Without addressing shape smoothness, gradient-based optimisation methods naturally amplify high-frequency shape components which can lead to poor convergence of the optimisation problem and a convergence rate exhibiting dependency on the fidelity of shape-control. Recent work by the authors demonstrated that this problem arises due to the discrete shape problem being ill-posed by naturally including geometries that are invalid both in physicality (shape) and discretisation (mesh), and a new shape control methodology has been developed which addresses this explicitly. The new approach uses two-dimensional control to recover shape-relevant displacements and surface gradient constraints to ensure smooth and valid iterates. The shape gradient constraints, approximating a $C^2$ continuity condition, are derived for two local shape control methods: discrete grid point control and cubic B-Splines. The local control methods provide high-fidelity control whereas the surface constraints exclude non-physical shapes from the design space making the shape problem well-posed. As a result, high-fidelity shape optimisation is possible at a reasonable computational cost. In this paper, further results are presented for the aerodynamic optimisation of two standard test cases defined by the AIAA aerodynamic design optimisation discussion group. A value of 7 drag counts is achieved on the inviscid case and 66 drag counts on the viscous case; to the authors’ knowledge these are the lowest results published for either case.

I. Introduction

Computational fluid dynamics (CFD) has matured rapidly over the last half-century and, in combination with increasing computing power, has become a standard tool for modelling fluid flow in industry. In addition to reducing the reliance of industry on expensive and lengthy ground and in-flight tests, this accessibility to CFD is also being increasingly applied in numerical optimal design: the coupling of optimisation methods to numerical analysis tools to improve on a design objective. Shape optimisation in particular has many applications across aerospace, with optimal design objectives such as aerodynamic drag, structural weight, aeroacoustic noise, mission endurance, electromagnetic stealth as well as various multidisciplinary objectives. A shape optimisation framework often comprises three main components: a numerical analysis code; a geometry representation and manipulation method; and a numerical optimisation method. The optimisation process involves successive calls to the analysis tool by the optimiser to iteratively improve upon the objective. Since the analysis method itself often involves an expensive iterative process, optimal design still remains an expensive process computationally. The choice of geometry control has particular influence on the number of optimisation iterations, and therefore computational cost, in shape optimisation since it acts as the interface between the search algorithm, which operates on the vector of design variables, and the analysis code, which is only concerned with the geometry. Hence the efficiency, robustness and final result of the shape optimisation problem is highly dependent on the shape control method.

A natural approach to shape control is provided by the individual vertices of the discretised analysis geometry, or mesh. In fact, for a fixed mesh resolution, this method provides the most comprehensive design space. Early work on aerodynamic shape optimisation [1, 2] avoided mesh point control however, due to the prohibitively large number of design variables and consequent high computational cost, and low dimension shape functions were adopted instead. As
aerodynamic sensitivity analysis matured in the form of adjoint flow solvers, which provide surface sensitivities at a greatly reduced cost [3], a more severe problem with surface mesh control was encountered: the rapid degeneration of surface smoothness and consequent poor efficiency of numerical optimisation methods.

Without addressing shape smoothness, gradient-based optimisation methods naturally amplify high-frequency shape components, producing non-smooth search directions. The implication for numerical optimisation is that in attempting to take a successful step the search algorithm is either forced to reduce the step size in order to avoid oscillatory shapes, or terminates at an oscillatory shape for which a sensitivity calculation is no longer possible. In the former case, the optimiser makes slow progress since more time is spent in the line search, and the overall number of major iterations required to reach the solution increases. Together these effects both lead to significant increases in function evaluations and therefore computational cost. Furthermore, convergence of the optimisation problem exhibits dependency on the resolution of the analysis mesh and fidelity of shape-control.

The source of these issues can be understood intuitively by recognising that the discrete shape problem is ill-posed by being insufficiently bounded; that is, it naturally includes geometries that are invalid both in physicality (shape) and discretisation (mesh). Several methods for shape control exist, which address the ill-posed nature of the problem either directly or indirectly; however such methods often a priori constrain the design space at low fidelities and perform poorly when shape control is refined due to insufficient regularisation of the shape problem or large increases in the conditioning of the numerical problem.

Recent work by the authors [4] has developed a new approach to shape control which, like existing methods, regularises the shape problem to maintain surface smoothness but without limitation on the resolution of shape control. The new approach uses two-dimensional control (x,z) to recovery shape-relevant displacements and surface gradient constraints to ensure smooth and valid iterates. Shape gradient constraints approximating a C2 continuity condition for discrete mesh points and cubic B-Spline curves were derived, and demonstrated on a standard test case. The resulting shape optimisation problem is well-posed, as the design space is sufficiently bounded, and well-conditioned since local control methods (mesh-points and B-Splines) are used. The local control methods with two-dimensional (x,z) control provide high-fidelity shape-relevant control, whereas the surface constraints exclude non-physical shapes from the design space. Moreover the constraint formulation also implicitly controls mesh validity by taking into account the in-plane surface variation. Significantly, the regularised shape problem is shown to have an optimisation convergence rate independent of shape control fidelity when using B-Splines, while still making use of increased control fidelity to achieve improved results. As a result, high-fidelity shape optimisation is possible at a reasonable computational cost.

