

Mixed Methods on Quadrilaterals
and Hexahedra

Philip T. Keenan

April, 1992
(revised May, 1992)

TR92-13

Mixed Methods on Quadrilaterals and Hexahedra

Philip T. Keenan*

May 20, 1992

Abstract

We describe a new family of discrete spaces suitable for use with mixed methods on general quadrilateral and hexahedral elements. The new spaces are natural in the sense of differential geometry, so all the usual mixed method theory, including the hybrid formulation, carries over to these new elements with proofs unchanged. Because transforming general quadrilaterals into squares introduces nonlinearity and because mixed methods involve the divergence operator, the new spaces are more complicated than either the corresponding Raviart-Thomas spaces for rectangles or corresponding finite element spaces for quadrilaterals. These new elements may be useful in topologically regular grids, where initially rectangular grids are deformed to match features of the physical region.

*Department of Mathematical Sciences, Rice University

Contents

1	Introduction	2
2	Mixed Methods	4
2.1	A model problem	4
2.2	The compatibility conditions	5
3	Differential Geometry	5
3.1	The Metric Tensor	6
3.2	Changing Coordinate Systems	8
3.3	Integration and Differentiation	10
3.4	Submanifolds and Normals	13
4	General Quadrilaterals	14
5	The New Spaces	18
5.1	Why the standard space does not work	18
5.2	The new spaces	18
6	Three Dimensional Elements	23
7	Hybrid Methods	26
8	Results	27

1 Introduction

Finite element methods for partial differential equations have long used nonrectangular elements including triangles and tetrahedra, as well as general quadrilaterals and hexahedra. Such elements are useful in situations where the shape of the computational domain is non-rectangular, or in situations where highly variable coefficients in the equations vary in non-rectangular patterns. This occurs for instance in the simulation of groundwater flow, because the porosity and permeability of the earth tend to vary sharply across layers, called strata, which are of varying thickness and slant. However, finite element methods are poorly suited to certain porous media computations because they do not conserve mass element by element; thus one might like to use mixed methods on quadrilateral elements, as mixed methods do conserve mass locally as well as globally.

Finite element methods for second order elliptic partial differential equations are easily formulated on quadrilaterals because they only involve gradient computations, which are the same in any coordinate system. Mixed methods however

have generally been restricted to parallelograms, triangles, and other shapes which are affine transformations of reference elements. Because transforming quadrilaterals to squares is a non-affine map, and because mixed methods involve divergence computations, which are *not* the same in all coordinate systems, generalizing mixed methods to quadrilaterals is somewhat more complicated, as we shall see below. However, the increased complication may be worthwhile in certain applications, such as groundwater flow, where the resulting cell by cell conservation of mass is important.

We describe a new family of discrete spaces suitable for use with mixed methods on general quadrilateral and hexahedral elements. We follow [10], constructing a mixed method by defining properly compatible discrete subspaces of L^2 and $H(\text{div})$. Such subspaces are commonly constructed for general collections of elements by mapping each element back to a standard reference element on which the spaces are defined. The mappings used to handle general triangular, rectangular and parallelogram elements are all affine, which allows one to handle transformations of vectors via the Piola transformation[3]. When one turns to general quadrilaterals, however, bilinear mappings arise. The resulting square reference element must be viewed as a manifold with an induced metric which is different from the usual Euclidean metric. This has a number of consequences. The unit outward normal can vary along edges. Distinction must be made between covariant and contravariant vector components. Most significantly the divergence of a vector field is no longer equal to the trace of its Jacobian matrix. As a result the usual Raviart-Thomas spaces no longer satisfy the compatibility conditions. We therefore define new spaces which are compatible. We deliberately avoid degeneracy in the new spaces to prevent conditioning difficulties with nearly rectangular elements, and we build in good approximation properties as well. This then allow quadrilaterals to be used like any other type of element, although the unusual metric must always be kept in mind when doing calculations.

We will also describe the extension of these ideas to three dimensions, which is straightforward.

We note that there are other approaches one could take in generalizing mixed methods to quadrilaterals. One approach is to attempt to define a method directly on the quadrilaterals, rather than through reference elements. Another is to try to construct “nonconforming” mixed methods which do not require the discrete velocity space to be a subspace of $H(\text{div})$. Work is in progress on some of these alternatives[1], but we will not consider them further in this paper.

The remainder of the paper is organized as follows. In section 2 we give a quick review of the formulation of mixed methods and the compatibility conditions required of the discrete spaces used in them. In section 3 we describe enough background material from differential geometry to enable us to apply this tool to general quadrilaterals in section 4. In section 5 we introduce the new family of discrete

spaces for two dimensional quadrilaterals. In section 6 we extend the results to three dimensional hexahedral elements. In section 7 we describe a hybrid formulation using these elements. Finally in section 8 we summarize our conclusions.

2 Mixed Methods

2.1 A model problem

We begin with a quick review of mixed methods via a model elliptic problem.

Consider solving the equation

$$-\Delta p = f \text{ in } \Omega,$$

with

$$p = 0 \text{ on } \partial\Omega,$$

on some domain $\Omega \in \mathcal{R}^n$. We set

$$u = -\nabla p,$$

whence

$$\nabla \cdot u = f.$$

We introduce the space

$$H(\operatorname{div}; \Omega) = \{v \in (L^2(\Omega))^n : \operatorname{div} v \in L^2(\Omega)\},$$

with norm

$$\|v\|_{H(\operatorname{div}; \Omega)} = \|v\|_{L^2(\Omega)} + \|\operatorname{div} v\|_{L^2(\Omega)}.$$

By the divergence theorem, the original problem is equivalent to solving

$$(u, v) = (p, \operatorname{div} v) \text{ for all } v \in H(\operatorname{div}; \Omega), \quad (1)$$

and

$$(\operatorname{div} u, w) = (f, w) \text{ for all } w \in L^2(\Omega). \quad (2)$$

Here

$$(a, b) = \int_{\Omega} ab \, dx$$

is the usual L^2 inner product.

2.2 The compatibility conditions

To define a mixed method we simply choose appropriate subspaces $V_h \subset H(\text{div}; \Omega)$ and $W_h \subset L^2(\Omega)$ and solve the above equations for $u_h \in V_h$ and $p_h \in W_h$, restricting the test functions to be in V_h and W_h respectively. That is, we seek $U \in V_h$ and $P \in W_h$ satisfying

$$(U, V) = (P, \text{div } V) \text{ for all } V \in V_h, \quad (3)$$

and

$$(\text{div } U, W) = (f, W) \text{ for all } W \in W_h. \quad (4)$$

This gives a finite dimensional linear system which can be solved for the unique approximate solution.

As described by Raviart and Thomas in [10], if the spaces V_h and W_h satisfy the so-called compatibility conditions that

$$v \in V_h \text{ and } (\text{div } v, w) = 0 \text{ for all } w \in W_h \text{ implies } \text{div } v = 0, \quad (5)$$

and that there exist an $\alpha > 0$ independent of h such that for all $w \in W_h$,

$$\sup_{v \in V_h} \frac{(\text{div } v, w)}{\|v\|_{H(\text{div}; \Omega)}} \geq \alpha \|w\|_{L^2(\Omega)}, \quad (6)$$

then the approximate method converges to the true solution and the error goes to zero like approximation error. That is, one has the following theorem.

