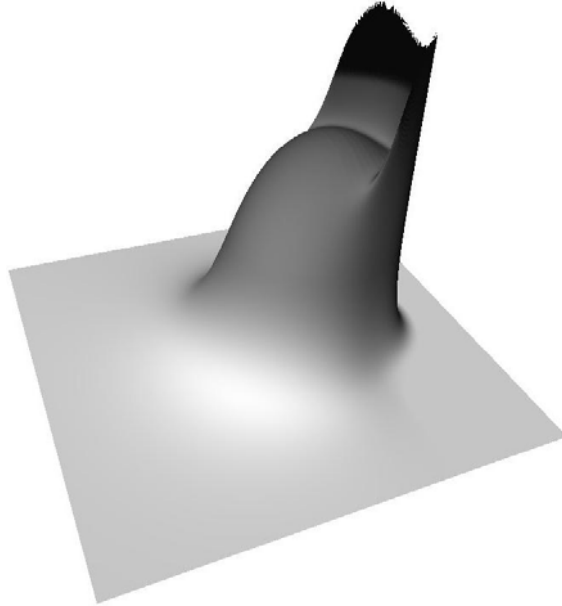
FIG. 1.1. *Typical solution for pressure across an EHL point contact.*

see, e.g., [14]. These problems vary from transient cases, with variable loading and contact speeds, through complicated rheological behavior, to solutions incorporating surface features and, ultimately, true roughness. This requires that the solvers being used must be designed to be extensible to meshes of tens or hundreds of thousands of points in each direction and that the solutions can be obtained sufficiently quickly and reliably for problems with the order of 10^7–10^8 variables. Part of the need for large meshes is that, even in one dimension, mesh spacings as small as 10^{-5} are needed to obtain converged solution values of a pressure spike [17]. The industrial requirement is to be able to make accurate predictions of friction and for this in the case of measured surface roughness large meshes are needed.

EHL problems possess three very different characteristics in different areas of the domain. The cavitation region is the area of the solution beyond the free boundary where the Reynolds equation is not valid. In the center of the domain is the *contact area*, where the pressure rises sharply to reach its maximum peak in a near Hertzian shape. EHL pressure profiles differ significantly from those in hydrodynamic cases with no surface deformation in that there is a ridge on the outflow side of the pressure peak. This can be seen in Figure 1.1 where a three-dimensional representation of the pressure profile is shown. Finally, in the noncontact region the pressure is very small compared to the contact region.

The deformation of the surface can be seen in Figure 1.2 which shows the shape of the contact along the centerline. The deformation of the original surface geometry is clearly visible. It can be seen that there is a constriction in the contact toward the outflow, which coincides with a position between the pressure spike and the cavitation boundary. At the outflow boundary a free boundary is formed representing the presence of a cavitation region where bubbles have formed in the lubricant film. The cases considered here always use a thin film approximation and are for the two-dimensional point contact case. In this work we shall restrict our discussion to the time dependent
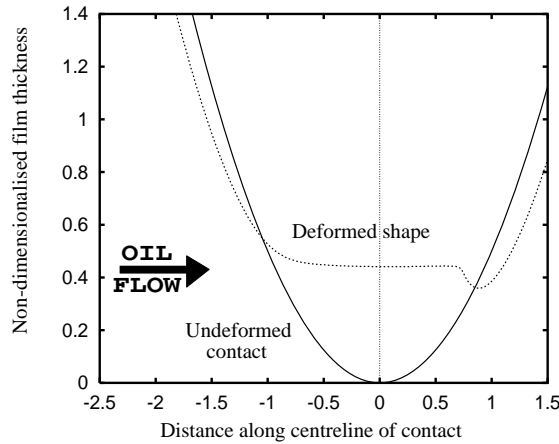
FIG. 1.2. *Typical solution for film thickness along the centerline of an EHL point contact.*

circular point contact cases and consider how to vary the timestep within the context of the multilevel multigrid solver, taking account of the spatial error present. The first results for adaptive timestepping in EHL solutions were presented by Goodyer et al. [13, 16]. More recently Watremetz et al. [39] have presented an alternative method of timestepping based on considering the physical situation rather than any measure of the truncation error.

As larger and larger EHL problems are solved or greater spatial accuracy is used it becomes more important to use a timestepping approach that reflects the spatial error present. This is particularly true for problems with very large numbers of points.

For example, in a benchmark parallel EHL calculation 250 million mesh points have been used by extending the techniques used in [18]. In this paper we thus look at the issue of developing an appropriate timestepping solver and error control method. In resolving this issue it is important to cast the discretized equations as a differential-algebraic system. This allows error estimation procedures from differential-algebraic equation (DAE) theory to be combined with both the multilevel multigrid solver and an estimate of the spatial error to obtain an efficient timesteping scheme. DAE theory also provides insight regarding which variables should be used in a local error test and which should be excluded. The DAE framework also makes it easier to understand how to choose the appropriate classes of time-integration methods and thus makes it easier to understand how to extend the approach described below to use high or variable order methods. In developing the approach described below we are aware of the tremendous advances made in the global error control of ODEs (see, e.g., [5, 9, 23, 26]) and think that such methods offer tremendous potential for the future. At present such methods are still sufficiently expensive that we will explore simpler methods based on attempting to keep the temporal error on a smaller scale than the spatial error [1, 22].

The rest of this paper is laid out as follows. In section 3 the equation system describing the problem is explained and the problem is reformulated as a DAE system. In section 4 a brief explanation of the spatial discretization and time-integration errors is supplied while section 5 describes the nonlinear equations solver employed. Section 6 describes the general approach we are using for the variable timestep approach, based on considering the DAE nature of the system. Section 7 describes in detail the error control mechanism employed in controlling the temporal error given the spatial error

present. Various examples of the use of variable timestepping are shown in section 8 where results for three very different cases, typical of the situations found in real transient EHL contacts, are presented, including consideration of the reduction in solution times.

2. Notation. The following notation of EHL variables will be used throughout the rest of the paper.

$\mathcal{A}$	surface roughness amplitude	a	halfwidth of Hertzian contact
$\mathcal{G}$	undeformed surface geometry	H	nondimensional film thickness
H_{00}	film thickness central offset	K	film thickness kernel matrix
P	nondimensional pressure	p_h	maximum Hertzian pressure
p_0	viscosity parameter	R_x	reduced radius of curvature
T	dimensionless time	u_s	sum of velocities of contacts
$\mathcal{W}$	surface roughness wavelength	X_d	surface roughness pattern center
X, Y	dimensionless coordinates	z	viscosity index
α	pressure viscosity index	ϵ	coefficient in Reynolds eqn.
λ	coefficient in Reynolds eqn.	η_0	viscosity at ambient pressure
$\bar{\eta}$	nondimensional viscosity	Φ	variable load factor
ψ	convergence test parameter	ρ_0	density at ambient pressure
$\bar{\rho}$	nondimensional density		

3. PDE system. The equations governing the EHL point contact are given, in nondimensional form, by the following six equations which are derived in standard texts such as [12, Chapter 4] and [38, Chapters 5 and 6]. The first is the Reynolds equation, which is based on fluid mass conservation and a thin film fluid flow approximation:

$$(3.1) \quad \frac{\partial}{\partial X}\left(\frac{\bar{\rho} H^3}{\bar{\eta}\lambda u_s(0)}\frac{\partial P}{\partial X}\right) + \frac{\partial}{\partial Y}\left(\frac{\bar{\rho} H^3}{\bar{\eta}\lambda u_s(0)}\frac{\partial P}{\partial Y}\right) - \frac{u_s(T)}{u_s(0)}\frac{\partial\left(\bar{\rho} H\right)}{\partial X} - \frac{\partial\left(\bar{\rho} H\right)}{\partial T} = 0,$$

where λ is given by

$$(3.2) \quad \lambda = \frac{6\,\eta_0\,R_x^2}{a^3\,p_h},$$

and where the other symbols are defined in section 2. The two rightmost terms of (3.1) are known as the wedge and squeeze terms, respectively. The boundary conditions are that the pressure is zero at all boundaries and at the outflow of the free boundary there are zero pressure derivatives; see [12, Chapter 11].

Second, the film thickness equation,

$$(3.3) \quad H(X,Y) = H_{00} + \mathcal{G}(X,Y,T) + \frac{2}{\pi^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{P(X',Y')\,\mathrm{d}X'\,\mathrm{d}Y'}{\sqrt{(X-X')^2+(Y-Y')^2}},$$

may be used to compute the contact shape for a known applied pressure distribution. This equation consists of three parts: the offset film thickness or mutual approach, H_{00}, which is the separation of the solids assuming zero deformation; given undeformed geometry $\mathcal{G}(X,Y,T)$, which is generally parabolic, although will also include any roughness features on the surfaces; and the deformation term. The deformation integral is derived by assuming linear, elastic deformation of the two semi-infinite half spaces.

Finally, the force balance equation,

$$(3.4) \qquad \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} P(X,Y)\, \mathrm{d}X\, \mathrm{d}Y = \frac{2\pi}{3},$$

is the constraint equation through which the applied force is implemented in non-dimensional form, as in (1.36)–(1.38) of [38]. By satisfying this conservation equation, the applied load of the case in question is guaranteed to be the calculated total pressure across the numerical solution.

The role of the force balance equation is quite complex in that without this extra equation there is one more degree of freedom than there are equations. The unknown variable is H_{00} which does not, however, appear directly in (3.4). A standard approach in EHL calculations (see, e.g., [38, equation (6.53)]) is to define an iteration for H_{00} using the residual of the discrete approximation to (3.4). This iteration is defined in (5.1) below, for example. A full explanation of how this iteration may be justified is given below in section 5. A brief explanation is that the discrete pressure integral in a discrete version of (3.4) is identical to an integral involving H_{00} and the film thicknesses. This identity may be seen by using the inverse of the discrete film thickness and pressure relationship defined by the discrete version of (3.3); see section 5 below.