Several benchmark optimisation cases are laid out by the AIAA Aerodynamic Design Optimisation Discussion Group (ADODG) to aid in the collaborative evaluation and comparison among shape optimisation practitioners. In this paper, an extension to work previously presented [4], two such test cases are considered here.

II. Background

A variety of shape control methods are encountered in aerodynamic shape optimisation all of which address shape smoothness in some way either directly or indirectly; commonly-used methods are presented in Table 1. A simple but effective approach to addressing smoothness during optimisation is to penalise bad shapes by augmenting the objective function with a weighted penalty term given by a function of some surface derivative such as curvature [22, 23] or alternatively on a calculated quantity such as pressure distribution [24]. This method is a form a regularisation which arises in a variety of fields for solving ill-posed problems. One way of viewing regularisation is the addition of prior knowledge in the form of assumptions that characterise the solution; the solution of the regularised problem is that which is most probable under the assumptions made. Within shape optimisation, penalisation of shape derivatives is equivalent therefore to an a priori characterisation of the differentiability of the solution shape. These methods are simple to implement since they only augment the objective with an explicit term, however the choice of weighting parameter, which represents an arbitrary trade-off between performance and smoothness, is not explicit but still directly influences the final result.

The most prevalent approach to addressing the challenges of shape optimisation is shape parameterisation, sometimes also referred to broadly as CAGD-based (Computer Aided Geometric Design) due to the origin of the methods; this refers to methods that describe the geometry using a lower-dimension representation with intuitive design controls. Examples include: spline surfaces; volume splines; and various other shape functions derived analytically (Hicks-Henne and shape-class functions), from physical considerations (fictitious loads ) or numerically (orthogonal modes). All such methods reduce to a linear combination of shape basis functions; different parameterisations vary only in the form and fidelity of the basis functions. Inherent in the reduction of dimensionality is the introduction of smoothing,
Table 1  Common shape control methods for optimisation

| Shape parameterisation | B-Spline surfaces [5]  
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Bezier surfaces [6]</td>
</tr>
<tr>
<td></td>
<td>Volume splines: Bezier[7], B-Spline[8], RBF[9]</td>
</tr>
<tr>
<td></td>
<td>Hicks-Henne [2]</td>
</tr>
<tr>
<td></td>
<td>CSTs [10]</td>
</tr>
<tr>
<td></td>
<td>Fictitious loads [11]</td>
</tr>
<tr>
<td></td>
<td>SVD [12]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Multilevel optimisation</th>
<th>Subdivision [13, 14]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B-Spline/Bezier knot insertion [15]</td>
</tr>
<tr>
<td></td>
<td>RBF [16]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sensitivity filtering</th>
<th>Implicit Sobolev [17]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Explicit Gaussian [18]</td>
</tr>
<tr>
<td></td>
<td>Multilevel preconditioner [19]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Shape filtering</th>
<th>Gaussian [20]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Vertex Morphing [21]</td>
</tr>
</tbody>
</table>

specifically the amplitude of higher order derivatives are implicitly bound by that of the basis functions. This advantage was emphasised by Braibant and Fleury when advocating the use of B-Splines for shape optimisation [5]. However the reduction in dimensionality inevitably corresponds to a reduction in the available design space and therefore attainable optimum; as such, the choice of parameterisation fidelity becomes important and this is investigated extensively by Masters et al. for several different parameterisation methods [25, 26]. High fidelity parameterisations often exhibit poor optimisation performance due to degraded design space conditioning or the generation of non-smooth shapes [26–28] as shape control locality approaches that of mesh-point control. In their studies, one parameterisation is shown to perform particularly well, orthogonal modes extracted by singular value decomposition (SVD). This method uses a database of aerofoils to extract shape modes, a subset of which is used as a basis for shape optimisation. The chosen SVD modes correspond to high predominance in the aerofoil library which also therefore represents filtering and removal of low-importance high-frequency shape information. An alternative approach to utilising high fidelity parameterisations, without the aforementioned issues, is hierarchical (or multilevel) parameterisations which progressively or adaptively enrich the parameterisation during successive optimisations [13–16, 29, 30].