Theorem 1 (Raviart-Thomas [10] based on Brezzi[2]) *If (5) and (6) hold, then there is a constant C independent of p , u and h such that (3) and (4) have a solution, which is unique and satisfies*

$$\begin{aligned} & \|u - U\|_{H(\text{div}; \Omega)} + \|p - P\|_{L^2(\Omega)} \leq \\ & C \left(\inf_{V \in V_h} \|u - V\|_{H(\text{div}; \Omega)} + \inf_{W \in W_h} \|p - W\|_{L^2(\Omega)} \right). \end{aligned} \quad (7)$$

3 Differential Geometry

Differential geometry may be thought of as a tool for making the integral theorems of vector calculus, such as the divergence theorem and the various Green's identities, hold in general coordinate systems. It explains how to calculate gradients, divergences and normals in any coordinate system. In many coordinate systems these rules reduce to the ordinary ones, but in non-affine coordinates they do not. Anyone who tries to apply the divergence theorem in non-affine coordinates using ordinary Euclidean vector calculus will find that it fails to hold. As mixed methods rely on

the divergence theorem for their very formulation, some knowledge of differential geometry is essential in generalizing them to quadrilateral elements.

In this section we define the notation we will use from differential geometry. Readers not familiar with differential geometry are referred to [9] and [7] for a comprehensive introduction and to such books as [6] for a quick summary of tensor analysis in general spaces. Books on orthogonal coordinate systems and cartesian tensors will *not* suffice. Readers familiar with the subject may skip this section, as I follow the standard notation closely.

We explain only as much differential geometry as is needed for our application. In particular we do not define smooth manifolds in their full generality but restrict attention to the following situation. Let $\Omega \subset \mathcal{R}^n$ be a simply connected open set with piecewise C^1 boundary. Let $f : \Omega \rightarrow \mathcal{R}^n$ be a C^1 injection. Then f defines a coordinate system on Ω in which the coordinates of any point $p \in \Omega$ are given by $f(p)$. Such an f is also called a *chart*. For instance when f is the identity map we have the usual Euclidean coordinate system. If $f' : \Omega \rightarrow \mathcal{R}^n$ is also a C^1 injection then f' defines a second coordinate system on Ω . We will write $x = (x^1, \dots, x^n) = f(p)$ and $x' = (x'^1, \dots, x'^n) = f'(p)$ to distinguish them.

The present section will use the techniques of differential geometry to study the effect of changing between general coordinate systems. The next section will specialize to the case of a quadrilateral transformed to look like the unit square.

3.1 The Metric Tensor

At each point p of our domain Ω is an n -dimensional space of tangent vectors, denoted $T_p(\Omega)$. If $c : [-1, 1] \rightarrow \Omega$ is a smooth curve, and $c(0) = p$, then $c'(0)$ denotes a specific element of $T_p(\Omega)$. Given a set of n basis vectors g_i for $T_p(\Omega)$ we may express any tangent vector $v \in T_p(\Omega)$ as a sum

$$v = v^i g_i,$$

for some numbers v^i . Here and henceforth we adopt the *summation convention* that repeated upper and lower indices denote summation from 1 to n , and that unrepeated indices imply that the equation holds for any choice of that index. We do not repeat two upper or two lower indices.

The coordinate system defines a natural choice of the basis vectors g_i . In particular g_i is the tangent vector to the curve $c_i(t) = f^{-1}(x^1, \dots, x^i + t, \dots, x^n)$ at $t = 0$.

We suppose that a symmetric inner product has been defined for the real vector space $T_p(\Omega)$. In particular, when f is the identity map we may use the usual Euclidean inner product. The choice of inner product is separate from the choice of coordinate system; together the two completely determine the intrinsic geometry of

an oriented Riemannian manifold. We define

$$g_{ij} = g_i \cdot g_j,$$

whence

$$g_{ij} = g_{ji}.$$

The numbers g_{ij} are called the components of the *metric tensor*. Knowing them is equivalent to knowing the inner product, since to compute the inner product of any two vectors $u = u^i g_i$ and $v = v^j g_j$ we simply use

$$u \cdot v = u^i v^j g_{ij}.$$

The set of linear functionals which map tangent vectors into real numbers is also an n -dimensional vector space at each point in the manifold. A natural basis for it are the functionals G^i defined by their action on the set of basis vectors according to

$$G^i(g_j) = \delta_j^i.$$

Here we use the Kronecker delta symbol δ_j^i , which is defined to be one when $i = j$ and zero otherwise.

By the Riesz representation theorem linear functionals correspond one to one with vectors, via the inner product operation. That is, to each G^i there corresponds a vector g^i such that

$$G^i(v) = g^i \cdot v,$$

for all tangent vectors v . Henceforth we identify linear functionals with tangent vectors. We thus have another “natural” choice for the set of basis vectors for the tangent space at each point, namely the g^i .

Let us expand the g^i in terms of the basis g_j . Suppose $g^i = g^{ij} g_j$ for some numbers g^{ij} . Forming the inner product with g_k yields

$$g^{ij} g_{jk} = g^i \cdot g_k = \delta_k^i.$$

Thus the g^{ij} are the components of the inverse of the metric tensor, thought of as a matrix. Similarly we deduce

$$g_i = g_{ij} g^j.$$

In addition,

$$g^{ij} = g^i \cdot g^j.$$

If v is any tangent vector we may express it in terms of either basis. We write

$$v = v^i g_i = v_i g^i.$$

The v^i are called the contravariant components of v , and the v_i are called the covariant components. Inner products with g^j or g_j show that the two types of components are related by

$$v^i = g^{ij} v_j,$$

and

$$v_i = g_{ij} v^j.$$

Thus knowing the metric tensor enables us to raise and lower indices as needed.

3.2 Changing Coordinate Systems

A different coordinate system f' will cause us to consider a different set of basis vectors $g_{i'}$. In this context we write $x = x(x') = f \circ f'^{-1}(x')$ and $x' = x'(x) = f' \circ f^{-1}(x)$ for short. Applying the chain rule for partial derivatives shows the $g_{i'}$ are related to the g_i by

$$g_{i'} = \beta_{i'}^i g_i,$$

where

$$\beta_{i'}^i = \frac{\partial x^i}{\partial x^{i'}}.$$

We also define

$$\beta_i^{i'} = \frac{\partial x^{i'}}{\partial x^i}.$$

Note that the chain rule for partial derivatives implies

$$\beta_i^{i'} \beta_{i'}^j = \delta_i^j,$$

and

$$\beta_{i'}^i \beta_i^{j'} = \delta_{i'}^{j'}.$$

Also,

$$g^{i'} = \beta_i^{i'} g^i.$$

In terms of the new basis a vector v can be written as $v = v^{i'} g_{i'}$. The primed components $v^{i'}$ are related to the unprimed components v^i through the equation $v^{i'} g_{i'} = v^i g_i$. Substitution of the transformation rule for $g_{i'}$ shows that

$$v^{i'} = \beta_i^{i'} v^i,$$

and

$$v_{i'} = \beta_{i'}^i v_i.$$

These equations show how the contravariant and covariant components of vectors transform. That is, a single entity, namely a tangent vector in $T_p(\Omega)$, can be written

in terms of 4 different sets of basis vectors, and the resulting components are related by the above equations, independent of the particular vector.

