Helsinki University of Technology Institute of Mathematics Research Reports
Teknillisen korkeakoulun matematiikan laitoksen tutkimusraporttisarja
Espoo 1999 A411

THE FIRST LOCKING-FREE PLANE-ELASTIC FINITE ELEMENT: *HISTORIA MATHEMATICA*

Juhani Pitkäranta

TEKNILLINEN KORKEAKOULU
TEKNISKA HÖGSKOLAN
HELSINKI UNIVERSITY OF TECHNOLOGY
TECHNISCHE UNIVERSITÄT HELSINKI
UNIVERSITE DE TECHNOLOGIE D'HELSINKI

Juhani Pitkäranta: The first locking-free plane-elastic finite element: *historia mathematica*; Helsinki University of Technology Institute of Mathematics Research Reports A411 (1999).

Abstract: *We consider a two-parameter family of conforming bilinear quadrilateral finite element formulations for plane elastic problems with homogeneous isotropic material. The approach, taken earlier by Belytschko et al., is based on numerical strain reductions and stabilization and may be viewed as an extension & reformulation of the classical Turner rectangle of 1956. We present the full theory of this approach in the energy norm framework. The analysis is based on splitting the error in two orthogonal components, the* approximation error *and the equilibrium or* consistecy error*. The theory covers the full parametric dependence of the plane elastic problem, emphasizing in particular the case of beam-like bodies with arbitrarily small aspect ratios. The theory shows that, depending on the mesh, the stabilized reduced-strain formulation is capable of avoiding some, or even all, the parametric locking effects characteristic to two-dimensional elastic problems. We analyze first these locking effects in the context of standard bilinear FEM.*

AMS subject classifications: 65N30, 73C02.

Keywords: finite elements, locking, elasticity.

ISBN 951-22-4409-8
ISSN 0784-3143
Edita, Espoo, 1999

Helsinki University of Technology
Department of Engineering Physics and Mathematics
Institute of Mathematics
P.O. Box 1100, 02015 HUT, Finland
email: *math@hut.fi*
downloadables: *http://www.math.hut.fi/*

author's email: *Juhani.Pitkaranta@hut.fi*

Contents

1 Introduction

In this paper we study an extension and reformulation of one of the earliest finite elements in plane elasticity, the *Turner rectangle*, designed in 1953 [6] and published in 1956 [30]. The element was originally thought of as a beam element and aimed for approximating the deformations of the supporting ribs and spars in airplane wing structures [30]. The design appears a natural extension of a simpler one-dimensional Timoshenko beam element – this was likewise introduced in [30] and was used as an alternative of the rectangle.

The third element introduced in [30], the (standard) linear triangle, was aimed for approximating the deformation of wing panels as a membrane. In retrospect, the triangle (proposed earlier by Courant [7] in a different context) had the greatest impact on the later breakthrough of finite element methodology in structural mechanics [6, 24]. Anyhow, the Turner rectangle – the second candidate for the "first finite element" in plane elasticity – has later created a vivid discussion as well. Again in retrospect, this was the first finite element that fully avoided the *parametric locking* effects in a plane elastic problem. Attempts to reformulate the element as a geometrically more flexible quadrilateral have been numerous.

In modern finite element terminology, the rectangular "direct stiffness" formulation of Turner et al. is viewed as a nonconforming displacement method. The key idea was to deform the (now) standard bilinear shape functions into problem specific quadratic shapes so as to exactly capture the two simplest bending modes when the element bends like a beam. As now is widely understood, the standard bilinear (or linear) element perfoms very poorly as a beam element, a numerical phenomenon known as overstiffness or *locking* [15, Ch 6]. Due to the quadratic shapes assumed, the Turner rectangle formally avoids the locking problem – as does the corresponding one-dimensional beam element in [30]. However, unlike in 1D, the two-dimensional shape functions are incompatible, so the otherwise natural extension from 1D to 2D is shadowed by a "variational crime". The convergence of such an approach, or extension to more general quadrilateral element shapes, is by no means obvious.

Alternative ways of improving the bending behavior of the standard bilinear rectangle, or the corresponding bilinear quadrilateral element, have been proposed by a number of authors over the years. The two classics among these are the hybrid formulation of Pian [18] and another nonconforming formulation by Wilson et al. [31, 29]. In general, these formulations have been most successful in the rectangular reference configuration or its simple affine images, less on general quadrilateral meshes [15, 16]. On the other hand, in the rectangular reference configuration there have been more formulations than actually new methods. For example, the formulations of Turner, Pian and Wilson all lead to the same stiffness matrix for the nodal displacements when the mesh is rectangular and aligned with the coordinates [8, 9].

More recently, a new systematic approach for deriving robust plane elastic quadrilaterals from the bilinear basis was taken by Belytschko et al. [2], cf. also [3, 15]. Here the idea is to combine numerical *strain reductions* (or strain projections, or underintegration) with *stabilization*. The approach involves free parameters and again extends the Turner rectangle, this time to a parameter dependent family of conforming quadrilateral elements [2].

The approach we take here is close to that in [2]. We introduce a parametric family of conforming bilinear quadrilaterals, derived by modifying the strain energy by projections (local averaging) and added "correction terms" involving two free parameters. Based on this formulation we present – for the first time – the full story of the first locking-free finite element in a mathematical frame. Our mission is not to seek for intuitive physical justification or derivation. Rather, we explain the formulation a posteriori by means of direct stability and error analysis. The analysis is carried out in the energy norm and attempts to be sharp, involving even lower error bounds. Besides explaining the underlying (perhaps more numerical than physical) ideas, this analysis shows also the limitations of the formulation.

The main conclusions from our theory are in agreement with those already drawn from engineering expertise [16]: The overstiffness in bending can be avoided (with appropriate choice of the free parameters) on rectangular grids, but cannot be avoided on general quadrilateral meshes by any of the assumed algorithms. On general meshes (on general polygonal domains) we can only show that the modified bilinear elements perform nearly as well – or as badly depending on case – as standard elements. There is one exception of this rule: In case of plane strain and nearly incompressible material, the modified bilinear elements turn out to be better than the standard ones also on more general meshes. As now is well known, standard elements are overstiff not only in bending but also when subject to volumetric constraints [10, 15]. In retrospect, it appears rather surprising that the Turner rectangle, which really was designed for capturing the beam bending asymptotics only, actually is able to handle the volumetric locking problem as well. Our error analysis describes the mathematical background of this additional "miracle". Here the existing finite element theory [12, 21] turns out to be relevant.

In short, the aim of the paper is to put the Turner rectangle in a mathematical frame and show the limits of its extended formulation. Besides this, the paper has a dual goal, which perhaps is more important. The aim is to use the early finite element formulation as a frame to popularize some of the recent advances of finite element theory more generally. As we claim, this theory is about to take its natural shape, while finally being able to explain the "first variational crime" in finite element engineering. Somewhat paradoxally, this early formulation was never given full blessing by finite element theory.

In modern finite element software, "variational crimes" are usually committed on purpose. The temptation arises in general in *parameter dependent* problems where uniform convergence of FEM with respect to a parameter would be desirable. In many of such problems in structural mechanics, standard low-order elements fail to converge uniformly. In the present problem, one of the critical parameters is related to the shape of the (two-dimensional) body considered. Suppose the body is a rectangle of length L and height H. Then one would wish the convergence of FEM to be uniform with respect to the aspect ratio $t = H/L$ when this ratio takes arbitrarily small (or large) values. In particular, if t is relatively small and the body is discretized by a single layer of, say bilinear elements of length $a = hL$, one would wish that the relative discretization error would only depend on h when $a \geq H$. This actually is the case if the body is in the *stretching state* of deformation. In the *bending state* of deformation, however, the error is amplified by factor t^{-1} and convergence fails completely. The aim of the "crime" is then to avoid the amplification and thus achieve the same convergence rate in both deformation states. The incompatible Turner rectangle apparently was the first formulation to achieve that.

In the alternative formulation of the Turner element considered here, the "crime" is committed within conforming FE framework. In our theory, however, the leading ideas are actually borrowed from the theory of nonconforming finite element methods. This theory was developed in the early 1970's – apparently as inspired by the existing incompatible approaches of finite element engineering by that time. So we end up closing the circle here.

In the nonconforming finite element theory, it became common practice to split the error in two components, the *approximation error* and the *consistency error*, using what later has been cited as the (second) Strang lemma [28, 27, 5]. This same principle extends naturally to any "variational crime" proceeding from (or interpreted as proceeding from) the standard energy/displacement framework. The resulting theory is rather straightforward, being free of mathematical artifacts like specific (mixed, hybrid, etc.) variational formulations or auxiliary fields. In [22, 23] it was demonstrated that the earlier (mixed) theory of reduced-strain plate-bending elements can be considerably simplified and extended by this approach. This theory also holds promises for shell problems where earlier FEM theory is largely missing.

In the present problem, the earlier theory was based – rather naturally – on nonconforming FE theory [13, 14, 25, 26]. However, the newer (extended) Wilson formulation [31, 29] is assumed here. One also assumes a fixed domain and makes further simplifications regarding either the finite element mesh [13] or the differential equations to be solved [14, 25, 26]. The result of this earlier theory is (roughly) that the Wilson element is no worse than the standard bilinear element when the body under consideration is "thick". Thus the theory so far gives no indication on why the "crime" in the Turner – Pian – Wilson element was actually made. Here we provide the missing pieces, by taking into account the full parametric dependence of the problem. On thin (rectangular) domains, our main concern is obviously in bending states of deformation, but our theory handles the "easy" stretching state as well (or any combination of the two states). – We underline that the ability to handle different asymptotic deformation states within the same formalism is one of the strong features of our theory. Such flexibility is of prime importance especially when approaching the asymptotic diversity of shell problems [20].

Besides being asymptotically flexible, our theory has the interesting feature of giving *lower* error bounds as well. This option emergres from the simple observation that the approximation and consistency errors are orthogonal error terms, and that it is relatively easy to obtain lower bounds for the consistency error. We use the option here to demonstrate that some of our upper error bounds are sharp. We can also predict in this way some total failures of convergence due to large consistency errors. In [23], lower error bounds of this type were obtained for certain low-order plate-bending elements.

In the practical finite element design, some preliminary checking of convergence has traditionally been included in a design principle called *patch test*. Here again, much of the original motivation came from nonconforming approaches [11]. In [15, Ch 5], the patch test is formulated as a two-fold test involving an "interpolation test" and an "equilibrium test". This resembles our error analysis philosophy, where approximation error analysis may be understood as a generalized interpolation test and consistency error analysis corresponds to equilibrium testing. However, we need both more precise and more quantitative formulation of these tests, since our aim is to resolve parametric dependence and goal to achieve uniform convergence with respect to a parameter. Facing such demands, the required analysis necessarily becomes somewhat abstract and thus less handy to use as an a priori design tool. Anyhow, we attempt to formulate the principles of our "mathematical patch test" as simply as possible without compromising in precision.

Our "equilibrium test" consists more precisely of estimating the seminorm of a *consistency error functional.* This arises from the "crime" made in the formulation. As in nonconforming theory, an essential step in the analysis is to find an appropriate *expansion formula* for this functional. The aim of the formula is to explicitly resolve the effect of the "crime" so that the "punishment" (= consistency error) can be judged correctly. Finding such expansions is sometimes tricky. For example, unexpected boundary terms may appear from seemingly harmless modifications and cause unwanted error growth, cf. [22, 23].

In our consistency error analysis, *stability* plays a central role (as in nonconforming theory). The minimal stability condition, as known from engineering design criteria [10, 15], is that there should arise no unphysical zero-energy modes (mechanisms, hourglass modes, spurious modes). In our error analysis, this condition again needs to be replaced by more quantitative stability estimates. Proving such estimates for our formulation turns out to be nontrivial and the result somewhat mesh dependent (Theorem 6.1 below). Our analysis indicates weaker stability when the elements have high aspect ratios – a characteristic of the original Turner rectangle as well. As thin elements are natural when discretizing thin bodies, we cannot rule them out. However, we can show that if the mesh is rectangular, the weakened stability can be partly compensated with better consistency, i. e., with a better expansion of the consistency error functional (Theorem 8.1). There remains a slight mesh dependent effect, however, and we demonstrate by a *lower* error bound that on extreme meshes, convergence can actually fail because of this effect. We demonstrate the whole interplay between stability and consistency in sections 7-8, where the error analysis in case of a fixed polygonal body is carried out.

In case of a "thick" body, the consistency error analysis of our extended quadrilateral formulation is rather straightforward, as far as possible error growth from badly shaped elements is tolerated (Theorem 7.1). The same holds for a thin body in a stretching state of deformation, whereas in the bending state, our analysis indicates possible "equilibrium locking", i. e., error amplification by factor t^{-1} due to consistency error. Further analysis shows that this actually is a true effect but can be avoided on specific meshes, including rectangular ones. A necessary condition in all meshes, however, is that one of the free parameters in the FE formulation is set to a proper value (Theorem 9.1).

The approximation error analysis is very straightforward in all cases where the standard bilinear element works well. Such "easy" deformation states arise when volumetric constraints are not imposed, and when the body under consideration is either "thick" or in a stretching state state of deformation when thin. In these cases the approximation error may be bounded by the usual interpolation-test argument, as in the standard FE theory [5, 4]. Instead when volumetric constraints are imposed (or nearly imposed), or when the body is thin and in the bending state of deformation, the approximation error analysis becomes more difficult. The main problem is to find, by mathematical construction, a clever "unlocking interpolant" that avoids the parametric growth of the approximation error in these cases. The best "interpolant" would in fact be the finite element solution itself, so the mathematical construction may be viewed as an attempt to simulate the FE algorithm. In a parametric situation, such a simulation must in general be guided by the asymptotics of the exact solution at the parametric limit where locking may occur. In the present problem we have two such parameters. These cause two different asymptotics, so we need also two different (in fact, unrelated) approaches when constructing the "unlocking" interpolants.

When dealing with volumetric constraints, it turns out that we can apply the earlier finite element theory [12, 21](Theorem 7.1). Instead when approximating the asymptotic beam bending mode of a thin body, no earlier theory is available. Our result here is rather specific: We find the desired interpolant in case of a rectangular mesh. This result (Theorem 10.1) is anyhow the main result of the paper. It actually leaves open only the case of parallelogram (or nearly parallelogram) meshes, since more general meshes are ruled out by lower error bounds (Theorem 4.1). In our construction, some of the ideas are borrowed from the error analysis of reduced-strain shell elements in [19].

In summary, our "generalized interpolation" test (Theorem 10.1) combined with the standard interpolation and "generalized equilibrium" tests (Theorems 7.1, 8.1 and 9.1) provides the final justification of the Turner – Pian – Wilson – et al. rectangle. In our formulation the Turner element is actually extended to a single-parameter family of locking-free rectangles, but if invariance under switching of coordinates is required, the formulation ends up being unique. In [2] this reformulation of the Turner rectangle was named QBI. For a "thick" polygonal body, or even for a thin body in a stretching state of deformation, our theory extends the Turner rectangle to a two-parameter family of quadrilateral elements (also found in [2]). These "unlock" the overstiffness in case of plane strain and nearly incompressible material (on restricted meshes at least), otherwise they perform nearly as well as the standard bilinear quadrilaterals in such problems. Finally when the bending of thin bodies is taken into account, the fully locking-free eight-dof quadrilateral of general shape is still the dream element – and perhaps has to remain such [16].

In what follows we want to present the full story and therefore start the finite element error analysis from standard bilinear elements. This prologue also gives us a way to somewhat "derive" the reduced-strain formulation that we are aiming at. The verification, however, comes a posteriori. The more detailed plan is the following.

In Section 2 we introduce the two-dimensional elastic problem to be considered. We assume homogeneous isotropic material, treating in parallel the cases of plane stress and plane strain. The body under consideration is assumed to be either of fixed polygonal shape, or of parametric rectangular shape with the aspect ratio varying arbitrarily. Basic regularity results/assumptions on the exact displacement field are stated that are needed later in the error analysis. These assumptions are collected in Hypotheses 2.1 and 2.2 at the end of the section. The physical as well as numerical reasoning behind these assumptions is clarified.

The convergence of the standard (isoparametric) bilinear scheme is studied in Section 3. Here and throughout the work we use relative error in the energy norm (possibly modified numerically) as the basic error indicator. The convergence results are summarized in Theorem 3.1. The two numerical locking effects, the volumetric locking and the "thickness locking" on thin bodies, are predicted.

In section 4, the locking effects of standard bilinear elements are analyzed in more detail in case of a thin rectangular body and a single-layer mesh. Our conclusions are very similar to those drawn in [15] from more qualitative reasoning. We first isolate the locking effect at small t to the leading asymptotic bending mode of the displacement field, and then trace the effect further back to two constrained approximation problems, each acting as an independent source of locking. The analysis is based on rewriting the energy density so that the strain terms causing locking appear explicitly. We detect the two locking effects, named *dilatation locking* and *shear locking* according to [15, Ch 6], by strict lower error bounds in case of a single-layer rectangular mesh. A third approximation failure, named *trapezoidal locking* in [15], is likewise detected by a lower error bound in case of a single-layer trapezoidal mesh. Finally, we detect the additional *volumetric locking* effect in a stretching state of deformation where other locking modes are absent.

In section 5 we "derive" the (stabilized) reduced-strain formulation based on the analysis of section 4. We introduce two equivalent formulations and show the historical connection to the Turner – Wilson – Pian rectangle (Theorem 5.1).

Section 6 marks the beginning of our main error analysis. We state here a basic stability theorem and introduce the leading error analysis principles forming the theoretical framework of the analysis that follows. The stability theorem (Theorem 6.1) may be considered the first main result of the paper. The rather technical proof is given in Appendix A. The error analysis principles are quite straightforward on a general level. It is demonstrated how the error can be split in two orthogonal error components, and how these components can be further bounded from above and from below.

In section 7 we carry out the approximation and consistency error analysis in the case where the body under consideration is of fixed polygonal shape. In the absence of volumetric constraints, the main concern here is in the consistency error. The analysis of this error term shows that the reduced-strain formulation performs essentially as well as the standard formulation, except for possible mesh geometric error amplification when the elements have high aspect ratios (Theorem 7.1). We then extend the error analysis to cover also the case of plane strain and nearly imcompressible material where volumetric constraints appear. In this case the approximation error analysis becomes the main issue – we cite the existing theory [12, 21].

In section 8, we give a further analysis of the mesh effect predicted in Theorem 7.1, in the special case of a rectangular mesh. Using a special expansion of the consistency error functional possible in that case, we can show that the mesh geometric error effect on a rectangular mesh is milder than predicted in Theorem 7.1. We further confirm by a counterexample that the sharpened bound (Theorem 8.1) is no more improvable.

In section 9 and in the associated Appendix B, we analyze the consistency error in case of parametric rectangular bodies. The main result here is that one of the free parameters in the formulation needs to be set to a specific value, and also the mesh has to be restricted severely, to avoid "equilibrium locking" in the bending state of deformation. We consider mainly the case of a rectangular mesh, where the avoidance of parametric error growth can be demonstrated relatively easily. Some further examples/counterexamples are given in Appendix B. The consistency error analysis is summarized in Theorem 9.1, the second main result of the paper.

The final obstacle, the approximation error analysis in case of parametric rectangular bodies, is met in section 10. Here we again assume a rectangular mesh. In that case we find a nonstandard "unlocking interpolant" showing that the approximation error is bounded uniformly with respect to t. A more laborious construction, given in Appendix C, further shows that the uniformity is preserved even in case of plane strain and nearly incompressible material, so that all locking effects are avoided simultaneously. The approximation error analysis is summarized in Theorem 10.1, the (third) main result of the paper.

2 The problem

Consider an elastic body occupying a region $\Omega \subset \mathbb{R}^2$ and deformable according to the laws of plane stress or plane strain. Denote the displacements along the coordinates x, y by u, v and the displacement vector field by $\mathbf{U}$,

$$\mathbf{U} = (u(x,y), v(x,y)), \quad (x,y) \in \Omega.$$

Assuming homogeneous material with Young modulus E and Poisson ratio ν, $0 \leq \nu < \frac{1}{2}$, the strain energy of the body is known to be proportional to the quadratic functional $\mathcal{A}(\mathbf{U},\mathbf{U})$ defined by

$$\mathcal{A}(\mathbf{U},\mathbf{U}) = \int_\Omega \left\{ \bar{\lambda}\Big(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\Big)^2 + 2\bar{\mu}\Big(\frac{\partial u}{\partial x}\Big)^2 + 2\bar{\mu}\Big(\frac{\partial v}{\partial y}\Big)^2 + \bar{\mu}\Big(\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x}\Big)^2 \right\} dxdy, \tag{2.1}$$

where we scale off the dimension of the Lamé coefficients

$$\lambda = \begin{cases} \dfrac{E\nu}{1-\nu^2} & \text{(plane stress)} \\[2ex] \dfrac{E\nu}{(1+\nu)(1-2\nu)} & \text{(plane strain)} \end{cases}$$

$$\mu = \frac{E}{2(1+\nu)}$$

by defining $\bar{\lambda} = \lambda/E$, $\bar{\mu} = \mu/E$.

We are mainly interested in the reference situation where the domain is a rectangle,

$$\Omega = \{(x,y) \mid 0 < x < L,\ -H/2 < y < H/2\},$$

possibly thin vertically so that the *dimensionless thickness* $t = H/L$ can take arbitrarily small values. As we may think of switching the coordinates when $H > L$, we will assume that $0 < t \leq 1$. In this parameter dependent setting, we are looking for finite element algorithms that converge *uniformly with respect to* t on the reference domains, under certain regularity assumptions on the exact solution that we make shortly. In addition, we want the scheme to extend to more general domains as well, so we allow Ω to be alternatively any "thick" domain of polygonal shape (Fig. 1).

The kinematical constraints possibly imposed at the boundary $\partial\Omega$ will play a role in some parts of our analysis. We allow any usual constraints where u, v, both or none are restricted on some parts of $\partial\Omega$. On the (thin) reference domains we make the more specific assumption that there are no constraints at $y = \pm H/2$. The body can then behave freely as a "bar" or as a "beam" when t is small. (At $x = 0, L$, any usual constraints are allowed.)

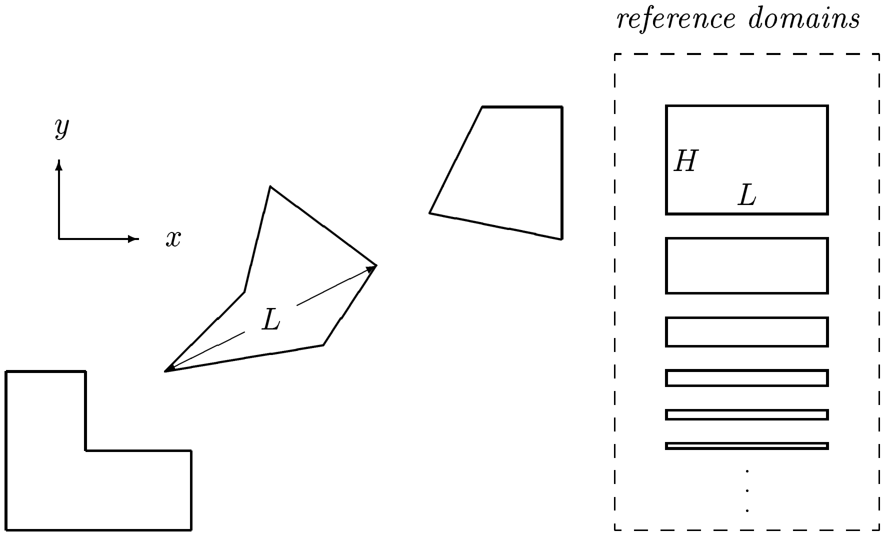

Figure 1: Possible shapes of Ω.

An important role in our analysis will be played by two kinds of constants, to be denoted by Q and C. These constants are *dimensionless*, positive and finite, and *independent of* t when concerning problems on the reference domains. First of all, we make below a number of regularity hypotheses on the exact solution and introduce in this context a constant denoted by Q. The actual value of Q depends on the regularity of the exact displacement field to be approximated, so Q is problem dependent. What matters in our error analysis is simply that Q is finite. We comment below on both the physical and the numerical interpretation of this basic regularity assumption. On the reference domains where t is a variable parameter, the regularity assumptions will be formulated so that constant Q is independent of t. We need to be rather careful here to find hypotheses that are both useful and realistic.

Another family of (dimensionless and positive) constants will be given the generic symbol C. Unlike Q, which is hypothetic and fixed when the problem setup is fixed, C is independent of the problem (and of t) and the value of C is computable (or can be chosen), but the value can be different in each different usage. Either, C is an absolute constant (independent of parameters), or it may depend on the finite element mesh via the angle condition on the mesh (see section 3 below), or it may depend on the scaled Lamé parameters $\bar{\lambda}, \bar{\mu}$. Some constants in our error analysis actually depend on the finite element mesh more severely than via the angle condition, but we then resolve the dependence explicitly. We use here further dimensionless parameters that characterize the geometry of the mesh (see section 3). When a constant C depends on the Lamé parameters, it will be proportional to $(\bar{\lambda} + \bar{\mu})^{1/2}$ or to $\bar{\lambda} + \bar{\mu}$, so this dependence is harmless except for the case of plane strain and nearly incompressible material where $\nu \to \frac{1}{2}$ implies $\bar{\lambda} \to \infty$. In the analysis that follows, we will first assume, at each step, that $\bar{\lambda} \leq C$ for some fixed C. We then extend the analysis to cover the parametric situation where $\bar{\lambda}$ can be arbitrarily large.

