M. O. BRISTEAU  
R. GLOWINSKI  
O. PIRONNEAU

**Numerical Solution of the Transonic Equation by the Finite Element Method via Optimal Control**

*Publications des séminaires de mathématiques et informatique de Rennes*, 1976, fascicule S5  
« Journées « éléments finis » », p. 1-12  
<http://www.numdam.org/item?id=PSMIR_1976___S5_A6_0>
NUMERICAL SOLUTION OF THE TRANSONIC EQUATION BY THE
FINITE ELEMENT METHOD VIA OPTIMAL CONTROL.

M.O. Bristeau, R. Glowinski, O. Pironneau

ABSTRACT
It is shown that the transonic equation for compressible potential
flow is equivalent to an optimal control problem of a linear distributed
parameter system. This problem can be discretized by the finite element
method and solved by a conjugate gradient algorithm. Thus a new class of
method for solving the transonic equation is obtained. It is particularly
well adapted to problems with complicate two or three dimensional geometries
and shocks.

1. PLAN
2. Introduction
3. Statement of the problem
4. Gelder's algorithm for subsonic flow
5. Formulation via optimal control
6. Discretization and numerical solution
7. Numerical results
8. Conclusion

2. INTRODUCTION
The transonic equation is a non linear partial differential equation
which has an elliptic behavior in the subsonic regions of the flow and a
hyperbolic behavior in the supersonic regions. At the interface the normal
component of the speed of the flow can be discontinuous (shocks). Some
finite difference methods have been successfully developed even for
flows around simple 3-D objects (Jameson (1974), Garabedian-Korn (1971)).
However the method of finite difference is not well suited to complicate geometries. An alternative approach using finite elements was studied by Gelder (1971), Norries & de Vries (1973), Periaux (1975) but their methods explode at supersonic speeds. Following Gelder's approach we shall replace the transonic equation by the minimization of a functional in an abstract space, a problem which can be solved by the methods of the theory of calculus of variations and optimal control theory.

3. STATEMENT OF THE PROBLEM

Stationary adiabatic monophasic compressible flows, in which the effects of viscosity are neglected, are well described by the set of equations

\[ \nabla \cdot (\rho \mathbf{u}) = 0 \]

\[ \rho = \rho_0 \left( 1 - \frac{\gamma - 1}{\gamma + 1} \left| \mathbf{u} \right|^2 \right)^{\frac{1}{\gamma - 1}} \]

\[ \mathbf{u} = \nabla \phi \quad (u_i = \frac{\partial \phi}{\partial x_i}, \ i = 1,2,3) \]

where \( \rho \) is the density, \( \mathbf{u} \) is the speed of the fluid and where \( \rho_0, c_s, \gamma \) and \( \gamma \) are constants (\( \gamma = 1.4 \) for di-atomic gas, see Landau-Lifchitz (1971)).

We shall denote \( k = \frac{\gamma - 1}{\gamma + 1} \frac{1}{c_s^2} \), \( \alpha = 1/\gamma - 1 \). Therefore, if \( \Omega \) is the region occupied by the fluid, one must solve the nonlinear partial differential equation:

\[ \nabla \cdot \left( (1-k|\nabla \phi|^2)^\alpha \nabla \phi \right) = 0 \text{ in } \Omega \]

with the boundary conditions

\[ \phi |_{\Gamma_1} = \phi_1 \]

\[ \frac{\partial \phi}{\partial n} |_{\Gamma_2} = g_2 \]
Where \( \Gamma_1 \) and \( \Gamma_2 \) are parts of the boundary \( \partial \Omega \) of \( \Omega \). We shall assume that 
\( \Gamma_1 \cup \Gamma_2 = \partial \Omega \) and \( \Gamma_1 \cap \Gamma_2 = \emptyset \). In addition, if there are shocks (i.e. 
lines or surfaces where the tangential speed of the flow is continuous 
but the speed normal to these lines or surfaces is discontinuous) then, 
across the shock:

\[
(3.7) \quad (\rho u)^+ = (\rho u)^- \quad \text{(Rankine-Hugoniot condition)}
\]

\[
(3.8) \quad \frac{u^+}{n} \leq \frac{u^-}{n} \quad \text{(entropy condition)}
\]

where it is understood that the particles of the fluid move from \(-\) to \(+\).