A tensor can be thought of as a multi-linear functional on tangent vectors. For instance a rank 2 tensor A maps pairs of tangent vectors into real numbers. It is defined by its action on basis vectors, so if

$$A(g_i, g_j) = A_{ij},$$

then for any $u = u^i g_i$ and $v = v^j g_j$,

$$A(u, v) = u^i v^j A_{ij}.$$

Notice that if we write

$$A = A_{ij} g^i g^j,$$

then

$$Auv = A_{ij} g^i g^j u^k v^l g_{kl} = u^i v^j A_{ij} = A(u, v).$$

Hence we may think of the A_{ij} as being the covariant components of the tensor A in this coordinate system. We find that in the primed coordinate system,

$$A_{i'j'} = \beta_{i'}^i \beta_{j'}^j A_{ij}.$$

Any entity with n^2 components in each coordinate system which transform according to this equation is a rank 2 tensor, and the components are its covariant components. We can define

$$A^i_{\cdot j} = g^{ik} A_{kj},$$

$$A_j^{\cdot i} = g^{ik} A_{jk},$$

and

$$A^{ij} = g^{ik} g^{jl} A_{kl}.$$

The dot may be omitted when the tensor is symmetric. One can verify that these transform according to

$$A^{i'}_{\cdot j'} = \beta_{i'}^i \beta_{j'}^j A^i_{\cdot j},$$

$$A_{j'}^{\cdot i'} = \beta_{i'}^i \beta_{j'}^j A_j^{\cdot i},$$

and

$$A^{i'j'} = \beta_{i'}^i \beta_{j'}^j A^{ij}.$$

Tensors of rank 3 and higher are defined analogously.

For instance the metric tensor in the primed coordinate system is defined by

$$g_{i'} \cdot g_{j'} = g_{i'j'},$$

which implies

$$g_{i'j'} = \beta_{i'}^i \beta_{j'}^j g_{ij},$$

which is exactly the transformation rule for the doubly covariant components of any two index tensor, so we are justified in calling it a tensor.

Similarly

$$g^{i'j'} = \beta_i^{i'} \beta_j^{j'} g^{ij}.$$

Note that the general elliptic operator $\text{div}(A \cdot \text{grad } p)$ may involve a tensor A , which is often taken to be diagonal in Euclidean coordinates. In the transformed coordinate system it may be that A will no longer be diagonal. However the identity tensor $A_j^i = \delta_j^i$ is unchanged in any coordinate system: $A_{j'}^{i'} = \beta_{j'}^j \beta_j^i \delta_j^i = \beta_{j'}^j \beta_j^i = \delta_{j'}^{i'}$.

In ordinary Euclidean coordinates the two basis sets g_i and g^i are in fact the same. In orthogonal coordinate systems, which are the subject of books on Cartesian tensors, the normalized versions of the g_i and g^i are also the same. Hence in both cases there is no distinction between covariant and contravariant components. In general curvilinear coordinate systems, however, the bases and the components do differ.

Because we have not made any assumptions about the two coordinate systems, all the above equations hold if all primed indices are unprimed and all unprimed indices are primed. This is one of the virtues of the tensor approach — one can write equations which are the same in any coordinate system, and hence are more likely to describe the essence of the situation.

3.3 Integration and Differentiation

We now turn to integration and differentiation of scalar and vector fields. Let $g = \det(g_{ij})$. For any scalar function ϕ , we define

$$\int_{\Omega} \phi = \int_{f(\Omega)} \phi \circ f^{-1}(x) \sqrt{g \circ f^{-1}(x)} dx. \quad (8)$$

Because of the \sqrt{g} factor, this definition is independent of the coordinate system f chosen in which to evaluate it. This follows from the usual change of variables theorem for integrals in multi-variable calculus.

We next denote partial derivatives by commas in index notation. That is, for any scalar field ϕ , we define

$$\phi_{,i} = \frac{\partial \phi}{\partial x^i}.$$

By the chain rule for partial derivatives,

$$\phi_{,i'} = \beta_{i'}^i \phi_{,i}.$$

Hence the scalars $\phi_{,i}$ are the covariant components of a vector field. We name that vector field the *gradient* of ϕ and write

$$\text{grad } \phi = \phi_{,i} g^i.$$

When differentiating vectors, however, the situation is more complex. If v is a vector field, neither the partial derivatives of the covariant components nor the partial derivatives of the contravariant components transform like tensor components. If the divergence of a vector field is to be meaningful, it must not depend on the coordinate system chosen for the computation. Therefore we consider the entire entity, in this case a vector, and differentiate it:

$$v_{,i} = (v^j g_j)_{,i} = v^j_{,i} g_j + v^j g_{j,i}.$$

We see that we need to know the derivatives of the basis vectors. These are given *formally* by

$$g_{i,j} = \Gamma_{ij}^k g_k,$$

where the Γ_{ij}^k are called the Christoffel symbols. For convenience we also define

$$\Gamma_{ijk} = g_{kl} \Gamma_{ij}^l.$$

The various Γ symbols are *not* tensors¹, but they do satisfy a number of useful properties. We will not derive them all here but will mention the most important ones. They are symmetric,

$$\Gamma_{ij}^k = \Gamma_{ji}^k,$$

they satisfy

$$g^i_{,j} = -\Gamma_{jk}^i g^k,$$

and most importantly they can be computed from the equation

$$\Gamma_{ijk} = \frac{1}{2}(g_{ik,j} + g_{kj,i} - g_{ji,k}).$$

We want to write the i 'th partial derivative of v as a vector:

$$v_{,i} = v^j_{,i} g_j,$$

where $v^j_{,i}$ denotes the j 'th contravariant component of the i 'th partial derivative of the vector v . Given the Christoffel symbols, we proceed to calculate

$$v_{,i} = v^j_{,i} g_j + v^j \Gamma_{ij}^k g_k.$$

¹For if they were tensors, the transformation law implies they would all be zero in every coordinate system, as they vanish in Euclidean coordinates.

Rewriting this as

$$v_{,i} = (v_{,i}^j + v^k \Gamma_{ik}^j) g_j,$$

shows that

$$v_{;i}^j = v_{,i}^j + v^k \Gamma_{ik}^j.$$

Similarly one can derive an expression for the j 'th covariant component of the i 'th partial derivative of a vector field:

$$v_{j;i} = v_{j,i} - v_k \Gamma_{ij}^k.$$

Just as the notation $,i$ denotes an ordinary partial derivative, we use $;i$ to denote this total derivative, which is also called a covariant derivative. For consistency we set

$$\phi_{;i} = \phi_{,i},$$

for scalar fields ϕ .

We can finally write the divergence of a vector field in coordinate invariant form. It is the scalar field corresponding to the trace of the covariant derivative matrix:

$$\operatorname{div} v = v_{;i}^i.$$

Applying the above definition yields

$$\operatorname{div} v = v_{,i}^i + v^k \Gamma_{ik}^i. \quad (9)$$

Alternatively, recalling $v^i = g^{ij} v_j$ we can write this in terms of the covariant components of v as

$$\operatorname{div} v = g^{ij} (v_{j,i} - v_k \Gamma_{ij}^k).$$

It turns out that in ordinary Euclidean space and in many other manifolds as well, all the Γ_{ij}^k are zero, whence we recover the ordinary definition of the divergence as the trace of the matrix of partial derivatives. Even in the case of a parallelogram mapped back to a square, the Γ_{ij}^k all vanish. But for general quadrilaterals they do *not* vanish, and the above general definition of the divergence must be used. That is to say, one wants the integral of the divergence times a test scalar field to be the same over the quadrilateral and over the reference square. Unless one uses this definition for *div*, the integrals will not in general be equal.