A common alternative to shape parameterisation is mesh point control which uses the individual coordinates of nodes in the discretised analysis mesh as design variables. The advantage of this method is that the design space is only limited by the analysis discretisation - the most expansive design space possible. Furthermore, it is trivial to implement compared to setting up shape parameterisation functions, and generalisable to arbitrary geometries, assuming a discrete surface mesh exists. However, despite the theoretical simplicity, practical implementation requires that either the gradients or surface perturbations be smoothed to maintain a smooth surface. Common within aerodynamic shape optimisation is use of a weighted Sobolev gradient as suggested by Jameson [17] and used in combination with design variables defined as normal perturbations to the initial geometry. This second-order implicit smoothing method is equivalent to applying the inverse of the mesh Laplacian [31] to the surface sensitivities and is most often combined with a simple steepest descent search forming the so-called Smoothed Descent Method [32]. Studies by Jameson notably conclude that the optimisation convergence of the smoothed steepest descent method is independent of the number of design variables [17]. Stück et al. [18] instead use a Gaussian kernel to explicitly smooth shape sensitivities and the resulting filter is shown to be first-order equivalent to the Sobolev method; however the implementation has improved flexibility and efficiency compared to the implicit Sobolev method which requires the solution of a partial differential equation. Explicit and implicit smoothing methods are studied by Jaworski and Müller [33] within a multilevel shape optimisation procedure. Vazquez et al. [19] draw analogy between sensitivity filtering and preconditioning of discrete
partial differential equations, and derive a multi-level filter to compensate the loss of regularity in the gradient update.

Sensitivity filtering alone however is specific to search algorithms where the search direction is a continuous function of the gradient. Such algorithms are limited to treating constraints indirectly via barrier functions and penalty methods. The more general class of non-linear programming algorithms select the search direction by solving a subproblem formed by the constraints and therefore smoothing the gradient will not correspond to a smooth search direction. This limitation can be overcome by including shape filtering in addition to sensitivity filtering. Le et al. [20] use two initially-identical meshes where one is constructed by smoothing the other with a Gaussian operator; a secondary in-plane correction is also used to maintain a regular discretisation. Hojjat et al. [21] use a similar approach but where mesh regularisation is performed in the same step as shape smoothing. This approach effectively splits the smoothing process between shape updates and sensitivity projection and is in-fact equivalent to a full-rank re-parameterisation, that is, a shape parameterisation which does not reduce dimensionality and in which the shape basis functions are derived from the smoothing kernel.

The methods discussed so far primarily control the regularity of the shape-relevant (normal) variation, however the mesh-relevant (in-plane) variation must also be addressed to maintain validity of the surface mesh. In the continuous problem, smoothness is an \textit{out-of-plane} phenomena related to the differentiability of normal variations. However in the discretised problem an additional \textit{in-plane} component is introduced which describes the variation of the discrete surface elements, \textit{i.e.} the mesh quality, see Figure 1. A ‘good’ discretisation could be described as one in which in-plane variations have minimal sensitivity compared to normal variations, \textit{i.e.} the discretisation error is small. Hence the normal component is sometimes referred to as the \textit{shape-relevant} component and the in-plane component as \textit{mesh-relevant}; both of these components need to exhibit smoothness during shape optimisation to ensure high quality shapes and meshes. In [32], where Sobolev smoothing is used with mesh-point control, the design variables are defined as perturbations normal to the initial geometry. Whereas this guarantees shape relevant updates are made initially, the normal direction changes during optimisation which reduces the effectiveness of the design variables as they become increasingly misaligned with the new shape-relevant direction. A common approach, which maintains surface validity indirectly, is to explicitly restrict or prescribe the degrees of freedom; for example, in aerofoil optimisation it is common to restrict freedom to the longitudinal (thickness) axis, similarly it is common to group parametric control points into specific freedoms such as sweep, twist and taper. A more general and complex approach is to perform a secondary correction in the in-plane direction after the normal surface update to maintain surface validity. Such in-plane updates can be explicit corrections [20] or involve the solution of a sub-problem minimising a mesh quality criteria [34] or physical analogy such as plane stress equilibrium [35] akin to variational surface fairing. However, performing normal perturbations followed by a non-linear in-plane correction requires limiting the normal perturbation step size [20] and also introduces dependency on the initial shape from which the normal directions are defined.

The objective of the gradient-limiting approach is to generalise the regularisation effects of the methods discussed above and remove dependence of the optimisation procedure on the shape control method. In [4] shape gradient constraints are derived and it is demonstrated that a sufficiently constrained problem space is all that is required for a well-posed and consistent shape optimisation problem without limitation on the numerical method used for solution. In the remainder of this paper, motivation for the gradient-limiting approach is first presented followed by the mathematical formulation and details on numerical implementation. Finally, in section IV the methodology is demonstrated on aerodynamic drag minimisation test cases in two dimensions using a gradient-based adjoint optimisation framework.
III. Shape Constraints

A. Motivation

Development of the gradient-limiting approach is motivated by four concerns in shape optimisation [4]:
1) Shape smoothness: the variation of shape components normal to the local surface;
2) Shape control conditioning: the numerical relationship between design variables and geometry definition;
3) Coordinate control: shape control freedom in more than one coordinate direction to allow shape relevant perturbations normal to the local surface;
4) Grid validity: the variation of discrete shape components in-plane to the local surface.