A slightly modified form of this equation will be used in section 8.3 in conjunction with time varying loads.

The lubricant model used is that of a generalized Newtonian fluid. The viscosity model is the Roelands equation [29],

$$(3.5) \qquad \overline{\eta}(P) = \exp\left\{ \frac{\alpha\, p_0}{z} \left[-1 + \left(1 + \frac{P p_h}{p_0} \right)^z \right] \right\}.$$

The density is defined using the Dowson and Higginson relation [6]:

$$(3.6) \qquad \overline{\rho}(P) = \left(1 + \frac{5.8 \times 10^{-10}\, P p_h}{1 + 1.7 \times 10^{-9}\, P p_h} \right).$$

It is worth remarking that the theoretical properties of the solution to these equations are not well understood. One notable exception is the finite element analysis of Wu and Oden [41, 42, 43] who considered lightly loaded EHL cases. One theoretical issue, addressed by Stenger, for example [33, 34], is that the film thickness $H(x,y)$ is uniquely defined by the pressure $P(x,y)$ via (3.3).

These equations are also precursors of equations involving rough surface geometries and/or non-Newtonian thermal effects; see, e.g., [10, 17].

3.1. Example reversal of entrainment. A typical example of a time dependent EHL problem is the test case of reversal of entrainment described in [30]. This situation corresponds to flow in one direction changing to the corresponding flow in the opposite direction over a short time period. This change is represented as a linear decrease in velocity. For the example from [30] it corresponds to oil entrainment changing from $0.05\mathrm{ms}^{-1}$ to $-0.05\mathrm{ms}^{-1}$ in 0.2s.

A key feature of this example is the saucer of viscous fluid that forms at the point of reversal (0.1s) and proceeds across the domain (toward the new outflow) before the deformation pattern readopts its characteristic horseshoe shape. All physical parameters are as given in [30], with the standard nondimensionalized quantities as given in Table 3.1. Figure 3.1 shows the central and minimum film thicknesses during this

TABLE 3.1
Nondimensionalized parameters for reversal example.

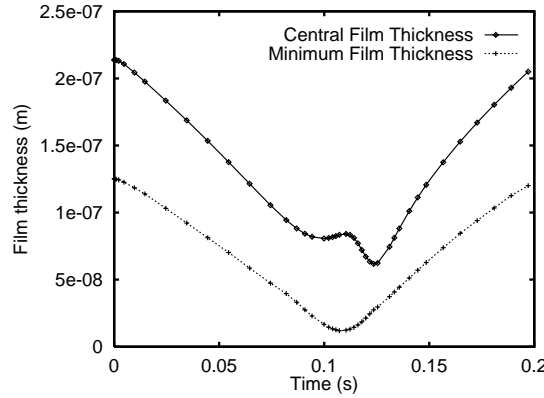| Parameter	Value
Viscosity index, α	2.1×10^{-8} Pa^{-1}
Maximum Hertzian pressure, p_h	0.45GPa
Initial sum of rolling speeds, $u_s(0)$	0.1ms^{-1}
Nondimensional coefficient, λ	0.056
Material parameter, G	2961
Load parameter, W	6.58×10^8
Speed parameter, U	1.47×10^9
Moes parameter, M	52.2
Moes parameter, L	6.9

FIG. 3.1. *Central and minimum film thicknesses for reversal of entrainment.*

example. The points on the line indicate where timesteps have been taken. This test case will be used to demonstrate the issues that arise during the development of the timestepping algorithms. Comparisons of numerical solutions against experimental results are given by Goodyer [13].

4. Spatial and temporal discretization. Spatial discretization of the EHL system defined by (3.1)–(3.6) leads to a system of DAEs [3]. Defining the vector of film thicknesses across the whole domain, $\underline{H}$, and its multiple by the density $\bar{\rho}$ in a pointwise manner by

$$(4.1) \qquad [\underline{H}]_k = H_{i,j} \text{ for } k = i + (j-1) \times N_X,$$
$$\left[\underline{\rho H}\right]_k = \rho_{i,j} H_{i,j}, \qquad i = 1, \ldots, N_X,$$
$$j = 1, \ldots, N_Y,$$

with the vector of pressures, $\underline{P}$, likewise, then the first order upwind finite difference in space, implicit Euler in time, discretization of the Reynolds equation (3.1) is given by

$$\frac{(\epsilon_{i-1,j}^n + \epsilon_{i,j}^n)(P_{i-1,j}^n - P_{i,j}^n) + (\epsilon_{i+1,j}^n + \epsilon_{i,j}^n)(P_{i+1,j}^n - P_{i,j}^n)}{2(\Delta X)^2}$$
$$+ \frac{(\epsilon_{i,j-1}^n + \epsilon_{i,j}^n)(P_{i,j-1}^n - P_{i,j}^n) + (\epsilon_{i,j+1}^n + \epsilon_{i,j}^n)(P_{i,j+1}^n - P_{i,j}^n)}{2(\Delta Y)^2}$$

$$(4.2) \qquad -\frac{u_s(T)}{u_{ref}} \frac{\rho_{i,j}^n H_{i,j}^n - \rho_{i-1,j}^n H_{i-1,j}^n}{\Delta X} - \frac{\rho_{i,j}^n H_{i,j}^n - \rho_{i,j}^{n-1} H_{i,j}^{n-1}}{\Delta T} = 0,$$

where the superscript n denotes values at time t_n. The boundary conditions are that all exterior boundaries have $P = 0$ and the line $j = 1$ is a symmetry condition in the Y direction. It is important to remark that once the entrainment direction has been reversed, the upwind direction of the discretization of the wedge term $\frac{\partial(\overline{\rho}H)}{\partial X}$ needs to be changed to reflect the changing flow direction.

The discretized forms of (3.3)–(3.6) are

$$(4.3) \qquad H_{i,j} = H_{00} + \mathcal{G}_{i,j} + \sum_{k=1}^{N_X} \sum_{l=1}^{N_Y} K_{i,j,k,l} P_{k,l},$$

$$(4.4) \qquad \frac{2\pi}{3} = \Delta X \Delta Y \sum_{i=1}^{N_X} \sum_{j=1}^{N_Y} P_{i,j},$$

$$(4.5) \qquad \overline{\rho}_{i,j} = \frac{0.59 \times 10^9 + 1.34 p_h P_{i,j}}{0.59 \times 10^9 + p_h P_{i,j}},$$

$$(4.6) \qquad \text{and} \quad \overline{\eta}_{i,j} = e^{\left\{ \frac{\alpha p_0}{z} \left[-1 + \left(1 + \frac{p_h P_{i,j}}{p_0} \right)^z \right] \right\}},$$

with all symbols as defined previously.

It is important to stress that the dependence of the film thickness $H_{i,j}$ at any point on the pressure values at every mesh point $P_{k,l}$ means that the sparsity pattern of the discrete nonlinear equations is dense in every case. When this dependence is combined with the nonlinear nature of the viscosity and density the result is a computationally challenging problem.

In (4.2) a first order upwind discretization of both first order derivatives is shown. Alternative schemes have been used by Venner and Lubrecht [38] and Wijnant [40]. Both methods again require the leading coefficients of both the wedge and squeeze terms to be one. Optimal use of these second order schemes rely on having a fixed relationship between the spatial and temporal timestep sizes, which, in an adaptive time-integration code, is no longer the case. These second order schemes are, in our experience, somewhat less stable when applied across a wide variety of cases.

4.1. DAE formulation. In order to apply DAE algorithms, theory, and error indicators it is helpful to recast the discretized EHL equations as DAEs. This is important as the form of the DAEs makes a significant difference to the algorithms that can be employed both for solution and for error estimation.

The differential-algebraic form of the discretized EHL equations may be written as

$$(4.7) \qquad \underline{F}_1(\underline{P}, \overline{\rho}\underline{H}) - \left[\overline{\rho}\underline{H} \right]' = 0,$$

the film thickness equation, (3.3), by

$$(4.8) \qquad \underline{F}_2(\underline{P}, \underline{H}) = 0,$$

and the density equation, (3.6), by

$$(4.9) \qquad \underline{F}_3(\underline{P}, \overline{\rho}) = 0,$$

where $'$ denotes differentiation in time. These can then be combined to define a DAE system for $\underline{U}^\mathrm{T} = (\underline{P}^\mathrm{T}, \underline{H}^\mathrm{T})$ by

$$\text{(4.10)} \qquad \underline{F}(\underline{U}, \underline{U}', t) = 0.$$

In the above case the DAE aspect arises from the lack of a temporal derivative of pressure in (4.7)–(4.9). Issues affecting the differences between solving ODE and DAE systems are discussed in detail in [3]. The authors have computational evidence that the EHL system is index one for regular meshes but need to show that $\frac{\partial \underline{F}_2}{\partial \underline{P}}$ is nonsingular in general.

The index of a DAE is defined as the minimum number of times that all or part of (4.10) must be differentiated with respect to T in order to determine U' as a continuous function of U and T [3]. Thus to find the index of the EHL DAE system given by (4.10), it is necessary to first differentiate (4.8) with respect to T to get

$$\text{(4.11)} \qquad \frac{\partial \underline{F}_2}{\partial \underline{P}} \frac{\mathrm{d}P}{\mathrm{d}T} + \frac{\partial \underline{F}_2}{\partial \underline{H}} \frac{\mathrm{d}H}{\mathrm{d}T} = 0.$$

Hence, if $(\frac{\partial \underline{F}_2}{\partial \underline{H}})^{-1}$ exists, this gives an expression for $\frac{\mathrm{d}H}{\mathrm{d}T}$. Similarly, differentiating (4.9) gives

$$\text{(4.12)} \qquad \frac{\partial \underline{F}_3}{\partial \underline{P}} \frac{\mathrm{d}P}{\mathrm{d}T} + \frac{\partial \underline{F}_3}{\partial \underline{\rho}} \frac{\mathrm{d}\overline{\rho}}{\mathrm{d}T} = 0,$$

which gives an explicit expression for $\frac{\mathrm{d}\overline{\rho}}{\mathrm{d}T}$ if $\frac{\partial \underline{F}_3}{\partial \underline{\rho}}$ is nonsingular.