For completeness we state the definition of the Laplacian operator in coordinate invariant form. Let $v = \operatorname{grad} \phi$. Then $\operatorname{lap} \phi = \operatorname{div} v$, whence a short calculation shows

$$\operatorname{lap} \phi = g^{ij} \phi_{;ij} = g^{ij} (\phi_{,ij} - \phi_{,k} \Gamma_{ij}^k).$$

3.4 Submanifolds and Normals

The divergence theorem requires a definition for the unit outward normal vector along the boundary of a manifold. For a smooth orientable manifold such as Ω , with a piecewise C^1 boundary, the inclusion map $i : \partial\Omega \rightarrow \Omega$ induces an injection $T(i) : T(\partial\Omega) \rightarrow T(\Omega)$ which allows us to identify $T_p(\partial\Omega)$ with a subspace of $T_p(\Omega)$, for each $p \in \partial\Omega$. The unit outward normal $\hat{n}(p)$ at p is defined to be a unit vector in $T_p(\Omega)$ which is perpendicular to all vectors in the subspace $T_p(\partial\Omega)$. Because the manifold was assumed to be orientable, there is a consistent continuous choice of sign over all of $\partial\Omega$ which we can call outward.

To calculate the contravariant components of the normal we must first parameterize the boundary. Let h_α be a chart on $\partial\Omega$ near p , and h_a a chart on Ω near p . In this section we will use Greek indices on $\partial\Omega$ and Latin indices on Ω . Given the natural basis vectors g_i on $T_p(\Omega)$ and g_μ on $T_p(\partial\Omega)$ induced by these charts, the injection $T(i)$ is defined in coordinates by

$$v^j = \beta_\mu^j v^\mu,$$

where $j = 1, \dots, m$ and $\mu = 1, \dots, m-1$, and

$$\beta_\mu^j = \frac{\partial(h_a \circ i \circ h_\alpha^{-1})^j}{\partial x^\mu}.$$

Now, let b_μ be the image of g_μ under $T(i)$. Note that b_μ is generally *not* the same as a g_i , as the coordinate systems on Ω and $\partial\Omega$ need not be related. In fact, writing

$$b_\mu = b_\mu^i g_i,$$

and noting that

$$g_\mu = \delta_\mu^\nu g_\nu,$$

we find that

$$b_\mu^i = \beta_\nu^i \delta_\mu^\nu = \beta_\mu^i,$$

are the contravariant components of b_μ .

The unit normal $\hat{n} = n^i g_i$. Its contravariant components satisfy

$$n^i n^j g_{ij} = 1,$$

and

$$n^i b_\mu^j g_{ij} = 0,$$

which say that \hat{n} is a unit vector which is perpendicular to the boundary. One can check that the solution to these equations defines a set of components which transform like contravariant vector components, and hence \hat{n} is well defined independent

of the choice of coordinates. Its representation in any particular coordinate system, however, does depend on the coordinate system.

In coordinate free form, we can now state the divergence theorem:

$$\int_{\Omega} \operatorname{div} v = \int_{\partial\Omega} v \cdot \hat{n},$$

which is true in any coordinate system since it is defined in coordinate independent form and is true, in particular, in Euclidean coordinates. Similarly we can rewrite it as Green's theorem:

$$\int_{\Omega} \phi_{;i} v^i = - \int_{\Omega} \phi v^i_{;i} + \int_{\partial\Omega} \phi v^i n^j g_{ij}.$$

To evaluate integrals over the boundary of the region, we need the Jacobian factor, which here is $\sqrt{g} = \sqrt{\det g_{\mu\nu}}$. Note that we can evaluate

$$g_{\mu\nu} = g_{\mu} \cdot g_{\nu} = b_{\mu} \cdot b_{\nu} = b_{\mu}^i b_{\nu}^j g_{ij}.$$

4 General Quadrilaterals

We now specialize the above theory to the case of a general quadrilateral. The computations are tedious but can be automated with a symbolic mathematics program — I used Maple V, available from Waterloo Maple Software[4]. Here we simply state the results.

In our application Ω will be a general quadrilateral in R^2 . We choose the initial coordinate system f to be the identity map composed if necessary with a translation and rotation so that one vertex of the quadrilateral is at the origin and another is at some point $(a, 0)$ on the positive x-axis. This does not alter any metric properties as the Euclidean metric is invariant under rotation and translation. There are 4 degrees of freedom remaining. We will think of a as the x-scale, and so let the vertex opposite $(a, 0)$ be (ab, c) , where c will be the y-scale. We use ab rather than b to express the new x-coordinate, because this makes b dimensionless and turns out to simplify the following formulas. If we had a parallelogram, the final vertex would be at $(a + ab, c)$, so we express the actual final vertex as a scaled shift away from this point. That is, we give it coordinates $(a + ab + ad, c + ce)$. We have now used up all the degrees of freedom. Figure 1 summarizes this labeling and also shows the reference element, which will be the unit square.

In the unprimed coordinate system $g_{ij} = \delta_{ij} = g^{ij}$, and all the $\Gamma^k_{ij} = 0$. Also $\sqrt{g} = 1$, so scalars, vectors, integrals and derivatives all obey the usual rules for calculus in R^2 .

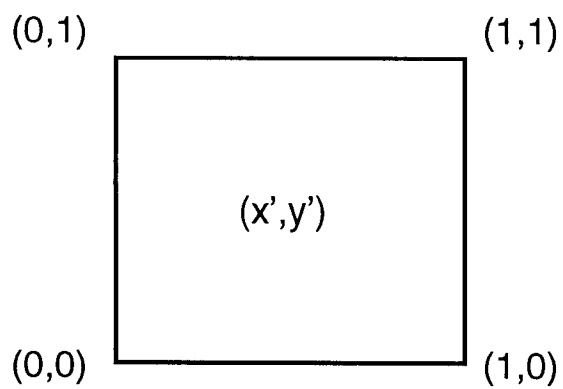
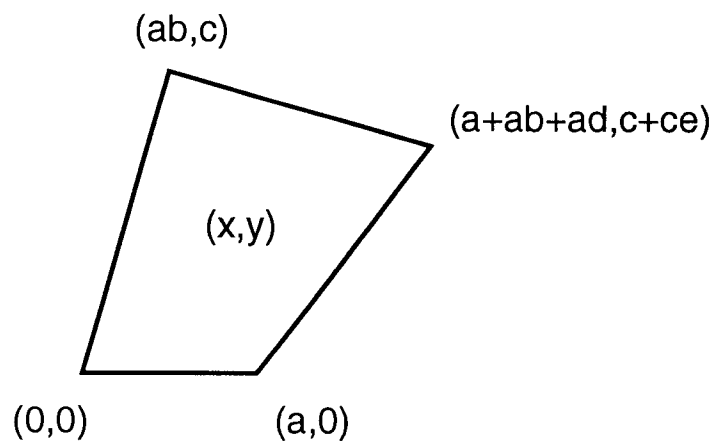


Figure 1: A general quadrilateral and its reference element

The primed coordinate system is designed to make Ω look like the unit square. In particular we take f' to be defined by $(x, y) = f'^{-1}(x', y')$, where

$$x = ax' + aby' + adx'y',$$

and

$$y = cy' + cex'y'.$$

Note that if the top and bottom edges are parallel, $e = 0$ and $y = cy'$. If the left and right edges are also parallel, then $d = 0$, and $x = ax' + aby'$, resulting in an affine transformation for parallelograms. For general quadrilaterals, however, the coefficients of the $x'y'$ cross terms do not in general vanish.