In the elastic problem considered, the convergence of standard FEM can fail, or slow down, basically for two kinds of reasons. First, the exact displacement field may simply be irregular causing standard approximation failure. Secondly, in parameter dependent situations as the one considered, there can arise asymptotic error amplification or *locking* as the parameter tends to a limit value, here as $t \to 0$ or as $\bar{\lambda} \to \infty$. Here we focus on these parametric effects. We will isolate the effects by assuming that the exact solution is sufficiently smooth, so that the lack of regularity is not a problem simultaneously. However, what "sufficiently smooth" means is still an issue in a two-dimensional elastic problem, since we want assumptions that are not only convenient but also realistic in some sense. Therefore, before we can even start, we have to translate some basics of classical linear elasticity into hypotheses that we will need. This is the subject of the remaining part of this section. Below we postulate first that $\bar{\lambda} \leq C$, so that t is the only active parameter. When $\bar{\lambda}$ is not bounded, we will need additional regularity hypotheses. These are stated at the end of the section.

On a fixed "thick" domain and when $\bar{\lambda} \leq C$, our basic regularity hypothesis is easily stated. Let $L = \mathrm{diam}(\Omega)$, denote by $\|\!| \cdot |\!\|$ the (scaled) *energy norm* defined as (see (2.1))

$$\|\!|\mathbf{U}|\!\| = \{\mathcal{A}(\mathbf{U},\mathbf{U})\}^{1/2} \tag{2.2}$$

and by $|\cdot|_{2,\Omega}$ a (Sobolev) seminorm measuring the regularity of $\mathbf{U} = (u, v)$ in terms of the second partial derivatives of u, v as

$$|\mathbf{U}|_{2,\Omega} = \left\{ \sum_{i+j=2} \int_\Omega \left[\left(\frac{\partial^2 u}{\partial x^i \partial y^j} \right)^2 + \left(\frac{\partial^2 v}{\partial x^i \partial y^j} \right)^2 \right] dxdy \right\}^{1/2}. \tag{2.3}$$

Then we assume that

$$L|\mathbf{U}|_{2,\Omega} \leq Q \|\!|\mathbf{U}|\!\|, \tag{2.4}$$

where Q is a finite (dimensionless, problem specific) constant. We may view Q as measuring the (lack of) regularity of the exact solution $\mathbf{U}$. Note that this measure is *scaling invariant.* Even if Q is not finite in a given situation, assumption (2.4) can still be justified when considering the "smooth part" of the exact displacement field. See the discussion at the end of this section.

We note that when no kinematical constraints are posed, $\|\!| \cdot |\!\|$ is actually a seminorm giving zero measure for rigid displacements. We assume such zero-energy modes to be eliminated, e. g. by symmetry conditions. The *energy space* where $\|\!| \cdot |\!\|$ acts as a norm is denoted by $\mathcal{U}$. As is well known, this is a Hilbert space supplied with the inner product (energy product) $\mathbf{U}, \mathbf{V} \to \mathcal{A}(\mathbf{U}, \mathbf{V})$.

On the parametric reference domains, assumption (2.4) is not realistic in general unless Q is allowed to grow without limit as $t \to 0$. As we want to resolve such growth explicitly, we need to study the behavior of the solution in more detail in this case. The first step is to split the solution as

$$\mathbf{U} = \mathbf{U}_s + \mathbf{U}_b, \tag{2.5}$$

where $\mathbf{U}_s = (u_s, v_s)$ and $\mathbf{U}_b = (u_b, v_b)$ are defined so that u_s, v_b are even and u_b, v_s are odd functions of y. This splitting is obviously always possible (and unique). On the reference domain, it is orthogonal in the energy norm, so that

$$|||\mathbf{U}|||^2 = |||\mathbf{U}_s|||^2 + |||\mathbf{U}_b|||^2. \tag{2.6}$$

If $\mathbf{U}_b = 0$, we may call the corresponding deformation state a pure *stretching state* and if $\mathbf{U}_s = 0$, a pure *bending state*.

We expand $\mathbf{U}_b$ in (2.5) further as

$$\begin{aligned} \mathbf{U}_b &= \big(y\theta(x),\ w(x) + \tfrac{1}{2}y^2\psi(x)\big) + \mathbf{U}_{br} \\ &= \mathbf{U}_{bm} + \mathbf{U}_{br}. \end{aligned} \tag{2.7}$$

Here we have simply separated the leading term in the *asymptotic expansion* of $\mathbf{U}_b$ for small t. Such expansions form the core of the dimension reduction theory in classical linear elasticity and apply in particular to the elastic strip under consideration. Here we note only that by the classical theory, the functions θ, w, ψ appearing in the leading term in (2.7) are related by

$$\theta + w' = 0, \quad \psi + \gamma\theta' = 0, \tag{2.8}$$

where

$$\gamma = \begin{cases} \nu & \text{plane stress} \\ \dfrac{\nu}{1-\nu} & \text{plane strain} \end{cases} \tag{2.9}$$

and that the remaining terms $\mathbf{U}_s = (u_s, v_s)$ and $\mathbf{U}_{br} = (u_{br}, v_{br})$ in (2.5), (2.7) can be further expanded as (cf. [17])

$$\begin{aligned} u_s &= \phi(x) + t^2\phi_{20}(x) + y^2\phi_{02}(x) + \cdots, \\ v_s &= y\varphi(x) + t^2 y\varphi_{21}(x) + y^3\varphi_{03}(x) + \cdots, \end{aligned} \tag{2.10}$$

and

$$\begin{aligned} u_{br} &= t^2 y\theta_{21}(x) + y^3\theta_{03}(x) + \cdots, \\ v_{br} &= t^2 w_{20}(x) + t^4 w_{40}(x) + t^2y^2\psi_{22}(x) + y^4\psi_{04}(x) + \cdots, \end{aligned} \tag{2.11}$$

where the dropped terms are (formally) of order $\mathcal{O}(t^2)$ smaller than the last terms shown. In a normal problem setup, when only t varies, we may assume that after proper rescaling of $\mathbf{U}_s$ and $\mathbf{U}_b$, functions θ, w, ψ in (2.7) and also $\phi, \phi_{ij}, \varphi, \varphi_{ij}, \theta_{ij}, w_{ij}, \psi_{ij}$ in expansions (2.10)–(2.11) are either independent of

t or admit further, simple expansions in terms of t [17]. This is the underlying assumption behind the regularity hypotheses ahead. The actual scaling of $\mathbf{U}_s$ and $\mathbf{U}_b$ is irrelevant for our analysis, since all our regularity assumptions and also the finite element error bounds will be scaling invariant and separate for $\mathbf{U}_s$ and $\mathbf{U}_b$.

Below we refer to $\mathbf{U}_{bm}$ as the *asymptotic bending mode*. As is well known, this mode (as determined by w in view of (2.8)) constitutes the simplest dimension reduction theory for a beam as $t \to 0$. Taking into account (2.8)–(2.9), the strain energy of the asymptotic bending mode can be evaluated from (2.1)–(2.2) as

$$|\!|\!|\mathbf{U}_{bm}|\!|\!|^2 = \tfrac{1}{12} D L^3 t^3 \int_0^L (\theta')^2 dx + \tfrac{1}{320}\bar{\mu}\gamma^2 L^5 t^5 \int_0^L (\theta'')^2 dx, \tag{2.12}$$

where

$$D = \begin{cases} 1 & \text{plane stress} \\ \dfrac{1}{1-\nu^2} & \text{plane strain} \end{cases} \tag{2.13}$$

In what follows we assume that $\theta' \neq 0$ whenever $|\!|\!|\mathbf{U}_b|\!|\!| \neq 0$, so that the leading term in (2.12) is dominant when t is small and θ is smooth enough. The assumption "smooth enough" is made precise by assuming that for some finite Q independent of t,

$$t^{k-2} L^{k-1} \left\{ \int_0^L (\theta^{(k)})^2 dx \right\}^{1/2} \leq Q \left\{ \int_0^L (\theta')^2 dx \right\}^{1/2}, \quad k = 2, 3. \tag{2.14}$$

(Here the case $k = 3$ is added for our later use.) Regarding the remaining terms in (2.5), (2.7), we assume finally that these obey the assumption (2.4) uniformly in t, so that

$$L|\mathbf{U}_s|_{2,\Omega} \leq Q|\!|\!|\mathbf{U}_s|\!|\!|, \qquad L|\mathbf{U}_{br}|_{2,\Omega} \leq Q|\!|\!|\mathbf{U}_b|\!|\!|. \tag{2.15}$$

That assumptions (2.15) are realistic for smooth fields, can be checked from expansions (2.7) and (2.10)–(2.11). Note first that since inequalities (2.14)–(2.15) are scaling invariant, we may assume any rescaling of $\mathbf{U}_s$ and $\mathbf{U}_b$ (separately) as t varies. Suppose that after such rescaling, θ in (2.7) and ϕ in (2.10) are independent of t and that $\theta' \neq 0$ and $\phi' \neq 0$. Assume also that these functions are smooth and that $\phi_{ij}, \varphi, \varphi_{ij}, \theta_{ij}, w_{ij}, \psi_{ij}$ in expansions (2.10)–(2.11) are likewise smooth uniformly with respect to t. Then by (2.12) and by (2.10)–(2.11), $|\mathbf{U}_s|_{2,\Omega} \sim |\!|\!|\mathbf{U}_s|\!|\!| \sim t^{1/2}$ and $|\mathbf{U}_{br}|_{2,\Omega} \sim |\!|\!|\mathbf{U}_b|\!|\!| \sim t^{3/2}$, so estimates (2.15) must hold with some finite Q independent of t.

Our regularity hypotheses are now all set for the case $\bar{\lambda} \leq C$. On thin domains, (2.12) and assumptions (2.14)–(2.15) can be condensed into two simpler hypotheses that will be more convenient for our analysis. We summarize these simpler hypotheses (involving a slightly different constant Q), as well as (2.4), in the following

Hypothesis 2.1 (Regularity assumptions) *Assume the existence of a finite dimensionless constant Q such that (a) on a fixed polygonal domain the exact displacement field $\mathbf{U}$ satisfies*

$$L|\mathbf{U}|_{2,\Omega} \leq Q|||\mathbf{U}|||,$$

where $L = diam(\Omega)$, and (b) on the parametrized rectangular domains with $0 < t \leq 1$, $\mathbf{U}$ satisfies (2.5)–(2.9), where

$$L|\mathbf{U}_s|_{2,\Omega} \leq Q|||\mathbf{U}_s|||,$$

and

$$\left\{\int_0^L \left[t^5L^7(\theta''')^2 + t^3L^5(\theta'')^2 + t^3L^3(\theta')^2\right] dx + L^2|\mathbf{U}_{br}|_{2,\Omega}^2\right\}^{1/2} \leq \sqrt{12}Q|||\mathbf{U}_b|||,$$

where Q is independent of t.

As to the practical relevance of the hypotheses made, we note that in practice, these assumptions may fail in basically in two different ways. First, on a fixed polygonal domain, constant Q may not be finite because of strong corner irregularities, or because of local irregularities do to the load or the kinematical constraints. Secondly, on the thin reference domains, the Saint Venant principle implies that there arises a boundary layer which decays exponentially from $x = 0$ and $x = L$ in a length scale proportional to $H = tL$ (c. f. [17]) In the presence of the layer, constant Q is typically not bounded uniformly as $t \to 0$, even if other irregularities were "turned off".

The inevitable conclusion is that our regularity assumptions are rather optimistic, or even unrealistic, except for benchmark situtations where the solution is known to be smooth (say, a polynomial) a priori. There is, however, another interpretation of our assumptions. Let us assume (quite realistically) that our exact displacement field can be decomposed as

$$\mathbf{U} = \mathbf{U}^{smooth} + \mathbf{U}^{singular} + \mathbf{U}^{layer}, \tag{2.16}$$

where $\mathbf{U}^{singular}$ contains the leading terms of, e. g. the corner irregularities, and the last term represents the Saint Venant boundary layer (or its leading part). Then the finite element scheme, being a linear projection method, may be viewed as projecting the three components of $\mathbf{U}$ separately. Adopting this view, we may consider our regularity assumptions realistic when estimating the projection error of the smooth part $\mathbf{U}^{smooth}$. We take this standpoint below, so that the "exact solution" actually means the smooth part of the exact solution. In the neglected error terms, locking plays less central role, so this simplification is justified. One should note, however, that in practice one can only see the combined effect of the three error terms. Also, when designing remedies to avoid locking, it could happen that a design based on the smooth part only (or on the asymptotic solution) actually causes unwanted error growth in some other part of the solution, such as the layer. Such issues, albeit relevant, will not be addressed here. The reader is referred to [22, 23] for an analysis of a conflict of this kind in a plate bending problem.

Let us finally state briefly the additional regularity hypotheses to be needed when extending our finite element error analysis beyond the assumption $\bar{\lambda} \leq C$. We consider then $\bar{\lambda}$ as another free parameter, so that a fixed-domain problem becomes parametric as well, and the family of thin-domain problems becomes doubly parametric. In such fully parametrized situations, we assume that $\bar{\lambda}$ varies in the range $1 \leq \bar{\lambda} < \infty$.

In case of a fixed domain, we postulate the splitting

$$\mathbf{U} = \mathbf{U}_0 + \bar{\lambda}^{-1}\mathbf{U}_1, \tag{2.17}$$

where $\mathbf{U}_0 = (u_0, v_0)$ satisifes

$$\frac{\partial u_0}{\partial x} + \frac{\partial v_0}{\partial y} = 0, \tag{2.18}$$

and $\mathbf{U}_1$ obeys the bound

$$L|\mathbf{U}_1|_{2,\Omega} \leq Q|||\mathbf{U}|||, \tag{2.19}$$

where now Q is assumed independent of $\bar{\lambda}$. Here constraint (2.18) is known as the *volumetric constraint.*

In case of parametric thin domains, we postulate a similar splitting for $\mathbf{U}_s$ and $\mathbf{U}_{sr}$ in (2.5) and (2.7):

$$\mathbf{U}_s = \mathbf{U}_{s,0} + \bar{\lambda}^{-1}\mathbf{U}_{s,1}, \qquad \mathbf{U}_{br} = \mathbf{U}_{br,0} + \bar{\lambda}^{-1}\mathbf{U}_{br,1}, \tag{2.20}$$

where $\mathbf{U}_{s,0} = (u_{s,0}, v_{s,0})$ and $\mathbf{U}_{br,0} = (u_{br,0}, v_{br,0})$ satisfy

$$\frac{\partial u_{s,0}}{\partial x} + \frac{\partial v_{s,0}}{\partial y} = 0, \qquad \frac{\partial u_{br,0}}{\partial x} + \frac{\partial v_{br,0}}{\partial y} = 0, \tag{2.21}$$

and $\mathbf{U}_{s,1}$, $\mathbf{U}_{br,1}$ obey the bounds

$$L|\mathbf{U}_{s,1}|_{2,\Omega} \leq Q|||\mathbf{U}_s|||, \qquad L|\mathbf{U}_{br,1}|_{2,\Omega} \leq Q|||\mathbf{U}_b|||, \tag{2.22}$$

where Q is assumed independent of both t and $\bar{\lambda}$.

The above assumptions can be formally justified by asymptotic expansion of the solution with respect to the parameter $\varepsilon = \bar{\lambda}^{-1/2}$. The assumptions are realistic when $\mathbf{U}$ is "smooth enough". In practice, one should again think in terms of the basic expansion (2.16) where the last two terms (or their leading parts) must be disregarded when disturbing.

We summarize the additional hypotheses in

Hypothesis 2.2 (Regularity assumptions when $\bar{\lambda} \gg 1$) *When considering a family of probems where $1 \leq \bar{\lambda} < \infty$ is a free parameter, assume that there exists a constant Q independent of both t and $\bar{\lambda}$ such that $\mathbf{U}$ satisfies Hypothesis 2.1. In addition, assume that (a) in case of a fixed polygonal domain, $\mathbf{U}$ satisfies* (2.17)–(2.19)*, and (b) in case of parametrized rectangular domains, $\mathbf{U}_s$ and $\mathbf{U}_{br}$ in* (2.5), (2.7) *satisfy* (2.20)–(2.22).

3 The standard bilinear scheme

In this section we consider the finite element approximation of the problem introduced above, assuming standard (conforming, isoparametric) bilinear quadrilateral elements. The reduced-strain formulation to be introduced later on (in section 5) will also be based on this simple finite element framework.

We will use K as a generic symbol for a quadrilateral ($K \subset \Omega$) taken from a finite element mesh on Ω. As usual, we assume that the quadrilaterals in the same mesh can only touch at a common vertex or along a common edge. We further assume the meshes to be (locally) *regular* so that any interior angle of any K satisfies

$$0 < \zeta \leq \text{angle}(K) \leq \pi - \zeta, \tag{3.1}$$

where ζ is an absolute constant. (The constants C below may sepend on ζ.) We characterize the shape of the quadrilaterals also by two other dimensionless parameters. Let a_K be length of the largest side of K and let c_K be the diameter of the smallest and d_K the diameter of the largest inscribed circle in K that touches three sides of K (there are at most two such circles). We define

$$m_K = a_K/d_K, \quad n_K = d_K/c_K \tag{3.2}$$

and assume that each K is non-degenerate, so that m_K, n_K are finite. In case of a rectangle or a parallelogram one has $n_K = 1$, whereas m_K (referred to as the *aspect ratio*), can obviously be arbitrarily large.

A given mesh that subdivides Ω will be characterized by three global parameters, which are likewise dimensionless. We define the dimensionless *mesh spacing* as

$$h = (\max_K a_K)/L, \tag{3.3}$$

where the max is taken over the mesh and L is the characteristic dimension of the domain (as in section 2). In our error analysis, h is the principal mesh parameter. We use h also as an index for a mesh and for the associated finite element solution. The remaining two mesh parameters are defined as

$$M_h = \max_K m_K, \quad N_h = \max_K n_K, \tag{3.4}$$

where m_K, n_K are defined by (3.2). Some of our error bounds depend on these parameters as well, and our aim is to resolve any such dependence explicitly.

Families of meshes such that M_h and N_h are uniformly bounded will be called *strongly regular* here. This is a standard assumption in finite element theories [5, 4, 10], but here it would be harmful on the parametric thin domains as it would prevent the possible growth of the global mesh aspect ratio M_h as

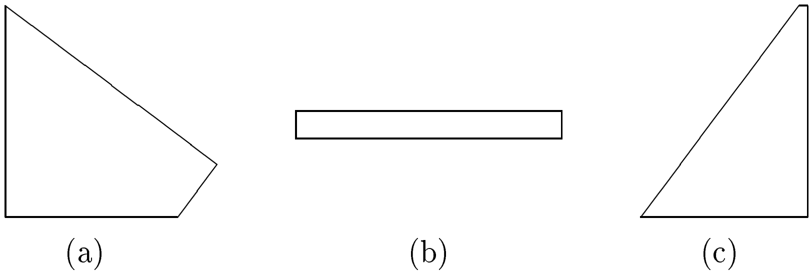

Figure 2: Different quadrilateral shapes: (a) $m_K = 1.9$, $n_K = 1.7$, (b) $m_K = 10$, $n_K = 1$, (c) $m_K = 2.4$, $n_K = 10$.

$t \to 0$. For example, in case of a single-layer rectangular mesh with a fixed number of elements as $t \to 0$, one necessarily has $M_h \sim t^{-1}$. Such meshes are obviously quite reasonable when t is small, so we do not want to rule them out.

We denote by $\mathcal{U}_h$ the finite element space of continuous piecewise bilinear fields associated to a given mesh with index h. This is a subspace of the energy space $\mathcal{U}$ introduced above. Let $\mathbf{U} = (u, v) \in \mathcal{U}$ be a displacement field (component) that satisfies Hypothesis 2.1. Then the finite element projection $\mathbf{U}_h = (u_h, v_h) \in \mathcal{U}_h$ of $\mathbf{U}$ is defined so that u_h interpolates u and/or v_h interpolates v at each node where the corresponding kinematical constraint is imposed in the original problem, and

$$\mathcal{A}(\mathbf{U}_h, \mathbf{V}) = \mathcal{A}(\mathbf{U}, \mathbf{V}), \quad \mathbf{V} \in \mathcal{U}_h^0, \tag{3.5}$$

where $\mathcal{U}_h^0 \subset \mathcal{U}_h$ is a subspace where homogeneous kinematical constraints are imposed.

We note that the right side of (3.5) may be viewed as the "generalized load" corresponding to field component $\mathbf{U}$. When summing up the the corresponding components of $\mathbf{U}_h$ as defined by (3.5), we then obtain the finite element formulation of the original problem where the right side of (3.5) equals the external load potential. The kinematical constraints come out correctly as well, by the assumed distribution of the constraints among the subproblems. Thus we may always think of expanding the finite element solution in this way, once the expansion of the exact solution is given so that the generalized loads are determined. Even though such expansions are rarely computed in practice, the mere thought helps in organizing our error analysis. We will assume the basic expansion (2.16) where we ignore the last two terms as noted. In the further expansion of $\mathbf{U}^{smooth}$ according to (2.5), (2.7), the projection principle (3.5) again applies.

We now turn to the error analysis of scheme (3.5). We will measure the error by the relative (scaling invariant) error indicator

$$e = \frac{\|\mathbf{U} - \mathbf{U}_h\|}{\|\mathbf{U}\|}, \tag{3.6}$$

where $\|\cdot\|$ is the energy norm defined by (2.2). The projection principle in (3.5) implies that $\mathbf{U}_h$ is the best approximation to $\mathbf{U}$ in the energy norm, under the assumed kinematical constraints. Since the constraints were interpolation constraints, it follows that

$$\|\mathbf{U} - \mathbf{U}_h\| \leq \|\mathbf{U} - \tilde{\mathbf{U}}\|, \tag{3.7}$$

where $\tilde{\mathbf{U}} \in \mathcal{U}_h$ is the (standard) interpolant of $\mathbf{U}$. By (2.1)–(2.2), the right side of (3.7) is bounded as

$$\|\mathbf{U} - \tilde{\mathbf{U}}\| \leq C|\mathbf{U} - \tilde{\mathbf{U}}|_{1,\Omega}, \tag{3.8}$$

where $C = \sqrt{2}(\bar{\lambda} + \bar{\mu})^{1/2}$, and the (Sobolev) seminorm $|\cdot|_{1,\Omega}$ is defined as

$$|\mathbf{U}|_{1,\Omega} = \left\{ \int_{\Omega} \left[\left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2 + \left(\frac{\partial v}{\partial x}\right)^2 + \left(\frac{\partial v}{\partial y}\right)^2 \right] \, dxdy \right\}^{1/2}. \tag{3.9}$$

Applying next standard FE approximation theory to bound the right side of (3.8), we get

$$\|\mathbf{U} - \tilde{\mathbf{U}}\| \leq ChL|\mathbf{U}|_{2,\Omega}, \tag{3.10}$$

where now constant C depends also on the finite element mesh, staying uniformly bounded on strongly regular meshes [5, 4]. Joining finally (3.7), (3.10) with Hypothesis 2.1, we conclude that on a fixed polygonal domain and on strongly regular meshes

$$e \leq CQh, \tag{3.11}$$

where C is proportional to $(\bar{\lambda} + \bar{\mu})^{1/2}$. For bilinear elements, this bound is not improvable.

When approaching thin domains, we should first note that estimates (3.10) and (3.11) actually remain valid on just regular meshes, i. e. , the constant C in (3.10) only depends on the parameter ζ in the angle condition (3.1). This sharper result – which is of importance to us – is somewhat beyond standard finite element approximation theory. We refer to [1], the lonely classic in the field of angle conditions. (The result we need is actually not stated in [1], but the main argument there is applicable.)

We start the analysis on a thin reference domain by splitting the finite element solution in analogy with (2.5) as

$$\mathbf{U}_h = \mathbf{U}_{sh} + \mathbf{U}_{bh}, \tag{3.12}$$

where $\mathbf{U}_{sh}, \mathbf{U}_{bh}$ are defined by the projection principle, as explained above. (If the mesh is symmetric with respect to the line $y = 0$, projection splitting is quivalent to parity splitting.) Defining

$$A_s = \frac{|||\mathbf{U}_s|||}{|||\mathbf{U}|||}, \quad A_b = \frac{|||\mathbf{U}_b|||}{|||\mathbf{U}|||}, \tag{3.13}$$

we may then estimate the error, using the triangle inequality, as

$$e \leq A_s e_s + A_b e_b,$$

where

$$e_s = \frac{|||\mathbf{U}_s - \mathbf{U}_{sh}|||}{|||\mathbf{U}_s|||}, \quad e_b = \frac{|||\mathbf{U}_b - \mathbf{U}_{bh}|||}{|||\mathbf{U}_b|||},$$

so that by the above reasoning

$$e \leq Ch\left(A_s \frac{L|\mathbf{U}_s|_{2,\Omega}}{|||\mathbf{U}_s|||} + A_b \frac{L|\mathbf{U}_b|_{2,\Omega}}{|||\mathbf{U}_b|||}\right).$$

Applying here the expansion (2.7) of $\mathbf{U}_b$, recalling (2.8) and using Hypothesis 2.1, we conclude that

$$\frac{L|\mathbf{U}_s|_{2,\Omega}}{|||\mathbf{U}_s|||} \leq CQ, \quad \frac{L|\mathbf{U}_b|_{2,\Omega}}{|||\mathbf{U}_b|||} \leq CQt^{-1}, \tag{3.14}$$

so we obtain the final error bound

$$e \leq CQh(A_s + A_b t^{-1}), \tag{3.15}$$

where C is again proportional to $(\bar{\lambda} + \bar{\mu})^{1/2}$.