As previously mentioned, using the discretised geometry alone for design variables results in an ill-posed problem which leads to poor optimisation performance. The ill-posedness arises since the length scale that characterises the numerical problem derives from that of the numerical grid, this is significantly smaller than that of the physical problem defined by the shape. In this way the search space not only contains the physical solution(s) but also non-physical solutions, which manifest as non-smooth shapes with wavelength of the numerical grid. This problem affects localised shape control methods, such as B-Splines, whereby as the number of design variables increases the shape functions become more localised, approaching the resolution of the numerical grid. Since the shape functions are more localised they exhibit non-smoothness and no-longer regularise the inverse problem sufficiently thereby resulting in non-smooth search directions, poor convergence and sub-optimal shapes. The poor performance of B-Splines at high fidelity matches that observed by Reuther and Jameson [27] and again by Castonguay et al. [28] who subsequently showed that this effect is eliminated when used with sensitivity filtering. Masters et al. [26] found that increasing the order of B-Splines, which increases the support and smoothness of the basis, was also able to alleviate some of the convergence issues.

In contrast, shape functions which exhibit a large zone of influence and remain smooth with refinement, the conditioning of the shape control generally increases significantly as the number of design variables increases. Whereas local basis functions retain linear independence as they are refined, shape basis functions with large support suffer from deteriorating linear independence such that increasing the number of design variables give diminishing additions to the design space. This means that similar shapes do not have similar control vectors, i.e. shapes that are geometrically similar do not lie close together within the design space. This is especially undesirable for gradient-based optimisers which generate a sequence of solutions by stepping through the design space, the effect being that although convergence is generally unaffected by increased control fidelity, the final attainable result of the optimisation is not improved. For example, in [36] high-order Beziers (96 knots) became ‘saturated’ as the design variables hit their respective upper bounds during optimisation and preventing further improvement.

Also of importance for shape optimisation, in addition to smoothness and conditioning, are the actual degrees of freedom available to the search algorithm. Among previous studies it is common for the parameterisation to be manually constrained such that only one ordinate is free to move [36], or that certain degrees of freedom (e.g. sweep, taper) are specified by hand [9], or combinations of both [37]. The parameterisation of a single coordinate, thickness, for aerofoils is convenient and commonly adopted since aerofoils are predominantly planar such that most of the surface normals are closely aligned with the vertical coordinate. However this has three drawbacks: first, the single coordinate parameterisation is restricted in its ability to produce local rotation of the surface; second, there is poorly defined control in areas where the local normals do not initially align with the vertical (e.g. the leading edge); and finally the notion of a dominant coordinate direction is not general and does not extend to three-dimensional geometries. In a similar way, manually specifying degrees of freedom relies on user-knowledge which, in the general case, does not necessarily know which controls are useful a priori. The implication is that to allow shape relevant perturbations (displacements normal to the current shape) at all stage of optimisation, shape control must be permitted in all coordinate directions. A side-effect of this is that consideration must be made for the grid-relevant, in-plane, variation of the discretised surface.

Based on this, the rationale for the gradient-limiting approach is to perform localised shape control in all coordinate directions and constrain the larger resulting design space by utilising a geometric constraint to maintain validity and continuity of the discrete surface. Constraint based continuity has been applied in [38] and [39] for structural topology optimisation where first spatial derivatives were bounded to enforce approximate $C^0$ on the solution. Xu et al. [40] use a non-linear constraint to enforce geometric continuity ($G^2$) between separate spline patches in their CAD-based shape optimisation. Similarly Cinquegrana et al. [23] include a penalty on derivatives of surface curvature to counter oscillations induced by a NURBS parameterisation. In the work here an approximate $C^2$ condition is applied across the entire shape to control both shape smoothness and mesh quality. In this section the approximate $C^2$ gradient constraint is presented for direct mesh point control.
B. Geometric smoothness

In this section the concept of smoothness is presented within the mathematical framework of a plane uniparametric curve (Figure 2(a)) in $\mathbb{R}^2$, described continuously by:

$$
\mathbf{x}(t) = \begin{bmatrix} x(t) \\ y(t) \end{bmatrix}, \quad t \in [a, b] \subset \mathbb{R} 
$$

where the component functions $x, y$ are differentiable with respect to the parameter $t$, denoted here by dot, such that the non-vanishing tangent vector is given by:

$$
\dot{\mathbf{x}}(t) = \begin{bmatrix} \dot{x}(t) \\ \dot{y}(t) \end{bmatrix} \neq 0, \quad \forall t 
$$

A local coordinate frame at the point $\mathbf{x}(t)$ is constructed exposing familiar geometric quantities; the unit tangent vector $T$ and unit normal vector $N$ are given by:

$$
T = \frac{\dot{\mathbf{x}}}{||\dot{\mathbf{x}}||}, \quad N = \frac{\dot{T}}{||\dot{T}||} 
$$

Any parametric curve can be re-parameterised naturally by the arc length parameterisation, $s(t)$:

$$
\dot{s}(t) = ||\dot{\mathbf{x}}|| 
$$

Which yields the following relationships:

$$
\mathbf{x}' = T, \quad \mathbf{x}'' = \kappa N 
$$

where $\kappa$ is the local curvature and prime $'$ denotes differentiation with respect to arc length.