If we cut the quadrilateral into two triangles and map each into half the unit square by an affine transformation, we avoid the nonlinear terms, but the resulting f is only piecewise smooth. This forces us to consider each of the two triangles separately and form a macro element from the pair. In three dimensions we need 5 tetrahedra to partition a cube. We will see below that while the nonlinear terms do complicate matters, the resulting spaces are more efficient in terms of number of unknowns than the tetrahedral macro element. In the rest of this section, therefore, we restrict attention to the smooth f introduced above.

The metric tensor is given by

$$\begin{aligned} g_{1'1'} &= a^2 + 2y'a^2d + (a^2d^2 + c^2e^2)y'^2, \\ g_{1'2'} &= a^2b + x'a^2d + (a^2db + c^2e)y' + (a^2d^2 + c^2e^2)x'y', \\ g_{2'1'} &= g_{1'2'}, \end{aligned}$$

and

$$g_{2'2'} = a^2b^2 + c^2 + (2c^2e + 2a^2db)x' + (a^2d^2 + c^2e^2)x'^2.$$

We now define $\bar{e} = d - eb$.

To integrate scalars in the primed coordinate system we use

$$\sqrt{\det g_{i'j'}} = acV,$$

where the normalized volume element V is the linear function

$$V(x', y') = 1 + ex' + \bar{e}y'.$$

To compute derivatives we need to know that all the $\Gamma_{i'j'}^{k'} = 0$ *except* $\Gamma_{1'2'}^{k'} = \Gamma_{2'1'}^{k'}$, which are given by

$$\Gamma_{1'2'}^{1'} = \bar{e}/V,$$

and

$$\Gamma_{1'2'}^{2'} = e/V.$$

We note in passing that the area of the quadrilateral is

$$\text{Area} = ac(1 + \frac{e}{2} + \frac{\bar{e}}{2}) = acV(\frac{1}{2}, \frac{1}{2}),$$

as may be checked by integrating 1 over the unit square in primed coordinates.

Applying (9), the divergence of a vector $v = v^i g_i$ is here given by

$$\text{div } v = v_{,1'}^{1'} + v_{,2'}^{2'} + v^{1'} \Gamma_{2'1'}^{2'} + v^{2'} \Gamma_{1'2'}^{1'},$$

which reduces to

$$\text{div } v = v_{,1'}^{1'} + v_{,2'}^{2'} + (ev^{1'} + \bar{e}v^{2'})/V. \quad (10)$$

The unit outward normal along any edge (x_0, y_0) to (x_1, y_1) can be computed by parameterizing this line segment as $[0, 1]$, for instance. Along the edge from $(0, 0)$ to $(a, 0)$, the unit normal at $(x', 0)$ turns out to be

$$\pm \frac{(b + dx', -1)}{c(1 + ex')},$$

in contravariant components. We choose the plus sign to make it outward. It then turns out that along this edge,

$$v \cdot \hat{n} = -v^{2'} c(1 + ex').$$

While we will not write out the normal for the other edges, one may verify that each is a rational function of position. In addition one may verify that there are four *linear* functions f_{x0} , f_{x1} , f_{y0} , and f_{y1} , which depend only on the geometry, such that

$$v \cdot \hat{n} = \begin{cases} (v^{1'} f_{x0})(y') & \text{along } x' = 0, \\ (v^{1'} f_{x1})(y') & \text{along } x' = 1, \\ (v^{2'} f_{y0})(x') & \text{along } y' = 0, \\ (v^{2'} f_{y1})(x') & \text{along } y' = 1. \end{cases}$$

We have

$$\begin{aligned} f_{x0} &= -\frac{ac(1 + y'\bar{e})}{\sqrt{a^2b^2 + c^2}}, \\ f_{x1} &= \frac{ac(1 + y'\bar{e} + e)}{\sqrt{a^2b^2 + c^2 + 2c^2e + 2a^2bd + a^2d^2 + c^2e^2}}, \\ f_{y0} &= -c(x'e + 1), \\ f_{y1} &= \frac{ac(x'e + 1 + \bar{e})}{\sqrt{a^2 + 2a^2d + a^2d^2 + c^2e^2}}. \end{aligned}$$

The Jacobian along each edge is just a constant factor. For example, along the edge $y' = 0$, using x' as the coordinate for the submanifold, we have $\sqrt{\bar{g}} = a$.

5 The New Spaces

5.1 Why the standard space does not work

The careful reader will now see several reasons why the usual Raviart-Thomas spaces on reference squares fail to satisfy the compatibility conditions when the original elements are general quadrilaterals. Consider for instance the lowest order Raviart-Thomas space on rectangles, in which W_h is the space of discontinuous piecewise constants and V_h the space of vectors whose x component is continuous piecewise linear in x and discontinuous piecewise constant in y , and whose y component is continuous piecewise linear in y and discontinuous piecewise constant in x . Here *piecewise* means element by element, when the domain of the PDE has been subdivided into more than one rectangular element.

Ordinarily the first compatibility condition (5) is trivially satisfied, as the divergence of any vector field in V_h is piecewise constant and hence is in W_h . Similarly (6) is satisfied because any $w \in W_h$ can be written as the divergence of some $v \in V_h$. In addition V_h is clearly a subspace of $H(\text{div})$.

For quadrilaterals mapped back to squares, however, several things go wrong. If two adjacent quadrilaterals have different shapes, the outward normals across the shared edge are different linear functions, and so V_h is no longer a subspace of $H(\text{div})$. Even when all the quadrilaterals are the same shape, the divergence of a vector in V_h is *not* in general piecewise constant. To see this, take $g_{i'}$ as the basis vectors for the velocity space and consider $v = v^{i'} g_{i'}$ with $v^{1'}(x', y') = c_1 x' + c_2$ and $v^{2'}(x', y') = c_3 y' + c_4$. By (10),

$$\text{div } v = c_1 + c_3 + (ec_1 x' + \bar{e}c_3 y' + (ec_2 + \bar{e}c_4))/V. \quad (11)$$

This is clearly not constant. In fact, the quantity $V \text{div } v$ is a completely general piecewise linear function. Investigating (5) cell by cell using (8) we find that on each element,

$$(\text{div } v, w) = \int_0^1 \int_0^1 acw(c_1 + c_3 + ec_2 + \bar{e}c_4 + e(2c_1 + c_3)x' + \bar{e}(c_1 + 2c_3)y') dx dy.$$

That is, we are integrating a linear function with three degrees of freedom against a constant test function, which is not enough to match it — the integral vanishing merely implies the average value of the linear function is zero. Hence the compatibility condition is no longer satisfied.