Note that (4.7) may also be written as

$$\text{(4.13)} \qquad \frac{\mathrm{d}\overline{\rho}H}{\mathrm{d}T} = \underline{H}\frac{\mathrm{d}\overline{\rho}}{\mathrm{d}T} + \overline{\rho}\frac{\mathrm{d}H}{\mathrm{d}T} = \underline{F}_1(\underline{P}, \overline{\rho}H);$$

then expressions for P', H', and $\overline{\rho}'$ have been obtained in one differentiation step. Hence, the index of the transient EHL problem is one, provided both the aforementioned inverses exist.

The calculation of the density, $\overline{\rho}$, is entirely local, leading to a diagonal Jacobian matrix, $\frac{\partial \underline{F}_3}{\partial \underline{P}} \frac{\mathrm{d}P}{\mathrm{d}T}$. This matrix is nonsingular provided that none of the diagonal entries given by

$$\text{(4.14)} \qquad \frac{\partial \overline{\rho}_{i,j}}{\partial P_{i,j}} = \frac{\partial}{\partial P_{i,j}} \left(\frac{0.59 \times 10^9 + 1.34 p_h P_{i,j}}{0.59 \times 10^9 + p_h P_{i,j}} \right)$$

$$\text{(4.15)} \qquad = \frac{0.2006 \times 10^9 p_h}{(0.59 \times 10^9 + p_h P_{i,j})^2}$$

are nonzero. For physical, nonnegative pressures, P, this matrix is invertible.

The invertibility of $\frac{\partial \underline{F}_2}{\partial P}$ is much harder to establish. This is because the kernel matrix K is dense, and hence simple analysis is not possible. A computational investigation into the properties of this matrix showed that it was numerically nonsingular. Unfortunately, due to the very large size of the matrices involved, it was only possible to numerically investigate cases up to grid size 33×33, which remains far short of the fine mesh cases, with more than 1000×1000 points.

5. Nonlinear equations solution method. The methods used for solving for the pressure, the film thickness, and the lubricant properties are now described in turn. In order to obtain solutions to EHL problems quickly it has proved necessary [36, 38] to use multilevel and multigrid methods. Standard texts such as Briggs [4] and Trottenberg, Oosterlee, and Schüller [35] provide descriptions of multigrid methods.

In solving the Reynolds equation it is important to make use of different numerical schemes to reflect the different physical situation in each of the three different regions [13]. In the noncontact region a Gauss–Seidel line relaxation scheme is used; in the contact region a Jacobi line scheme is employed (see Nurgat, Berzins, and Scales [27]) rather than the distributed scheme of Venner [36]. In the cavitation region the Christopherson approach is used [7] to set all calculated negative pressures to zero.

The film thickness calculation, (3.3) once discretized, has the form given in (4.3). Hence for every mesh point, (i, j), the deformation term is a multi-summation of the pressures at all the other points in the computational domain. This calculation is therefore extremely expensive to compute exactly as it is $\mathcal{O}(N_X^2 N_Y^2)$ operations. We therefore use the multilevel multi-integration technique of Brandt and Lubrecht [2], which is described fully by Venner and Lubrecht [38], to reduce this calculation to $\mathcal{O}(N_X N_Y \ln N_X \ln N_Y)$ operations.

The calculation of H_{00} in (4.3) is accomplished by relaxation of the force balance equation (3.4), according to

$$(5.1) \qquad H_{00} \leftarrow H_{00} - c\left(\frac{2\pi}{3} - \Delta X \Delta Y \sum_{i=1}^{N_X} \sum_{j=1}^{N_Y} P_{i,j}\right),$$

for all mesh points (i, j), where c is a small relaxation parameter; see [38, p. 148] and [12, p. 365].

At first sight it is unclear exactly how this equation may be justified. Write (4.3) and (4.4) in vector form as

$$(5.2) \qquad \underline{H} = H_{00}\,\underline{e} + \underline{G} + K\underline{P},$$

$$(5.3) \qquad \frac{2\pi}{3} = \Delta X \Delta Y \underline{e}^T \underline{P},$$

where $\underline{e}$ is the vector whose components are all one, $\underline{G}$ is the vector whose components are $\mathcal{G}_{i,j}$, and K is the matrix whose entries are composed of the elements $K_{i,j,k,l}$. Define the vector $\underline{y}$ as the solution of the system of equations

$$(5.4) \qquad K\underline{y} = \underline{e}.$$

Hence

$$(5.5) \qquad \underline{y}^T = \underline{e}^T\,K^{-1}.$$

It is worth remarking that although we do not know if K has an inverse we can easily attempt to compute $\underline{y}$ numerically. It then follows that the constraint equation (5.3) may be written by using (5.2) and (5.5) as

$$(5.6) \qquad \frac{2\pi}{3} = \Delta X \Delta Y\,\underline{y}^T(\underline{H} - H_{00}\,\underline{e} - \underline{G}).$$

Although it would be possible to store and save $\underline{y}^T$ and $\underline{e}$, it is easier to use an iterative method. Applying a standard Newton–Jacobi-type procedure to calculate H_{00} from this equation with relaxation factor λ then gives

$$(5.7) \qquad H_{00} \leftarrow H_{00} - c \left(\frac{2\pi}{3} - \Delta X \Delta Y \underline{y}^T (\underline{H} - H_{00}\underline{e} - \underline{G}) \right),$$

where

$$(5.8) \qquad c = \frac{\lambda}{\Delta X \Delta Y \underline{y}^T \underline{e}}.$$

Equations (5.1) and (5.7) are easily seen to be equivalent using (5.2). Computational experience suggests that c must be small and be chosen with some care.

One smoothing cycle is said to be the sequence of updating all the pressures, $P_{i,j}$, and film thickness values, $H_{i,j}$, along with the corresponding density and viscosity values.

In any implicit time-integration process it is important to ensure that the starting guess on a timestep is a good prediction. We therefore use linear extrapolation of the pressure, film thickness, and H_{00} solutions from the previous two steps. This prediction typically brings the root mean square residual at the start of the timestep down by an order of magnitude and possibly suggests that higher order methods might make an impact in time integration. Although it is possible to calculate $\underline{H}^{(0)}(t_n)$ directly from the predicted pressure solution, $\underline{P}^{(0)}(t_n)$, this does not lead to any noticeable improvements.

The solution method employed, described in [27] and [13], iteratively solves for pressure before updating the film thickness. One addition for the time dependent case is that the method for testing for convergence of the nonlinear iteration is the well-known (see, e.g., [3]) convergence test of [31]. In this test the iteration cycle on a timestep is continued until

$$(5.9) \qquad \frac{\psi}{1 - \psi} \left\| \underline{H}^{(m+1)}(t_n) - \underline{H}^{(m)}(t_n) \right\| < 0.33 tol,$$

where tol is the time-integration error tolerance, and m is the iteration number. This also requires the definition of an estimate for the rate of convergence, ψ,

$$(5.10) \qquad \psi = \left(\frac{\left\| \underline{H}^{(m+1)}(t_n) - \underline{H}^{(m)}(t_n) \right\|}{\left\| \underline{H}^{(1)}(t_n) - \underline{H}^{(0)}(t_n) \right\|} \right)^{\frac{1}{m}},$$

with a suitable norm, $\| \cdot \|$, which we define below. An additional source of difficulty for EHL solvers is that it is often necessary to solve a steady state problem to calculate the initial conditions.

6. Local error estimation. For transient EHL calculations the choice of correct timestep size is critical. If the timestep is too large then important physical features may be missed should they fall between successive steps. Equally, choosing a very small stepsize may, at best, lead to a large amount of computational work for very small changes in the solution and, at worst, result in solutions diverging, for example, due to the magnification of temporal gradients. This is due not to the stability of the problem but to the convergence properties of the nonlinear solver described in detail in [13]. The strategy used here comes from [13] but a similar approach has been

used by Gaskell et al. [11] more recently. There are also issues unique to EHL-based multigrid solvers. For example, should the timestep become very small, then any corrections made may amplify, rather than reduce, the errors in the solution unless very small underrelaxation parameters are used in the multigrid solver.

6.1. Choice of variables. One of the central issues in the numerical solution of DAEs is the proper estimation of the error and which variables to include or exclude in the error test [3]. In this case there is a choice as to whether it is the errors in $\underline{P}$ or in $\underline{H}$ which are controlled. Intuitively, because the system given by (4.7) is solved for $\underline{P}$ and (4.8) is solved for $\underline{H}$ using this $\underline{P}$, it seems sensible to control the errors in $\underline{P}$. However, it is the contact region where the most change is taking place, and this is dominated by the wedge and squeeze terms in (3.1) which involve derivatives of $\underline{H}$. Experiments have confirmed that controlling these errors requires significantly less work per timestep and fewer timesteps are required. The error tests will therefore be formulated for variable $\underline{H}$. The difference between the magnitudes of the local errors in H and P is shown in the next subsection.

6.2. Calculating the local error estimate. Once the convergence test of (5.9) has been satisfied, a local error calculation is undertaken to establish a new timestep size. The size and nonlinearity of the systems of nonlinear equations to be solved at each timestep preclude the use of LU linear equation solvers to estimate the DAE local error by means of an extra backsolve [3, p. 147]. Instead, a multigrid-based iterative approach will be used to estimate the local error over the step. In estimating the local error we rely on the types of methods used by Brenan et al. [3] in our software, while recognizing that there are many alternatives.