One way to mathematically describe smoothness is differentiability. The definition of $\mathbf{x}(t)$ above states that the curve is at least once differentiable with respect to the parameter $t$ and hence this curve has parametric continuity 1, denoted $C^1$, which implies continuously varying tangent vectors. Parametric curves that are twice differentiable with respect to their current parameterisation are $C^2$ and will be curvature continuous. However this definition is dependent on the particular parameterisation and indeed a curve can have continuity of curvature while not being $C^2$. Hence the more general class of curves is defined as Geometrically continuous $G^2 \supset C^2$ which contains those curves that are twice differentiable with respect to the invariant arc length parameterisation. Using subscript $+$ and $-$ to represent the left and right limits respectively, then the conditions for $C^2$ and the more general $G^2$ can be expressed as:

$$
C^2 : \quad \ddot{x}_+(t) - \ddot{x}_-(t) = 0 
$$

$$
G^2 : \quad \mathbf{x}''_+(s) - \mathbf{x}''_-(s) = 0 
$$

The difference between the stricter $C^2$ condition and the more general $G^2$ can be explained by expanding the term $\ddot{x}$ using the chain rule:

$$
\ddot{x} = \mathbf{x}''s^2 + \mathbf{x}'s 
$$
Substituting the relations of equation 5 this becomes:

\[ \ddot{x} = \kappa \dot{s}^2 N + \dot{s} T \]  

The parametric derivative \( \dot{x} \) contains a normal component corresponding to the local curvature, and an in-plane component representing the ‘acceleration’ of the parameterisation. Notice that the condition for \( G^2 \) (equation 7), which uses the arc length derivative \( x'' \), is concerned only with the normal component (for curvature continuity) whereas the stricter \( C^2 \) condition (equation 6), defined using the parametric derivative \( \dot{x} \), also requires continuity of the parameterisation; this is illustrated in Figure 3 where ‘velocity’ refers to the rate of change in the parameter and ‘acceleration’ refers to the rate of change in velocity. i.e. Parametric continuity is a special case of geometric continuity. This makes intuitive sense since only the normal component is visible (shape-relevant) and so the class of visually smooth curves should not exclude those with discontinuous parameterisations hence the classification of geometric continuity \( G^k \). Therefore all geometrically continuous curves permit an arclength parameterisation \( t \rightarrow s \) (equation 4), however the same cannot be said about the inverse process, formulating a \( C^k \) parameterisation \( s \rightarrow t \) for a \( G^k \) curve.

In this work, the parametric derivative will be used to enforce the stricter \( C^2 \) condition since, by suitable choice of the parameter \( t \), the in-plane component can be used to represent the smoothness of the discretisation. Hence by enforcing the \( C^2 \) condition both the shape relevant and mesh relevant components can be controlled simultaneously.

C. Gradient constraint for mesh point control

Consider now a mesh made up of vertices sampled at discrete parameter locations (Figure 2(b)) described by:

\[ x_i = x(t_i), \quad i \in [0, n] \subset \mathbb{Z} \]  

In this discrete setting differentiation with respect to the parameterisation \( t \) (denoted by dot) is approximated by discrete difference in the logical domain such that the facets of the mesh are given by:

\[ \dot{x} = \Delta x_i = x_{i+1} - x_i \neq 0, \quad i \in [1, n] \subset \mathbb{Z} \]  

Similarly the relations of equation 3 give approximations to the unit tangent and normal vectors defined at the surface elements and vertices respectively by:

\[ T_i = T(\Delta x_i) = \frac{\Delta x_i}{||\Delta x_i||}, \quad N_i = N(x_i) = \frac{T_{i+1} - T_i}{||T_{i+1} - T_i||} \]  

and the arc length parameterisation is replaced by an approximation using chord length:

\[ s_i = \sum_{i=1}^{i} ||\Delta x_i|| \]  

To enforce some smoothness restriction on the mesh, an analogous condition to that of \( C^2 \) (equation 6), based on the parametric derivative \( \dot{x} \) is sought. Since the parameter \( t \) corresponds to the mesh point index \( i \), then the second parametric derivative \( \ddot{x} \) is simply the uniform mesh Laplacian given by:

\[ \ddot{x}_i = \Delta x_{i+1} - \Delta x_i = x_{i+1} - 2x_i + x_{i-1} \]
D. Numerical implementation

A consequence of bounding the parametric derivative is that the resulting smoothness constraints for mesh point control and B-Splines are linear with respect to the coordinate design variables and can hence be implemented in following form:

\[
\begin{bmatrix}
D \\
-D
\end{bmatrix} \alpha \leq \epsilon
\]

(16)

Here \( \alpha \) is the design variable vector consisting of stacked mesh point or control polygon coordinates for mesh point and B-Spline shape control respectively and \( \epsilon \) is the constraint bound which is henceforth given by:

\[
\epsilon = \sigma h^3
\]

(17)

where \( h \) is the mean mesh spacing of the initial grid and \( \sigma \) is the constraint parameter related to the minimum length scale imposed by the constraint. Extensive studies have been performed to analyse the range for the \( \sigma \) value including inverse geometric design, as applied in [26].