5.2 The new spaces

There are many possible ways to remedy the situation. Each way involves tradeoffs between power and efficiency. One approach is to abandon the compatibility condi-

tions and invent a new theory of non-conforming mixed methods in which the normal fluxes of a velocity vector field can be discontinuous across element boundaries; see [1]. Alternatively, we can use differential geometry to construct new compatible spaces which work within the old theory. We take this latter approach here, gaining simplicity from differential geometry at the cost of a restriction on the types of quadrilateral elements we can connect together.

Returning to the example of the lowest order Raviart-Thomas spaces, we fix compatibility by letting the pressure space consist of discontinuous piecewise linears rather than discontinuous piecewise constants. In order to keep the method conforming, we restrict attention to topologically regular grids where an initially rectangular grid is deformed globally by a single bilinear transformation – this insures that the normal fluxes all match. Figure 2 shows such a grid. We would prefer to be able to use irregular grids like the one in Figure 3; however, this introduces nonconformity. In [1] we will show that this restriction can actually be lifted, by developing a theory of non-conforming mixed methods; at present however it appears that one loses a power of h in accuracy if one allows non-conforming elements.

Another restriction on the quadrilaterals arises from the need to avoid degeneracy. Since the divergence on parallelograms is just a constant (because $e = \bar{e} = 0$), we only need discontinuous piecewise constants for pressure in that case. Therefore, on elements which are *almost* parallelograms, the resulting system of three equations will be ill-conditioned. This follows from (11), which we re-write as

$$V \operatorname{div} v = c_1 + c_3 + e c_2 + \bar{e} c_4 + e(2c_1 + c_3)x' + \bar{e}(c_1 + 2c_3)y'.$$

The condition number of this system with respect to a basis of linear polynomials depends inversely on e and \bar{e} . Thus we restrict to quadrilaterals which are far from being parallelograms, in the sense that they satisfy

$$\min\{|e|, |\bar{e}|\} > c_0 > 0, \tag{12}$$

for some fixed constant c_0 .

We also need the Jacobian V to be non-singular. Write the Euclidean coordinates of the fourth point of the quadrilateral as

$$\alpha(ab, c) + \beta(a, 0).$$

We see that $\alpha = 1 + e$ and $\beta = 1 + \bar{e}$, yielding an explanation for the prevalence of the symbols e and \bar{e} in the above formulas. Non-degeneracy is equivalent to requiring $\alpha > 0$, $\beta > 0$, and $\alpha + \beta > 1$. To prevent ill-conditioning, we therefore require that

$$1 + \min\{e, \bar{e}, e + \bar{e}\} > c_1 > 0, \tag{13}$$

for some fixed constant c_1 . Figure 4 illustrates the situation.

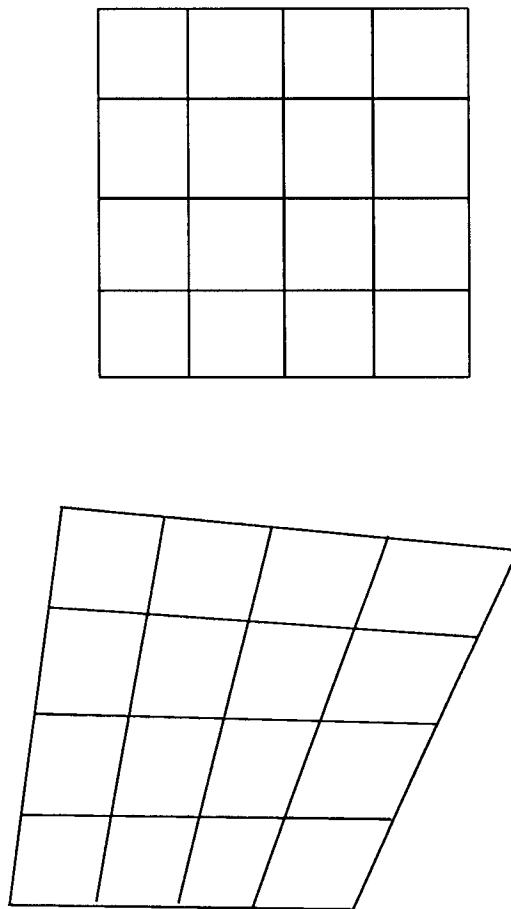


Figure 2: A rectangular mesh and its image under a bilinear map

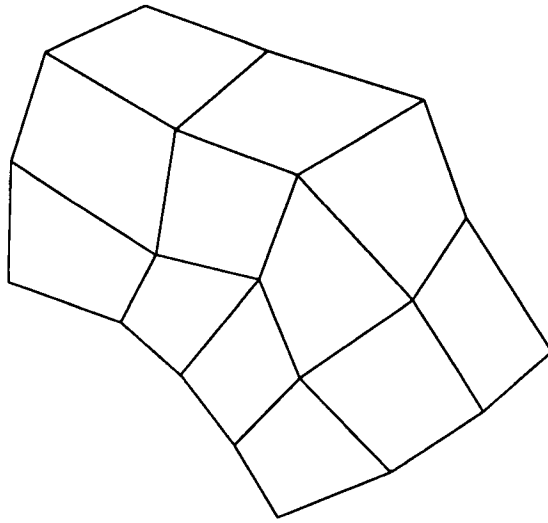


Figure 3: An irregular grid of quadrilaterals

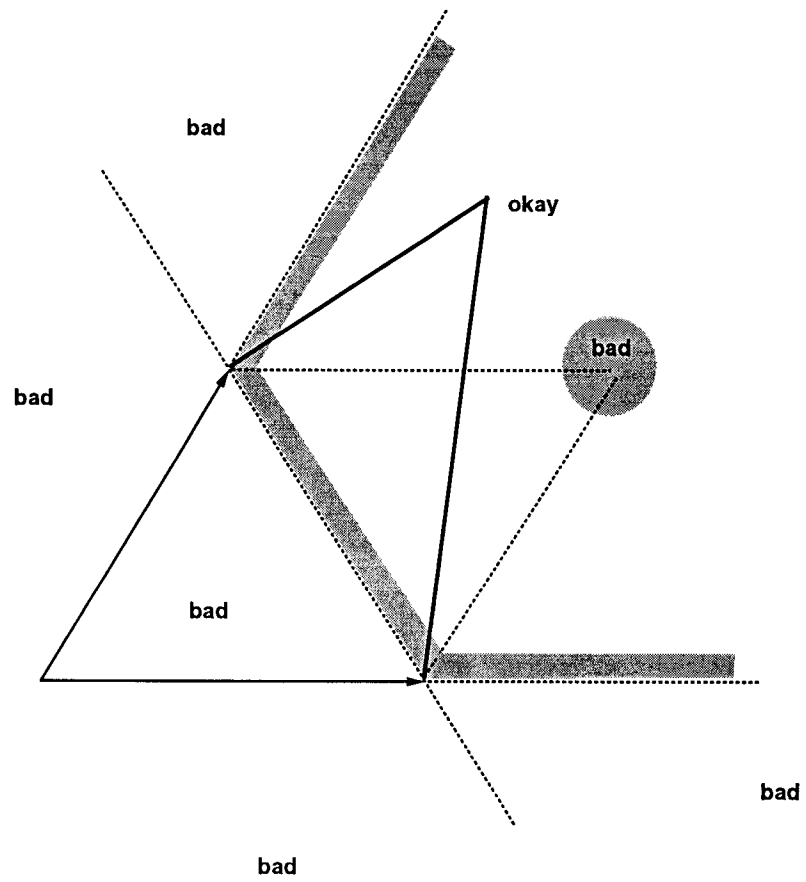


Figure 4: The non-degenerate region for the 4'th vertex of a quadrilateral

Let

$$P_{k,l} = \left\{ f(x', y') = \sum_{i=0}^k \sum_{j=0}^l f_{ij}(x')^i (y')^j \right\}.$$

Note that the dimension of $P_{k,l}$ is $(k+1)(l+1)$.