First let us consider the Reynolds equation (4.7). Following the derivation given in [3, section 5.4], the system given by (4.7) can be considered, for the implicit Euler scheme used here, as

$$(6.1) \qquad \frac{\overline{\rho_M}(\underline{P}_{n+1})\underline{H}_{n+1} - \overline{\rho_M}\left(\underline{P}_n\right)\underline{H}_n}{\Delta T} - \underline{F}_1(\underline{P}_{n+1}, \underline{H}_{n+1}) = 0,$$

where $\overline{\rho_M}\left(\underline{P}\right)$ is a diagonal matrix with the ith diagonal entry consisting of the function $\overline{\rho}(.)$ evaluated using the ith pressure value. The true solution at timestep $n + 1$ satisfies

$$\frac{\overline{\rho_M}\left(\underline{P}(t_{n+1})\right)\underline{H}\left(t_{n+1}\right) - \overline{\rho_M}\left(\underline{P}(t_n)\right)\underline{H}(t_n) - \underline{TE}}{\Delta T} - \underline{F}_1(\underline{P}(t_{n+1}), \underline{H}(t_{n+1})) = 0,$$
$$(6.2)$$

where $\underline{TE}$ is the local truncation error, defined by

$$(6.3) \qquad \underline{TE} = \left[\overline{\rho_M}\left(\underline{P}\right)\underline{H}\right]'' \frac{(\Delta T)^2}{2}.$$

In estimating the truncation error, $\underline{TE}$, our experiments have shown that the variation of the density is not as important as that of the film thickness. One may therefore consider treating the truncation error simply as $\overline{\rho_M}\underline{H}''$ and estimate it using the usual difference between predicted and corrected values; see [3]. Thus the estimated value denoted by $\underline{TE}_a$ is

$$(6.4) \qquad \underline{TE}_a = \frac{\overline{\rho_M}(\underline{P}_{n+1})}{2}(\underline{H}_{n+1} - \underline{H}_{n+1}^{(0)}).$$

In estimating the growth in local error from t_n to t_{n+1} we assume that the error at t_n is zero and consequently that $\underline{P}_n = \underline{P}(t_n)$ and that $\underline{H}_n = \underline{H}(t_n)$. The local errors in pressure and film thickness are defined by $\underline{le\,H}$ and $\underline{le\,P}$ as

$$\underline{le\,P} = \underline{P}_{n+1} - \underline{\tilde{P}}_{n+1},$$
$$(6.5) \qquad \underline{le\,H} = \underline{H}_{n+1} - \underline{\tilde{H}}_{n+1},$$

where $\underline{\tilde{P}}_{n+1}$ and $\underline{\tilde{H}}_{n+1}$ are the values of the local solution to the equations with the computed values at t_n as the initial values.

This therefore leads to the following equation for the local errors $\underline{le\,H}$ and $\underline{le\,P}$:

$$\frac{\overline{\rho_M}(\underline{\tilde{P}}(t_{n+1}))\underline{\tilde{H}}(t_{n+1}) - \overline{\rho_M}(\underline{\tilde{P}}(t_n))\underline{\tilde{H}}(t_n)}{\Delta T} - \underline{F}_1(\underline{\tilde{P}}(t_{n+1}), \underline{\tilde{H}}(t_{n+1})) = \underline{TE}_a,$$
$$(6.6) \qquad\qquad\qquad \underline{F}_2(\underline{\tilde{P}}(t_{n+1}), \underline{\tilde{H}}(t_{n+1})) = 0.$$

This equation may then be solved for $\underline{\tilde{P}}$ and $\underline{\tilde{H}}$ using the standard EHL multigrid algorithm with right-hand side $\frac{\overline{\rho}}{2\Delta T}(\underline{H}_{n+1} - \underline{H}_{n+1}^{(0)})$.

In the case of the film thickness equation (4.3) since there is no transient term present, it can easily shown that neither the geometry nor H_{00} is present in the local error calculation, hence giving

$$(6.7) \qquad\qquad \underline{le\,H} - K\underline{le\,P} = 0.$$

This gives us a relationship between the local errors in $\underline{H}$ and $\underline{P}$, where K is the film thickness integration kernel matrix.

The cost of this local error calculation on a timestep is that two or three more multigrid V-cycles are carried out to obtain solutions, $\underline{\tilde{P}}$ and $\underline{\tilde{H}}$, to the local error problem. This number of additional cycles is about one fifth of the total number of multigrid cycles.

Once these new solutions are calculated, an estimate of the total local error in $\underline{H}$ may be defined as

$$(6.8) \qquad\qquad \|\underline{le\,H}\|_\omega = \|\underline{\tilde{H}}_{n+1} - \underline{H}_{n+1}\|_\omega.$$

The chosen norm, $\|\cdot\|_\omega$, is a weighted root mean square L_2-norm, as used in the software package DASSL, described in [3], defined by

$$(6.9) \qquad\qquad \|\underline{H}\|_\omega = \sqrt{\frac{1}{N_X N_Y} \sum_{i=1}^{N_X} \sum_{j=1}^{N_y} \left(\frac{H_{i,j}}{\omega_{i,j}}\right)^2},$$

with weights $\omega_{i,j}$ defined by

$$(6.10) \qquad\qquad \omega_{i,j} = ATOL + H_{i,j}^{(0)} RTOL.$$

These weights are themselves, therefore, given in terms of the predicted solution at that mesh point on that timestep, $H_{i,j}^{(0)}$, and the absolute and relative error tolerances, $ATOL$ and $RTOL$, which will be defined in the next section.

Returning to the relationship between the local errors in P and H given in (6.7) and taking norms give

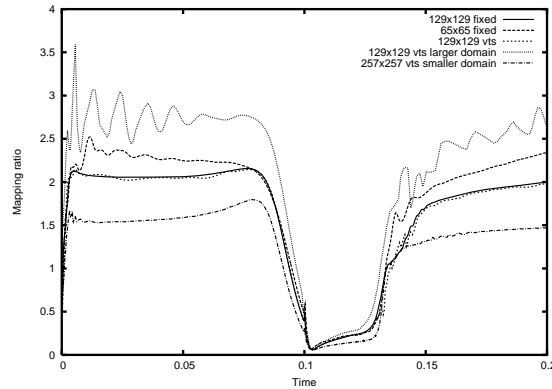$$(6.11) \qquad\qquad \|\underline{le\,H}\|_\omega \leq \|K\| \, \|\underline{le\,P}\|_\omega$$

FIG. 6.1. *Ratio of* $\|\underline{le\,H}\|_\omega / \|\underline{le\,P}\|_\omega$.

for an appropriate matrix norm $\|.\|$. Examples of estimates of the values of this norm are shown below for the problem described in section 3.1. Figure 6.1 shows that the ratio of these errors satisfies

$$(6.12) \qquad \|\underline{le\,H}\|_\omega / \|\underline{le\,P}\|_\omega \le 4$$

for calculations such as those in section 8. It is straightforward to monitor the ratio defined by the left-hand side of (6.12) during the calculation.

The local error test employed is then

$$(6.13) \qquad \|\underline{\tilde{H}}_{n+1} - \underline{H}_{n+1}\|_\omega \le 1.$$

Once this is true the stepsize change ratio is calculated by

$$(6.14) \qquad r = \left(2\|\underline{le\,H}\|_\omega\right)^{\frac{-1}{k+1}},$$

with k being the order of the method [32]. The new stepsize should then be chosen to be

$$(6.15) \qquad \Delta T_{n+1} = r\Delta T_n,$$

subject to some standard limitations concerning maximum permissible stepsize. For example, if the error is small, e.g., $r > 1.5$ in (6.14), then the stepsize may be increased for the next timestep. If $r < 0.9$, then either the stepsize is reduced for the following step or, if the current step is considered to have failed, the current timestep may be retaken with a new stepsize. In between these extremes the stepsize is left unchanged. Stepsize increases are allowed only every two steps, but decreases are allowed on every step as is common in a number of time-integration codes. The main issue now is how to choose $RTOL$ and $ATOL$ so that they reflect the underlying spatial discretization errors.

7. Space and time error estimation and balancing. The general issue of controlling the global error in time dependent PDE calculations remains problematic despite many recent successful algorithms based upon adjoint methods; see [5, 26]. Adjoint methods are also important for EHL problems in that they provide a means of controlling the error in quantities of interest such as friction [20]. The key issue

for large scale engineering calculations is the cost of the estimation algorithms. The counterargument is that without incurring this cost there is no sure way of understanding the accuracy of the solution. Nevertheless for problems for which even computing any solution is a formidable challenge and an expensive use of computer time the option of using the adjoint-based error control algorithms cost is unattractive. For example, the algorithm of [5] requires one preliminary integration with a coarse tolerance, storage of solution values and then two adjoint integration techniques before the main integration is started.

In this section we consider less expensive options which take into account the magnitude of the spatial error based on the approach of [1, 22] but modified still further to reduce computational costs.