From (3.15) we conclude that when $\bar{\lambda} \leq C$, we have the uniformly optimal error bound $e = \mathcal{O}(h)$ in the pure stretching state ($A_b = 0$) or more generally in deformation states such that $A_b = \mathcal{O}(t)$ as $t \to 0$. In the pure bending state ($A_s = 0$), we can only quarantee that that $e = \mathcal{O}(h/t)$, which is also the worst-case error bound for general deformation states. In practice the relative magnitude of A_s, A_b (note that $A_s^2 + A_b^2 = 1$ by (3.13) and (2.6)) varies from case to case, depending on the loading and kinematical constraints, cf. [17].

We summarize the convergence results in

Theorem 3.1 (Convergence of standard FEM) *Let* $\mathbf{U}$ *be a displacement field satisfying Hypothesis 2.1 and let* $\mathbf{U}_h$ *be the finite element approximation of* $\mathbf{U}$ *based on standard isoparametric bilinear elements on a regular mesh with dimensionless mesh spacing* h*. Then the relative error in the energy norm, as defined by* (3.6)*, satisfies either* (3.11) *(fixed polygonal domain) or* (3.15) *(thin reference domain), where* A_s, A_b *are the relative amplitudes in the splitting* (2.5)*, as defined by* (3.13)*.*

The error amplification predicted by estimate (3.15) when $A_b \sim 1$ is a well known (locking) phenomenon. The engineering literature on this phenomenon is numerous and, as noted, essentially starts from the very first finite element formulations in structural mechanics. Our aim so far has been to mathematically describe this locking effect. As we have seen, the error amplification by factor t^{-1} arises from the second inequality in (3.14) (which is not improvable). This inequality states that in the pure bending state of deformation, the energy norm (arising from the physical formulation of the problem) scales unfavorably as compared with the seminorm $|\cdot|_{2,\Omega}$ (arising from polynomial interpolation theory). On the other hand, the first inequality in (3.14) tells that this conflict does *not* arise in the pure stretching state of deformation.

We note finally that estimates (3.11) and (3.15) predict parametric error amplification also when $\bar{\lambda} \to \infty$, since constant C in these estimates was proportional to $(\bar{\lambda} + \bar{\mu})^{1/2}$. As is well known, this is also a true phenomenon, often named *volumetric locking*. In the next section we show that the "thickness locking" at small t actually involves two independent locking modes. Thus we can actually count (at least) *three* different locking effects in problems where both t and $\bar{\lambda}$ are active parameters.

4 Locking on thin domains: lower error bounds

In this section we isolate in more detail the sources of the locking effect predicted by estimate (3.15) when t is small. In this context we prove also *lower* error bounds for the case of a single-layer finite element mesh on the reference domain (Theorem 4.1 below). The lower error bounds underline that the locking effects we isolate are true phenomena.

The analysis in this section is quite parallel with the more descriptive treatment given in [15, Ch 6], see also [16]. Albeit reproducing the big picture from these references, our twin analysis turns out to be useful in it self. It gives us some guidelines on how to "unlock" the FE scheme by means of reduced strains. Below we adopt some of the language in [15] when describing the locking phenomena.

So far we have seen that locking, i. e., error amplification by factor t^{-1}, arises from the term $\mathbf{U}_b$ in (2.5). We can immediately sharpen the picture from the further expansion (2.7). By assumption (2.15) together with the splitting & interpolation error arguments of section 3, the projection error of the remainder term in (2.7) cannot contribute to the locking effect. Hence we may drop this remainder from our analysis and assume that we are approximating an asymptotic bending mode $\mathbf{U}_{bm}$ only. Thus let the field $\mathbf{U} = (u, v)$ to be approximated be of the form

$$\begin{aligned} u &= y\theta(x), \\ v &= w(x) + \tfrac{1}{2}y^2\psi(x), \end{aligned} \tag{4.1}$$

where θ, w, ψ satisfy constraints (2.8).

We may rewrite constraints (2.8) in terms of u, v in (4.1) as

$$\gamma\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} = 0, \tag{4.2}$$

$$\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x} = \tfrac{1}{2}y^2\psi'(x). \tag{4.3}$$

Here the right side of (4.3) is of relative order $\mathcal{O}(t^2)$ on Ω, so when $t \to 0$, (4.3) takes the asymptotic form $\partial u/\partial y + \partial v/\partial x = 0$, which is known as the (asymptotic) *shear constraint*. We may call (4.2) a *dilatation constraint*. Both of these constraints have a role in the locking of the standard bilinear element, as we shall show by an example.

We will consider the approximation of the simplest nontrivial bending mode, where θ, w, ψ in (4.1) are defined by

$$\begin{aligned} \theta(x) &= \theta_0\,(x - \tfrac{1}{2}L), \\ w(x) &= \tfrac{1}{2}\theta_0\, x(L - x), \\ \psi(x) &= -\gamma\theta_0, \end{aligned} \tag{4.4}$$

where $\theta_0 \neq 0$ is a constant coefficient. (With proper θ_0, this is known as the unit bending mode.) Then $\partial u/\partial y + \partial v/\partial x = 0$, and the deformation energy (2.1) takes the simple form

$$\|\mathbf{U}\|^2 = D \int_\Omega \Big(\frac{\partial u}{\partial x}\Big)^2 dxdy, \tag{4.5}$$

where D is defined by (2.13).

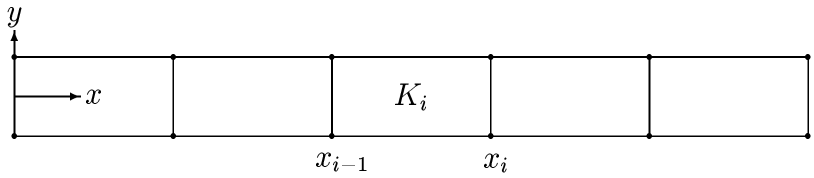

Figure 3: A single-layer rectangular mesh.

We now approach the above problem numerically assuming that Ω is subdivided by just a single layer of rectangular elements

$$K_i = (x_{i-1} < x < x_i, \quad -H/2 < y < H/2),$$

where $0 = x_0 < x_1 < \cdots < x_I = L$. We assume further the subdivision to be uniform so that $x_i = ai, \quad i = 0 \ldots I, \quad a = L/I$ (Fig. 3). For parity reasons, the finite element solution $\mathbf{U}_h = (u_h, v_h) \in \mathcal{U}_h$ must then be of the form

$$\begin{aligned} u_h &= y\theta_h(x), \\ v_h &= w_h(x), \end{aligned} \tag{4.6}$$

where θ_h, w_h are continuous piecewise linear functions on the subdivision of $[0, L]$ imposed by points x_i.

We aim at proving lower bounds for $\|\mathbf{U} - \mathbf{U}_h\|/\|\mathbf{U}\|$ in the above example case. The first step is to rewrite the strain energy (2.1) in terms $\partial u/\partial x$ and the strain components appearing in (4.2), (4.3) (recall also (2.9)) as

$$\mathcal{A}(\mathbf{U}, \mathbf{U}) = \int_\Omega \Big\{ D\Big(\frac{\partial u}{\partial x}\Big)^2 + \frac{\bar{\lambda}}{\gamma}\Big(\gamma\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\Big)^2 + \bar{\mu}\Big(\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x}\Big)^2 \Big\} dxdy. \tag{4.7}$$

In view of (4.1)–(4.4) and (4.7), the above finite element approximation problem may be formulated qualitatively as: Find $(u_h, v_h) \in \mathcal{U}_h$ of the form (4.6) such that

$$\begin{aligned} \frac{\partial u_h}{\partial x} &\approx \frac{\partial u}{\partial x}, \\ \gamma\frac{\partial u_h}{\partial x} + \frac{\partial v_h}{\partial y} &\approx \gamma\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} = 0, \\ \frac{\partial u_h}{\partial y} + \frac{\partial v_h}{\partial x} &\approx \frac{\partial u}{\partial y} + \frac{\partial v}{\partial x} = 0. \end{aligned} \tag{4.8}$$

Note that if we impose no kinematical constraints in this example case, then the three approximation problems (4.8) are trivial each alone: Simply choose $u_h = u$ in the first problem (possible since u is bilinear) and $u_h = v_h = 0$ in the others. Instead, when looking for a minimal energy solution where all the three conditions are given the same weight in the least squares sense, good approximability is less obvious. In fact, our aim is to prove that the approximability is rather poor. To this end, we split (4.8) in two subproblems.

Consider first the constrained approximation problem

$$\frac{\partial u_h}{\partial x} \approx \frac{\partial u}{\partial x}, \quad \gamma\frac{\partial u_h}{\partial x} + \frac{\partial v_h}{\partial y} \approx 0, \tag{4.9}$$

which is part of (4.8). This corresponds to energy projection where the last term of the energy density in (4.7) is dropped. Consider then a given element K_i and write

$$\begin{aligned} \theta_h(x) &= b_0 + b_1(x - x_{i-\frac{1}{2}}), \\ w_h(x) &= c_0 + c_1(x - x_{i-\frac{1}{2}}), \quad x_{i-1} < x < x_i, \end{aligned} \tag{4.10}$$

where b_0, b_1, c_0, c_1 are constants and $x_{i-\frac{1}{2}} = \frac{1}{2}(x_{i-1} + x_i)$. Then by (4.1), (4.4), (4.2) and (4.10),

$$\begin{aligned} \int_{K_i} &\Big\{ D\Big[\frac{\partial(u-u_h)}{\partial x}\Big]^2 + \frac{\bar\lambda}{\gamma}\Big[\gamma\frac{\partial(u-u_h)}{\partial x} + \frac{\partial(v-v_h)}{\partial y}\Big]^2\Big\}\, dxdy \\ &= \int_{K_i} \Big\{ D\Big(\frac{\partial u}{\partial x} - \frac{\partial u_h}{\partial x}\Big)^2 + \bar\lambda\gamma\Big(\frac{\partial u_h}{\partial x}\Big)^2\Big\}\, dxdy \\ &= \tfrac{1}{12}aH^3[D(\theta_0 - b_1)^2 + \bar\lambda\gamma b_1^2] \\ &\geq \tfrac{1}{12}D\gamma^2 aH^3\theta_0^2 \\ &= D\gamma^2\int_{K_i}\Big(\frac{\partial u}{\partial x}\Big)^2 dxdy, \end{aligned}$$

so that summing over K_i and using (4.5),

$$\begin{aligned} \int_{\Omega} &\Big\{ D\Big[\frac{\partial(u-u_h)}{\partial x}\Big]^2 + \frac{\bar\lambda}{\gamma}\Big[\gamma\frac{\partial(u-u_h)}{\partial x} + \frac{\partial(v-v_h)}{\partial y}\Big]^2\Big\}\, dxdy \\ &\geq D\gamma^2\int_{\Omega}\Big(\frac{\partial u}{\partial x}\Big)^2 dxdy \\ &= \gamma^2\|\mathbf{U}\|^2. \end{aligned} \tag{4.11}$$

Next consider the constrained approximation problem (the second part of (4.8))

$$\frac{\partial u_h}{\partial x} \approx \frac{\partial u}{\partial x}, \quad \frac{\partial u_h}{\partial y} + \frac{\partial v_h}{\partial x} \approx 0, \tag{4.12}$$

interpreted as the energy projection problem where now the middle term in (4.7) is dropped. Using (4.1), (4.4), (4.3) and (4.10), we then get

$$\begin{aligned}\int_{K_i} \Big\{ D\Big[\frac{\partial(u-u_h)}{\partial x}\Big]^2 + \bar{\mu}\Big[\frac{\partial(u-u_h)}{\partial y} + \frac{\partial(v-v_h)}{\partial x}\Big]^2 \Big\}\, dxdy \\ &= \tfrac{1}{12} D a H^3 (\theta_0 - b_1)^2 + \tfrac{1}{12}\bar{\mu} a^3 H b_1^2 + \bar{\mu} a H (b_0 + c_1)^2 \\ &\geq \frac{1}{12}\,\frac{D\bar{\mu} a^3 H^3}{\bar{\mu} a^2 + D H^2}\,\theta_0^2 \\ &= \frac{D a^2}{a^2 + (D/\bar{\mu})H^2}\int_{K_i}\Big(\frac{\partial u}{\partial x}\Big)^2\, dxdy,\end{aligned}$$

so that again summing over K_i and using (4.5),

$$\begin{aligned}\int_{\Omega} \Big\{ D\Big[\frac{\partial(u-u_h)}{\partial x}\Big]^2 + \bar{\mu}\Big[\frac{\partial(u-u_h)}{\partial y} + \frac{\partial(v-v_h)}{\partial x}\Big]^2 \Big\}\, dxdy \\ &\geq \frac{a^2}{a^2 + (D/\bar{\mu})H^2}\,|\!|\!|\mathbf{U}|\!|\!|^2. \qquad (4.13)\end{aligned}$$

To sum up, we conclude that estimates (4.11) and (4.13) hold for any $(u_h, v_h) \in \mathcal{U}_h$ of the form (4.6) and thus also for the the finite element projection $\mathbf{U}_h$. Therefore, and since the left sides of (4.11) and (4.13) are strict lower bounds for $|\!|\!|\mathbf{U} - \mathbf{U}_h|\!|\!|^2$, we come to the following absolute lower error bound:

$$e \geq \max\Big\{\gamma,\ \Big[\frac{a^2}{(a^2 + (D/\bar{\mu})H^2}\Big]^{\frac{1}{2}}\Big\}. \qquad (4.14)$$

As is obvious from the above reasoning, the two lower bounds in (4.14) arise, because there exist no accurate solutions to the constrained approximation problems (4.9) or (4.12). The failure in (4.12), which gave rise to the second lower bound in (4.14), is usually referred to as *shear locking*. The failure in (4.9) causing the first lower bound in (4.14) we may name *dilatation locking* [15]. This locking mode is not present if $\gamma = 0\,(\nu = 0)$, because the finite element solution (4.6) automatically satisfies constraint (4.2) in that case. More generally, if $\gamma = 0$, dilatation locking is absent on vertically aligned layered meshes but still present on more general meshes.

As to the further interpretation of the lower bound (4.14), we note that if $a \geq H$, so that $h = \max\{a/L,\ H/L\} \geq H/L = t$, (4.14) gives

$$e \geq (1 + D/\bar{\mu})^{-\frac{1}{2}},$$

showing that the upper error bound (3.15) cannot be improved in the pure bending state when $h/t \sim 1$. (If $a < H$ so that $h = t$, estimate (3.15) is neither improvable when $\gamma \neq 0$.) When $a \gg H$ so that $h \gg t$, estimates (4.14) and (3.15) imply that $C^{-1} \leq e \leq Ch/t$ for a finite constant C, so there remains a slight gap. Although closing the gap may not be of great practical interest, let us note that the estimates can still be sharpened in two possible ways.

First, we note that the lower error bound derived above was not necessarily sharp if kinematical constraints were imposed at $x = 0, L$. For example, suppose we had (built-in) constraints $u_h = u$, $v_h = v$ at the four corners of Ω in the above example case. Then if we choose $a = L/2\,(h = 1/2)$, the finite element solution above is computable as

$$\begin{aligned} u_h &= u, \\ v_h &= \tfrac{1}{8}\theta_0\,(L^2 - \gamma H^2) - \tfrac{1}{4}\theta_0 L\,|x - L/2|, \end{aligned}$$

giving

$$e = \left[\bar{\lambda}\gamma/D + (\bar{\mu}/D)(h/t)^2\right]^{1/2},$$

so the upper bound (3.15) is sharp in this case. Note as well that here $e \to \infty$ also when $\bar{\lambda} \to \infty$, so the dilatation/volumetric locking effect is stronger than predicted in (4.14).

There are also situations where the lower bound (4.14) remains sharp for $h \gg t$. This happens in particular when the kinematical constraints are either missing or happen to come out as homogeneous constraints when applied to (u, v) (as given by (4.1)). Then the finite element solution $\mathbf{U}_h$ is a true projection with maximal error occurring when $\mathbf{U}_h = 0$. Hence the gap between (3.15) and (4.14) is closed this time by the additional upper error bound $e_b \leq 1$.

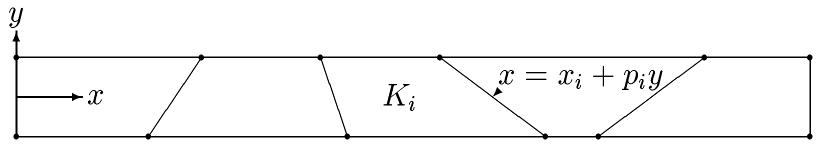

Figure 4: A single-layer trapezoidal mesh.

So far we assumed rectangular mesh. On more general meshes, there arises a locking effect of a different (third) kind when approximating the asymptotic bending mode. As an example, let us consider a single-layer mesh composed of the quadrilaterals (see Fig. 4)

$$K_i = \left\{(x, y) \mid -H/2 < y < H/2, \quad x_{i-1} + p_{i-1}y < x < x_i + p_i y\right\}. \quad (4.15)$$

In this case it turns out that when $p_i - p_{i-1} \neq 0$ (i. e. , when the elements are not parallelograms), there arises locking already in the unconstrained approximation $\partial u_h/\partial x \approx \partial u/\partial x$ in (4.8). Namely, a straightforward computation shows that when $u = \theta_0\, y\,(x - \frac{1}{2}L)$ (as above) and when u_i is any isoparametric bilinear function on K_i, one has the (sharp) lower bound

$$\int_{K_i} \left[\frac{\partial(u - u_i)}{\partial x}\right]^2 dxdy \geq c_i \int_{K_i} \left(\frac{\partial u}{\partial x}\right)^2 dxdy,$$

where c_i, defined in terms of

$$\delta_i = \tfrac{1}{2}|p_i - p_{i-1}| \, \frac{H}{x_i - x_{i-1}} \qquad (0 \le \delta_i < 1) \tag{4.16}$$

as

$$\begin{aligned} c_i &= 1 - \tfrac{2}{3}\delta_i^3 (1 - \tfrac{1}{3}\delta_i^2)^{-1} \Big[\log \frac{1+\delta_i}{1-\delta_i} - 2\delta_i \Big]^{-1} \\ &= 1 - \Big[1 + \sum_{k=1}^{\infty} \tfrac{4k}{(2k+1)(2k+3)} \delta_i^{2k} \Big]^{-1} \end{aligned}$$

satisfies

$$\tfrac{1}{5}\delta_i^2 \le c_i < 1.$$

Hence if all elements are uniformly trapezoidal so that $\delta_i \ge \delta > 0$ for all i, we must have

$$\int_\Omega \Big[\frac{\partial(u - \tilde{u})}{\partial x} \Big]^2 dxdy \ge \tfrac{1}{5}\delta^2 \int_\Omega \Big(\frac{\partial u}{\partial x} \Big)^2 dxdy$$

for any piecewise bilinear $\tilde{u}$, so that

$$\begin{aligned} |||\mathbf{U} - \tilde{\mathbf{U}}|||^2 &\ge D \int_\Omega \Big[\frac{\partial(u - \tilde{u})}{\partial x} \Big]^2 dxdy \\ &\ge \tfrac{1}{5}\delta^2 D \int_\Omega \Big(\frac{\partial u}{\partial x} \Big)^2 dxdy \\ &= \tfrac{1}{5}\delta^2 |||\mathbf{U}|||^2 \end{aligned} \tag{4.17}$$

for any $\tilde{\mathbf{U}} = (\tilde{u}, \tilde{v}) \in \mathcal{U}_h$. This approximation failure, named *trapezoidal locking* in [15], obviously has nothing to do with constraints (4.2), (4.3). These constraints rather contribute to further error amplification when taken into account.

We have now resolved the modes of "thickness locking" appearing on thin domains when $t \to 0$. Let us finally confirm that there arises parametric error amplification also when $\bar{\lambda} \to \infty$. This *volumetric locking* was predicted both in (3.11) (fixed domain) and in (3.15) (thin domain), since $C \sim (\bar{\lambda} + \bar{\mu})^{1/2}$ in these estimates. In the present context, the volumetric locking effect is most easily demonstrated in the stretching state of deformation where other locking modes are not present.

Consider the simple stretching state

$$u = \tfrac{1}{2}\phi_0 \, x(L - x), \quad v = \phi_0 \, y \, (x - \tfrac{1}{2}L), \tag{4.18}$$

where $\phi_0 \neq 0$. Here $\partial u/\partial x + \partial v/\partial y = 0$, so by (2.1)–(2.2), $\mathcal{A}(\mathbf{U}, \mathbf{U}) = |||\mathbf{U}|||^2$ stays bounded as $\bar{\lambda} \to \infty$. (In fact, $\mathbf{U}$ satisfies Hypothesis 2.2.) Assume a single-layer rectangular mesh as above (Fig. 3). Then the finite element solution $\mathbf{U}_h = (u_h, v_h)$ takes the form

$$u_h = \phi_h(x), \quad v_h = y\varphi_h(x),$$

where ϕ_h, φ_h are continuous piecewise linear functions. Assume in addition that no kinematic constraints are imposed at the boundary. Then by (2.1)–(2.2) and by the projection principle,

$$\int_\Omega \bar{\lambda}\Big(\frac{\partial u_h}{\partial x} + \frac{\partial v_h}{\partial y}\Big)^2 \, dxdy \leq \|\mathbf{U}_h\|^2 \leq \|\mathbf{U}\|^2.$$

Here the right side is uniformly bounded, so it follows that

$$\frac{\partial u_h}{\partial x} + \frac{\partial v_h}{\partial y} = \phi'_h + \varphi_h \to 0 \quad \text{as } \bar{\lambda} \to \infty,$$

which is possible only if $\phi'_h \to c$ and $\varphi_h \to -c$ for some constant c. But we must have $c = 0$ by symmetry, so it follows that $\|\mathbf{U}_h\| \to 0$, hence $e = \|\mathbf{U} - \mathbf{U}_h\|/\|\mathbf{U}\| \to 1$ as $\bar{\lambda} \to \infty$. Thus the error behaves nonuniformly with respect to $\bar{\lambda}$, as predicted.

We summarize the analysis in this section in

Theorem 4.1 (Lower error bounds) *(a) Let $\mathbf{U}_h$ be the standard bilinear finite element approximation of the asymptotic bending mode* (4.1), (4.4) *on a single-layer mesh of quadrilaterals K_i, $i = 1 \ldots I$, defined by* (4.15), *where $x_i = ai$, $a = L/I$. Then if $p_i = 0$ for all i so that the elements are rectangular, the lower error bound* (4.14) *holds for the relative error in the energy norm. If the elements are uniformly trapezoidal so that $\delta_i \geq \delta > 0, i = 1 \ldots I$, where δ_i is defined by* (4.16), *then the additional approximation-error lower bound* (4.17) *holds for any $\tilde{\mathbf{U}} = (\tilde{u}, \tilde{v}) \in \mathcal{U}_h$. (b) Assume a single-layer rectangular mesh as in (a), and let $\mathbf{U}_h$ be the standard bilinear FE approximation of the of the pure stretching mode* (4.18). *Then if no kinematic constraints are imposed at the boundary, the relative error in the energy norm satisfies $e \to 1$ as $\bar{\lambda} \to \infty$.*

Regarding thin domains, our expedition has now reached its first goal. We obtained an error bound, estimate (3.15), and we verified by counterexamples that this bound is essentially sharp when $\bar{\lambda} \leq C$. We also saw that the error amplification by factor $\sim t^{-1}$, in the pure bending case, was actually due to the asymptotic bending mode $\mathbf{U}_{bm}$. Here we further isolated the source of numerical locking in the two constrained approximation problems (4.9) and (4.12). Finally, we detected the additional volumetric locking in the stretching state of deformation when $\bar{\lambda} \to \infty$.

Having now characterized the basic locking phenomena mathematically, we set a new goal with some engineering flavor as well. We look for remedies.

5 The reduced-strain formulation

When the standard finite element scheme suffers from parametric locking, a remedy may be attempted by modifying the strain energy numerically as

$$\mathcal{A}(\mathbf{U},\mathbf{U}) \hookrightarrow \mathcal{A}_h(\mathbf{U},\mathbf{U}), \tag{5.1}$$

where the subscript indicates that the modification is mesh dependent. We consider modifications of this kind within the conforming bilinear FE framework, i. e., the finite element spaces are not changed.