We now define the new family of spaces more formally. To create the k' th order member of this family, $k = 1, 2, \dots$, we follow the Raviart-Thomas pattern for velocities. That is,

$$V_h = P_{k,k-1} \times P_{k-1,k}.$$

Recall that we have assumed all the quadrilaterals come from the same transformation, so velocity fluxes are continuous across edges, and are in fact polynomials of degree $k-1$ in one primed variable along each edge, times fixed flux functions such as f_{x0} . Thus the velocity space contains $2k^2 + 2k$ degrees of freedom total, consisting of $4k$ values shared across edges and $2k^2 - 2k$ internal ones.

Considering divergences, we find

$$\begin{aligned} & \{V \operatorname{div} v : v \in V_h\} \\ &= \{(1 + ex' + \bar{e}y')P_{k-1,k-1} + eP_{k,k-1} + \bar{e}P_{k-1,k}\} \end{aligned} \quad (14)$$

$$= P_{k,k-1} + P_{k-1,k}. \quad (15)$$

provided both e and \bar{e} are nonzero. Therefore we take

$$W_h = P_{k,k-1} + P_{k-1,k}.$$

This yields $k^2 + 2k$ pressures.

In the lowest order case, we have 4 velocity degrees of freedom and 3 pressure degrees of freedom per element. Since this is a conforming method, velocity fluxes are continuous across edges, giving 5 unknowns net per element, in two dimensions, compared to 3 in the Raviart-Thomas case on parallelograms. When $k = 2$, there are 12 velocities and 8 pressures, with 8 of the velocities shared across edges, for a total of 16 unknowns net per element, compared with 12 in the Raviart-Thomas case on parallelograms. Figure 5 illustrates the two lowest order Raviart-Thomas spaces, and Figure 6 shows the corresponding two lowest order quadrilateral spaces on the reference element.

We also note that the above spaces have good approximation properties. This is clear since the spaces are just familiar spaces of polynomials, in the transformed coordinate system (which is as valid as any other for computing norms of errors in).

6 Three Dimensional Elements

In three dimensions the construction of the new spaces is analogous. We will not write out the general transformation, though one can easily work it out with a

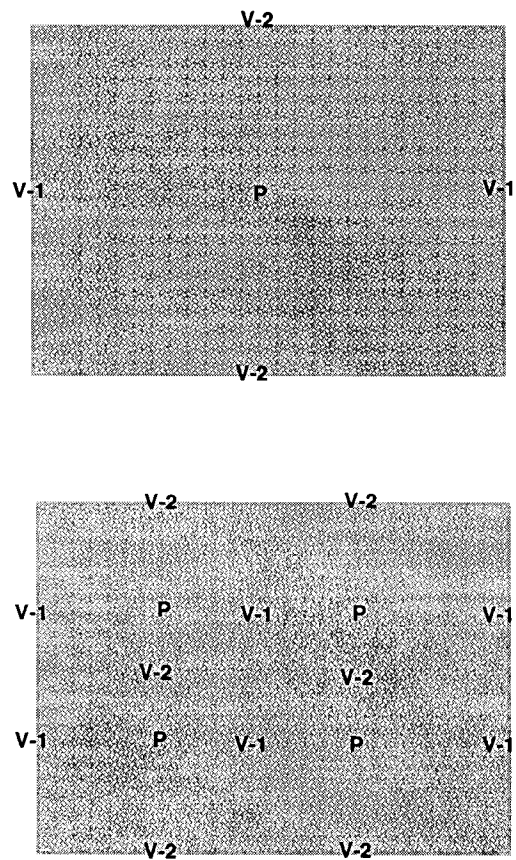


Figure 5: The two lowest order Raviart-Thomas spaces

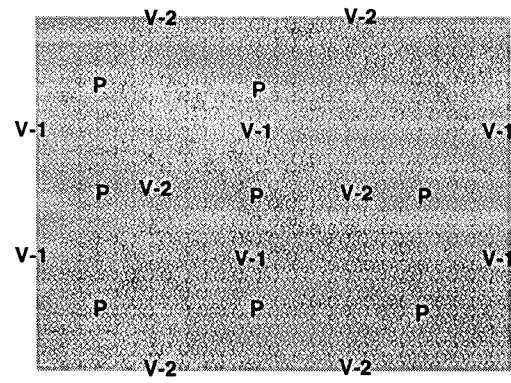
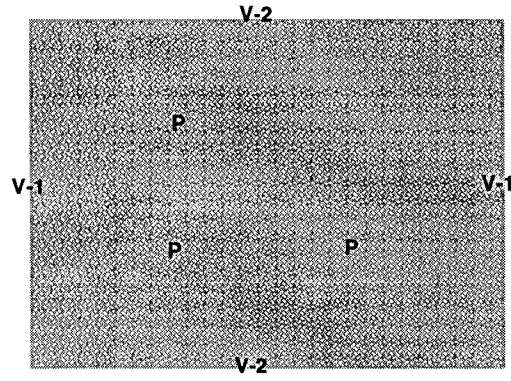


Figure 6: The two lowest order quadrilateral spaces

symbolic calculator program. A general hexahedron has 6 planar faces, each of which is a quadrilateral. A plane is determined by 3 points, and we take the first four vertices to be the same as in the two dimensional case, in the $z = 0$ plane. We now add a fifth point which would be the point on the z -axis if we had a cube. It has three degrees of freedom. Together with the origin and the point $(a, 0, 0)$ it determines a plane, so the fourth vertex on that face has only 2 degrees of freedom remaining. The same holds with the point $(ab, c, 0)$. The final vertex is constrained to lie at the intersection of three planes previously formed, and hence is uniquely determined. Thus we have 12 degrees of freedom altogether.

The velocity fluxes still satisfy conditions such as

$$v \cdot \hat{n} = v^1 f_{x0},$$

along $x' = 0$, but the f functions are now functions of two variables. To see that $v^{2'}$ and $v^{3'}$ do not enter this equation, recall that $v = v^{i'} g_{i'}$ and by construction, the normal to the face $x' = 0$ satisfies $g_{2'} \cdot \hat{n} = g_{3'} \cdot \hat{n} = 0$. Similarly for the other faces.

In the k 'th order case, we take

$$V_h = P_{k,k-1,k-1} \times P_{k-1,k,k-1} \times P_{k-1,k-1,k}.$$

Then

$$W_h = P_{k,k-1,k-1} + P_{k-1,k,k-1} + P_{k-1,k-1,k}.$$

For instance, in the lowest order case ($k = 1$), there are 6 velocity degrees of freedom, one per face, and 4 pressures. This yields 7 unknowns net per element. In comparison, subdividing the cube into 5 tetrahedra requires 16 velocities and 5 pressures, with 4 of the velocities internal to the element, yielding 15 unknowns net per cube.