Let the continuous equation system, defined by (3.1)–(3.6), have an exact solution $u(x, y, t)$ together with a vector of values of this solution at spatial grid points at time t $\underline{u}(t)$, and the discretized equation system, defined by (4.10), have exact solution $\underline{U}(t)$. If, at time t, the numerical approximation to the solution of the system is $\underline{\tilde{U}}(t)$, then the total error, $E(t)$, is defined by

$$
\begin{aligned}
\underline{E}(t) &= \underline{u}(t) - \underline{\tilde{U}}(t) \\
&= (\underline{u}(t) - \underline{U}(t)) + (\underline{U}(t) - \underline{\tilde{U}}(t)) \\
&= \underline{e}(t) + \underline{g}(t),
\end{aligned}
$$

(7.1)

where $\underline{e}(t) = \underline{u} - \underline{U}$ represents the vector of spatial discretization errors on the mesh points, and $\underline{g}(t) = \underline{U} - \underline{\tilde{U}}$ is the global error in the time integration. Given that a solution has been discretized in space to a particular degree of accuracy, $\underline{e}(t)$, it is not worth solving the transient part to a much higher degree of accuracy, but equally this transient error $\underline{g}(t)$ must not degrade the spatial accuracy.

The control of the global error using adjoint methods is an active area; see [5, 9, 23, 26]. The optimal choice of timestep is governed by successfully relating the spatial error of the solution, with the time error. It has been shown (see, e.g., [1, 22]) that controlling the local (temporal) error per step, so that the local growth in time of the spatial error dominates, provides a practical reliable algorithm for balancing these errors. The disadvantage is that the strong theoretical basis of the adjoint methods is lacking. This form of error control, in all the experiments we have conducted, has done a good job of balancing the spatial and temporal errors. However, in common with the results of [1], it is the case that it is relatively expensive in terms of the number of timesteps taken.

In order to reduce computation times a modified approach is thus taken so that instead of controlling the local error in time so that it is a fraction of the growth in the spatial error we control the local error so that it is a fraction of the spatial discretization error. The approach used is like the extrapolation approach of section 3.1 of [22] in that the fine mesh solution is used to estimate the error in a coarser mesh and then extrapolation is used to estimate the error on the fine mesh. Using the approach in these papers yields a simple extrapolation-based estimate of the first order spatial error in the fine mesh solution given by

(7.2)
$$
e_{i,j}(t) \approx \tilde{H}_{i,j}(t) - \tilde{H}_{I,J}(t),
$$

where (i,j) is the fine grid point coincident with the coarse grid point (I,J).

The error control algorithm of [22, equation (3.10)] is then given by defining

(7.3)
$$
\|\underline{le\,H}\|_\omega < \varepsilon \left\|\underline{e}_{n+1} - \underline{e}_n\right\|_\omega,
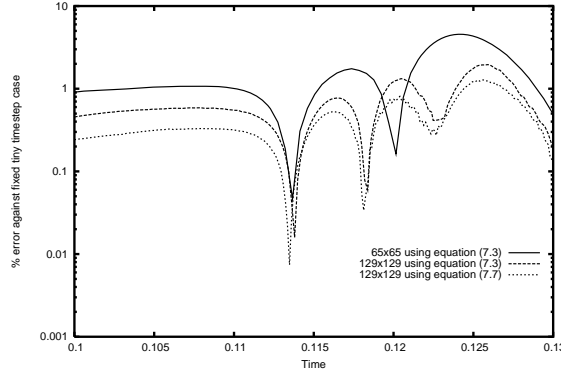$$

FIG. 7.1. *Comparison of errors in the minimum film thickness through reversal using different error control strategies.*

where $\| \cdot \|_\omega$ is a weighted root mean square L_2-norm, as used in (6.9). Using the approximation

$$(7.4) \qquad \|\underline{e}_{n+1} - \underline{e}_n\|_\omega \approx \Delta t \, \left\| \frac{\partial \underline{e}}{\partial t} \right\|_\omega$$

allows the error control approximation to be written as a local error per unit step. Assuming that the spatial error may be expanded in powers of Δx and that $\Delta x = \Delta y$,

$$(7.5) \qquad \left\| \frac{\partial \underline{e}}{\partial t} \right\|_\omega \approx \Delta x^p \, \left\| \frac{\partial \underline{\tilde{E}}}{\partial t} \right\|_\omega,$$

where $p = 1$ is the power of Δx in the spatial error; thus the error control can be approximated, as in [1, 22], by

$$(7.6) \qquad \|\underline{le\,H}\|_\omega < \varepsilon_0 \, \Delta x^p \, \Delta t \left\| \frac{\partial \underline{\tilde{E}}}{\partial t} \right\|_\omega,$$

where $\varepsilon_0 = 0.1$ or 0.3. Although the efficiency could be increased by using a larger value of ε_0, a possible alternative to (7.3) is then given by

$$(7.7) \qquad \|\underline{le\,H}\|_\omega < \varepsilon_1 \, \Delta x^p \, \|\underline{\tilde{E}}\|_\omega.$$

This approach adapts the timestep so that the local time error is a fraction of the spatial error estimate defined as in (7.7).

In order to assess the effectiveness of this approach an experiment was conducted using the two error controls on the test problem described in section 3.1. In this experiment a high-accuracy solution is computed using meshes of both 65×65 and 129×129 points and a small uniform timestep of $\Delta X^f/32$, where ΔX^f is from the finer resolution case. In comparing the methods the values of $\varepsilon_0 = 0.1$ and $\varepsilon_1 = 0.1$ were used on the 129×129 case. A comparison solution with the 65×65 point case is also included as this shows the effect of increasing mesh resolution. Figure 7.1 shows the percentage "errors" against this high accuracy solution in the key quantity of H_{cen}, the minimum film thickness, as the contact goes through the nonlinear reversal

TABLE 7.1
Computational performance comparison between different error control methods.

Grid dimension	Error control method	Timesteps taken	Iterations required
129×129	Fixed - small ΔT	36998	324065
129×129	Variable - using (7.7)	36628	111580
129×129	Variable - using (7.3)	140	2640
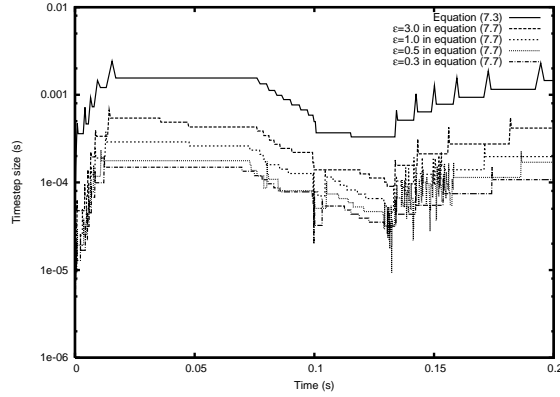
FIG. 7.2. *Comparison of timestep sizes for reversal of entrainment.*

phase of the calculation. The results show that in the first half of the simulation the errors using (7.7) are about 50% smaller than those using (7.3) relative to the tiny timestep case, but in the second half they are comparable. The figure concentrates on the important region around reversal. The computational cost of the three methods is demonstrated in Table 7.1. This table shows that the modified error control defined by (7.3) is a fraction of the cost of the approach defined by (7.7). This cost differential is reflected by the timesteps used by the two approaches as is shown in Figure 7.2. This figure compares the timestep sizes for the method of (7.3) (denoted as our method) with $\varepsilon = 0.1$ with the method of (7.7) with values of ε between 0.3 and 3. These values of ε are slightly larger than those used in the error balancing experiments of [1, 22]. The significant feature of these comparison experiments is that the error in a key quantity of interest, the central film thickness, is computed equally accurately in all cases, showing that the temporal error has been kept small. It is interesting to also see the variation in the timestep as the parameter ε in (7.3) is modified. This is shown in Figure 7.3 where it can be seen that the proposed method of step selection gives similar patterns of change. We have seen that sometimes using (7.7) may result in periods of relatively small steps without gaining additional accuracy. This seems to be related to the interplay between spatial and temporal error resulting in convergence difficulties for the multigrid solver.

The error control algorithm is implemented in the choice of the tolerances $RTOL$ and $ATOL$ being chosen at time t to be dynamically defined by

$$(7.8) \qquad ATOL = \frac{1}{10} \sqrt{\frac{1}{N_X^C N_Y^C} \sum_{I=1}^{N_X^C} \sum_{J=1}^{N_Y^C} \left[\tilde{H}_{i,j}(t) - \tilde{H}_{I,J}(t) \right]^2}$$
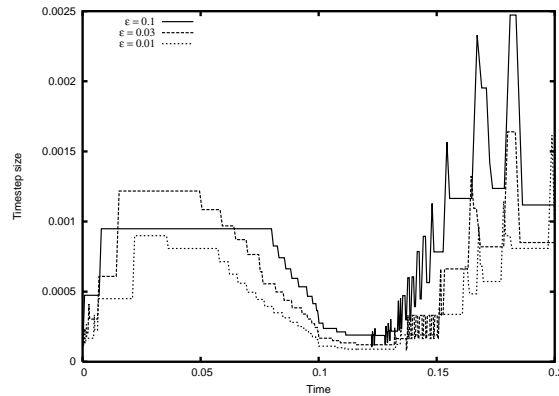
FIG. 7.3. *Comparison of timestep sizes when changing ε in (7.3).*

and

$$(7.9) \qquad\qquad\qquad RTOL = ATOL$$

for fine mesh points (i, j) with coincident coarse points (I, J).

8. Examples. In this section we present examples of some typical, challenging, transient EHL cases. The first example in section 8.1 is that of reversal of entrainment. Second in section 8.2 is the overrolling of a single surface roughness feature. In section 8.3 we consider the situation where the load is varying. Finally, in section 8.4 the presence of a measured surface roughness pattern is simulated.