In the approach to be taken, modification (5.1) will be done in two steps. The first and primary step is to impose numerical *strain reduction* operators on selected strain components. The aim is to slightly weaken the energy norm so that numerical locking is avoided in the weakened (semi)norm. The second step is to supply the strain energy with additional (again purely numerical) "correction terms", each multiplied by a free parameter. The added terms are chosen so that they have little effect in the approximation error. The free parameters can then be selected on the basis of stability and consistency error analysis, as will be shown.

In our approach, modification (5.1) starts from the alternative strain energy formulation (4.7), where we know that the locking is due to the last two terms in the energy density. We then leave the first term untouched and modify the critical strain terms as

$$\begin{aligned} \gamma\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} &\hookrightarrow R_h\Big(\gamma\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\Big), \\ \frac{\partial u}{\partial y} + \frac{\partial v}{\partial x} &\hookrightarrow S_h\Big(\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x}\Big). \end{aligned} \tag{5.2}$$

Here R_h and S_h are reduction operators to be found, with the minimal requirement that both operators should be true reductions, i. e., weaker than identity when acting on strains in the finite element space. Indeed, by the analysis of section 4, both constraints in (4.8) need to be strictly weakened to avoid locking. The actual "design" of the reduction operators could be based on spying the error analysis ahead, but our choice is actually very simple: We choose both R_h and S_h (henceforth $= R_h$) to be local averaging operators so that for each quadrilateral K of a given mesh

$$R_h\,\sigma(x,y) = \frac{1}{\mathrm{area}(K)}\int_K \sigma\,dx'dy', \quad (x,y)\in K. \tag{5.3}$$

Obviously then $S_h = R_h$ acts as a true reduction in (5.2) when u, v are piecewise bilinear.

The second step in (5.1) consists of adding two supplementary terms in the modified strain energy functional. These are chosen to be

$$
\begin{aligned}
\mathcal{B}_1(\mathbf{U},\mathbf{U}) &= 2\beta_1\bar{\mu}\int_\Omega\Big[\Big(\frac{\partial u}{\partial x}\Big)^2 - \Big(R_h\frac{\partial u}{\partial x}\Big)^2\Big]\,dxdy, \\
\mathcal{B}_2(\mathbf{U},\mathbf{U}) &= 2\beta_2\bar{\mu}\int_\Omega\Big[\Big(\frac{\partial v}{\partial y}\Big)^2 - \Big(R_h\frac{\partial v}{\partial y}\Big)^2\Big]\,dxdy,
\end{aligned}
\tag{5.4}
$$

where β_1, β_2 are so far arbitrary (dimensionless and real) parameters. The additional terms will again be justified by the error analysis ahead, which also imposes some restrictions on the two free parameters. At this point we note only that since $R_h - Identity = \mathcal{O}(h)$ when acting on smooth functions, the added terms are expected to be relatively harmless by the usual (non-parametric) equilibrium-test reasoning [15].

In what follows, it is convenient to pass to a shorthand notation involving L_2 inner products and norms. As usual, $L_2(\Omega)$ denotes the space of square integrable (scalar) functions supplied with the inner product and norm

$$
(f,g) = \int_\Omega fg\,dxdy, \quad \|f\| = (f,f)^{1/2}.
$$

Using this notation, the modified strain energy as described above takes the form

$$
\begin{aligned}
\mathcal{A}_h(\mathbf{U},\mathbf{U}) = D\Big\|\frac{\partial u}{\partial x}\Big\|^2 &+ \frac{\bar{\lambda}}{\gamma}\Big\|R_h\Big(\gamma\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\Big)\Big\|^2 + \bar{\mu}\Big\|R_h\Big(\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x}\Big)\Big\|^2 \\
&+ 2\beta_1\bar{\mu}\Big(\Big\|\frac{\partial u}{\partial x}\Big\|^2 - \Big\|R_h\frac{\partial u}{\partial x}\Big\|^2\Big) + 2\beta_2\bar{\mu}\Big(\Big\|\frac{\partial v}{\partial y}\Big\|^2 - \Big\|R_h\frac{\partial v}{\partial y}\Big\|^2\Big).
\end{aligned}
\tag{5.5}
$$

Note that here R_h is interpreted globally as the L_2 projection (orthogonal projection) into the space of piecewise constant-valued functions on the finite element mesh.

The modified strain energy can also be written in an alternative form that closely resembles the original form (2.1) of the strain energy. As readily verified, this equivalent form is

$$
\begin{aligned}
\mathcal{A}_h(\mathbf{U},\mathbf{U}) = \bar{\lambda}\Big\|R_h\Big(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\Big)\Big\|^2 &+ 2\bar{\mu}\Big\|R_h\frac{\partial u}{\partial x}\Big\|^2 + 2\bar{\mu}\Big\|R_h\frac{\partial v}{\partial y}\Big\|^2 \\
&+ \bar{\mu}\Big\|R_h\Big(\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x}\Big)\Big\|^2 \\
+ 2\alpha_1\bar{\mu}\Big(\Big\|\frac{\partial u}{\partial x}\Big\|^2 - \Big\|R_h\frac{\partial u}{\partial x}\Big\|^2\Big) &+ 2\alpha_2\bar{\mu}\Big(\Big\|\frac{\partial v}{\partial y}\Big\|^2 - \Big\|R_h\frac{\partial v}{\partial y}\Big\|^2\Big),
\end{aligned}
\tag{5.6}
$$

where

$$
\alpha_1 = \beta_1 + \gamma + 1, \quad \alpha_2 = \beta_2. \tag{5.7}
$$

By (5.6), we could have derived our modified scheme directly from (2.1), by first averaging each strain term and then adding the supplementary terms. This approach (taken in [2]) would in fact be natural when programming. Our more laborious derivation is natural for the theory instead: We need formulation (5.5) when dealing with the locking problem on thin domains (sections 9 and 10 ahead). In the stability analysis (section 6 & Appendix A) and when bounding the error in case of a fixed polygonal body (sections 7 and 8), formulation (5.6) is somewhat more convenient and will be used.

At this point, let us interpret the main results of our analysis ahead in view of formulation (5.6). First, the stability analysis shows that the scheme based on (5.6) is stable if and only if (Theorem 6.1 ahead)

$$\alpha_1 > 0, \quad \alpha_2 > 0. \tag{5.8}$$

Hence the added terms in (5.6) have the role of *stabilizing* terms. However, there is more: The analysis in section 9 shows that one actually needs to set $\alpha_1 = \gamma + 1$ (corresponding to $\beta_1 = 0$ in (5.5) by (5.7)) to prevent the consistency error from being amplified by factor t^{-1} on thin domains. Only when this choice is made, the scheme based on (5.6) converges unifromly in the beam bending state (on rectangular meshes, see Theorems 9.1 and 10.1). The result still holds for a family of algorithms, since parameter α_2 is only restricted by the stability constraint (5.8). A rather natural choice, however, is $\alpha_2 = \alpha_1 = \gamma + 1$, in which case (recall (2.9), (2.13))

$$2\alpha_1\bar{\mu} = 2\alpha_2\bar{\mu} = D = \begin{cases} 1 & \text{(plane stress)} \\ \dfrac{1}{1-\nu^2} & \text{(plane strain)} \end{cases} \tag{5.9}$$

When this choice is made, formulation (5.6) (and hence also formulation (5.5) & (5.7)) is invariant under switching of the coordinates, so it then applies to horizontally thin rectangular bodies as well. To handle a more general body consisting of thin rectangular parts, it is necessary to align the coordinates locally in each part that is neither horizontal nor vertical. A rather natural approach computationally is to impose the numerical modifications on the reference element level [2].

In [2], formulation (5.6) & (5.9) is named QBI. It turns out that this is just an extension & reformulation of the classical Turner rectangle.

Theorem 5.1 (Equivalence result) *When the element is a rectangle aligned with the coordinates x, y and when the parameters in* (5.6) *are chosen according to* (5.9)*, the element stiffness matrix arising from* (5.6) *is identical with the stiffness matrix of the Turner–Pian–Wilson rectangle.*

In the remaining part of the paper we raise into daylight the error analysis tools that have so far played a background role only. We state first the main principles and then apply them to the a posteriori verification of our new-old finite element formulation (5.5)–(5.7).

6 Stability. Error estimation principles

When numerically modifying the deformation energy according to (5.1), the first risk is to lose stability, i.e., to loose the positive definiteness of the energy (modulo rigid displacements). In practice, the loss of stability is detected by the appearance of unphysical zero-energy modes or "hourglass" modes, especially when no kinematical constraints are imposed at the boundary. In the present situation, it turns out that condition (5.8) in (5.6) is both necessary and sufficient for preventing such spurious modes. This follows from the following more precise stability theorem, which states the first of our main results concerning formlation (5.6). The reader is referred to Appendix A for (the main incredients of) the proof.

In the theorem, we refer to the modified energy (semi)norm defined as

$$|||\mathbf{V}|||_h = \{\mathcal{A}_h(\mathbf{V},\mathbf{V})\}^{1/2}, \quad \mathbf{V} \in \mathcal{U}. \tag{6.1}$$

Theorem 6.1 (Stability estimates) *Let $\mathcal{A}_h$ be defined by* (5.6) *where $\alpha_1 > 0$, $\alpha_2 > 0$ and let $|||\cdot|||_h$ be defined by* (6.1)*. Further let $\delta_1 = \min\{\alpha_1, 1\}$, $\delta_2 = \min\{\alpha_2, 1\}$ and let $\delta = \min\{\delta_1, \delta_2\}$. Then for any $\mathbf{V} = (r, s) \in \mathcal{U}$*

$$\left\|\frac{\partial r}{\partial x}\right\| \le \Big(\frac{1}{2\delta_1\bar{\mu}}\Big)^{1/2} |||\mathbf{V}|||_h,$$
$$\left\|\frac{\partial s}{\partial y}\right\| \le \Big(\frac{1}{2\delta_2\bar{\mu}}\Big)^{1/2} |||\mathbf{V}|||_h,$$

and for any $\mathbf{V} = (r, s) \in \mathcal{U}_h$

$$\left\|\frac{\partial r}{\partial y} + \frac{\partial s}{\partial x}\right\| \le C M_h (1 + \log N_h)^{1/2} (\bar{\mu}\delta)^{-1/2} |||\mathbf{V}|||_h,$$

where C is an absolute constant an M_h, N_h are the geometric mesh parameters defined by (3.4), (3.2)*. These estimates cannot be improved on a general quadrilateral or even uniform rectangular mesh. Finally, if $\alpha_1 \le 0$ or $\alpha_2 \le 0$, there arises in general spurious modes $\mathbf{V} \in \mathcal{U}_h$ with $\mathcal{A}_h(\mathbf{V},\mathbf{V}) \le 0$, so condition* (5.8) *is also necessary for stability.*

In what follows we strengthen the stability condition (5.8) by assuming that for some given constant C,

$$C^{-1} \le \alpha_1, \alpha_2 \le C. \tag{6.2}$$

This condition is obviously passed by the recommended choice (5.9).

Under assumption (6.2), the dependence on parameter δ in the estimates of Theorem 6.1 may be disregarded. When combined, these estimates then imply

$$|||\mathbf{V}|||_h \ge C_h^{-1} \|\mathbf{V}\|, \quad \mathbf{V} \in \mathcal{U}_h, \tag{6.3}$$

where C_h is a mesh dependent constant bounded by

$$C_h \leq CM_h(1 + \log N_h)^{1/2}. \tag{6.4}$$

By (6.3), $\mathbf{V} \in \mathcal{U}_h$ & $|||\mathbf{V}|||_h = 0$ imply $|||\mathbf{V}||| = 0$, (since C_h is finite by (6.4)), so rigid displacements are the only zero-energy modes in $\mathcal{U}_h$. In case of strongly regular meshes where C_h is uniformly bounded, the modified energy norm $|||\cdot|||_h$ and the original energy norm $|||\cdot|||$ are uniformly equivalent norms in the finite element spaces. In the entire space $\mathcal{U}$, $|||\cdot|||_h$ is a seminorm, see Appendix A.

We establish next the error analysis principles to be used in the sequel. We first introduce a new error indicator, defined as

$$e = \frac{|||\mathbf{U} - \mathbf{U}_h|||_h}{|||\mathbf{U}|||}, \tag{6.5}$$

where $|||\cdot|||_h$ is defined by (6.1), (5.6) (under assumption (6.2)). Note that we need to "soften" the error indicator from (3.6) to be able to detect any advantage of the modified formulation. (The standard scheme would give the least error if indicator (3.6) were used.)

At this point we should comment on an issue that arises with the modified error indicator (6.5). This indicator obviously depends on how the norm $|||\cdot|||_h$ on $\mathcal{U}_h$ (which defines the finite element scheme) is extended to a seminorm of $\mathcal{U}$. The extension, however, is not unique, since we could add to $\mathcal{A}_h$ any positive semidefinite functional that vanishes in $\mathcal{U}_h$. This relates to hidden ambiguity in the above derivation of the finite element scheme: We could have derived the same scheme from a number of different modifications of type (5.1). Thus in this sense the error indicator (6.5) is up to our choice. The analysis ahead anyhow confirms that the choice we already made works, with one exception: At the final step of our error analysis in section 10, we need to slightly weaken the seminorm $|||\cdot|||_h$ to get through. That this change actually is necessary (as we show) indicates that the convergence characteristics of the scheme studied are somewhat delicate on thin domains. Below we assume definition (6.1), (5.6) until the beginning of section 10, where the slightly altered definition (on a rectangular mesh) will be introduced.

When the strain energy is modified according to (5.1), the finite element approximation of $\mathbf{U} \in \mathcal{U}$ satisfies (3.5) with $\mathcal{A}_h$ replacing $\mathcal{A}$ on the left side, i. e.,

$$\mathcal{A}_h(\mathbf{U}_h, \mathbf{V}) = \mathcal{A}(\mathbf{U}, \mathbf{V}), \quad \mathbf{V} \in \mathcal{U}_h^0. \tag{6.6}$$

Since the two bilinear forms here are different, (6.6) no more defines a projection method, i. e., we have committed a "variational crime".

The error analysis for scheme (6.6) will be based on splitting the finite element solution in two parts as

$$\mathbf{U}_h = \tilde{\mathbf{U}}_h + \mathbf{Z}_h, \tag{6.7}$$

where we impose on $\tilde{\mathbf{U}}_h$ the same kinematical constraints as on $\mathbf{U}_h$ (the same as in (3.5)), so that $\mathbf{U}_h - \tilde{\mathbf{U}}_h \in \mathcal{U}_h^0$ and also $\mathbf{Z}_h \in \mathcal{U}_h^0$. We define then $\tilde{\mathbf{U}}_h$ as the best approximation to $\mathbf{U}$ in the seminorm $\|\cdot\|_h$ under the assumed kinematical constraints, so that

$$\mathcal{A}_h(\tilde{\mathbf{U}}_h, \mathbf{V}) = \mathcal{A}_h(\mathbf{U}, \mathbf{V}), \quad \mathbf{V} \in \mathcal{U}_h^0. \tag{6.8}$$

By (6.6)–(6.8), the remainder $\mathbf{Z}_h \in \mathcal{U}_h^0$ then satisfies

$$\mathcal{A}_h(\mathbf{Z}_h, \mathbf{V}) = (\mathcal{A} - \mathcal{A}_h)(\mathbf{U}, \mathbf{V}), \quad \mathbf{V} \in \mathcal{U}_h^0. \tag{6.9}$$

We note that by (6.8), $\mathcal{A}_h(\mathbf{U} - \tilde{\mathbf{U}}_h, \mathbf{Z}_h) = 0$, so (6.7) corresponds to an orthogonal splitting of the error:

$$\|\mathbf{U} - \mathbf{U}_h\|_h^2 = \|\mathbf{U} - \tilde{\mathbf{U}}_h\|_h^2 + \|\mathbf{Z}_h\|_h^2.$$

We rewrite the last identity equivalently as

$$e^2 = e_A^2 + e_C^2, \tag{6.10}$$

where e is defined by (6.5), e_A is the *approximation error* defined by

$$e_A = \frac{\|\mathbf{U} - \tilde{\mathbf{U}}_h\|_h}{\|\mathbf{U}\|}, \tag{6.11}$$

and e_C is the *consistency error* defined by (see also (6.9))

$$e_C = \frac{\|\mathbf{Z}_h\|_h}{\|\mathbf{U}\|} = \sup_{\mathbf{V} \in \mathcal{U}_h^0, \|\mathbf{V}\|_h \neq 0} \frac{(\mathcal{A} - \mathcal{A}_h)(\mathbf{U}, \mathbf{V})}{\|\mathbf{U}\| \|\mathbf{V}\|_h}. \tag{6.12}$$

The idea of the error analysis is now to bound e_A and e_C separately (either from above or from below) and then use (6.10) to bound the total error. Here we can apply the following simple principles.

The approximation error principles.

A1. For any $\tilde{\mathbf{U}} \in \mathcal{U}_h$ satisfying the same kinematical constraints as $\mathbf{U}_h$, $e_A \leq \|\mathbf{U} - \tilde{\mathbf{U}}\|_h / \|\mathbf{U}\|$.

A2. If $\|\mathbf{U} - \tilde{\mathbf{U}}\|_h / \|\mathbf{U}\| \geq \delta_h$ for all $\tilde{\mathbf{U}} \in \mathcal{U}_h$, then $e_A \geq \delta_h$.

The consistency error principles.

C1. If $|(\mathcal{A} - \mathcal{A}_h)(\mathbf{U}, \mathbf{V})| \leq \varepsilon_h \|\mathbf{U}\| \|\mathbf{V}\|_h$ for all $\mathbf{V} \in \mathcal{U}_h^0$, then $e_C \leq \varepsilon_h$.

C2. If there exists $\mathbf{V} \in \mathcal{U}_h^0$, $\|\mathbf{V}\|_h \neq 0$, such that $|(\mathcal{A} - \mathcal{A}_h)(\mathbf{U}, \mathbf{V})| \geq \delta_h \|\mathbf{U}\| \|\mathbf{V}\|_h$, then $e_C \geq \delta_h$.

Here Principle A1 follows from (6.11) and the fact that $\tilde{\mathbf{U}}_h$ is the best approximation to $\mathbf{U}$ in the seminorm $\|\cdot\|_h$. Principle A2 is trivial, and principles C1 and C2 follow immediately from the second expression of the consistency error in (6.12).

In the remaining three sections, we carry out the error analysis of the scheme (6.6), (5.6) based on the principles above. Before turning to the details, let us summarize briefly the main features of the analysis ahead, in view of the framework just stated.

First, when looking for a sharp upper bound for the consistency error, we obviously want the (nearly) smallest ε_h such that Principle C1 is applicable. Here the role of stability is obvious: The weaker the norm $\|\cdot\|_h$ in $\mathcal{U}_h^0$, the larger the consistency error is expected to be in general. Indeed, if there exist (spurious) modes $\mathbf{V} \in \mathcal{U}_h^0$ such that $\|\mathbf{V}\|_h = 0$, then $e_C = \infty$, unless $(\mathcal{A} - \mathcal{A}_h)(\mathbf{U}, \mathbf{V}) = \mathcal{A}(\mathbf{U}, \mathbf{V}) = 0$ for all such modes. Even in the absence of spurious modes, we still need stability estimates relating $\|\cdot\|_h$ with more standard norms. In our analysis, Theorem 6.1 serves that purpose, and will be needed heavily. Note that by (6.3), growth of the consistency error is expected on meshes such that constant C_h is large, as then $\|\cdot\|_h$ is predicted to be a relatively weak norm in $\mathcal{U}_h^0$.

Another, more hidden feature of the consistency error analysis in practice is the need of a (problem specific) *expansion formula* for the functional that appears on the right side of (6.9). Thinking of $\mathbf{U}$ to be fixed, we call this the *consistency error functional* (a linear functional on $\mathcal{U}_h^0$). The aim of the expansion formula is simply to resolve the effect of modification (5.1) so that the seminorm of the functional, as defined in (6.12), can be estimated in a sharp way. Special care is needed here to detect any, possibly mesh dependent, error cancellation. In the present problem, it turns out that on rectangular meshes, a special expansion of $(\mathcal{A} - \mathcal{A}_h)(\mathbf{U}, \mathbf{V})$ is possible that partly compensates the weaker stability on meshes with large mesh aspect ratio M_h. This leads to an improved consistency error bound on such meshes (Theorem 8.1 ahead) as compared with the bound on general quadrilateral meshes (Theorem 7.1). We will also show, using Principle C2 above, that the bound in case of rectangular meshes is sharp.

In case of a fixed, "thick" body, the consistency error analysis ahead (section 7) is relatively straightforward, apart from the mentioned mesh geometric effect. Instead when the body under consideration is thin, special care is again needed, this time to detect the possible parametric error growth or "equilibrium locking". We analyze this effect in section 9 and in the related Appendix B. Using Principle C2 we demonstrate that equilibrium locking does appear in general in case of a thin rectangular body. We also show that the effect can be avoided if (and only if) (a) $\alpha_1 = \gamma + 1$ in (5.6) ($\beta_1 = 0$ in (5.5)) and (b) the meshes are rectangular (or otherwise very specific, see Appendix B).

Regarding the approximation error e_A, let us first quote a *lower* error bound due to Theorems 4.1 and 6.1. Note that in general, stability gives no help when bounding e_A, rather the opposite. Here it follows from the first inequality in Theorem 6.1 and from (6.2) that for any $\tilde{\mathbf{U}} = (\tilde{u}, \tilde{v}) \in \mathcal{U}_h$

$$\left\| \frac{\partial u}{\partial x} - \frac{\partial \tilde{u}}{\partial x} \right\| \leq C \| \mathbf{U} - \tilde{\mathbf{U}} \|_h. \tag{6.13}$$

(This remains valid also after the change of $\|\cdot\|_h$ in section 10.) Then by Theorem 4.1 and Principle A2 above, we must conclude that *trapezoidal locking is unavoidable* within the assumed formulation, so we can hope for uniform convergence with respect to t at most on parallelogram meshes. (A similar conclusion was drawn in [16].) Despite accepting this "reduced dream" ahead, we will attempt sharp error analysis with minimal assumptions on the mesh at each step.

Finally, there remains the problem of finding a sharp upper bound for the approximation error. Applying Principle A1, the most straightforward "interpolation test" is to bound e_A with the standard interpolant of $\mathbf{U}$ as $\tilde{\mathbf{U}}$. Here this choice is sufficient except when dealing with the leading term in expansion (2.17) (at large $\bar{\lambda}$) or when approximating the leading asymptotic bending mode in expansion (2.7) (at small t). In the former problem we can cite the existing theory (section 7), whereas the treatment of the asymptotic bending mode requires a new approach. We devote section 10 to this final approximation problem.

7 Error analysis on a fixed domain

Assume a general polygonal domain Ω and a family of regular quadrilateral meshes on it. Bounding the approximation error (6.11) in this case is easy. We need only to note that

$$|||\mathbf{V}|||_h \leq C|||\mathbf{V}|||, \quad \mathbf{V} \in \mathcal{U}, \tag{7.1}$$

as follows from (6.1), (5.6), (6.2) and the fact that $\|R_h\sigma\| \leq \|\sigma\|$, $\sigma \in L_2(\Omega)$. The standard interpolation test in Principle A1 above then gives

$$e_A \leq CQh. \tag{7.2}$$

(assuming that $\bar{\lambda} \leq C$). To bound the consistency error (6.12), we proceed stepwise.

Step 1. Expand the consistency error functional as

$$(\mathcal{A} - \mathcal{A}_h)(\mathbf{U}, \mathbf{V}) = F(\mathbf{V}) + G(\mathbf{V}), \quad \mathbf{V} \in \mathcal{U}, \tag{7.3}$$

where for $\mathbf{V} = (r, s)$,

$$F(\mathbf{V}) = (\sigma_1 - R_h\sigma_1, \frac{\partial r}{\partial x}) + (\sigma_2 - R_h\sigma_2, \frac{\partial s}{\partial y}), \tag{7.4}$$

$$G(\mathbf{V}) = (\tau - R_h\tau, \frac{\partial r}{\partial y} + \frac{\partial s}{\partial x}), \tag{7.5}$$

where further

$$\begin{aligned} \sigma_1(\mathbf{U}) &= \bar{\lambda}(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}) + 2(1-\alpha_1)\bar{\mu}\frac{\partial u}{\partial x}, \\ \sigma_2(\mathbf{U}) &= \bar{\lambda}(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}) + 2(1-\alpha_2)\bar{\mu}\frac{\partial v}{\partial y}, \\ \tau(\mathbf{U}) &= \bar{\mu}(\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x}). \end{aligned} \tag{7.6}$$

This expansion follows easily from (2.1), (5.6), using only the identity $(R_hf, R_hg) = (R_hf, g),\ f, g \in L_2(\Omega)$.