7 Hybrid Methods

The hybrid form of mixed methods adds Lagrange multiplier unknowns to certain edges or faces in the set of elements, decoupling the resulting system in a way which is suitable for use with domain decomposition[8, 5]. The problem is formulated by replacing equations (1) and (2) with the equivalent system for p , u , and the new unknown $\lambda \in L^2(\cup_E \partial E)$

$$(u, v)_E = (p, \text{div } v)_E - \langle \lambda, v \cdot \hat{n} \rangle_{\partial E} \text{ for all } v \in H(\text{div}, E), \quad (16)$$

$$(\text{div } u, w)_E = (f, w)_E \text{ for all } w \in L^2(E), \quad (17)$$

and

$$\sum_E \langle u \cdot \hat{n}, \mu \rangle_{\partial E} = 0 \text{ for all } \mu \in L^2(\cup_E \partial E). \quad (18)$$

Here

$$(a, b)_E = \int_E ab,$$

is the usual L^2 inner product element by element, and

$$\langle a, b \rangle_{\partial E} = \int_{\partial E} ab,$$

is the integral over the boundary.

We have written it with Lagrange multipliers on every edge or face; one can do domain decomposition by applying them to just the edges or faces of subdomains by thinking of each E as a subdomain rather than an element.

To discretize a hybrid method we proceed as usual, using subspaces $W_h(E) \subset L^2(E)$ and $V_h(E) \subset H(\text{div}, E)$, and $\Lambda_h \subset L^2(\bigcup_E \partial E)$. The result is that we use the same spaces as before, except that velocities across subdomain boundaries do not share the same degree of freedom explicitly. This continuity is then reimposed by requiring Λ_h to be dual to the space of velocity fluxes.

In two dimensions, the k 'th order variant of the new spaces requires k multipliers shared across each cell edge, plus the $2k^2 + 2k$ velocities and the $k^2 + 2k$ pressures.

In the lowest order case this yields 2 multiplier unknowns per element, plus local problems of 7 unknowns and 7 equations on each element.

In three dimensions, the lowest order case uses 3 multipliers per element, plus 10 by 10 local problems on each element.

In contrast, using triangles to subdivide quadrilaterals also requires 2 multipliers per cell and 7 by 7 local problems in two dimensions, and using tetrahedra to subdivide hexahedra requires 6 multipliers per macro element and has 21 equations and 21 unknowns in each local problem.

8 Results

Quadrilateral and hexahedral elements arise naturally in topologically regular meshes, where one seeks to combine the flexibility of general geometry with the computational advantages of rectangular arrays. They can also be used in adaptive mesh situations, where vertices in the mesh are perturbed from their initially rectangular layout to optimize the fit to the solution. In the previous two sections we have seen a new family of discrete spaces for pressure and velocity suitable for use with mixed methods on such elements. Because mixed methods can be written in coordinate independent form, and because the compatibility conditions can be as well, and because our constructed spaces satisfy the compatibility conditions, we immediately obtain from Theorem 1:

Theorem 2 *If the k 'th order variant of the above new spaces are used for V_h and W_h to solve an elliptic equation on a domain using quadrilateral or hexahedral elements,*

then there is a constant C independent of p , u and h such that (3) and (4) have a solution, which is unique and satisfies

$$\begin{aligned} & \|u - U\|_{H(\text{div}; \Omega)} + \|p - P\|_{L^2(\Omega)} \leq \\ & C \left(\inf_{V \in V_h} \|u - V\|_{H(\text{div}; \Omega)} + \inf_{W \in W_h} \|p - W\|_{L^2(\Omega)} \right). \end{aligned} \quad (19)$$

We must check that the usual polynomial approximation theory results still hold in the primed coordinates. While approximation of functions works the same in any coordinate system, we also need to approximate the divergence, which is trickier. However, the extra terms in the divergence formula occur both in the divergence of a function and in the divergence of any approximation for it, so everything works. We immediately obtain:

Theorem 3 *Let Ω be a quadrilateral(hexahedral) domain. Let f be a bilinear(trilinear) map of the unit square(cube) onto Ω , as described above. Given a constant $C_2 \geq 1$, let the unit square(cube) be partitioned into a rectangular array of N^2 sub-rectangles(N^3 sub-cubes) of maximum diameter $h = C_2/N$, and let their images under f be the computational elements within Ω . If the k 'th order variant of the above new spaces are used for V_h and W_h to solve an elliptic equation on this domain using these elements, then there is a constant C depending on the PDE and on the geometry of the domain, and on the constants c_0 and c_1 in (12) and (13), but not depending on h , such that*

$$\|u - U\|_{H(\text{div}; \Omega)} + \|p - P\|_{L^2(\Omega)} \leq Ch^k.$$

While differential geometry relieves us from having to invent new methods and new proof techniques, the above result does come with the price tag of increased complexity of calculation. However, the increased complication may be worthwhile in certain applications, such as groundwater flow, where the resulting cell by cell conservation of mass is important. Numerical comparisons of these new elements with various competitors, including standard rectangular meshes for the same domain and triangulations or tetrahedral decompositions, will all be necessary to evaluate how substantial a penalty this really is.

References

- [1] T. ARBOGAST, L. COWSAR, P. T. KEENAN, AND M. F. WHEELER, *Non-conforming mixed methods*, tech. report, Dept. of Mathematical Sciences, Rice University, in preparation.

- [2] F. BREZZI, *On the existence, uniqueness and approximation of saddle-point problems arising from lagrangian multipliers*, R.A.I.R.O. Anal. Numer., 8 (1974), pp. 129–151.
- [3] F. BREZZI AND M. FORTIN, *Mixed and Hybrid Finite Element Methods*, Springer-Verlag, 1991.
- [4] B. CHAR, K. GEDDES, ET AL., *Maple V Language Reference Manual*, Wiley-Interscience, 1991.
- [5] L. COWSAR AND M. F. WHEELER, *Parallel domain decomposition method for mixed finite element methods for elliptic partial differential equations*, Tech. Report 90-37, Dept. of Mathematical Sciences, Rice University, 1990.
- [6] P. DIRAC, *General Theory of Relativity*, Wiley-Interscience, 1975.
- [7] L. EISENHART, *Riemannian Geometry*, Princeton University Press, 1966.
- [8] R. GLOWINSKI AND M. F. WHEELER, *Domain decomposition and mixed finite element methods for elliptic problems*, Tech. Report 87-11, Dept. of Mathematical Sciences, Rice University, 1987.
- [9] R. LASHOF AND A. LIULEVICIUS, *Topology and geometry of locally euclidean spaces*, lecture notes in mathematics, University of Chicago, Department of Mathematics, 1973.
- [10] P. A. RAVIART AND J. M. THOMAS, *A mixed finite element method for second order elliptic problems*, in Mathematical Aspects of the Finite Element Method, G. I and E. Magenes, eds., Springer-Verlag, 1977.
- [11] J. M. THOMAS, *Sur l'analyse numérique des méthodes d'éléments finis hybrides et mixtes*, PhD thesis, Université Pierre et Marie Curie, Paris, 1977.