The matrix \( D \) is either a third difference matrix for mesh point control or a spline derivative matrix for B-Spline control. For mesh point control the third difference matrix forms a linear approximation to the coordinate components of \( \tilde{x}_{i+1} - \tilde{x}_i \) in equation 15. In this work, simple first-order finite-differences have been used, however higher order methods may also be used. Both the mesh point differences and B-Spline derivatives are notoriously susceptible to numerical noise due to the repeated differencing, however for the application in this work it is the relative magnitude of the surface derivative that is of concern and not the local accuracy. Moreover since the condition is implemented as a linear constraint, the generation of feasible iterates during optimisation can be handled robustly and efficiently by existing methods (linear/quadratic programming).

E. Constrained optimisation: SNOPT

The SNOPT [41] (Sparse Non-linear Optimiser) package is used here for gradient-based optimisation. This package implements a Sequential Quadratic Programming (SQP) algorithm for solving general non-linear constrained optimisation problems. The power of this package lies in its ability to efficiently and robustly handle large problems (\( \approx 1000 \)s of variables and constraints) while allowing precise constraint satisfaction.
The SQP algorithm operates iteratively whereby successive search directions are found from the solution of a quadratic programming (QP) sub-problem and a line-search is used to determine step length. In this work a non-derivative line-search is chosen. The sub-problems are formed from quadratic approximations to the augmented objective function (Lagrangian) and linearisations of the constraints. The quadratic approximation is initialised with an identity matrix and BFGS updates are used to approach the Hessian of the Lagrangian. The quadratic sub-problems are well-posed and the function values and derivatives thereof are easily evaluated; difficulty is not usually encountered when solving these sub-optimisation tasks which, within SNOPT, are handled by the SQOPT semi-definite QP solver [42].

**IV. Two Dimensional Drag Minimisation**

**A. Flow discretisation and analysis**

For this work the Stanford University Unstructured (SU$^2$) [43] flow solver is adopted. This open-source software was constructed with aerodynamic shape optimisation in mind and hence has both continuous and discrete adjoint implementations [44]. The main flow solver implements both the compressible Euler and RANS equations using an unstructured finite volume method. Multigrid acceleration is available as well as MPI parallel processing. The SU$^2$ suite also includes other modules for tasks such as shape parameterisation, mesh adaption and mesh deformation, however only the CFD module is used here for obtaining flow solutions and flow sensitivities. During optimisation, the primal and adjoint solutions are converged down to an absolute residual of $10^{-8}$ on density (and corresponding dual), except in cases where the primal has restarted from a previous solution when a tighter tolerance of $10^{-10}$ is used.

High quality volume meshes for the initial surface geometry are generated here by a conformal mapping approach to produce O-Grids for inviscid calculations, and by transfinite interpolation with smoothing to produce C-Grids for viscous calculations [45]. During shape optimisation, mesh deformation is used to produce new meshes for the displaced surface geometry from the initial volume mesh. Not only is this computationally cheaper than regenerating a mesh for each geometry iteration but it also maintains consistency of the discretisation error which is highly desirable during iterative numerical optimisation. Therefore robustness and preservation of mesh quality are of prime importance for mesh deformation schemes so that no secondary effects are able to influence the optimisation process.

In this work interpolation using multiscale radial basis functions (RBFs) [46] is used. Interpolation using radial basis functions (RBFs) has recently become a prominent mesh deformation method boasting excellent robustness and quality-preserving characteristics [47–49]. Moreover, the method is completely generic, operating on point-clouds alone, and is perfectly parallel due to the localised surface rotation. The multiscale RBF method [46] is particularly effective, both increasing the computational efficiency and improving the system conditioning over conventional or reduced datapoint RBF methods, by using variable length scales depending on boundary point locations.

**B. Test case 1: Symmetric transonic drag reduction of NACA0012**

This case is defined as minimising the inviscid drag at Mach 0.85 and 0 deg incidence subject to the constraint that the solution shape must lie outside the profile of the initial shape, a NACA0012 aerofoil:

$$\begin{align*}
\min_\alpha & \quad C_D \\
\text{s.t.} & \quad y \geq y_0 \\
& \quad 0.85 \text{ Mach} \\
& \quad 0 \text{ incidence}
\end{align*}$$

(18)

This case is chosen since it has been extensively studied in literature [36, 50, 51] and most recently within the Bristol Fluids group by Masters et al. [26]. The case has several characteristics that make it a particularly difficult problem for numerical optimisation such as asymmetric solutions, Mach-hysteresis and non-unique solutions [51]. Moreover the geometric problem is demanding, requiring a high-fidelity and flexible design space as well as a robust and quality-preserving mesh deformation method.