8.1. Reversal of entrainment. A clear example of the benefits of using variable timesteps can be seen in the case of the reversal of entrainment problem defined in section 3.1. In Figure 8.1(a) the points on the line indicate where timesteps have been taken. It can be seen that these are clustered around reversal and just after, as was desired. The actual timestep sizes can be seen in Figure 8.1(b) alongside the values of r, the stepsize change ratio. It was explained above that there can be advantages in not changing the stepsize too often. Choosing the new stepsize based on an a priori error test cannot guarantee that the new stepsize will be valid for more than one timestep. Thus having the range of values for r in (6.14) where the stepsize
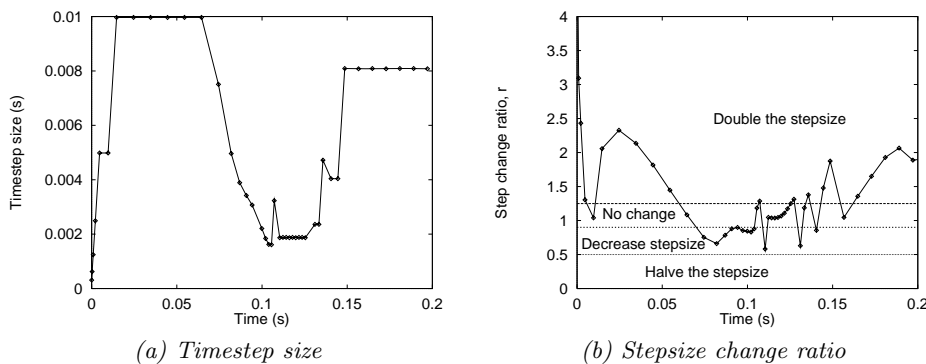

(a) *Timestep size*

(b) *Stepsize change ratio*

FIG. 8.1. *Timestep sizes and stepsize change ratios for reversal.*

TABLE 8.1
Computational performance comparison between fixed and variable timestepping codes.

Grid dimension	Fixed/ variable	Time taken (s)	Iterations required
65×65	Fixed	677	3386
65×65	Variable	458	1763
129×129	Fixed	1948	5139
129×129	Variable	1260	1604

TABLE 8.2
Film thickness (FT) comparisons at two reference times between fixed and variable timestepping codes.

Film thickness $\times 10^{-8}$ m		$t = 0.1$ s		$t = t_{min} \approx 0.108$ s	
		Min FT	Cen FT	Min FT	Cen FT
65×65	fixed	1.71	9.34	1.24	9.66
65×65	variable	1.81	9.41	1.28	9.80
129×129	fixed	2.51	9.29	1.97	9.87
129×129	variable	2.52	9.33	1.99	9.95

remains unchanged is important. The size of this region also governs how often the stepsize can change because if it is too small then the stepsize may be successively increased and decreased. This "chattering" effect, well known in the ODE community, may cause instabilities in the solution. This region is considered, for example, by Shampine and Gordon [32] and Hairer, Nørsett, and Wanner [19]. It is the range of values calculated for r in the error test, for which r should be set to 1 in (6.15).

The use of variable timestepping may require more cycles per timestep, but through the larger stepsizes these are being done considerably less often. It is also possible to limit the number of iterations per timestep if the convergence of individual steps is failing to satisfy the convergence test (5.9) quickly enough.

One of the stated aims for the use of variable timestepping was a reduction in the required computational time. Table 8.1 shows the computational time comparison between run times for fixed step ($\Delta T = \Delta X$) and variable timestep sizes. To confirm the validity of the results obtained, central and minimum film thicknesses are compared in Table 8.2 at two reference times: the point of reversal ($t = 0.1$ s) and the time that the minimum film thickness is achieved ($t = t_{min} \approx 0.108$ s). At these times the minimum film thickness is an order of magnitude less than the initial steady state. It is clear that the significant saving in computational work described produces results of similar accuracy.

8.2. Surface features. Following the experimental work of Kaneta, Sakai, and Nishikawa [21], Venner and Lubrecht [37] solved this transient example of a transverse ridge proceeding from left to right through the domain. For the local error control method to be considered effective it should be able to capture periods of rapid change in surface roughness cases where the relative importance of the asperity changes as it progresses through the contact. The undeformed geometry is given by

$$(8.1) \qquad \mathcal{G}(X, Y, T) = \frac{X^2}{2} + \frac{Y^2}{2} - \mathcal{A} \times 10^{-10\left(\frac{X-X_d}{\mathcal{W}}\right)^2} \cos\left(2\pi \frac{X - X_d}{\mathcal{W}}\right),$$

TABLE 8.3
Parameters used for transverse ridge example, after Venner and Lubrecht [37].

Parameter	Value	Parameter	Value
E'	1.17×10^{11} Pa	Moes parameter, M	233
α	2.2×10^{-8} Pa^{-1}	Moes parameter, L	5.42
η_0	1.22 Pa s	material parameter, G	2.62×10^3
u_s	0.0215 m s^{-1}	speed parameter, U	8.8×10^{-12}
R_x	0.0127 m	load parameter, W	2.0×10^{-6}
a	1.84×10^{-4} m	ridge amplitude, $\mathcal{A}$	0.075
p_h	0.54×10^9 Pa	ridge wavelength, $\mathcal{W}$	0.7

(a) *Timestep size*   (b) *Stepsize change ratio, r*
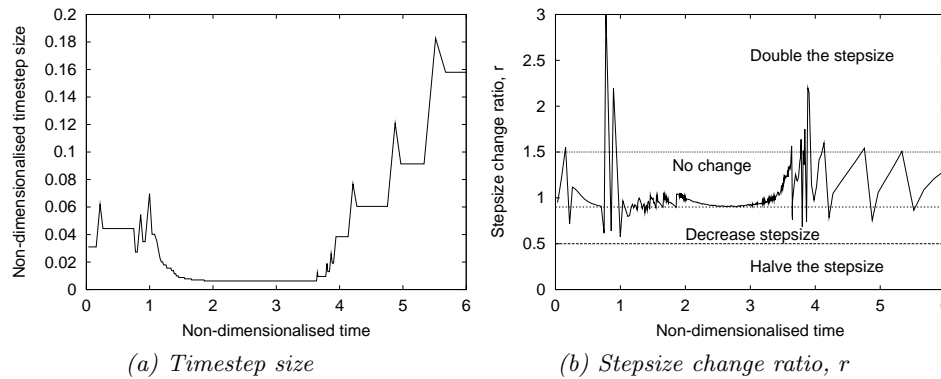
FIG. 8.2. *Timestep sizes and stepsize change ratio during overrolling of a transverse ridge using variable timestepping.*

where $\mathcal{A}$ is the dimensionless amplitude of the ridge, $\mathcal{W}$ is the dimensionless wavelength of the ridge, and $X_d(T)$ is the dimensionless position of the ridge given by

$$(8.2) \qquad X_d(T) = X_d(0) + \frac{2u_2 T}{u_s}.$$

The parameters are as given in Table 8.3 with the ridge entering from the left-hand side of the domain, with $X_d(0) = -2.5$, and then progressing through the domain. They correspond to a ridge of height 0.2 μm and a width of 0.07 mm, approximately what was used by Kaneta, Sakai, and Nishikawa, and precisely what was used by Venner and Lubrecht. Initially the ridge is outside the area of influence on the solution, and the timestep is expected to be larger than the optimal timestep as the ridge passes through the contact region. Once the ridge has entered the cavitation region no further transient effects should be present, and hence the solutions should return to the steady state conditions, and the timestep size should increase.

The physical solutions obtained were as shown by Venner and Lubrecht [37], and hence the only part of interest is the timestep information. In Figure 8.2(a) the timestep size is shown. As expected it does rise to a maximum which then falls as the ridge enters the contact area at $T = 1.5$. This remains fairly constant until the ridge enters the cavitation region just after $T = 3.5$. Once out of the contact area the solution returns to steady state conditions and the timestep size increases dramatically. Consideration of the stepsize change parameter, r, in Figure 8.2(b) reveals how satisfactory the selected timestep size is, as the ridge passes through the contact area. The desire to increase the stepsize both initially and after overrolling is contrasted against the reduction in timestep size between $T = 1$ and $T = 2$.
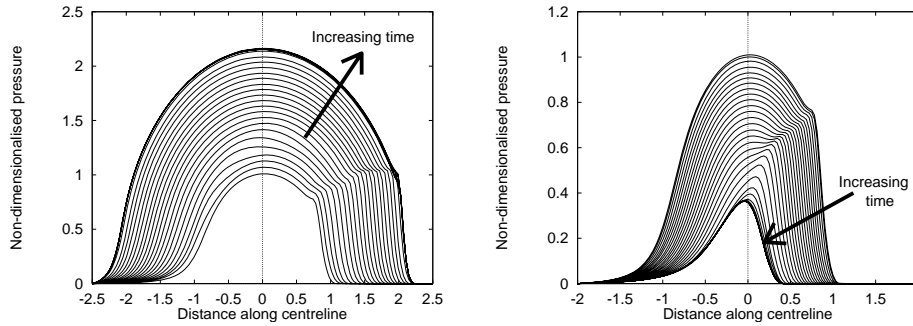
FIG. 8.3. (i) *Shock loading from* 0.45 *to* 0.90 *GPa in* 0.02 *s and* (ii) *unloading from* 0.45 *to* 0.22 *GPa in* 0.02 *s.*

8.3. Shock loading. The example of shock loading models the kind of change to the physical conditions experienced by the teeth on gears during meshing. Two kinds will be demonstrated in this section: shock loading and shock unloading. The load changes are of much greater amplitude than the oscillating loads investigated by Wijnant [40], and the timescale is very short.