Step 2. Assuming that $\mathbf{V} \in \mathcal{U}_h^0$, bound $F(\mathbf{V})$ and $G(\mathbf{V})$ in (7.3). First applying the Cauchy-Schwarz inequality and then Theorem 6.1 (recall also (6.2)), we obtain

$$\begin{aligned} |F(\mathbf{V})| &\leq \|\sigma_1 - R_h\sigma_1\| \left\|\frac{\partial r}{\partial x}\right\| + \|\sigma_2 - R_h\sigma_2\| \left\|\frac{\partial s}{\partial y}\right\| \\ &\leq C\big(\|\sigma_1 - R_h\sigma_1\|^2 + \|\sigma_2 - R_h\sigma_2\|^2\big)^{1/2} |||\mathbf{V}|||_h. \end{aligned} \tag{7.7}$$

This holds in fact for any $\mathbf{V} \in \mathcal{U}$. Similarly for any $\mathbf{V} \in \mathcal{U}_h$

$$\begin{aligned} |G(\mathbf{V})| &\leq \|\tau - R_h\tau\| \left\|\frac{\partial r}{\partial y} + \frac{\partial s}{\partial x}\right\| \\ &\leq C_h\|\tau - R_h\tau\| |||\mathbf{V}|||_h, \end{aligned} \tag{7.8}$$

where C_h is bounded according to (6.4).

Step3. Bound $\|\sigma_i - R_h\sigma_i\|$ and $\|\tau - R_h\tau\|$. Knowing that $R_h\sigma$ is the best piecewise constant approximation of σ in $L_2(\Omega)$, this is classical approximation theory. In our notation (recall (3.3)), the usual bounds are

$$\|\sigma_i - R_h\sigma_i\| \le ChL|\sigma_i|_{1,\Omega},\ i = 1, 2, \tag{7.9}$$

$$\|\tau - R_h\tau\| \le ChL|\tau|_{1,\Omega}, \tag{7.10}$$

where

$$|\sigma|_{1,\Omega} = \Big\{\int_\Omega \Big[\Big(\frac{\partial\sigma}{\partial x}\Big)^2 + \Big(\frac{\partial\sigma}{\partial y}\Big)^2\Big]\, dxdy\Big\}^{1/2}. \tag{7.11}$$

Step 4. Bound $|\sigma_i|_{1,\Omega}$ and $|\tau|_{1,\Omega}$. By (7.11), (7.6), (2.3) and Hypothesis 2.1 we have (assuming that $\bar\lambda \le C$)

$$L|\sigma_i|_{1,\Omega} \le CL|\mathbf{U}|_{2,\Omega} \le CQ|\!|\!|\mathbf{U}|\!|\!|,\ i = 1, 2. \tag{7.12}$$

Similarly,

$$L|\tau|_{1,\Omega} \le CQ|\!|\!|\mathbf{U}|\!|\!|. \tag{7.13}$$

Step 5. Combine the estimates and apply the consistency error principle C1. Upon inserting (7.7)–(7.10) and (7.12) –(7.13) in (7.3) we have

$$|(\mathcal{A} - \mathcal{A}_h)(\mathbf{U}, \mathbf{V})| \le C_hQh|\!|\!|\mathbf{U}|\!|\!||\!|\!|\mathbf{V}|\!|\!|_h, \quad \mathbf{V} \in \mathcal{U}_h,$$

so by Principle C1

$$e_C \le C_hQh. \tag{7.14}$$

Estimates (7.2) and (7.14) formally complete our error analysis, but so far both of these estimates still predict severe error growth when $\bar\lambda \to \infty$. Namely, $C \sim (\bar\lambda + \bar\mu)^{1/2}$ in (7.2), because we needed estimate (3.8), and $C_h \sim \bar\lambda + \bar\mu$ in (7.14), because we needed (7.12) where $C \sim \bar\lambda + \bar\mu$ in the first inequality (by (7.6)). Our aim next is to show that the parametric dependence in these estimates is (almost) removable under Hypothesis 2.2. So let us consider now the parametrized problem where $\bar\lambda$ can vary freely in the range $1 \le \bar\lambda < \infty$. Below we denote by $\bar C$ a constant independent of $\bar\lambda$.

The improvement of the consistency error bound under Hypothesis 2.2 is straightforward: We have by assumptions (2.17)–(2.18) and by (7.6),

$$L|\sigma_i|_{1,\Omega} \le \bar CL(|\mathbf{U}_0|_{2,\Omega} + |\mathbf{U}_1|_{2,\Omega}), \quad i = 1, 2.$$

By (2.17), the triangle inequality, and by assumptions (2.4) and (2.19), the right side here is bounded as

$$L(|\mathbf{U}_0|_{2,\Omega} + |\mathbf{U}_1|_{2,\Omega}) \le L(|\mathbf{U}|_{2,\Omega} + 2|\mathbf{U}_1|_{2,\Omega}) \le 3Q|\!|\!|\mathbf{U}|\!|\!|, \tag{7.15}$$

so it follows that

$$L|\sigma_i|_{1,\Omega} \le \bar{C}Q\|\mathbf{U}\|, \quad i = 1, 2.$$

This removes the parametric dependence from (7.12), so we conclude that (7.14) holds uniformly with respect to $\bar{\lambda}$ under Hypothesis 2.2.

To prove that the approximation error is likewise uniformly bounded when $\mathbf{U}$ satisfies Hypothesis 2.2, is much harder. The main difficulty here is to show that the following constrained approximation problem is solvable.

Problem. Given $\mathbf{U}_0 = (u_0, v_0)$ such that (2.18) holds, find $\tilde{\mathbf{U}}_0 = (\tilde{u}_0, \tilde{v}_0) \in \mathcal{U}_h$ satisfying

$$R_h\Big(\frac{\partial \tilde{u}_0}{\partial x} + \frac{\partial \tilde{v}_0}{\partial y}\Big) = 0 \tag{7.16}$$

and approximating $\mathbf{U}_0$ so that

$$|\mathbf{U}_0 - \tilde{\mathbf{U}}_0|_{1,\Omega} \le \bar{C}hL|\mathbf{U}_0|_{2,\Omega}, \tag{7.17}$$

where $|\cdot|_{1,\Omega}$ is defined by (3.9).

Note that when $\mathbf{U}_0$ satisfies (2.18) and $\tilde{\mathbf{U}}_0$ satisfies (7.16), $\|\mathbf{U}_0 - \tilde{\mathbf{U}}_0\|_h$ is independent of $\bar{\lambda}$ (recall (5.6), (6.1)), so if $\tilde{\mathbf{U}}_0$ exists, (7.17) implies that

$$\|\mathbf{U}_0 - \tilde{\mathbf{U}}_0\|_h \le \bar{C}|\mathbf{U}_0 - \tilde{\mathbf{U}}_0|_{1,\Omega} \le \bar{C}hL|\mathbf{U}_0|_{2,\Omega}. \tag{7.18}$$

This improves the standard interpolation error bound (3.10) (valid when $\bar{\lambda} \le C$) to a uniformly optimal bound with respect to $\bar{\lambda}$. Apparently such an improvement is possible only under constraint (7.16).

Assume now the splitting (2.17), and suppose that $\tilde{\mathbf{U}}_0 \in \mathcal{U}_h$ exists such that (7.16)–(7.17) hold. In addition, assume that $\tilde{\mathbf{U}}_0$ interpolates the kinematical constraints on $\mathbf{U}_0$ at the boundary. Then set $\tilde{\mathbf{U}} = \tilde{\mathbf{U}}_0 + \bar{\lambda}^{-1}\tilde{\mathbf{U}}_1$ where $\tilde{\mathbf{U}}_1$ is the standard interpolant of $\mathbf{U}_1$, and apply the triangle inequality, estimate (7.18), a standard interpolation error bound dampened by factor $\bar{\lambda}^{-1}$, and finally (7.15), to obtain

$$\begin{aligned}\|\mathbf{U} - \tilde{\mathbf{U}}\|_h &\le \|\mathbf{U}_0 - \tilde{\mathbf{U}}_0\|_h + \bar{\lambda}^{-1}\|\mathbf{U}_1 - \tilde{\mathbf{U}}_1\|_h \\ &\le \bar{C}hL\left(|\mathbf{U}_0|_{2,\Omega} + |\mathbf{U}_1|_{2,\Omega}\right) \\ &\le \bar{C}Qh\|\mathbf{U}\|.\end{aligned}$$

It follows (Principle A1) that (7.2) holds uniformly with respect to $\bar{\lambda}$. Thus the critical question is, whether the above constrained approximation problem is solvable.

Let us warn at this point of a possible false conjecture. We see that the volumetric constraint (2.18) is the limit of the dilatation constraint (4.2) as $\bar{\lambda} \to \infty$ ($\gamma \to 1$). Then it appears that making a distinction between the two constraints is irrelevant, and that we could assume (2.17)–(2.19) with

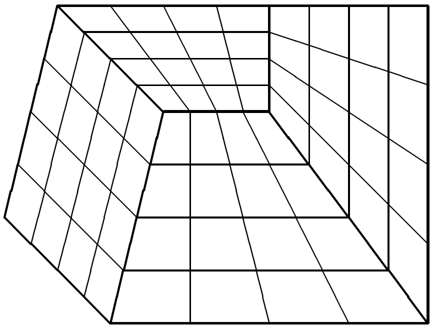

Figure 5: Restricted quadrilateral mesh (Ref. [21]).

the dilatation constraint replacing (2.18). This actually is correct. But then it further appears that the volumetric locking problem is simply contained in the problem of locking at small t, so that only the latter problem needs to be solved. This is false, because when dealing with the beam bending state where the dilatation and shear constraints appear simultaneously, we can also utilize the specific asymtotic form (4.1) of the solution in that case. In the approximation error analysis ahead (section 10), we actually utilize this information heavily. Then this analysis does not solve the above approximation problem where $\mathbf{U}_0$ (the asymptotic solution at large $\bar{\lambda}$) is of different nature.

That $\tilde{\mathbf{U}}_0$ satisfying (7.16)–(7.17) anyhow exists (on restricted meshes at least), so that formulation (5.6) does avoid volumetric locking on a fixed domain, is one of the most curious "miracles" of finite element methodology in all times. Here again, practice came first [10, Ch 4]: It was found empirically that when $\bar{\lambda}$ is large, the simple averaging in the leading term of (2.1) (arising originally from a mixed formulation with piecewise constant pressure), significantly improves the performance of the standard bilinear quadrilateral. In fact, other modifications in (5.6) are rather irrelevant on a fixed domain, causing mainly just additional consistency error terms, as indicated by the above analysis. Note also that the constrained approximation problem above relates to the mentioned single "trick" only.

In the theory, the existence of $\tilde{\mathbf{U}}_0 \in \mathcal{U}_h$ satisfying (7.16)–(7.17) was shown first in case of a rectangular mesh [12], then for a mesh composed of 4 by 4 quadrilateral macroelements such that the local mesh in each macroelement is uniform [21] (Fig. 5). Whether or not $\tilde{\mathbf{U}}_0$ actually exists on a general (strongly) regular quadrilateral mesh, is still an open problem. The theory so far also sets constraints on how to impose the kinematic conditions at the boundary: On the restricted quadrilateral meshes, the usual interpolation conditions are allowed at the vertices of the macroelements, whereas at all the other nodes (including those at the boundary), nonstandard interpolation conditions must be chosen so as to enforce constraint (7.16) [21].

In practice, the above limitations of the theory are hardly observed (or obeyed), so there remains a slight gap between theory and practice when $\bar{\lambda}$ is large. We leave this gap as it is, summarizing the theory so far in

Theorem 7.1 (Convergence theorem on a fixed domain) *Assume a family of regular quadrilateral meshes on a polygonal domain Ω. Let $\mathbf{U}$ be the* *exact solution satisfying Hypothesis 2.1 and let* $\mathbf{U}_h \in \mathcal{U}_h$ *be the finite element solution defined according to* (6.6)*, where* $\mathcal{A}_h$ *is defined by* (5.6)*, with the parameters* α_1, α_2 *satisfying* (6.2)*. Then the total error* (6.5) *satisfies* $e \leq C_h Q h$*, where* C_h *is a mesh dependent constant. The constant is bounded according to* (6.4)*, where* M_h *and* N_h *are defined by* (3.2)*,* (3.4)*. Under Hypothesis 2.2, this error bound extends further to a uniform bound with respect to parameter* $\bar{\lambda}$ *in the range* $1 \leq \bar{\lambda} < \infty$*, assuming restricted quadrilateral meshes and modified interpolation constraints on* $\mathbf{U}_h$ *at the boundary.*

An obvious corollary of Theorems 7.1 and 3.1 is that in the non-parametric situation where the domain is fixed, $\bar{\lambda}$ is uniformly bounded, and the meshes are strongly regular, the modified FE scheme (6.6), (5.6) performs essentially as well as (or no worse than) the standard bilinear scheme. A similar result was proved previously in [13], [14], [25] for the nonconforming Wilson quadrilaterals.

8 Consistency error on a rectangular mesh

In this section we raise the question, whether the error bound of Theorem 7.1 is improvable on meshes with large C_h. An equivalent question obviously is, whether the consistency error bound (7.14) is improvable. We give an affirmative answer to this question in the special case where the mesh is rectangular and aligned with coordinates x, y (possible on restricted domains, see Fig. 6). To what extent estimate (7.14) possibly is improvable on general quadrilateral meshes, is left as an open problem.

On a rectangular mesh one has $N_h = 1$ in (6.4) but M_h can still be large. We show that when M_h is large, estimate (7.14) is improvable as

$$e_C \leq C(Q + Q_1 M_h^{1/2} h^{1/2})\, h, \tag{8.1}$$

where Q_1 is another dimensionless constant depending on $\mathbf{U}$ and finite when $\mathbf{U}$ is smooth enough. (We need slightly enhanced smoothness here as compared with Hypothesis 2.1). The more precise result will be stated in Theorem 8.1 ahead. Note that (8.1) predicts error amplification (from the optimal rate $\mathcal{O}(h)$) when $M_h \gg h^{-1}$, whereas from (7.14) we expect it already when $M_h \gg 1$. We will show by a counterexample that (8.1) is no more improvable, so the mild error amplification predicted here (when $M_h \gg h^{-1}$) is a true phenomenon.

Recall that the mesh dependent constant in (7.14) arises form (7.8), where we needed the third stability estimate of Theorem 6.1. This estimate was not improvable, and neither is the Cauchy-Schwarz inequality applied in (7.8) improvable in general. However, when writing

$$\frac{|G(\mathbf{V})|}{\|\tau - R_h\tau\| \|\!|\mathbf{V}|\!\|_h} = \Big[\frac{|G(\mathbf{V})|}{\|\tau - R_h\tau\| \|\frac{\partial r}{\partial y} + \frac{\partial s}{\partial x}\|}\Big] \Big[\frac{\|\frac{\partial r}{\partial y} + \frac{\partial s}{\partial x}\|}{\|\!|\mathbf{V}|\!\|_h}\Big],$$

it is still possible that when $\mathbf{V} \in \mathcal{U}_h^0$, weak stability (the second factor large) is compensated by better consistency (the first factor small). To resolve such effects, we need to study the functional $G(\mathbf{V})$ in more detail to detect any possible (global) error cancellation. The vague idea is to find a better expansion of $G(\mathbf{V})$ where the Cauchy-Schwarz inequality and the stability estimates can be combined without losing information. Below we show that in case of an aligned rectangular mesh, $G(\mathbf{V})$ can be expanded in such a way that the use of the third stability estimate of Theorem 6.1 is actually avoided. This leads to a sharp consistency error bound, as will be seen.

Suppose the boundary of Ω consists of lines parallel with the coordinate lines and assume that the finite element mesh on Ω consists of rectangles

$$K_{ij} = \{x_{i-1} < x < x_i,\ y_{j-1} < y < y_j\}, \quad (i,j) \in \mathcal{I}_h. \tag{8.2}$$

We introduce two anisotropic local averaging operators that will be needed frequently below when the mesh is rectangular. The operators, denoted by

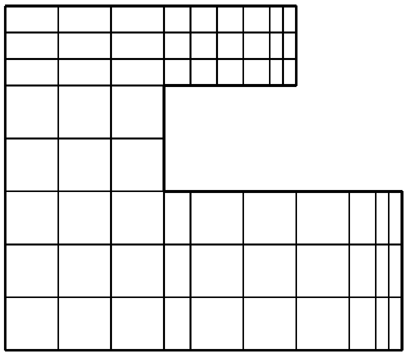

Figure 6: A rectangular mesh.

Π_h and Λ_h, are defined so that for each $K_{ij} \subset \Omega$,

$$(\Pi_h f)(x,y) = \frac{1}{x_i - x_{i-1}} \int_{x_{i-1}}^{x_i} f(x', y)\, dx', \quad (x,y) \in K_{ij}, \tag{8.3}$$

$$(\Lambda_h f)(x,y) = \frac{1}{y_j - y_{j-1}} \int_{y_{j-1}}^{y_j} f(x, y')\, dy', \quad (x,y) \in K_{ij}. \tag{8.4}$$

We note that if $\mathbf{V} = (r,s) \in \mathcal{U}_h$, then $R_h \partial r/\partial y = \Pi_h \partial r/\partial y$ and $R_h \partial s/\partial x = \Lambda_h \partial s/\partial x$. Therefore and since $(\partial/\partial y)\Pi_h = \Pi_h \partial/\partial y$ and $(\partial/\partial x)\Lambda_h = \Lambda_h \partial/\partial x$, we may rewrite the error functional (7.5) as

$$\begin{aligned} G(\mathbf{V}) &= (\tau, \frac{\partial r}{\partial y} - R_h \frac{\partial r}{\partial y}) + (\tau, \frac{\partial s}{\partial x} - R_h \frac{\partial s}{\partial x}) \\ &= \big(\tau, \frac{\partial}{\partial y}(r - \Pi_h r)\big) + \big(\tau, \frac{\partial}{\partial x}(s - \Lambda_h s)\big). \end{aligned}$$

Integrating by parts, noting that $r - \Pi_h r$ is continuous in y and $s - \Lambda_h s$ is continuous in x, we further get

$$\begin{aligned} G(\mathbf{V}) &= \Big[-\big(\frac{\partial \tau}{\partial y}, r - \Pi_h r\big) - \big(\frac{\partial \tau}{\partial x}, s - \Lambda_h s\big) \Big] \\ &\quad + \int_{\partial\Omega} \tau [n_y (r - \Pi_h r)\, dx + n_x (s - \Lambda_h s)\, dy] \\ &= G_1(\mathbf{V}) + G_2(\mathbf{V}), \quad \mathbf{V} = (r,s) \in \mathcal{U}_h, \end{aligned} \tag{8.5}$$

where (n_x, n_y) is the outward unit normal to $\partial\Omega$. We use now this expansion to sharpen estimate (7.8).

To first bound $G_1(\mathbf{V})$ in (8.5), note that for $(r,s) \in \mathcal{U}_h$,

$$\int_{x_{i-1}}^{x_i} (r - \Pi_h r)^2\, dx = \frac{1}{12}(x_i - x_{i-1})^2 \int_{x_{i-1}}^{x_i} \Big(\frac{\partial r}{\partial x}\Big)^2 dx,$$

$$\int_{y_{j-1}}^{y_j} (s - \Lambda_h s)^2\, dy = \frac{1}{12}(y_j - y_{j-1})^2 \int_{y_{j-1}}^{y_j} \Big(\frac{\partial s}{\partial y}\Big)^2 dy,$$

so knowing that $x_i - x_{i-1} \le hL$, $y_j - y_{j-1} \le hL$, we have

$$\|r - \Pi_h r\| \le \frac{hL}{\sqrt{12}} \left\| \frac{\partial r}{\partial x} \right\|, \quad \|s - \Lambda_h s\| \le \frac{hL}{\sqrt{12}} \left\| \frac{\partial s}{\partial y} \right\|.$$

Then by the previous reasoning (Steps 2 and 4 above) we conclude

$$\begin{aligned} |G_1(\mathbf{V})| &\le \left\| \frac{\partial \tau}{\partial y} \right\| \|r - \Pi_h r\| + \left\| \frac{\partial \tau}{\partial x} \right\| \|s - \Lambda_h s\| \\ &\le CQh \|\mathbf{U}\| \|\mathbf{V}\|_h, \end{aligned} \tag{8.6}$$

so there arises no error amplification from this term.

To bound the second term in (8.5), consider a rectangle (8.2) with a horizontal edge Γ_i on $\partial\Omega$. Then since $\partial r/\partial x$ is constant on Γ_i, we get integrating by parts and using the Cauchy-Schwarz inequality

$$\begin{aligned} \Big| \int_{\Gamma_i} n_y \tau (r - \Pi_h r)\, dx \Big| &= \Big| \int_{\Gamma_i} \frac{1}{2}(x - x_{i-1})(x_i - x) \frac{\partial \tau}{\partial x} \frac{\partial r}{\partial x}\, dx \Big| \\ &\le \frac{1}{8}\, (x_i - x_{i-1})^2 \Big[\int_{\Gamma_i} \Big(\frac{\partial \tau}{\partial x} \Big)^2 dx \Big]^{1/2} \Big[\int_{\Gamma_i} \Big(\frac{\partial r}{\partial x} \Big)^2 dx \Big]^{1/2}. \end{aligned}$$

Here $\partial r/\partial x = c_0 + c_1(y - y_{j-1}) = \rho(y)$ on K_{ij}, where c_0, c_1 are some constants, so

$$\begin{aligned} \Big[\int_{\Gamma_i} \Big(\frac{\partial r}{\partial x} \Big)^2 dx \Big]^{1/2} &\le (x_i - x_{i-1})^{1/2} \max\{\, |\rho(y_{j-1})|,\ |\rho(y_j)|\, \} \\ &\le 2\, (x_i - x_{i-1})^{1/2} (y_j - y_{j-1})^{-1/2} \Big[\int_{y_{j-1}}^{y_j} \rho^2\, dy \Big]^{1/2} \\ &= 2\, (y_j - y_{j-1})^{-1/2} \Big[\int_{K_{ij}} \Big(\frac{\partial r}{\partial x} \Big)^2 dxdy \Big]^{1/2}. \end{aligned}$$

Hence using $(x_i - x_{i-1})^2 (y_j - y_{j-1})^{-1/2} \le M_h^{1/2} (hL)^{3/2}$, we have

$$\begin{aligned} &\Big| \int_{\Gamma_i} n_y \tau (r - \Pi_h r)\, dx \Big| \\ &\qquad \le \frac{1}{4} M_h^{1/2} (hL)^{3/2} \Big[\int_{\Gamma_i} \Big(\frac{\partial \tau}{\partial x} \Big)^2 dx \Big]^{1/2} \Big[\int_{K_{ij}} \Big(\frac{\partial r}{\partial x} \Big)^2 dxdy \Big]^{1/2}. \end{aligned}$$

Similarly if K_{ij} has a vertical edge Γ_j on $\partial\Omega$, we have

$$\begin{aligned} &\Big| \int_{\Gamma_j} n_x \tau (s - \Lambda_h s)\, dy \Big| \\ &\qquad \le \frac{1}{4} M_h^{1/2} (hL)^{3/2} \Big[\int_{\Gamma_j} \Big(\frac{\partial \tau}{\partial y} \Big)^2 dy \Big]^{1/2} \Big[\int_{K_{ij}} \Big(\frac{\partial s}{\partial y} \Big)^2 dxdy \Big]^{1/2}. \end{aligned}$$

Summing then over all K_{ij} that touch $\partial\Omega$, using the Cauchy-Schwarz inequality for the sum, assuming that for some finite Q_1

$$L^{3/2}\Big[\int_{\partial\Omega}|n_y|\Big(\frac{\partial\tau}{\partial x}\Big)^2\,dx+|n_x|\Big(\frac{\partial\tau}{\partial y}\Big)^2\,dy\Big]^{1/2}\le Q_1|\!|\!|\mathbf{U}|\!|\!|, \tag{8.7}$$

and finally applying Theorem 6.1, we get

$$\begin{aligned}|G_2(\mathbf{V})| &\le CM_h^{1/2}(hL)^{3/2}Q_1L^{-3/2}|\!|\!|\mathbf{U}|\!|\!|\Big\{\int_\Omega\Big[\Big(\frac{\partial r}{\partial x}\Big)^2+\Big(\frac{\partial s}{\partial y}\Big)^2\Big]\,dxdy\Big\}^{1/2}\\ &\le CM_h^{1/2}h^{3/2}Q_1|\!|\!|\mathbf{U}|\!|\!|\,|\!|\!|\mathbf{V}|\!|\!|_h.\end{aligned} \tag{8.8}$$

In view of (8.5), (8.6), (8.8) and the consistency error principle C1, we have verified (8.1) with Q_1 such that (8.7) holds. Note that Q_1 is not necessarily finite under Hypothesis 2.1, so (8.7) is an additional regularity assumption.