Note that (3.4) multiplied by \( w \in C^1(\Omega) \) and integrated by parts, 
leads to

\[
(3.9) \quad \int_{\Omega} (1-k|\nabla \phi|^2)^{\alpha} \nabla \phi \cdot \nabla w \, dx = \int_{\Gamma} (1-k|\nabla \phi|^2) \ g_2 w \, d\Gamma_2 \\
\quad \forall w \in C^1(\Omega) \text{ s.t. } w|_{\Gamma_1} = 0 ; \phi|_{\Gamma_1} = 0
\]

If the notion of derivative is extended and the space \( C^1(\Omega) \) is replaced by 
\( H^1(\Omega) = \{w \in L^2(\Omega) | \nabla w \in (L^2(\Omega))^3 \} \) then (3.9) is called a weak formulation 
of (3.4)-(3.6). Note that it contains (3.7).

4. GELDER'S ALGORITHM FOR SUBSONIC FLOW

For notational convenience we suppose \( g_2|_{\Gamma_2} = 0 \). Consider the 
functional

\[
(4.1) \quad E_0(\phi) = - \int_{\Omega} (1-k|\nabla \phi|^2)^{\alpha+1} \, dx
\]

we shall say that \( \phi \) is a stationary point of \( E_0 \) on 

\[
(4.2) \quad H^1_{0,1}(\Omega) = \{ \phi \in H^1(\Omega) | \phi|_{\Gamma_1} = 0 \}
\]

if

\[
\delta E_0 = E_0(\phi + \delta \phi) - E_0(\phi) = o(\delta \phi) \quad \forall \delta \phi \in H^1_{0,1}(\Omega)
\]
Since, from (4.1)

\[ (4.3) \quad \delta E_0 = \int \Omega \left[ 2k(a+1)(1-k|V\phi|^2)\alpha \nabla \phi \nabla \delta \phi \right] dx + o(\delta \phi) \]

any stationary point of \( E_0 \) on \( H^1_{o1}(\Omega) \) satisfies

\[ \int \Omega \left[ (1-k|V\phi|^2)\alpha \nabla \phi \nabla w \right] dx = 0 \quad \forall w \in H^1_{o1}(\Omega) \]

Thus all stationary points of \( E_0 \) on \( H^1_{o1}(\Omega) \) such that \( \phi|\Gamma_1 = \phi_0|\Gamma_1 \) and which satisfy (3.8) are solutions of our problem.

Let us look at

\[ \frac{d^2}{d\lambda^2} (E(\phi+\lambda \delta \phi)) \bigg|_{\lambda=0} = 2k(a+1)\int \Omega \left[ (1-k|V\phi|^2)\alpha \nabla \delta \phi \nabla \phi - \frac{2k\alpha(V\phi \cdot V\delta \phi)^2}{(1-k|V\phi|^2)} \right] dx \]

with our notation the mach number is such that

\[ M^2 = 2k\alpha \left( 1-k|V\phi|^2 \right)^{-1} |V\phi|^2 \]

therefore, if \( \theta \) is the angle between \( V\phi \) and \( V\delta \phi \);

\[ \frac{d^2 E}{d\lambda^2} = -2k(a+1)\int \Omega \rho(1-M^2 \cos^2 \theta) |V\delta \phi|^2 dx \]