The examples to be considered here both start from the initial conditions of the reversal example, shown in Table 3.1. The loading will take place by increasing the maximum Hertzian pressure to double its initial value for loading, or by halving it for the unloading case. Since this calculation is done using nondimensionalized versions of the equations as explained in section 3, it is the initial load against which the parameters are defined. In the nondimensionalization process the applied load is normally assumed to be constant, and thus, in the case of a varying load, a change to the conservation equation, (3.4), is necessary. This entails a modified right-hand side, hence changing the target sum of pressures. Thus (3.4) becomes

$$(8.3) \qquad \frac{2\pi}{3}\Phi = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} P(X,Y)\,\mathrm{d}X\mathrm{d}Y,$$

where the variable load, Φ, is given by

$$(8.4) \qquad \Phi = \left\{ \begin{array}{ll} 1, & t \leq 0\mathrm{s}, \\ 1 + 450t, & 0 \leq t \leq 0.02\mathrm{s}, \\ 10, & t > 0.02\mathrm{s}, \end{array} \right\} \text{for increasing loading,}$$

$$(8.5) \qquad \Phi = \left\{ \begin{array}{ll} 1, & t \leq 0\mathrm{s}, \\ 1 - 45t, & 0 \leq t \leq 0.02\mathrm{s}, \\ \frac{1}{10}, & t > 0.02\mathrm{s}, \end{array} \right\} \text{for decreasing loading.}$$

Pressure solutions for the two cases are very different. The centerline pressure is shown at selected timesteps in Figure 8.3, with the arrow indicating the direction of increasing time of solution plots. The increase in both the size of the contact area and maximum pressure is evident.

The unloading example has a very different behavior, as shown in the right side of Figure 8.3, because the problem has dropped out of the EHL regime to just being hydrodynamically lubricated instead. This means that the pressure spike has completely disappeared, and the deformation is now minimal, with no sidelobes present at all.
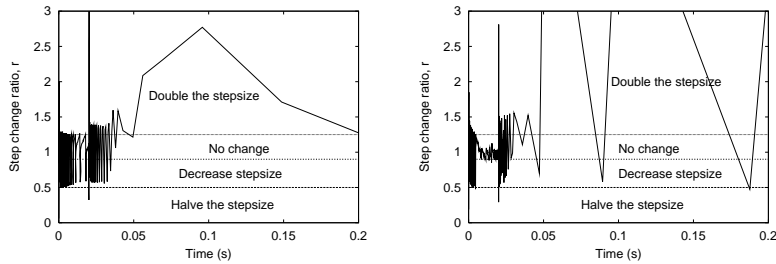
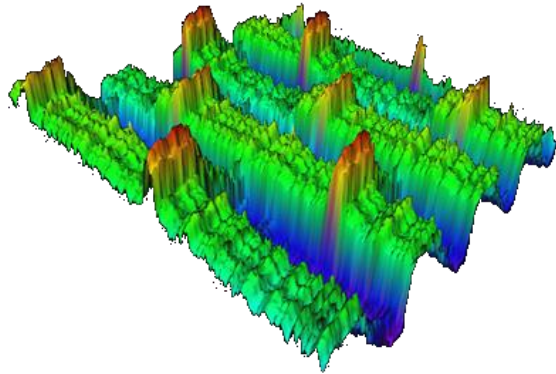FIG. 8.4. *Shock loading and unloading: Stepsize change ratios.*

FIG. 8.5. *Roughness pattern applied across contact.*

The timestep size is chosen based on the local error estimate which in turn depends on derivatives of the solution. It is therefore expected that when the system has reached the final steady state solution beyond 0.02 s the timestep should be much larger than in the initial stages. For the shock loading example the stepsize change ratio, r, is shown on the left of Figure 8.4. It can be clearly seen that the timestep does not increase until after the end of the loading period.

The stepsize change ratios for the shock unloading example is shown on the right side of Figure 8.4. Similar behavior to the shock loading case can be observed, although with a significantly less smooth increase in the timestep size.

8.4. Real surface roughness. A recent facet of EHL research has been the solution of problems with real surface roughness by a number of research groups; see, e.g., [8]. On a microscopic scale real surfaces of contacts are not defined by parabolas, nor are any features, such as bumps and dents, perfectly geometric shapes. In fact the surface is covered in small asperities. In an EHL contact the deformation of the solid will seek to smooth out these irregularities, and hence the pressure profile will be related to the local surface geometry. In fact for each individual asperity being flattened it is possible to see that it has its own EHL contact problem—just on a much smaller physical scale.

For this example we have imposed the measured roughness pattern shown in Figure 8.5 onto one of the surfaces. This therefore rolls through the contact over time, in the same manner as the surface feature simulated in section 8.2. An example, with the amplitude of the roughness applied at a nominal 25%, is shown in Figure 8.6.
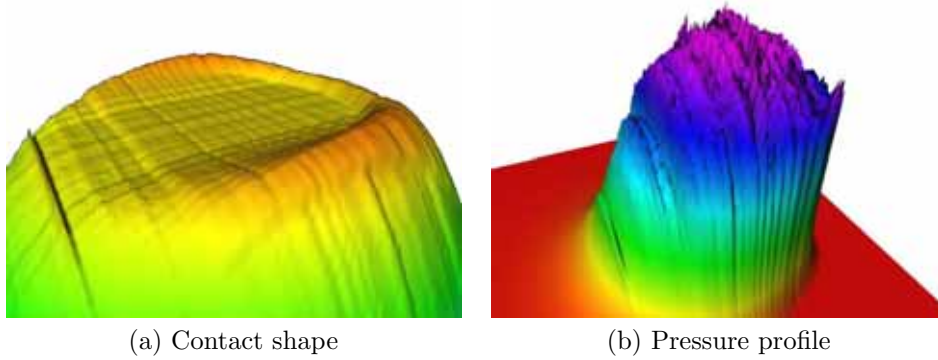
(a) Contact shape   (b) Pressure profile

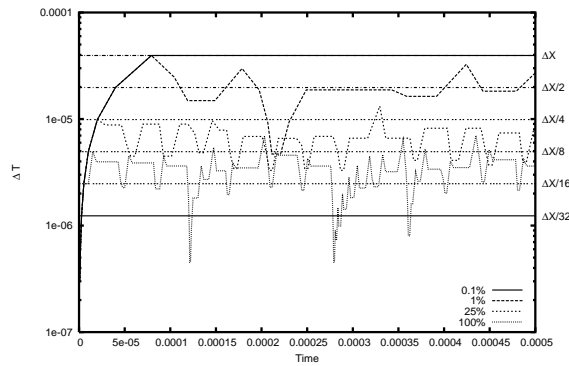FIG. 8.6. *Roughness pattern applied across contact.*

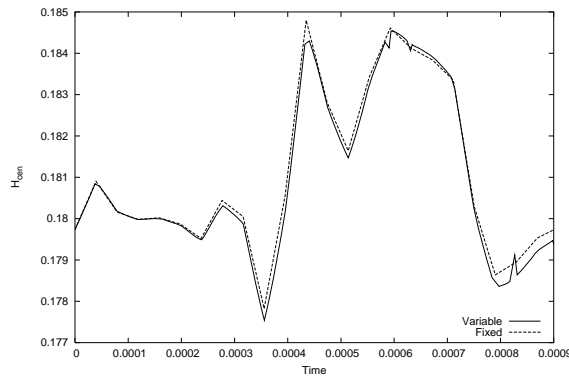FIG. 8.7. *Timestep size for different amplitudes of roughness.*

FIG. 8.8. *Central film thickness using fixed and variable timestepping approaches.*

In the previous examples it has been shown how the use of variable timestepping increases the efficiency of the solver without damaging the accuracy. For surface roughness cases the situation is slightly different in that variable timestepping will allow the tracking of the progression of each asperity through the domain. This is shown well in Figure 8.7 where the timestep size is monitored for increasing roughness

amplitudes. It can be seen how the selected timestep size diminishes as the contacts become less smooth, as we would expect.

The increase in accuracy that may be attained through the use of temporal adaptivity is shown in Figure 8.8, which shows the contrast in film thickness solution values at the center of the contact region. It is clearly seen how the fixed timestep approach both diverges from the variable timestepping solution and also misses out on the subtimestep scale features.

9. Conclusions. In this paper we have shown how a variable timestep approach has been used to provide an efficient method for solving transient EHL problems. It has been explained how the adaptive methods proposed fit into the standard multilevel schemes used for numerical solutions to EHL problems.

Variable timestepping has been shown, by experiments, to substantially reduce the required work while maintaining the same level of solution accuracy. The overhead in calculating new stepsizes is small, relative to the increase in performance. Changing the stepsize away from ΔT, within predefined limits, has been seen to pose no problems for the solver.

Examples of representative transient EHL situations have been considered, covering three different physical changes. The problem of reversal of entrainment was showcased and compared against the experimental results of Scales et al. [30]. The progression of the disc of highly viscous lubricant through the center of the contact just after reversal was automatically picked up by the variable timestep approach, and the results were shown to be just as accurate as for fixed timesteps.

The problem of surface roughness has been considered. It was seen how the overrolling of a single surface feature required small steps while the asperity passes through the center of the contact. One conclusion to be drawn from this example could be that for cases with real surface roughness, the timestep could be expected to be constant throughout. In the example of a full surface roughness pattern it has been seen how this level of constant timestep is related to the amplitude of the roughness pattern and may be much smaller than the mesh size. It is in cases like this where grid adaptation [15, 17] will be very beneficial.

The other case considered was that of shock loading, where the loading on the contact changed dramatically over a very short length of time. Again it was seen how the timesteps were very small as the major changes in the loading changed, but as the system reached a steady state the allowable timestep size increased.

The grid adaptation techniques shown to be beneficial in [15, 17] have not been included, hence allowing the advantages of the variable timestepping methods to be considered independently of other factors. Similarly the advantages of parallelism [14, 18] have also been neglected since they have no impact on the accuracy of the solution obtained.

Future work in this area should consider the use of adjoint-based approaches with regard to the direct estimation and control of errors in quantities such as the central and minimum film thicknesses, or other quantities of interest such as the friction [17, 20].

REFERENCES

[1] M. BERZINS, *Temporal error control for convection-dominated equations in two space dimensions*, SIAM J. Sci. Comput., 16 (1994), pp. 558–580.