From the above proof we see that the second term in estimate (8.1) is a pure edge effect where the interior elements play no role, insofar as these are all rectangles aligned with the coordinates. As will be shown shortly, this bound admits no further improvement in general. However, in a given problem with given mesh and kinematical constraints, slight improvement may still be possible as follows. Define a function q_h on $\partial\Omega$ so that q_h takes piecewise constant values according to the rule

$$q_h(x,y)=\begin{cases}h^{-1}L^{-3/2}(x_i-x_{i-1})^2(y_j-y_{j-1})^{-1/2}, & \text{if } (x,y)\in\Gamma_i\subset\partial K_{ij}\cap\partial\Omega,\\[2ex] h^{-1}L^{-3/2}(y_j-y_{j-1})^2(x_i-x_{i-1})^{-1/2}, & \text{if } (x,y)\in\Gamma_j\subset\partial K_{ij}\cap\partial\Omega,\end{cases} \tag{8.9}$$

where Γ_i denotes a horizontal and Γ_j a vertical edge of K_{ij} as above. Let further Γ_0 denote the part of $\partial\Omega$ where the tangential displacement is not restricted by the kinematical constraints, so that for $\mathbf{V}=(r,s)\in\mathcal{U}_h^0$, n_yr-n_xs does not vanish on Γ_0. Then defining a mesh dependent constant Q_h so that

$$L^{3/2}\Big\{\int_{\Gamma_0}q_h^2\Big[|n_y|\Big(\frac{\partial\tau}{\partial x}\Big)^2\,dx+|n_x|\Big(\frac{\partial\tau}{\partial y}\Big)^2\,dy\Big]\Big\}^{1/2}\le Q_h|\!|\!|\mathbf{U}|\!|\!|, \tag{8.10}$$

the above analysis implies

Theorem 8.1 (Consistency error on a rectangular mesh) *Suppose the boundary of Ω consists of lines parallel with the coordinate lines and let the finite element mesh on Ω be rectangular. Let $\mathbf{U}$ be a displacement field on Ω that satisfies Hypothesis 2.1, assume further that* (8.7) *holds for some finite Q_1, let q_h be defined by* (8.9) *on $\partial\Omega$, and let Q_h be a mesh dependent constant such that estimate* (8.10) *holds, where Γ_0 denotes the part of $\partial\Omega$ where the tangential displacement is not restricted by kinematical constraints. Then the consistency error of scheme* (6.6), (5.6) *satisfies*

$$e_C\le C(Q+Q_h)\,h\le C(Q+Q_1M_h^{1/2}h^{1/2})\,h.$$

Note that by Theorem 8.1 and by (8.9)–(8.10), error growth due to large Q_h is actually possible only if the kinematical constraints, the exact solution $\mathbf{U}$, and the mesh are such that (i) Γ_0 is non-empty, and for some non-empty subset $\Gamma_1 \subset \Gamma_0$, (ii) the integrand in (8.10) is non-vanishing on Γ_1, and (iii) the rectangles K_{ij} that touch Γ_1 are all thin in the normal direction to Γ_1.

Let us finally show by a counterexample that the consistency error bound of Theorem 8.1 is sharp. To this end, let (x_i, y_j) be the nodes of a rectangular mesh on the reference domain, with $0 = x_0 < \ldots < x_I = L$, $-H/2 = y_0 < \ldots < y_J = H/2$, $H = tL$, $0 < t \leq 1$. Assume that $I \geq 2$ is even, and let $x_i - x_{i-1} = a = hL$ when i is even and $x_i - x_{i-1} = a/2$ when i is odd. Let further $J \geq 2$ and let $y_j - y_{j-1} \leq a$ when $j = 1 \ldots J-1$ and $y_j - y_{j-1} = b \ll a$ when $j = J$. Finally let the exact solution be given by

$$u(x, y) = u_0\,(x - L/2)\,y, \quad v(x, y) = 0,$$

where $u_0 \neq 0$ is a constant. Choosing then $\mathbf{V} = (r, s) \in \mathcal{U}_h$ so that $s = 0$ and $r(x_i, y_j) = (-1)^i$ when $j = J$ and $r(x_i, y_j) = 0$ when $0 \leq j \leq J-1$, we find by direct computation that

$$(\mathcal{A} - \mathcal{A}_h)(\mathbf{U}, \mathbf{V}) = \delta_h \|\!|\mathbf{U}|\!\|\,\|\!|\mathbf{V}|\!\|_h,$$

where

$$\delta_h = \frac{(a/b)^{1/2}\,t^{-1/2}\,h^{3/2}}{(48 + 16\alpha_1 + 24\bar{\lambda}/\bar{\mu})^{1/2}\,[1 + (2 + \bar{\lambda}/\bar{\mu})t^2]^{1/2}}.$$

By Principle C2 of section 6 we have then $e_C \geq \delta_h$. Note also that the total error is $e = e_C$, since $\mathbf{U}$ is bilinear.

In the above case, $M_h = a/b$ and estimate (8.7) holds for Q_1 no less than $24(\bar{\mu}/t)^{1/2}$, so we conclude that second bound in Theorem 8.1 is sharp in this case. We also see that in the example case, the convergence (in the sense of error indicator (6.5)) actually fails on a family of extreme meshes such that $b/a \leq C(a/L)^3 = Ch^3$ as $h \to 0$.

Let us note finally that the above counterexample does not work if we choose $x_i - x_{i-1} = a$ for all i, since then $\delta_h = 0$. This indicates further error cancellation, which is possible more generally when the (rectangular) mesh is tangentially uniform on $\partial\Omega$ (or on Γ_0, see Theorem 8.1). To see that there still remains a mesh effect, let the index I be odd in the above example, and let $x_i - x_{i-1} = a = hL$ for $i = 1 \ldots I$. Then choosing $\mathbf{V}$ as above gives $\delta_h \sim (a/b)^{1/2} t^{-1/2} h^{5/2}$, so the consistency error is at least of order $e_C \sim M_h^{1/2} h^{5/2}$ in that case.

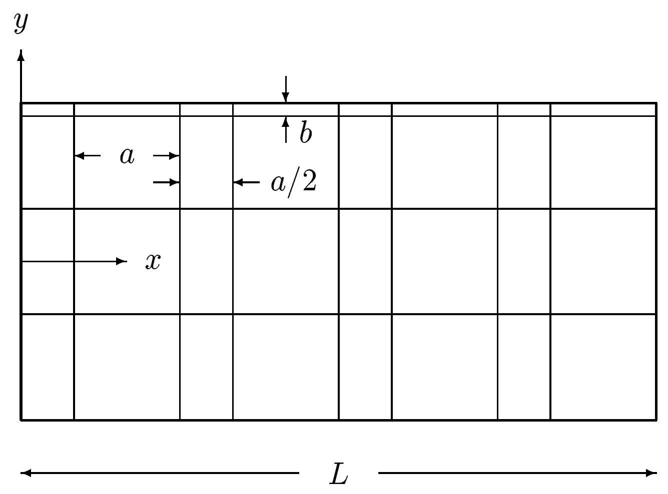

Figure 7: A consistency error anomaly: Convergence fails if $b/a \leq Ch^3$ as $h = a/L \rightarrow 0$. Exact solution: $u = u_0\,(x - L/2)\,y,\ v = 0$.

9 Thin domains: Consistency error

In the remaining part of the paper, we study the rectangular reference domains with $t = H/L$ as a variable parameter. Our aim is to prove uniform error bounds with respect to t as $t \to 0$, whenever possible. In this section we start from the "easy part", the consistency error term.

In the pure stretching state of deformation, the (consistency) error bounds of sections 7-8 extend immediately to uniform bounds with respect to t. Indeed, estimate (2.4) holds uniformly in t in the pure stretching state (by Hypothesis 2.1), so the consistency error bound (7.14) is uniform in t as well. The bound of Theorem 8.1 is likewise uniform in t, provided estimate (8.7) holds with Q_1 independent of t. This is a reasonable (though an additional) assumption in view of expansion (2.10), so we may consider the case of pure stretching closed.

The case of pure bending is less straightforward. We proceed again as in section 7, but now we use expression (5.5) of $\mathcal{A}_h$ and rewrite σ_1, σ_2 in (7.6) as

$$\begin{aligned}\sigma_1(\mathbf{U}) &= \bar{\lambda}\Big(\gamma\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\Big) - 2\beta_1\bar{\mu}\frac{\partial u}{\partial x} \\ \sigma_2(\mathbf{U}) &= \frac{\bar{\lambda}}{\gamma}\Big(\gamma\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\Big) - 2\beta_2\bar{\mu}\frac{\partial v}{\partial y}.\end{aligned} \tag{9.1}$$

(By (2.9) and (5.7), expressions (7.6) and (9.1) are equivalent.) Recalling (2.7), (2.8) we have here

$$\sigma_1(\mathbf{U}) = \sigma_1(\mathbf{U}_{bm}) + \sigma_1(\mathbf{U}_{br}) = -2\beta_1\bar{\mu}\, y\theta'(x) + \sigma_1(\mathbf{U}_{br}), \tag{9.2}$$

$$\sigma_2(\mathbf{U}) = \sigma_2(\mathbf{U}_{bm}) + \sigma_2(\mathbf{U}_{br}) = 2\beta_2\bar{\mu}\gamma\, y\theta'(x) + \sigma_2(\mathbf{U}_{br}). \tag{9.3}$$

By the reasoning of section 7 (Steps 2–3), the consistency error functional associated to σ_1 is bounded as

$$|(\sigma_1 - R_h\sigma_1, \frac{\partial r}{\partial x})| \le ChL|\sigma|_{1,\Omega}\|\mathbf{V}\|_h. \tag{9.4}$$

By (9.2) and Hypothesis 2.1, we have here

$$L|\sigma_1(\mathbf{U}_{bm})|_{1,\Omega} \le C(1 + |\beta_1|t^{-1})Q\|\mathbf{U}\|, \tag{9.5}$$

$$L|\sigma_1(\mathbf{U}_{br}\ )|_{1,\Omega} \le CQ\|\mathbf{U}\|, \tag{9.6}$$

so that combining (9.4)–(9.6)

$$|(\sigma_1 - R_h\sigma_1, \frac{\partial r}{\partial x})| \le CQ(1 + |\beta_1|t^{-1})h\|\mathbf{U}\|\|\mathbf{V}\|_h. \tag{9.7}$$

Estimate (9.7) predicts error amplification by factor t^{-1} unless $\beta_1 = 0$. In Appendix B we show that such an "equilibrium locking" is a true effect even on a rectangular mesh, so we conclude that *choosing* $\beta_1 = 0$ *is necessary* for uniform convergence with respect to t. Note that in view of (5.7), this choice is legitimate within constraint (6.2).

The consistency error term due to σ_2 is similarly bounded as

$$|(\sigma_2 - R_h\sigma_2, \frac{\partial s}{\partial y})| \le ChL|\sigma_2|_{1,\Omega}\|\mathbf{V}\|_h, \tag{9.8}$$

where by (9.3) and Hypothesis 2.1,

$$L|\sigma_2(\mathbf{U}_{bm})|_{1,\Omega} \le C(1 + |\beta_2|\gamma t^{-1})Q\|\mathbf{U}\|, \tag{9.9}$$

$$L|\sigma_2(\mathbf{U}_{br}\ )|_{1,\Omega} \le CQ\|\mathbf{U}\|, \tag{9.10}$$

so that combining (9.8)–(9.10)

$$|(\sigma_2 - R_h\sigma_2, \frac{\partial s}{\partial y})| \le CQ(1 + |\beta_2|\gamma t^{-1})h\|\mathbf{U}\|\|\mathbf{V}\|_h. \tag{9.11}$$

This predicts again error amplification unless $\beta_2 = 0$, or unless it happens that $\gamma = 0$. In Appendix B we show that this is again a correct prediction in general. The problem then is that this time we cannot escape by choosing $\beta_2 = 0$, since that would violate stability (see (5.7) and Theorem 6.1).

We are now at the first critical point where restricting the mesh helps to avoid parametric locking. Namely, it turns out that the error amplification in (9.11) is a mesh dependent effect, unlike that in (9.7). The most general mesh where we can easily improve estimate (9.11) is a rectangular mesh. For some other examples/counterexamples, see Appendix B.

Suppose the mesh on Ω is rectangular with nodes (x_i, y_j). Then $R_h\partial s/\partial y = \Pi_h\partial s/\partial y$ for any piecewise bilinear s, where Π_h is defined by (8.3). Using this and the orthogonal projection property of both R_h and Π_h, we have

$$\Big(\sigma_2 - R_h\sigma_2, \frac{\partial s}{\partial y}\Big) = \Big(\sigma_2, \frac{\partial s}{\partial y} - \Pi_h\frac{\partial s}{\partial y}\Big) = \Big(\sigma_2 - \Pi_h\sigma_2, \frac{\partial s}{\partial y}\Big).$$

Here, since $x_i - x_{i-1} \le hL$, we have by standard 1D approximation theory

$$\|\sigma_2 - \Pi_h\sigma_2\| \le \frac{hL}{\pi}\left\|\frac{\partial\sigma_2}{\partial x}\right\|.$$

This allows us to first improve (9.8) as

$$\left|\Big(\sigma_2 - R_h\sigma_2, \frac{\partial s}{\partial y}\Big)\right| \le ChL\left\|\frac{\partial\sigma_2}{\partial x}\right\|\|\mathbf{V}\|_h,$$

and then, by (9.3) and Hypothesis 2.1, to improve (9.11) as

$$\left|\Big(\sigma_2 - R_h\sigma_2, \frac{\partial s}{\partial y}\Big)\right| \le CQh\|\mathbf{U}\|\|\mathbf{V}\|_h, \tag{9.12}$$

so there arises no error amplification on a rectangular mesh.

Combining estimates (9.7) and (9.11) or (9.12) we have a bound for the term $F(\mathbf{V})$ in (7.3), so it remains to bound the second term $G(\mathbf{V})$ there. This is more straightforward. We have

$$\tau(\mathbf{U}) = \tau(\mathbf{U}_{bm}) + \tau(\mathbf{U}_{br}) = -\tfrac{1}{2}\bar{\mu}\gamma\, y^2\theta''(x) + \tau_{br}, \tag{9.13}$$

so estimate (7.13) holds by Hypothesis 2.1. Hence the error contribution from $G(\mathbf{V})$ in (7.14) is uniform with respect to t on general meshes.

For the improved estimate (8.1) to hold uniformly in case of rectangular meshes, we again need to assume that (8.7) holds with Q_1 independent of t. Since

$$\begin{aligned} L^3 &\int_{\partial\Omega} \Big[|n_y|\Big(\frac{\partial\tau}{\partial x}\Big)^2 dx + |n_x|\Big(\frac{\partial\tau}{\partial y}\Big)^2 dy\Big] \\ &\le 2\bar{\mu}^2\gamma^2\Big\{\frac{1}{32}t^4L^7\int_0^L(\theta''')^2\,dx + \frac{1}{12}t^3L^6[(\theta''(0))^2 + (\theta''(L))^2]\Big\} \\ &\qquad + 2L^3\int_{\partial\Omega}\Big[|n_y|\Big(\frac{\partial\tau_{br}}{\partial x}\Big)^2 dx + |n_x|\Big(\frac{\partial\tau_{br}}{\partial y}\Big)^2 dy\Big] \end{aligned}$$

by (9.13), it is sufficient to assume that the estimate

$$\begin{aligned} &\frac{1}{16}t^4L^7\int_0^L(\theta''')^2\,dx + \frac{1}{6}t^3L^6[(\theta''(0))^2 + (\theta''(L))^2] \\ &\qquad + 2L^3\int_{\partial\Omega}\Big[|n_x|\Big(\frac{\partial\tau_{br}}{\partial y}\Big)^2 dy + |n_y|\Big(\frac{\partial\tau_{br}}{\partial x}\Big)^2 dx\Big] \le Q_1^2|||\mathbf{U}_b|||^2 \end{aligned} \tag{9.14}$$

holds with Q_1 independent of t. This assumption again slightly enhances the regularity assumed previously in Hypothesis 2.1. But in view of expansion (2.11), the assumption is realistic.

Let us finally note that in view of (9.1), estimates (9.6) and (9.10) hold uniformly with respect to $\bar{\lambda}$ under Hypothesis 2.2. On the other hand when $\mathbf{U} = \mathbf{U}_{sm}$ or $\mathbf{U} = \mathbf{U}_{bm}$, $\sigma_1(\mathbf{U})$ and $\sigma_2(\mathbf{U})$ are independent of $\bar{\lambda}$, so all the above bounds are uniform in $\bar{\lambda}$ under Hypothesis 2.2.

We summarize the results of this section (see also Appendix B) in

Theorem 9.1 (Consistency error on thin domains) *Assume that* $\mathbf{U}$ *satisfies Hypothesis 2.1 on a rectangular reference domain with* $t = H/L$ *variable, and assume a family of rectangular meshes. Assume further that* $\alpha_1 = \gamma + 1$ *in* (5.6). *Then in both the pure stretching and the pure bending state of deformation, the consistency error of scheme* (6.6), (5.6) *satisfies estimate* (7.14) *uniformly in* t *when* $0 < t \le 1$. *In the former case this holds even on general meshes. On rectangular meshes, estimate* (8.1) *likewise holds uniformly in* t*, provided the enhanced regularity estimates* (8.7) *(stretching state) and* (9.14) *(bending state) hold with* Q_1 *independent of* t. *On more* *general meshes, the consistency error is amplified in general by factor t^{-1} in the bending state of deformation, except when $\gamma = 0$. If $\alpha_1 \neq \gamma + 1$, the amplification takes place on any mesh and for any γ. The stated bounds are uniform with respect to $\bar{\lambda}$ under Hypothesis 2.2.*

As a final curiosity, let us point out that when $\beta_1 = 0$ in (9.1), $\mathbf{U} = \mathbf{U}_{bm}$, and $\theta'' = 0$ (the simplest bending mode), then $\sigma_1 = \tau = \partial\sigma_2/\partial x = 0$, so when the mesh is rectangular, the consistency error vanishes. In the next section we see that even the approximation error vanishes in this special case.

10 Thin domains: Approximation error

In this section, we bound the approximation error (6.11) in case of parametrized rectangular domains. This final step of the error analysis is apparently rather decisive, as the ultimate motivation of modification (5.1) was to dampen the large approximation error of the standard bilinear FEM.

As noted above (see section 6), the modified scheme (6.6) still suffers from trapezoidal locking, so the parametric error growth in e_A can be avoided at most on parallelogram (or nearly parallelogram) meshes. For simplicity, we consider here only the case of a rectangular mesh with nodes (x_i, y_j), $0 = x_0 < x_1 < \cdots < x_I = L$, $-H/2 = y_0 < y_1 < \cdots < y_J = H/2$.

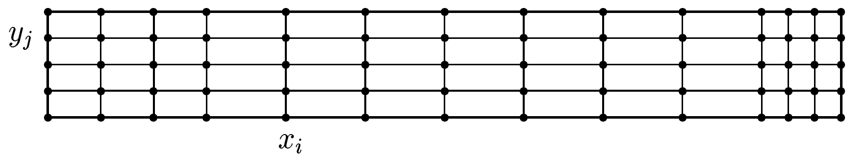

Figure 8: Rectangular mesh on the reference domain.

In addition to restricting the mesh, we will also weaken the seminorm $|||\cdot|||_h$ slightly from the previous definition (6.1), (5.6). That this is necessary, is seen by a simple counterexample. Assume a single-layer rectangular mesh as in section 4, let again the exact solution be given by (4.1), (4.4), and let $\tilde{\mathbf{U}}_h = (\tilde{u}_h, \tilde{v}_h) \in \mathcal{U}_h$ be the best approximation to $\mathbf{U}$ defined by (6.8). Then $\tilde{\mathbf{U}}_h$ must be of the form (4.6), so that $\partial\tilde{v}_h/\partial y = 0$. By (4.1), (4.4), (4.5) we have also $|||\mathbf{U}||| = \gamma^{-1}\|\partial v/\partial y\|$ in this case, so $\|\partial(v - \tilde{v}_h)/\partial y\|/|||\mathbf{U}||| = \gamma$. However, by the second estimate in Theorem 6.1 and by (6.2), $\|\partial(v - \tilde{v}_h)/\partial y\| \leq C|||\mathbf{U} - \tilde{\mathbf{U}}_h|||_h$, so it follows that $e_A \geq \gamma/C$. Hence there is no convergence in this case unless $\gamma = 0$.

As the example shows, the second stability estimate of Theorem 6.1 is critical. If this holds for all $\mathbf{V} \in \mathcal{U}$, then it is not possible to bound the approximation error uniformly optimally with respect to t. Thus we should modify the seminorm $|||\cdot|||_h$ in such a way that this critical estimate only remains valid for $\mathbf{V} \in \mathcal{U}_h$. We do this by rewriting the second functional in (5.4) in terms of the horizontal averaging operator Π_h (see (8.3)) as

$$\mathcal{B}_2(\mathbf{U}, \mathbf{U}) = 2\beta_2\bar{\mu}\int_\Omega \left[\left(\frac{\partial v}{\partial y}\right)^2 - \left(\Pi_h\frac{\partial v}{\partial y}\right)^2\right] dxdy \tag{10.1}$$

and by modifying the last terms in (5.5), (5.6) accordingly. Then since $\|\Pi_h f\| \geq \|R_h f\|$, $f \in L_2(\Omega)$, and since $\Pi_h \partial v/\partial y = R_h \partial v/\partial y$ when v is piecewise bilinear, the effect of the modification is to weaken $|||\cdot|||_h$ as a seminorm of $\mathcal{U}$ while leaving it unchanged as a norm of $\mathcal{U}_h$. Below we assume this new definition of $\mathcal{A}_h$ in (6.1) when bounding the approximation error (6.11). The consistency error bounds above remain valid, since e_C only depends on how $|||\cdot|||_h$ is defined in $\mathcal{U}_h$ (see (6.12)).

The approximation errors from $\mathbf{U} = \mathbf{U}_s$ in (2.5) and from $\mathbf{U} = \mathbf{U}_{br}$ in (2.7) are easily bounded. In both cases the uniformly optimal approximability (in the sense of estimate (7.2)) is immediate from (7.1) and from Hypothesis 2.1 when $\bar{\lambda} \leq C$, and the extension to uniform bounds with respect to $\bar{\lambda}$ follows the reasoning of section 7. So we may concentrate on estimating the approximation error of the asymptotic bending mode $\mathbf{U} = \mathbf{U}_{bm}$. We rewrite this now as

$$\begin{aligned} u &= y\theta(x), \\ v &= w(x) + \rho(y)\psi(x), \end{aligned} \tag{10.2}$$

where $\rho(y) = y^2/2$, so that (recall (2.8))

$$y - \rho' = 0, \quad \psi + \gamma\theta' = 0, \quad \theta + w' = 0. \tag{10.3}$$

When bounding e_A, we apply Principle A1 of section 6, looking for $\tilde{\mathbf{U}} = (\tilde{u}, \tilde{v}) \in \mathcal{U}_h$ that "simulates" $\tilde{\mathbf{U}}_h$, i. e., approximately minimizes $\|\mathbf{U} - \tilde{\mathbf{U}}\|_h$. Guided by (10.2), we assume that such a $\tilde{\mathbf{U}}$ can be found in the form

$$\begin{aligned} \tilde{u} &= y\tilde{\theta}(x), \\ \tilde{v} &= \tilde{w}(x) + \tilde{\rho}(y)\tilde{\psi}(x), \end{aligned} \tag{10.4}$$

where $\tilde{\theta}, \tilde{w}, \tilde{\psi}, \tilde{\rho}$ are continuous piecewise linear functions on the one-dimensional grids induced by the rectangular mesh. – Note that we could not try (10.4) on a general mesh.

We start by choosing $\tilde{\rho}$ and $\tilde{\psi}$ to be standard interpolants, so that

$$\tilde{\rho}(y_j) = \rho(y_j), \quad j = 0 \dots J, \tag{10.5}$$

$$\tilde{\psi}(x_i) = \psi(x_i), \quad i = 0 \dots I. \tag{10.6}$$

Interpolation conditions would also be natural for defining $\tilde{\theta}$ and $\tilde{w}$ in (10.4). However, it turns out that this choice would not lead to an optimal bound of e_A (see below), so we impose the interpolation conditions only at the endpoints, where they may be forced by the kinematical constraints:

$$\tilde{\theta}(x) = \theta(x), \quad \tilde{w}(x) = w(x), \quad x = 0, L. \tag{10.7}$$

Together with (10.2) and (10.4)–(10.6), (10.7) assures that $\tilde{\mathbf{U}}(x_i, y_j) = \mathbf{U}(x_i, y_j)$ for $i = 0, I$ and $j = 0 \dots J$, so any possible constraints at $x = 0, L$ are taken care of.

Assuming constraints (10.5) - (10.7) in (10.4), it remains to set $\tilde{\theta}(x_i)$, $\tilde{w}(x_i)$, $i = 1 \dots I-1$, so that $\|\mathbf{U} - \tilde{\mathbf{U}}\|_h^2 = \mathcal{A}_h(\mathbf{U} - \tilde{\mathbf{U}}, \mathbf{U} - \tilde{\mathbf{U}})$ is approximately minimized. In what follows, we use expression (5.5) of $\mathcal{A}_h(\mathbf{U} - \tilde{\mathbf{U}}, \mathbf{U} - \tilde{\mathbf{U}})$ and consider each of the five terms there separately. Our aim is to show that with a proper definition of $\tilde{\mathbf{U}}$, each term can be bounded by $CQ^2h^2\|\mathbf{U}\|^2$, where

Q is the constant in Hypothesis 2.1. Then it follows that $\|\mathbf{U}-\tilde{\mathbf{U}}\|_h \le CQh$, hence $e_A \le CQh$ by Principle A1.