Masters et al. compared six different parameterisations on the benchmark problem [26] as well as applying a multilevel subdivision parameterisation [14]. The authors achieved a result of 27.8 drag counts on a $257 \times 257$ mesh using single-level B-Splines (16 degrees of freedom) and a result of 15.7 drag counts on the same mesh using multilevel subdivision curves (129 degrees of freedom); when performed on a $1025 \times 513$ mesh the multilevel subdivision scheme was able achieve an impressive 4.2 drag counts, the lowest published result so far. Two distinct solution shapes were
observed by the authors differing in boat-tail angle and thickness, where the thicker shapes with larger-boat tail angles corresponded to different trailing-edge flow structures with lower overall inviscid drag. Two techniques found to be particularly effective by Masters for this test case are also used here: first, the explicit imposition of a symmetry boundary condition on the volume mesh to stabilise convergence of the flow; and second, the use of flow-restarts during optimisation to minimise the influence of flow hysteresis during line-searches.

Since shape control is permitted in all axes, the constraint on lying outside the initial profile requires some extra consideration. A linearisation is made at each point on the surface and constraints applied such that the displacement of each point must be above the initial local tangent plane, i.e. the dot product of displacement and initial normal must be positive. An advantage of this is that the constraint can be implemented linearly and is hence guaranteed to be satisfied (when using SQP). Whereas this constraint is only geometrically approximate, it is found to be sufficiently accurate for well-discretised surface meshes. Moreover this can be considered ‘conservative’ for convex shapes in the sense that when the approximate constraint is active, the real constraint remains strictly feasible.

![Fig. 5 Convergence of symmetric transonic drag minimisation using gradient constrained mesh point control.](image)

The transonic test case described above was run on a 257 × 257 half-mesh with a farfield distance of 50 chords. Shape control is implemented using the two ordinates of the 257 vertices of the surface mesh such that there are 514 design variables. The surface gradient constraint is applied at each point (equation 16) as a linear constraint in SNOPT and multiple optimisations have been performed for varying values of the continuity parameter $\sigma$, summarised in Table 2. In this work, minor iterations refer to those of the line-searches performed at each major iteration. The number of minor iterations is equal to the sum of the number of flow solutions and adjoint solutions (where one adjoint solution is required at each major iteration) and hence gives an indication of computational work.

<table>
<thead>
<tr>
<th>$\sigma$</th>
<th>Drag (counts)</th>
<th>Major iterations</th>
<th>Minor iterations</th>
</tr>
</thead>
<tbody>
<tr>
<td>25</td>
<td>18.99</td>
<td>43</td>
<td>145</td>
</tr>
<tr>
<td>50</td>
<td>11.04</td>
<td>52</td>
<td>179</td>
</tr>
<tr>
<td>100</td>
<td>7.68</td>
<td>60</td>
<td>229</td>
</tr>
<tr>
<td>150</td>
<td>7.13</td>
<td>57</td>
<td>223</td>
</tr>
<tr>
<td>200</td>
<td>7.92</td>
<td>65</td>
<td>254</td>
</tr>
</tbody>
</table>

Figure 5 shows the optimisation convergence histories for the sweep of parameter values and Figure 6 shows the resulting optimised profiles. All cases recover similar profiles, specifically those exhibiting large boat-tail angles and a maximum thickness at approximately fifty percent chord as found in [26]. The optimised pressure distributions, shown in Figure 7, clearly show the removal of the strong shock and the tendency towards a flat distribution with minimal variation at the leading edge acceleration and trailing edge compression. The same result is also clear in the Mach and pressure contours shown in Figure 8.

In the case of the two tightest gradient constraints considered here, $\sigma = 25$ and $\sigma = 50$, the higher optimal drag values indicate that the respective design spaces may be over-constrained for the shape problem compared to larger
values of $\sigma$ which achieved better results. In the resulting profiles (Figure 6), the lesser performing cases have not achieved as tight a corner at the leading edge and have lower overall thicknesses indicating that the gradient constraint has restricted the short wavelength variations required in these areas. In contrast, where $\sigma > 100$, all cases achieved below 8 drag counts with variations of less than a drag count between them. As the gradient constraint is loosened (larger $\sigma$) the attainable result improves, corresponding to a tighter leading edge corner, a maximum thickness closer to 50% and a smaller boat-tail angle. The improved results for larger $\sigma$ also come at increased computational cost of optimisation (more function calls), however the increase is not substantial and is expected since the feasible design space is larger for larger values of $\sigma$.