[2] A. BRANDT AND A. A. LUBRECHT, *Multilevel matrix multiplication and fast solution of integral equations*, J. Comput. Phys., 90 (1990), pp. 348–370.

[3] K. E. BRENAN, S. L. CAMPBELL, AND L. R. PETZOLD, *Numerical Solution of Initial-Value problems in Differential-Algebraic Equations*, Classics Appl. Math. 14, SIAM, Philadelphia, 1995.

[4] W. L. BRIGGS, V. E. HENSEN, AND S. MCCORMICK, *A Multigrid Tutorial*, 2nd ed., SIAM, Philadelphia, 2000.

[5] Y. CAO AND L. PETZOLD, *A posteriori error estimation and global error control for ordinary differential equations by the adjoint method*, SIAM J. Sci. Comput., 26 (2004), pp. 359–374.

[6] D. DOWSON AND G. R. HIGGINSON, *Elasto-hydrodynamic Lubrication, The Fundamentals of Roller and Gear Lubrication*, Permagon Press, Oxford, Great Britain, 1966.

[7] D. DOWSON AND C. M. TAYLOR, *Cavitation in bearings*, in Annual Review of Fluid Mechanics Vol. 11, Annual Reviews, Palo Alto, CA, 1979, pp. 35–66.

[8] C. D. ELCOATE, H. P. EVANS, T. G. HUGHES, AND R. W. SNIDLE, *Transient elastohydrodynamic analysis of rough surfaces using a novel coupled differential deflection method*, Proc. Instn. Mech. Engrs. J. Engineering Tribology Part J, 215 (2001) pp. 319–337.

[9] D. ESTEP, *A posteriori error bounds and global error control for approximation of ordinary differential equations*, SIAM J. Numer. Anal., 32 (1995), pp. 1–48.

[10] R. FAIRLIE, C. E. GOODYER, M. BERZINS, AND L. E. SCALES, *Numerical modelling of thermal effects in elastohydrodynamic lubrication solvers*, in Tribological Research and Design for Engineering Systems, Proceedings of the 29th Leeds-Lyon Symposium on Tribology, D. Dowson et al., eds., Elsevier, Amsterdam, 2003, pp. 675–683.

[11] P. H. GASKELL, P. K. JIMACK, M. SELLIER, AND H. M. THOMPSON, *Efficient and accurate time-adaptive multigrid simulations of droplet spreading*, Internat. J. Numer. Methods Fluids, 45 (2004) pp. 1161–1186.

[12] R. GOHAR, *Elastohydrodynamics*, 2nd ed., Imperial College Press, London, 2001.

[13] C. E. GOODYER, *Adaptive Numerical Methods for Elastohydrodynamic Lubrication*, Ph.D. thesis, University of Leeds, Leeds, UK, 2001.

[14] C. E. GOODYER, M. BERZINS, P. K. JIMACK, AND L. E. SCALES, *Grid-based numerical optimisation in a problem solving environment*, in Proceedings of the All Hands Meeting 2003, S. Cox, ed., EPSRC, Swindon, UK, 2003, pp. 854–861.

[15] C. E. GOODYER, R. FAIRLIE, M. BERZINS, AND L. E. SCALES, *Adaptive mesh methods for elastohydrodynamic lubrication*, in ECCOMAS CFD 2001: Computational Fluid Dynamics Conference Proceedings, Institute of Mathematics and Its Applications, Minneapolis, MN, 2001.

[16] C. E. GOODYER, R. FAIRLIE, M. BERZINS, AND L. E. SCALES, *Adaptive techniques for elastohydrodynamic lubrication solvers*, in Tribology Research: From Model Experiment to Industrial Problem, Proceedings of the 27th Leeds-Lyon Symposium on Tribology, G. Dalmaz et al., eds., Elsevier, Amsterdam, 2001, pp. 709–719.

[17] C. E. GOODYER, R. FAIRLIE, D. E. HART, M. BERZINS, AND L. E. SCALES, *Calculation of friction in steady-state and transient EHL simulations*, in Transient Processes in Tribology, Proceedings of the 30th Leeds-Lyon Symposium on Tribology, G. Dalmaz et al., eds., Elsevier, Amsterdam, 2004, pp. 579–590.

[18] C. E. GOODYER, J. WOOD, AND M. BERZINS, *A parallel grid based PSE for EHL problems*, in Applied Parallel Computing, Proceedings of PARA'02, Lecture Notes in Comput. Sci. 2367, J. Fagerholm et al., eds., Springer-Verlag, New York, 2002, pp. 523–532.

[19] E. HAIRER, S. P. NØRSETT, AND G. WANNER, *Solving Ordinary Differential Equations* I: *Nonstiff Problems*, Springer-Verlag, New York, 1980.

[20] D. E. HART, M. BERZINS, C. E. GOODYER, P. K. JIMACK, AND L. E. SCALES, *Adjoint error estimation for EHL-like models*, Internat. J. Numer. Methods Fluids, 47 (2005), pp. 1069–1075.

[21] M. KANETA, T. SAKAI, AND H. NISHIKAWA, *Optical interferometric observations of the effects of a bump on point contact EHL*, ASME J. Tribology, 114 (1992), pp. 779–784.

[22] J. LAWSON, M. BERZINS, AND P. M. DEW, *Balancing space and time errors in the method of lines for parabolic equations*, SIAM J. Sci. Statist. Comput., 12 (1991), pp. 573–594.

[23] A. LOGG, *Multi-adaptive time integration*, Appl. Numer. Math., 48 (2004), pp. 339–354.

[24] A. A. LUBRECHT, W. E. TEN NAPEL, AND R. BOSMA, *Multigrid, an alternative method of calculating film thicknesses and pressure profiles in elastohydrodynamically lubricated line contacts*, ASME J. Tribology, 108 (1986), pp. 551–556.

[25] A. A. LUBRECHT, W. E. TEN NAPEL, AND R. BOSMA, *Multigrid, an alternative method of solution for two-dimensional elastohydrodynamically lubricated point contact calculations*, ASME J. Tribology, 109 (1987), pp. 437–443.

[26] K.-S. MOON, A. SZEPESSY, R. TEMPONE, AND G. ZOURARIS, *Hyperbolic differential equations and adaptive numerics*, in Theory and Numerics of Differential Equations (Durham, 2000), Universitext, Springer-Verlag, Berlin, 2001, pp. 231–280.

[27] E. NURGAT, M. BERZINS, AND L. E. SCALES, *Solving EHL problems using iterative, multigrid and homotopy methods*, ASME J. Tribology, 121 (1999), pp. 28–34.

[28] A. I. PETRUSEVICH, *Fundamental conclusions from the contact-hydrodynamic theory of lubrication*, Izv. Akad. Nauk. SSSR (OTN), 2 (1951), p. 209.

[29] C. J. A. ROELANDS, *Correlational Aspects of the Viscosity-Temperature-Pressure Relationship of Lubricating Oils*, Ph.D. thesis, Technische Hogeschool Delft, Delft, The Netherlands, 1966.

[30] L. E. SCALES, J. E. RYCROFT, N. R. HORSWILL, AND B. P. WILLIAMSON, *Simulation and observation of transient effects in elastohydrodynamic lubrication*, SP-1182, in SAE International Fuels and Lubricants Meeting, Dearborn, MI, 1996, pp. 23–34.

[31] L. F. SHAMPINE, *Implementation of implicit formulas for the solution of ODEs*, SIAM J. Sci. Statist. Comput., 1 (1980), pp. 103–118.

[32] L. F. SHAMPINE AND M. K. GORDON, *Computer Solution of Ordinary Differential Equations: The Initial Value Problem*, W. H. Freeman and Co., San Francisco, CA, 1975.

[33] F. STENGER, *Numerical Methods Based on Sinc and Analytic Functions*, Springer Ser. Comput. Math. 20, Springer-Verlag, New York, 1993.

[34] F. STENGER, *private communication*, 2004.

[35] U. TROTTENBERG, C. OOSTERLEE, AND A. SCHÜLLER, *Multigrid*, Academic Press, New York, 2001.

[36] C. H. VENNER, *Multilevel Solution of the EHL Line and Point Contact Problems*, Ph.D. thesis, University of Twente, Endschende, The Netherlands, 1991.

[37] C. H. VENNER AND A. A. LUBRECHT, *Numerical simulation of a transverse ridge in a circular EHL contact under rolling/sliding*, ASME J. Tribology, 116 (1994), pp. 751–761.

[38] C. H. VENNER AND A. A. LUBRECHT, *Multilevel Methods in Lubrication*, Elsevier, Amsterdam, 2000.

[39] B. WATREMETZ, F. COLIN, C. H. VENNER, AND A. A. LUBRECHT, *Optimum time step in a transient EHL contact*, in Transient Processes in Tribology, Proceedings of the 30th Leeds-Lyon Symposium on Tribology, G. Dalmaz et al., eds., Elsevier, Amsterdam, 2004, pp. 591–600.

[40] Y. WIJNANT, *Contact Dynamics in the Field of Elastohydrodynamic Lubrication*, Ph.D. thesis, University of Twente, Endschende, The Netherlands, 1998.

[41] S. R. WU, *A penalty formulation and numerical approximation of the Reynolds-Hertz problem of elastohydrodynamic lubrication*, Internat. J. Engrg. Sci., 24 (1986), pp. 1001–1013.

[42] S. R. WU AND J. T. ODEN, *A note on some mathematical studies on elastohydrodynamic lubrication*, Internat. J. Engrg. Sci., 25 (1987), pp. 681–690.

[43] S. R. WU AND J. T. ODEN, *Convergence and error estimates for finite element solutions of elastohydrodynamic lubrication*, Comput. Math. Appl., 13 (1987), pp. 583–593.