We need the pointwise bounds

$$
\begin{aligned}
|\tilde{\rho}(y)| &\le \frac{1}{4}H^2 = \frac{1}{4}t^2L^2, \\
|\tilde{\rho}'(y)| &\le \max_{-\frac{H}{2}\le y\le \frac{H}{2}} |\rho'(y)| = tL, \\
|(\rho-\tilde{\rho})(y)| &\le \frac{1}{4}hL \max_{-\frac{H}{2}\le y\le \frac{H}{2}} |\rho'(y)| = \frac{1}{4}thL^2,
\end{aligned}
\tag{10.8}
$$

and the L_2 bounds

$$
\int_0^L (\psi-\tilde{\psi})^2\,dx \le \pi^{-2}h^2L^2\int_0^L (\psi')^2\,dx, \tag{10.9}
$$

$$
\int_0^L (\psi-\tilde{\psi})^2\,dx \le \pi^{-4}h^4L^4\int_0^L (\psi'')^2\,dx, \tag{10.10}
$$

$$
\int_0^L (\psi'-\tilde{\psi}')^2\,dx \le \pi^{-2}h^2L^2\int_0^L (\psi'')^2\,dx. \tag{10.11}
$$

These are classical bounds derivable from the interpolation conditions (10.5)–(10.6) when $x_i - x_{i-1} \le hL$ and $y_j - y_{j-1} \le hL$. We note that the constants in these estimates are optimal and that the interpolant actually is the best approximation in (10.11). So far we have not defined $\tilde{\theta}$, but we assume the definition to be such that an estimate analogous to (10.11) holds, viz.

$$
\int_0^L (\theta'-\tilde{\theta}')^2\,dx \le Ch^2L^2\int_0^L (\theta'')^2\,dx. \tag{10.12}
$$

We now start estimating $\mathcal{A}_h(\mathbf{U}-\tilde{\mathbf{U}},\mathbf{U}-\tilde{\mathbf{U}})$ from (5.5). First, by (10.2), (10.4), (10.12) and Hypothesis 2.1,

$$
\begin{aligned}
D\left\|\frac{\partial}{\partial x}(u-\tilde{u})\right\|^2 &= \tfrac{1}{12}Dt^3L^3\int_0^L (\theta'-\tilde{\theta}')^2\,dx \\
&\le Ch^2t^3L^5\int_0^L (\theta'')^2\,dx \le CQ^2h^2\|\mathbf{U}\|^2,
\end{aligned}
\tag{10.13}
$$

so the first term in (5.5) is clarified (so far as (10.12) holds).

The second term in (5.5) we first expand as

$$
\begin{aligned}
R_h\Big[\gamma\frac{\partial}{\partial x}(u-\tilde{u}) &+ \frac{\partial}{\partial y}(v-\tilde{v})\Big] \\
&= -R_h[\gamma\rho'(y)\tilde{\theta}'(x) + \tilde{\rho}'(y)\tilde{\psi}(x)] \\
&= -R_h[\gamma\tilde{\rho}'(y)\tilde{\theta}'(x) + \tilde{\rho}'(y)\tilde{\psi}(x)] \\
&= -\tilde{\rho}'(y)\Pi_h[\gamma\tilde{\theta}'(x) + \tilde{\psi}(x)] \\
&= \tilde{\rho}'(y)\Pi_h[\gamma(\theta'-\tilde{\theta}')(x) + (\psi-\tilde{\psi})(x)].
\end{aligned}
\tag{10.14}
$$

Here we used $y = \rho'$, $\psi = -\gamma\theta'$ from (10.3) and applied the identity

$$R_h[(\rho' - \tilde{\rho}')(y)\,\tilde{\theta}'(x)] = \Lambda_h(\rho' - \tilde{\rho}')(y)\,\tilde{\theta}'(x) = 0,$$

where the first part follows since $\tilde{\theta}'$ is piecewise constant and the second part is due to the interpolation conditions (10.5). (Here and below, Π_h and Λ_h are understood as the 1D equivalents of operators (8.3)–(8.4).) From (10.14) we can now proceed using estimates (10.8), (10.12) and (10.9). Recalling once more that $\psi = -\gamma\theta'$ and applying Hypothesis 2.1 we have

$$\begin{aligned}\frac{\bar{\lambda}}{\gamma}\left\|R_h\Big[\gamma\frac{\partial}{\partial x}(u-\tilde{u}) + \frac{\partial}{\partial y}(v-\tilde{v})\Big]\right\|^2 &\leq C\bar{\lambda}\gamma h^2 t^3 L^5 \int_0^L (\theta'')^2\,dx \\ &\leq CQ^2h^2|\!|\!|\mathbf{U}|\!|\!|^2. \qquad (10.15)\end{aligned}$$

(assuming $\bar{\lambda} \leq C$). Thus the second term in (5.5) is clarified. Note also that in case of a single-layer mesh the error term (10.14) actually vanishes, since then $\tilde{\rho}' = 0$.

The third term in (5.5) is expanded as

$$\begin{aligned}R_h\Big[\frac{\partial}{\partial y}(u-\tilde{u}) + \frac{\partial}{\partial x}(v-\tilde{v})\Big] &= \Pi_h[(\theta-\tilde{\theta})(x) + (w'-\tilde{w}')(x)] \\ &\quad + R_h[(\rho-\tilde{\rho})(y)\psi'(x) + \tilde{\rho}(y)(\psi'-\tilde{\psi}')(x)]. \qquad (10.16)\end{aligned}$$

To proceed, we define now $\tilde{\theta}, \tilde{w}$ so that the first term on the right side of (10.16) vanishes. Since $\theta + w' = 0$ by (10.3), this is achievable by defining $\tilde{\theta}, \tilde{w}$ so that

$$\Pi_h(\tilde{\theta} + \tilde{w}') = \Pi_h\tilde{\theta} + \tilde{w}' = 0. \qquad (10.17)$$

To impose this constraint in practice, assume that $\tilde{\theta}$ is defined first and let then $\tilde{w}$ be defined in terms of $\tilde{\theta}$ as

$$\tilde{w}(x) = w(0) - \int_0^x (\Pi_h\tilde{\theta})(x')\,dx'. \qquad (10.18)$$

Then $\tilde{w}$ is continuous and piecewise linear as required, (10.17) holds, and $\tilde{w}(0) = w(0)$ as required by (10.7). Finally since

$$(w-\tilde{w})(x_i) = \int_0^{x_i} (\Pi_h\tilde{\theta} - \theta)\,dx = \int_0^{x_i} (\tilde{\theta} - \theta)\,dx, \quad i = 0\ldots I, \qquad (10.19)$$

by (10.3) and (10.18), we see that the remaining kinematical constraint on $\tilde{w}$ can be imposed by assuming

$$\int_0^L (\theta - \tilde{\theta})\,dx = 0. \qquad (10.20)$$

It then remains to specify $\tilde{\theta}$ so that the three constraints in (10.7), (10.20) are satisfied. This is apparently possible whenever $I \geq 2$. Assuming this, let us define $\tilde{\theta}$ by minimizing the left side of (10.12) under the mentioned constraints. By treating (10.20) with a Lagrange multiplier, the solution of this constrained minimization problem is found to be

$$\tilde{\theta} = \breve{\theta} + \breve{\xi}, \tag{10.21}$$

where $\breve{\theta}$ is the standard interpolant of θ, and $\breve{\xi}$ interpolates

$$\xi = \xi_0 x(L - x), \tag{10.22}$$

where the coefficient ξ_0 is defined so that (10.20) holds, i. e., so that

$$\int_0^L \breve{\xi}\,dx = \int_0^L (\theta - \breve{\theta})\,dx. \tag{10.23}$$

It remains to check that (10.12) holds. Here we have to first exclude anomalous meshes where constant C in (10.12) can actually grow without limit. Namely, this happens if one of the mesh intervals $[x_{i-1}, x_i]$ covers nearly the whole interval $[0, L]$ (so that $h \approx 1$). For example, let $\theta = \theta_0 x(L - x)$ and choose $x_1 = \delta L$, $x_2 = (1 - \delta)L$ for $I = 3$, where $\delta \ll 1$. Then $\breve{\theta} \sim \delta\theta_0 L^2$ is nearly zero, so by constraint (10.23) and by symmetry, $\breve{\xi}(x_1) = \breve{\xi}(x_2) \approx \frac{1}{6}\theta_0 L^2$. It follows that $|\breve{\xi}'(x)| \approx \frac{1}{6}\delta^{-1}|\theta_0|L$ on intervals $[x_0, x_1]$ and $[x_2, x_3]$, so estimate (10.12) can only hold with C proportional to δ^{-1} in this case. Below we simply exclude situations like this by imposing the specific coarse-mesh condition

$$\min_{i=1\ldots I-1} [x_i(L - x_i)]^{-1} \le CL^{-2}, \tag{10.24}$$

where C is a fixed constant. Note that this only rules out some unusual coarse meshes.

Assuming (10.24) we can now prove estimate (10.12). Note first that if the minimum in (10.24) is achieved at index i, then by (10.24) and (10.22) and since $\breve{\xi}$ interpolates ξ,

$$|\int_0^L \breve{\xi}\,dx| \ge \tfrac{1}{2}|\xi(x_i)|L \ge \tfrac{1}{2}C^{-1}|\xi_0|L^3 = 3C^{-1}|\int_0^L \xi\,dx|.$$

(where C is as in (10.24)). Using this and (10.23) we have now

$$\begin{aligned}\int_0^L (\breve{\xi}')^2\,dx \le \int_0^L (\xi')^2\,dx = 12L^{-3}\Big[\int_0^L \xi\,dx\Big]^2 \le CL^{-3}\Big[\int_0^L \breve{\xi}\,dx\Big]^2 \\ = CL^{-3}\Big[\int_0^L (\theta - \breve{\theta})\,dx\Big]^2 \le CL^{-2}\int_0^L (\theta - \breve{\theta})^2\,dx,\end{aligned}$$

so that by an interpolation error bound similar to (10.10)

$$\int_0^L (\breve{\xi}')^2\,dx \le Ch^4L^2\int_0^L (\theta'')^2\,dx. \tag{10.25}$$

Recall finally that when $\tilde{\theta} = \breve{\theta}$, the bound (10.12) holds with $C = \pi^{-2}$ (optimal constant). Then if $\tilde{\theta}$ is defined by (10.21), it follows from estimate (10.25) and from the triangle inequality that (10.12) holds with $C = \pi^{-2}(1+\mathcal{O}(h^2))$, i. e., nearly optimally. Thus we have finished the construction of $\tilde{\mathbf{U}}$, and the assumptions made above have been verified under the mild additional condition (10.24) on the meshes.

To finish bounding the third term in (5.5), we now apply (10.8) and (10.11) together with Hypothesis 2.1 to bound the non-vanishing terms in (10.16). We get

$$\begin{aligned}\bar{\mu}\left\| R_h\Big[\frac{\partial}{\partial y}(u-\tilde{u}) + \frac{\partial}{\partial x}(v-\tilde{v})\Big]\right\|^2 &\le Ch^2\int_0^L [t^3L^5(\theta'')^2 + t^5L^7(\theta''')^2]\,dx \\ &\le CQ^2h^2|\!|\!|\mathbf{U}|\!|\!|^2, \end{aligned} \tag{10.26}$$

so this term is clarified.

The fourth term in (5.5) is of the same order as the leading term, so the bound (10.13) applies to this term as well when $|\beta_1| \le C$. Thus choosing $\beta_1 \ne 0$ does not cause “interpolation locking” (compare with Theorem 8.1).

Finally we come to the last term in (5.5), which now is rewritten according to (10.1). Here we apply first the identity

$$\int_\Omega [\sigma^2 - (\Pi_h\sigma)^2]\,dxdy = \int_\Omega (\sigma^2 - \Pi_h\sigma)^2\,dxdy, \tag{10.27}$$

valid because Π_h is an L_2 projection. Upon expanding here $\sigma = \partial(v-\tilde{v})/\partial y$ as

$$\begin{aligned}\sigma &= \rho'(y)\psi(x) - \tilde{\rho}'(y)\tilde{\psi}(x) \\ &= (\rho' - \tilde{\rho}')(y)\,\psi(x) + \tilde{\rho}'(y)\,(\psi - \tilde{\psi})(x),\end{aligned}$$

we conclude

$$\begin{aligned}\sigma - \Pi_h\sigma = (\rho' - \tilde{\rho}')(y)\,(\psi - \Pi_h\psi)(x) \\ + \tilde{\rho}'(y)[(\psi-\tilde{\psi})(x) - \Pi_h(\psi - \tilde{\psi})(x)].\end{aligned} \tag{10.28}$$

Note that replacing here Π_h by R_h on the left side would produce an extra term $(\Pi_h - R_h)\sigma = (\rho' - \Lambda_h\rho')(y)\Pi_h\psi(x)$ on the right. This could not be bounded in the desired way, so here is the (only) point where the modified definition (10.1) is needed.

Combining now (10.27), (10.28), using estimates (10.8) and (10.9) together with

$$\int_0^L (\psi - \Pi_h\psi)^2\,dx \le \pi^{-2}h^2L^2\int_0^L (\psi')^2\,dx,$$

inserting $\psi = -\gamma\theta'$, and finally applying Hypothesis 2.1, we get

$$2|\beta_2|\bar{\mu}\left(\left\|\frac{\partial}{\partial y}(v-\tilde{v})\right\|^2 - \left\|\Pi_h\frac{\partial}{\partial y}(v-\tilde{v})\right\|^2\right) \le CQ^2h^2|\!|\!|\mathbf{U}|\!|\!|^2, \tag{10.29}$$

so the final term of $\mathcal{A}_h(\mathbf{U}-\tilde{\mathbf{U}},\mathbf{U}-\tilde{\mathbf{U}})$ is clarified.

Upon summing up estimates (10.13), (10.15), (10.26), the mentioned bound for the fourth term in (5.5), and (10.29), we get $\|\mathbf{U}-\tilde{\mathbf{U}}\|_h \le CQh\|\mathbf{U}\|$, so $e_A \le CQh$ by Principle A1.

Let us point out that we actually only needed estimates (2.14) from Hypothesis 2.1 in the above analysis. It follows in particular that $e_A = 0$ when $\theta'' = 0$ (the simplest bending mode), since (2.14) then holds with $Q = 0$. In this case one has also $\tilde{\theta} = \theta$, so $\tilde{w}$ interpolates w by (10.19), and hence the above generalized interpolation reduces to standard interpolation. We conclude that $\tilde{\mathbf{U}}_h = \tilde{\mathbf{U}} =$ standard interpolant, and $e_A = 0$, when $\theta'' = 0$.

In Appendix C we extend the above construction to cover the doubly parametric situation where $\bar{\lambda}$ can be arbitrarily large and t arbitrarily small simultaneously. The above analysis fails in that case (except when $\theta'' = 0$), since constant C in estimate (10.15) is proportional to $\bar{\lambda}$. However, we can find a better approximation of $\tilde{\mathbf{U}}_h$ when $\bar{\lambda}$ is large, by constructing $\tilde{\mathbf{U}}$ in such a way that the whole term (10.14) vanishes, while the other estimates above are still preserved. The key of this alternative construction is the constraint $\Pi_h\tilde{\psi}+\gamma\tilde{\theta}' = 0$, which can be imposed simultaneously with (10.17) when $I \ge 3$ (see Appendix C). With such a definition of $\tilde{\mathbf{U}}$, the error bound $e_A \le CQh$ holds uniformly with respect to $\bar{\lambda}$, showing that volumetric locking is avoided as well.

Before summarizing these results, let us go through briefly also the standard "interpolation test" for bounding e_A. So let $\tilde{\mathbf{U}}$ be the standard interpolant of $\mathbf{U}$, as defined by (10.4) where now $\tilde{\rho}, \tilde{\theta}, \tilde{w}, \tilde{\psi}$ are all set by interpolation conditions. Then $\Pi_h(\theta' - \tilde{\theta}') = 0$ and $\Pi_h(w' - \tilde{w}') = 0$, so (10.14) and the leading term in (10.16) are simplified. Upon bounding these terms by estimates (10.9)–(10.10) and proceeding otherwise as above, we get

$$\frac{\|\mathbf{U}-\tilde{\mathbf{U}}\|_h}{\|\mathbf{U}\|} \le CP_2h + C\min\{P_2, P_3h\}\bar{\lambda}^{1/2}h + CP_2t^{-1}h^2, \tag{10.30}$$

where C is an absolute constant and P_k are defined by (compare with assumption (2.14))

$$P_k = L^{k-1}\left\{\int_0^L (\theta^{(k)})^2dx\right\}^{1/2}\left\{\int_0^L (\theta')^2dx\right\}^{-1/2}, \; k = 2,3. \tag{10.31}$$

In the special case where $\theta = \theta_0x(L-x)(x-L/2)$, $\theta_0 \ne 0$, and the mesh is a uniform double-layer mesh with $y_0 = -H/2$, $y_1 = 0$, $y_2 = H/2$, and $x_i = ihL$, $i = 0\dots I$, we have $P_2 = \sqrt{60}$, $P_3 = 12\sqrt{5}$, and

$$\frac{\|\mathbf{U}-\tilde{\mathbf{U}}\|_h^2}{\|\mathbf{U}\|^2} = \tfrac{9}{2}(\gamma/D)\bar{\lambda}h^4 + 6(\bar{\mu}/D)t^{-2}h^4 + \mathcal{O}(h^2+t^2+t^{-2}h^4),$$

so the leading error terms are as predicted by (10.30).

The conclusion from estimate (10.30) is that the standard interpolant as $\tilde{\mathbf{U}}$ leads to the best possible bound for e_A if (and in general, only if) either $\theta'' = 0$ or if the parameters $h, t, \bar{\lambda}$ are restricted so that $h/t \leq C$ and $\min\{1, P_3 h/P_2\}\bar{\lambda}^{1/2} \leq C$ for some fixed C. Beyond these constraints, the generalized interpolants constructed above and in Appendix C are needed. These constructions may then be viewed as simulating the actual mechanism by which the modified FE scheme avoids parametric error growth when $t \to 0$ or $\bar{\lambda} \to \infty$.

We summarize the analysis of this section & Appendix C in

Theorem 10.1 (The main approximation theorem) *Assume a rectangular reference domain and a rectangular mesh with nodes* (x_i, y_j), $i = 0 \ldots I \geq 2$, $j = 0 \ldots J$, *where* x_i *are chosen so that condition* (10.24) *holds with some given constant* C. *Let then* $\tilde{\mathbf{U}}_h$ *be defined according to* (6.8), *where* $\mathcal{A}_h$ *is defined by* (5.5)/(5.6) *with* $0 < \alpha_1, \alpha_2 \leq C$ *and where the exact solution* $\mathbf{U}$ *is the asymptotic bending mode* (10.2) *satisfying* (10.3) *and Hypothesis 2.1. Let further* $\tilde{\mathbf{U}} \in \mathcal{U}_h$ *be defined by* (10.4), *where* $\tilde{\rho}, \tilde{\psi}$ *satisfy* (10.5), (10.6), $\tilde{w}$ *is defined by* (10.18), *and* $\tilde{\theta}$ *is defined by* (10.21)–(10.23) *where* $\theta \to \check{\theta}$ *denotes piecewise linear interpolation. Assume finally that when defining the seminorm* $\|\!|\cdot|\!\|_h$ *in* (6.11) *according to* (6.1), *the last term of* $\mathcal{A}_h$ *is redefined according to* (10.1). *Then the approximation error* (6.11) *satisfies*

$$e_A \leq \frac{|\!|\!|\mathbf{U} - \tilde{\mathbf{U}}|\!|\!|_h}{|\!|\!|\mathbf{U}|\!|\!|} \leq CQh.$$

With an alternative definition of $\tilde{\mathbf{U}}$ *given in Appendix C, constant* C *in this estimate is independent of* $\bar{\lambda}$, *assuming* $I \geq 3$ *and excluding anomalous coarse meshes. Moreover, in case of the simplest bending mode* ($\theta'' = 0$) *one has* $\tilde{\mathbf{U}}_h = \tilde{\mathbf{U}} =$ *standard interpolant, and* $e_A = 0$. *More generally, the standard interpolation error is bounded according to* (10.30)–(10.31).

Let us finally recall the remark made at the end of section 9: When approximating the simplest bending mode on a rectangular mesh, the consistency error likewise vanishes, provided $\alpha_1 = \gamma + 1$ in (5.6). Thus under such assumptions, the modified FE approximation according to (6.6) reduces to standard interpolation, i.e., the FE solution is exact at the nodal points. As expected from Theorem 5.1, the Turner rectangle shares this same property, and indeed, that was the very idea of the pioneering formulation in [30].

Our mathematical story of the first locking-free plane-elastic finite element ends here.

A Proof of Theorem 6.1

Let us first show that the estimates stated in Theorem 6.1 cannot be improved. To this end, assume a rectangular domain aligned with coordinates x, y and a uniform rectangular mesh with nodes (x_i, y_j), $x_i - x_{i-1} = a$, $y_j - y_{j-1} = b$. Let then $\mathbf{V} = (r, s) \in \mathcal{U}_h$ be such that $r(x_i, y_j) = c_1(-1)^{i+j}$, $s(x_j, y_j) = c_2(-1)^{i+j}$ for some constants c_1, c_2. Then the averaged strains all vanish, so by (5.6)

$$\mathcal{A}_h(\mathbf{V}, \mathbf{V}) = 2\alpha_1\bar{\mu} \left\| \frac{\partial r}{\partial x} \right\|^2 + 2\alpha_2\bar{\mu} \left\| \frac{\partial s}{\partial y} \right\|^2 .$$

Choosing $c_2 = 0$, resp. $c_1 = 0$, we conclude: (1) The first two inequalities in Theorem 6.1 reduce to equalities for the assumed $\mathbf{V}$ when $\alpha_1, \alpha_2 \leq 1$. (2) If assumption (5.8) is violated, we find a spurious mode of the assumed form such that $\mathcal{A}_h(\mathbf{V}, \mathbf{V}) \leq 0$. Thus the first two estimates are essentially optimal, and condition (5.8) is necessary for stability.

Next let $\alpha_1 = \alpha_2 = \alpha > 0$ and choose $c_2 = 0$ if $a > b$ or $c_1 = 0$ if $a \leq b$. Then

$$\begin{aligned} \left\| \frac{\partial r}{\partial y} + \frac{\partial s}{\partial x} \right\|^2 / \mathcal{A}_h(\mathbf{V}, \mathbf{V}) = \frac{1}{2\alpha\bar{\mu}} \frac{a^2c_1^2 + b^2c_2^2}{b^2c_1^2 + a^2c_2^2} &= \frac{1}{2\alpha\bar{\mu}} \max\left\{ \left(\frac{a}{b}\right)^2, \left(\frac{b}{a}\right)^2 \right\} \\ &= \frac{1}{2\alpha\bar{\mu}} M_h^2, \end{aligned}$$

so the third inequality in Theorem 6.1 is optimal on a rectangular mesh.

Consider finally a quadrilateral mesh obtained by subdividing each rectangle in the above mesh into two quadrilaterals of equal shape, with a line that passes through the vertical sides of the rectangle. Assume that $a > b$, and let the vertical sides of the quadrilaterals be of length c and $b - c$, where $0 < c \leq b/2$. Then $M_h \sim a/b$ and $N_h \sim b/c$. Further let $\mathbf{V} = (r, s) \in \mathcal{U}_h$ be such that $s = 0$ and $r(x_i, y_j) = (-1)^i$ at the nodes of the original mesh and $r = -(-1)^i$ at the new nodes at $x = x_i$. Then the averaged strains again vanish and we get for $\alpha_1 = \alpha_2 = \alpha > 0$

$$\left\| \frac{\partial r}{\partial y} + \frac{\partial s}{\partial x} \right\|^2 / \mathcal{A}_h(\mathbf{V}, \mathbf{V}) \sim \frac{1}{\alpha\bar{\mu}} \left(\frac{a}{b}\right)^2 \left(1 + \log\frac{b}{c}\right) \sim \frac{1}{\alpha\bar{\mu}} M_h^2(1 + \log N_h),$$

so the third inequality of Theorem 6.1 is not improvable on this mesh.

It remains to prove that the asserted inequalities actually hold. The first two estimates are easy: Since $0 \leq \|R_h\sigma\| \leq \|\sigma\|$, we have

$$\|R_h\sigma\|^2 + \alpha(\|\sigma\|^2 - \|R_h\sigma\|^2) \geq \min\{\alpha, 1\}\|\sigma\|^2.$$

Applying this in (5.6), we conclude that for any $\mathbf{V} \in \mathcal{U}$

$$\mathcal{A}_h(\mathbf{V}, \mathbf{V}) \geq 2\delta_1\bar{\mu} \left\| \frac{\partial r}{\partial x} \right\|^2 + 2\delta_2\bar{\mu} \left\| \frac{\partial s}{\partial y} \right\|^2 , \tag{A.1}$$

where $\delta_i = \min\{\alpha_i, 1\}$. The first two inequalities follow.

The proof of the third inequality is more technical. The core is the following lemma.