The best result is obtained for $\sigma = 150$ with an objective value of 7.13 drag counts. The previous best result for this test case, at the same mesh resolution ($257 \times 257$), is 15.7 drag counts by Masters et al. using multilevel subdivision curves [14]. The improvement achieved here is primarily attributed to the use of coordinate control (control over both x and y coordinates) which has allowed unrestricted representation of the blunt leading and trailing edges observed in the optimised results. Coordinate control (of mesh points or spline control points) is made practical by the gradient constraints which, as well as regularising the localised shape control, introduce consideration for surface connectivity by their additional in-plane components. As a result the deformed surface and volume meshes are of higher quality than would be produced by unconstrained coordinate control or thickness modifications alone.
Fig. 6 Initial and optimised shape of symmetric transonic drag minimisation using gradient-limited mesh point control
Fig. 7  Optimised pressure distributions from gradient-limited mesh-point control (top), showing detail at the trailing edge (bottom).

Fig. 8  Mach and pressure fields for baseline and optimised profiles.
C. Test case 2: Transonic viscous drag reduction of lifting RAE2822

This case is defined as the drag minimisation of the RAE 2822 aerofoil in transonic viscous flow subject to constraints on volume ($V$), lift ($C_L$) and moment ($C_M$). In this work a modified form has been used in which the lift constraint is relaxed to an inequality:

$$\min_{\alpha} C_D \quad \text{s.t.} \quad C_L \geq 0.824 \quad V \geq V_{\text{initial}} \quad C_M \geq -0.092 \quad \text{Mach no. 0.734} \quad \text{Reynolds no. } 6.5 \times 10^6$$

(19)

Multiblock C-Grids extending to 100 chords are generated by transfinite interpolation with improved orthogonality and smoothness [45]. Three grid densities are considered here, summarised in Table 3, and shown in Figure 9 for the lowest density.

Fig. 9 (257 + 130) x 97 C-Grid around RAE2822 aerofoil, near-field view.

Table 3 RAE 2822, $M=0.734$, $R=6.5E6$

<table>
<thead>
<tr>
<th>Grid size</th>
<th>$C_L$</th>
<th>$C_D$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(n_{\text{surf}} + n_{\text{wake}}) \times n_f$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$(257 + 130) \times 97$</td>
<td>0.8240</td>
<td>0.0195</td>
</tr>
<tr>
<td>$(385 + 194) \times 145$</td>
<td>0.8240</td>
<td>0.0173</td>
</tr>
<tr>
<td>$(513 + 258) \times 193$</td>
<td>0.8240</td>
<td>0.0160</td>
</tr>
</tbody>
</table>

Based on the results of the initial parameter sweep a value of $\sigma = 50$ was chosen and optimisation is performed on the $(513 + 258) \times 193$ grid. Shape control is again implemented using the two ordinates of the vertices of the surface mesh plus an additional one for angle of incidence giving 1025 design variables.

Figure 10 shows the convergence history during optimisation. Note that the first iterate is not trimmed and is hence not feasible. The majority of the drag reduction is made in the first five iterations along with reductions in infeasibility as the lift constraint is gradually met and optimisation terminates after 12 major iterations due to lack of improvement in the objective. The final drag value attained is 66.5 counts at a slightly higher lift coefficient of 0.830, which is feasible under the relaxed inequality constraint, and which is, to the authors’ knowledge, the lowest value achieved for this test case.

Figures 11 and 12 show the performance of the resulting optimised profile compared with that of the trimmed RAE 2822. There is a noticeable shift in area from fore to aft (Figure 11(b)), with corresponding change to the point of maximum thickness, resulting in the replacement of the shock with a gradual compression (Figure 11(a)). These geometric changes are in line with those achieved by other authors [52]. A more subtle modification is made at the leading edge whereby combined $x$ and $y$ displacement has increased the radius of curvature at the leading edge, leading the significant increase in peak pressure and slight off-loading of the lower surface.
Fig. 10  Optimisation convergence

Fig. 11  Comparison of baseline and optimised profiles

Fig. 12  Mach flow fields for baseline and optimised profiles
V. Conclusions

Further results have been presented for the novel gradient-limiting shape control methodology applied to aerodynamic optimisation. The methodology utilises shape constraints to ensure surface smoothness and mesh validity when using local shape control for optimisation; unlike existing shape control methods the constraint-based approach does not a priori constrain the design space or require the manual definition of shape functions and deformation modes. The resulting shape optimisation problem is well-posed, as the design space is sufficiently bounded, and well-conditioned since local control methods are used. The local control methods used in two-dimensions produce a high-fidelity shape-relevant design space, whereas the surface constraints exclude non-physical shapes. Moreover, the constraint formulation also implicitly controls mesh validity by taking into account the in-plane surface variation. The constraint-based approach has been demonstrated on two standard aerodynamic drag minimisation problems, achieving the lowest published results so far. The constraint bound corresponds to enforcing a minimum length scale on the shape problem and thereby introduces regularisation such that grid-point shape control generates smooth shapes while still taking advantage of the large design space. As a result, high-fidelity shape optimisation is possible at a reasonable computational cost.

Acknowledgments

This work was carried out using the computational facilities of the Advanced Computing Research Centre, University of Bristol - http://www.bristol.ac.uk/acrc. This work was funded through an EPSRC Doctoral Training Grant.

References