Lemma A.1 *Let K be a convex quadrilateral satisfying the angle condition* (3.1) *and let $\Xi = \Xi(x, y)$ be the isoparametric bilinear function on K that vanishes at the midpoint of every edge of K and attains the value ± 1 at the vertices. Then for any $c_1, c_2 \in \mathcal{R}$*

$$\int_K \left(c_1 \frac{\partial \Xi}{\partial y} + c_2 \frac{\partial \Xi}{\partial x}\right)^2 dxdy \le C_K \int_K \left[c_1^2 \left(\frac{\partial \Xi}{\partial x}\right)^2 + c_2^2 \left(\frac{\partial \Xi}{\partial y}\right)^2\right] dxdy,$$

where C_K is a finite constant depending on the shape of K. The constant can be further bounded as $C_K \le C m_K^2 (1 + \log n_K)$, where m_K, n_K are the geometric shape factors of K as defined in (3.2), *and C is an absolute constant.*

Proof. We need to bound the largest eigenvalue $\lambda_{\max}$ of $B^{-1}A$, where $A = (a_{ij})$ and $B = (b_{ij})$ are symmetric 2 by 2 matrices defined by

$$\begin{aligned}
a_{11} &= \int_K \left(\frac{\partial \Xi}{\partial y}\right)^2 dxdy, \quad a_{22} = \int_K \left(\frac{\partial \Xi}{\partial x}\right)^2 dxdy, \\
a_{12} &= \int_K \frac{\partial \Xi}{\partial x} \frac{\partial \Xi}{\partial y} dxdy = a_{21}, \\
b_{11} &= a_{22}, \quad b_{22} = a_{11}, \quad b_{12} = b_{21} = 0.
\end{aligned}$$

The claim is that $\lambda_{\max} \le C_K = C m_K^2 (1 + \log n_K)$.

Without loss of generality, we may assume that K has one vertex at the origin. Then $\lambda_{\max}$ depends on the six coordinates x_i, y_i of the remaining three vertices. Since the asserted inequality is invariant under scaling (i. e., $\lambda_{\max}$ depends only on the shape of K), we may assume that, say, $|x_i| \le 1$, $|y_i| \le 1$. The assumed angle conditions impose further restrictions on x_i, y_i. Obviously, $\lambda_{\max}$ is a continuous function of x_i, y_i on the resulting (compact) set, except at points where matrix B becomes singular. This can only happen at degenerate shapes of K with $m_K = \infty$ or $n_K = \infty$, so if we asssume that $m_K \le M$ and $n_K \le N$ for some given M, N, then $\lambda_{\max}$ is uniformly bounded by a constant depending on M, N and ζ. The first conclusion is then that on strongly regular meshes, constant C_K is uniformly bounded.

The second conclusion concerns aligned rectangular meshes. Let $K = \{(x, y) \mid -a < x < a,\ -b < y < b\}$, so that $m_K = \max\{a/b, b/a\}$ and $\Xi(x, y) = (x/a)(y/b)$. Then $\lambda_{\max} = m_K^2$, so the assertion of the lemma holds (and is not improvable) in this case.

So far the proof is incomplete, but we already covered the most important cases: On strongly regular meshes C_K is uniformly bounded, and on aligned rectangular meshes $C_K = m_K^2$ for each K(optimal value). We stop here and skip the general case, where the details of the proof get more technical and less interesting. □

Accepting Lemma A.1, we can now finish the proof of Theorem 6.1. Let r, s be any isoparametric bilinear functions on a given (convex and nondegenerate) quadrilateral K. Then we may split r, s as

$$\begin{aligned} r(x,y) &= r_0(x,y) + c_1\Xi(x,y) = r_0(x,y) + r_1(x,y), \\ s(x,y) &= s_0(x,y) + c_2\Xi(x,y) = s_0(x,y) + s_1(x,y), \end{aligned}$$

where c_1, c_2 are constants and r_0, s_0 are polynomials of degree 1. Since Ξ is linear along each edge of K and vanishes at the midpoints of the edges, it follows that

$$\int_K \frac{\partial \Xi}{\partial x}\, dxdy = \int_K \frac{\partial \Xi}{\partial y}\, dxdy = 0.$$

Hence the strains related to r_1, s_1 have zero averages. On the other hand, the strains related to r_0, s_0 are constant, so the assumed splitting is L_2 - orthogonal on straints. In particular,

$$\begin{aligned} \int_K \Big(\frac{\partial r}{\partial y} + \frac{\partial s}{\partial x}\Big)^2 dxdy &= \int_K \Big(\frac{\partial r_0}{\partial y} + \frac{\partial s_0}{\partial x}\Big)^2 dxdy + \int_K \Big(\frac{\partial r_1}{\partial y} + \frac{\partial s_1}{\partial x}\Big)^2 dxdy \\ &= \int_K \Big[R_h\Big(\frac{\partial r}{\partial y} + \frac{\partial s}{\partial x}\Big)\Big]^2 dxdy + \int_K \Big(\frac{\partial r_1}{\partial y} + \frac{\partial s_1}{\partial x}\Big)^2 dxdy, \end{aligned}$$

where the second identity follows since $\partial r_0/\partial y + \partial s_0/\partial x$ is constant and $\partial r_1/\partial y + \partial s_1/\partial x$ has zero average. Here we can now use Lemma A.1 to bound the second term on the right, thus obtaining

$$\begin{aligned} \int_K \Big(\frac{\partial r}{\partial y} + \frac{\partial s}{\partial x}\Big)^2 dxdy \le{}& \int_K \Big[R_h\Big(\frac{\partial r}{\partial y} + \frac{\partial s}{\partial x}\Big)\Big]^2 dxdy \\ &+ C_K \int_K \Big[\Big(\frac{\partial r_1}{\partial x}\Big)^2 + \Big(\frac{\partial s_1}{\partial y}\Big)^2\Big] dxdy. \end{aligned}$$

Finally, by the strain orthogonality,

$$\int_K \Big[\Big(\frac{\partial r_1}{\partial x}\Big)^2 + \Big(\frac{\partial s_1}{\partial y}\Big)^2\Big] dxdy \le \int_K \Big[\Big(\frac{\partial r}{\partial x}\Big)^2 + \Big(\frac{\partial s}{\partial y}\Big)^2\Big] dxdy.$$

Combining the last inequalities and using $C_K \le CM_h^2(1 + \log N_h)$, we get

$$\begin{aligned} \int_K \Big(\frac{\partial r}{\partial y} + \frac{\partial s}{\partial x}\Big)^2 dxdy \le{}& \int_K \Big[R_h\Big(\frac{\partial r}{\partial y} + \frac{\partial s}{\partial x}\Big)\Big]^2 dxdy \\ &+ CM_h^2(1 + \log N_h) \int_K \Big[\Big(\frac{\partial r}{\partial x}\Big)^2 + \Big(\frac{\partial s}{\partial y}\Big)^2\Big] dxdy. \end{aligned}$$

Upon summing this inequality over K and using (5.6) and (A.1), we finally get

$$\begin{aligned}\left\|\frac{\partial r}{\partial y}+\frac{\partial s}{\partial x}\right\|^2 &\le C_{h,\delta,\bar\mu}\left\{\bar\mu\left\|R_h\left(\frac{\partial r}{\partial y}+\frac{\partial s}{\partial x}\right)\right\|^2+2\bar\mu\delta\left\|\frac{\partial r}{\partial x}\right\|^2+2\bar\mu\delta\left\|\frac{\partial s}{\partial y}\right\|^2\right\}\\ &\le 2C_{h,\delta,\bar\mu}|||\mathbf{V}|||_h^2,\end{aligned}$$

where $\delta=\min\{\delta_1,\delta_2\}$ and

$$C_{h,\delta,\bar\mu}=\bar\mu^{-1}+CM_h^2\left(1+\log N_h\right)\left(2\bar\mu\delta\right)^{-1}\le CM_h^2\left(1+\log N_h\right)\left(\bar\mu\delta\right)^{-1}.$$

This proves the third inequality in Theorem 6.1 and hence the whole theorem. □

Let us finally note that unlike the first two inequalities in Theorem 6.1, which hold in the entire energy space $\mathcal{U}$, the third inequality does not admit any meaningful extension beyond the finite element space. Assume, for example, a rectangular domain and rectangular mesh with nodes (x_i, y_j), and set

$$\begin{aligned} r(x,y)&=(2y-y_j-y_{j-1})^2/(y_j-y_{j-1})^2, &&(x,y)\in\Omega, && y_{j-1}<y<y_j,\\ s(x,y)&=(2x-x_i-x_{i-1})^2/(x_i-x_{i-1})^2, &&(x,y)\in\Omega, && x_{i-1}<x<x_i,\end{aligned}$$

Then $\mathbf{V}=(r,s)\in\mathcal{U}$ and $\partial r/\partial y+\partial s/\partial x\neq 0$, but $|||\mathbf{V}|||_h=0$, so $|||\cdot|||_h$ is just a seminorm of $\mathcal{U}$.

B Equilibrium locking

We complete here the proof of Theorem 9.1, by showing that the asserted consistency error amplification does occur in general, when either $\beta_1 \neq 0$ in (5.5) or when the mesh is not restricted.

Assume first that the mesh is rectangular but $\beta_1 \neq 0$. Consider again the example case of section 4 where $\mathbf{U} = \mathbf{U}_{bm}$, given by (4.1), (4.4), is approximated on a single-layer rectangular mesh. Assuming no kinematical constraints at $x = 0, L$, we can choose in this case $\mathbf{V} = (r, s) \in \mathcal{U}_h^0$ so that

$$r(x,y) = -\mathrm{sgn}(\beta_1)\, y\theta(x), \quad s(x,y) = \mathrm{sgn}(\beta_1) \int_0^x (\Pi_h\theta)(x')\, dx',$$

where Π_h is defined by (8.3). Then $\partial s/\partial y = 0$, $R_h \partial r/\partial x = 0$, and $R_h(\partial r/\partial y + \partial s/\partial x) = 0$, so

$$|||\mathbf{V}|||_h = (1 + 2\beta_1 \bar{\mu}/D)^{1/2} |||\mathbf{U}|||.$$

Moreover

$$(\mathcal{A} - \mathcal{A}_h)(\mathbf{U}, \mathbf{V}) = 2|\beta_1|(\bar{\mu}/D)|||\mathbf{U}|||^2,$$

so by Principle C2 of section 6

$$e_C \geq 2|\beta_1|\,(\bar{\mu}/D)\,(1 + 2\beta_1\bar{\mu}/D)^{-1/2}.$$

Thus choosing $\beta_1 \neq 0$ causes the consitency error to be of order $e_C \sim 1$ (at least) in general. When $h \sim t$, this means error amplification by factor $\sim t^{-1}$ from the uniformly optimal rate.

In the second counterexample we consider the case of a single-layer trapezoidal mesh, again as in section 4. So let Ω be subdivided into quadrilaterals K_i, $i = 1 \ldots I$, as defined by (4.15) with $x_i = ai$, and let the exact solution $\mathbf{U}$ be as above, so that $\sigma_2(\mathbf{U}) = 2\beta_2\gamma\theta_0\, y$. Then if $\mathbf{V} = (r, s) \in \mathcal{U}_h$, we find by direct computation that

$$\begin{aligned}
&(\sigma_2 - R_h\sigma_2, \frac{\partial s}{\partial y}) = \\
&= \tfrac{1}{12}\beta_2\gamma\theta_0 H^2 \sum_{i=0}^{I} [(p_{i-1} - p_{i+1})\,(s_i^+ - s_i^-) - (H/a)\,(p_{i-1}^2 - 2p_i^2 + p_{i+1}^2)\,(s_i^- + s_i^+)],
\end{aligned}$$

where $s_i^\pm = s(x_i \pm p_i H/2, \pm H/2)$ are the nodal values of s, and we set $p_i = 0$ for $i < 0$ or $i > I$. Choosing then $\mathbf{V}$ so that $r = 0$ and assuming that either $a \sim H$ or $a > H$, we have

$$\begin{aligned}
|||\mathbf{V}|||_h^2 &\leq C\left(\left\|\frac{\partial s}{\partial x}\right\|^2 + \left\|\frac{\partial s}{\partial y}\right\|^2\right) \\
&\leq CH^{-2}\|s\|^2 \leq C(a/H) \sum_{i=0}^{I} [\,(s_i^-)^2 + (s_i^+)^2\,],
\end{aligned}$$

hence choosing $s_i^{\pm} = 0$ for $i = 0, I$ and

$$s_i^+ - s_i^- = p_{i-1} - p_{i+1}, \quad s_i^- + s_i^+ = -(H/a)\,(p_{i-1}^2 - 2p_i^2 + p_{i+1}^2)$$

for $i = 1 \ldots I-1$, we get

$$\frac{(\mathcal{A} - \mathcal{A}_h)(\mathbf{U}, \mathbf{V})}{\|\mathbf{U}\| \|\mathbf{V}\|_h} = \frac{(\sigma_2 - R_h\sigma_2, \frac{\partial s}{\partial y})}{\|\mathbf{U}\| \|\mathbf{V}\|_h} \geq \delta,$$

where

$$\delta = C^{-1}\gamma\beta_2 H(aL)^{-1/2}\Big\{\sum_{i=1}^{I-1}[\,(p_{i-1} - p_{i+1})^2 \;+\; (H/a)^2(p_{i-1}^2 - 2p_i^2 + p_{i+1}^2)^2\,]\Big\}^{1/2}.$$

On an unrestricted mesh, we can have $\delta \sim 1$. For example, if $p_i = 0$ when i is even and $p_i = a/H$ when i is odd, this is the case. We have then again $e_C \geq \delta \sim 1$ by Principle C2.

The above example also indicates, how to construct sepcific meshes so that equilibrium locking can possibly be avoided: Choose either $p_i = p$ or $p_i = (-1)^i p$, for $i = 1 \ldots I-1$, or for all but a fixed number of indices i. Then most of the terms in the above sum disappear. It seems that on such meshes, or even on horizontally multi-layered meshes of this type, equilibrium locking is actually avoided. (But this would require some further analysis.)

C Extension of Theorem 10.1

We complete here the analysis of section 10 by an alternative construction of $\tilde{\mathbf{U}} \in \mathcal{U}_h$ in such a way that volumetric locking is avoided as well. We assume again Ansatz (10.4), with $\tilde{\rho}$ still defined by (10.5) but with $\tilde{\theta}, \tilde{w}, \tilde{\psi}$ now defined differently.

Assume that $1 \leq \bar{\lambda} < \infty$ (so that γ is bounded away from zero), let $\tilde{\theta}$ be defined in terms of $\tilde{\psi}$ by

$$\tilde{\theta}(x) = \theta(0) - \gamma^{-1} \int_0^x (\Pi_h \tilde{\psi})(x')\, dx', \tag{C.1}$$

and $\tilde{w}$ again in terms of $\tilde{\theta}$ by (10.18). Then $\Pi_h(\gamma\tilde{\theta}' + \tilde{\psi}) = 0$ by (C.1), so the whole term (10.14) vanishes.

By (C.1), (10.18) and (10.3), the kinematical constraints (10.7) hold if

$$\int_0^L (\psi - \tilde{\psi})\, dx = 0, \tag{C.2}$$

and in addition, constraint (10.20) is satisified. The latter constraint can also be expressed in terms of $\tilde{\psi}$. Namely, using (C.1) and (10.3), integrating by parts and using the L_2 -projection property of Π_h together with (C.2), one has

$$\begin{aligned}
\int_0^L (\theta - \tilde{\theta})\, dx &= \gamma^{-1} \int_0^L \int_0^x (\Pi_h \tilde{\psi} - \psi)(x')\, dx' dx \\
&= \gamma^{-1} L \int_0^L (\Pi_h \tilde{\psi} - \psi)\, dx - \gamma^{-1} \int_0^L x(\Pi_h \tilde{\psi} - \psi)\, dx \\
&= -\gamma^{-1} \int_0^L x(\Pi_h \tilde{\psi} - \psi)\, dx,
\end{aligned}$$

so constraint (10.20) is imposed by the condition

$$\int_0^L x\, \Pi_h \tilde{\psi}\, dx = \int_0^L x\psi\, dx. \tag{C.3}$$

Thus we can define $\tilde{\psi}$, e. g. by minimizing the left side of (10.11) under conditions (C.2) and (C.3) together with the kinematical constraints $\tilde{\psi}(x) = \psi(x)$ at $x = 0, L$. The solution to this constrained minimization problem exists if $I \geq 3$ and is then given by

$$\tilde{\psi} = \breve{\psi} + \breve{\eta},$$

where $\breve{\psi}$ is the standard interpolant of ψ, and $\breve{\eta}$ interpolates

$$\eta = x(L - x)(\eta_0 + \eta_1 x),$$

where η_0, η_1 are coefficients determined so that conditions (C.2), (C.3) hold.

With $\tilde{\theta}, \tilde{w}, \tilde{\psi}$ defined in the above manner, approximation error bounds similar to (10.9)–(10.12) hold with nearly optimal constants, except on anomalous coarse meshes where two mesh intervals $[x_{i-1}, x_i]$ nearly cover the whole interval $[0, L]$ (so that $h \approx \frac{1}{2}$ or $h > \frac{1}{2}$). Upon excluding such meshes with the condition

$$\min_{1 \le i < k \le I-1} [x_i(x_k - x_i)(L - x_k)]^{-1} \le CL^{-3},$$

estimates (10.13), (10.26) and (10.29) then remain valid, and since the error contribution from the second term in (5.5) now vanishes, the bound of Theorem 10.1 holds uniformly with respect to $\bar{\lambda}$.

References

[1] I. Babuŝka and A.K. Aziz, On the angle condition in the finite element method, SIAM J. Numer. Anal. 13 (1976) 214-226.

[2] T. Belytschko and W.E. Bachrach, Efficient implementation of quadrilaterals with high coarse-mesh accuracy, Comput. Methods Appl. Mech. Engr. 54 (1986) 279-301.

[3] T. Belytschko, W.K. Liu and B.E. Engelman, The gamma-elements and related developments, in: T.J.R. Hughes and E. Hinton, ed., Finite Element Methods for Plate and Shell Problems, Vol 2 (Pineridge Press, 1986) 316-347.

[4] S. Brenner and L.R. Scott, The Mathematical Theory of Finite Element Methods (Springer, 1994).

[5] P.G. Ciarlet, The Finite Element Method for Elliptic Problems (North-Holland, 1978).

[6] R.W. Clough, Original formulation of the finite element method, Finite Elements in Analysis and Design 7 (1990) 89-101.

[7] R. Courant, Variational methods for the solution of problems of equilibrium and vibration, Bull. Amer. Math. Soc. 49 (1943) 1-43.

[8] M. Fröier, L. Nilsson and A. Samuelsson, The rectangular plane stress element by Turner, Pian and Wilson, Int. J. Numer. Methods Engrg. 8 (1974) 433-437.

[9] R.H. Gallagher, Finite Element Analysis Fundamentals (Prentice-Hall, 1975).

[10] T.J.R. Hughes, The Finite Element Method: Linear Static and Dynamic Finite Element Analysis (Prentice-Hall, 1987).

[11] B.M. Irons and A. Razzaque, Experience with patch test for convergence of finite elements, in: A.K. Aziz, ed., The Mathematical Foundations of the Finite Element Method with Applications to Partial Differential Equations (Academic Press, 1972) 557-587.

[12] C. Johnson and J. Pitkäranta, Analysis of some mixed finite element methods related to reduced integration, Math. Comp. 38 (1982) 375-400.

[13] P. Lesaint, On the convergence of Wilson's nonconforming element for solving the elastic problem, Comput. Methods Appl. Mech. Engrg. 7 (1976) 1-16.

[14] P. Lesaint and M. Zlámal, Convergence of the nonconforming Wilson element for arbitrary quadrilateral meshes, Numer. Math. 36 (1980) 33-52.

[15] R.H. MacNeal, Finite Elements: Their Design and Performance (Marcel Dekker, 1994).

[16] R.H. MacNeal, A theorem regarding the locking of tapered four-node membrane elements, Int. J. Numer. Meth. Engrg. 24 (1987) 1793-1799.

[17] O. Ovaskainen and J. Pitkäranta, An energy method approach to the problem of elastic strip, SIAM J. Appl. Math. 58 (1998) 999-1021.

[18] T.H.H. Pian, Derivation of element stiffnes matrices by assumed stress distributions, AIAA J. 2 (1964) 1333-1336.

[19] J. Pitkäranta, The problem of membrane locking in finite element analysis of cylindrical shells, Numer. Math. 61 (1992) 523-542.

[20] J. Pitkäranta, Y. Leino, O. Ovaskainen and J. Piila, Shell deformation states and the finite element method: A benchmark study of cylindrical shells, Comput. Meth. Appl. Mech. Engrg. 133 (1996) 157-182.

[21] J. Pitkäranta and R. Stenberg, Error bounds for the approximation of the Stokes problem using bilinear/constant elements on irregular quadrilateral meshes, in: J.R. Whiteman, ed., The Mathematics of Finite Elements and Applications V (MAFELAP 1984) (Academic Press, 1985) 325-334.

[22] J. Pitkäranta and M. Suri, Design principles and error analysis for reduced-shear plate-bending finite elements, Numer. Math. 75 (1996) 223-266.

[23] J. Pitkäranta and M. Suri, Upper and lower error bounds for some (low-orer) plate-bending finite elements. Preprint (1998), Department of Mathematics and Statistics, University of Maryland Baltimore County.

[24] J. Robinson, The Life And Work of Early FEM Pioneers (Robinson and Associates, 1985).

[25] Z.C. Shi, A convergence condition for the quadrilateral Wilson element, Numer. Math. 44 (1984) 349-361.

[26] Z.C. Shi, B. Jiang and W. Xue, A new superconvergence property of Wilson nonconforming finite element, Numer. Math. 78 (1997) 259-268.

[27] G. Strang, Variational crimes in the finite element method, in A.K. Aziz, ed., The Mathematical Foundations of the Finite Element Method with Applications to Partial Differential Equations (Academic Press, 1972) 689-710.

[28] G. Strang and G.J. Fix, An Analysis of the Finite Element Method (Prentice-Hall, 1973).

[29] R.L. Taylor, P.J. Beresford and E.L. Wilson, A nonconforming element for stress analysis, Int. J. Numer. Meth. Engrg. 10 (1976) 1211-1219.

[30] M.J. Turner, R.W. Clough, H.C. Martin and L.J. Topp, Stiffness and deflection analysis of complex structures, J. Aeronaut. Sci. 23 (1956) 805-823.

[31] E.L. Wilson, R.L. Taylor, W. Doherty and J. Ghaboussi, Incompatible displacement models, in: S.J. Fenves et al. ,eds. , Numerical and Computer Methods in Structural Mechanics (Academic Press, 1973) 43-57.

(continued from the back cover)

A402 Saara Hyvönen
Growth of resolvents of certain infinite matrice, Nov 1998

A399 Otso Ovaskainen
Asymptotic and Adaptive Approaches to thin Body Problems in Elasticity

A398 Jukka Liukkonen
Uniqueness of Electromagnetic Inversion by Local Surface Measurements, Aug 1998

A397 Jukka Tuomela
On the Numerical Solution of Involutive Ordinary Differential Systems, 1998

A396 Clement Ph., Gripenberg G. and Londen S-O
Hölder Regularity for a Linear Fractional Evolution Equation, 1998

A395 Matti Lassas and Erkki Somersalo
Analysis of the PML Equations in General Convex Geometry, 1998

A393 Jukka Tuomela and Teijo Arponen
On the numerical solution of involutive ordinary differential equation systems, 1998

A392 Hermann Brunner, Arvet Pedas, Gennadi Vainikko
The Piecewise Polynomial Collocation Method for Nonlinear Weakly Singular Volterra Equations, 1997

A391 Kari Eloranta
The bounded Eight-Vertex Model, 1997

A390 Kari Eloranta
Diamond Ice, 1997

A389 Olavi Nevanlinna
Growth of Operator Valued Meromorphic Functions, 1997

A388 Jukka Tuomela
On the Resolution of Singularities of Ordinary Differential Systems, 1997

A387 Gennadi Vainikko
Fast Solvers of the Lippman-Schwinger Equation, 1997

A386 Ville Turunen and Gennadi Vainikko;
On Symbol Analysis of Periodic Pseudodifferential Operators, 1997

HELSINKI UNIVERSITY OF TECHNOLOGY INSTITUTE OF MATHEMATICS
RESEARCH REPORTS

The list of reports is continued inside. Electronical versions of the reports are available at *http://www.math.hut.fi/reports/* .

A412 Marko Huhtanen
Ideal GMRES can be bounded from below by three factors, Jan 1999

A411 Juhani Pitkranta
The first locking-free plane-elastic finite element: historia mathematica, Jan 1999

A410 Kari Eloranta
Bounded Triangular and Kagomé Ice, Jan 1999

A409 Jukka Tuomela
On the numerical solution of involutive ordinary differential systems: Boundary value problems, Dec 1998

A408 Ville Turunen
Commutator Characterization of Periodic Pseudodifferential Operators, Dec 1998

A403 Saara Hyvönen and Olavi Nevanlinna
Robust bounds for Krylov method, Nov 1998

ISBN 951-22-4409-8
ISSN 0784-3143
Edita, Espoo, 1999