This shows that if in some part of the fluid \( M>1 \), \( E \) is not convex and the solution of (3.4)-(3.8) is only a saddle point of \( E \). On the other hand, if \( M<1 \) in \( \Omega \) then \( E \) is convex and the solution of (3.4)-(3.8) is a minimum of \( E \). This fact was utilized by Gelder (1971) and Periaux (1975) for constructing a solution of (3.4)-(3.8). The functional \( E \) is minimized by a gradient method with respect to the \( H^1(\Omega) \)-norm; i.e. \( \{\phi_n\}_{n\geq 2} \) is constructed by solving for \( \phi_{n+1} \in H^1(\Omega) : \)

\[ \int \Omega \nabla \phi_{n+1} \cdot \nabla w d\Omega = 0 \quad \forall w \in H^1_{o1}(\Omega) \]

This method works very well (less than 15 iterations in most cases) and it is desirable to construct a method as near to it as possible, for supersonic flows,
5. FORMULATION VIA OPTIMAL CONTROL

Along the line of §5 we shall look for functionals which have the solution of (3.4)-(3.8) for minimum. Several functional where studied in Glowinski-Pironneau (1975) and Glowinski-Periaux-Pironneau (1976). In this presentation we shall study the following functional:

\[
E(\xi) = \int_{\Omega} \rho(\|\nabla \xi\|^2) \|\nabla(\phi - \xi)\|^2 \, dx, \quad \rho(\|\nabla \xi\|^2) = (1-k\|\nabla \xi\|^2) a
\]

where \( \phi = \phi(\xi) \) is the solution in \( H^1(\Omega) \) of

\[
\int_{\Omega} \rho(\|\nabla \xi\|^2) \nabla \phi \nabla w \, dx = 0 \quad \forall w \in H^1_0(\Omega), \quad \phi|_{\Gamma_1} = \phi_1
\]

Proposition 1

Given \( \varepsilon > 0 \) small the problem

\[
\min \{ E(\xi) | \xi \in \Xi \}
\]

where \( \Xi = \{ \xi \in H^1(\Omega) | \xi|_{\Gamma_1} = \phi_1, \|\nabla \xi(x)\| \leq k^{-1/2}(1-\varepsilon) \text{ a.e. } x \in \Omega \} \) has at least one solution and if \( \Delta \xi(x) < +\infty \forall x \in \Omega \), it is a solution of (3.4)-(3.8) if it exists.

Proof

Let \( \{\xi_n\} \) be a minimizing sequence of \( E \) then \( \xi_n \in \Xi \) implies that

\[
\|\nabla \xi_n\|^2 < k^{-1}(1-\varepsilon)^2 \int_{\Omega} dx,
\]

therefore a subsequence (denoted \( \{\xi_n\} \) also) converging towards a \( \xi \in \Xi \) can be extracted.

From the definition of \( \rho \), and \( \bar{\phi} \),

\[
\varepsilon^\alpha \|\nabla(\phi_n - \bar{\phi})\|^2 \leq \int_{\Omega} \rho_n \nabla(\phi_n - \bar{\phi}) \nabla(\phi_n - \bar{\phi}) \, dx = \int_{\Omega} (\rho_n - \bar{\rho}) \nabla \bar{\phi} \nabla(\phi_n - \bar{\phi}) \, dx
\]

\[
\leq \|\nabla(\phi_n - \bar{\phi})\| \int_{\Omega} (\rho_n - \bar{\rho}) \nabla \bar{\phi} \, dx
\]

But \( \rho_n \to \bar{\rho} \) weakly therefore \( \phi_n \to \bar{\phi} \) strongly in \( H^1(\Omega) \). The functional \( E \) is convex and continuous in \( \xi - \phi \) therefore it is weakly l.s.c. so that
Therefore (5.3) has at least one solution. Since any solution of (3.4)-(3.8) is a solution of 5.3 with \( \phi = \xi \) and \( E(\xi) = 0 \), then \( E(\tilde{\xi}) = 0 \) and \( \tilde{\xi} = \tilde{\phi} \), therefore \( \tilde{\xi} \) (and \( \tilde{\phi} \)) is a solution of (3.4)-(3.7) condition (3.8) can be rewritten: \( \nabla \cdot u < +\infty \), hence \( \Delta \xi < +\infty \).

**Proposition 2**

If \( \xi|_{\Gamma_1} = \phi_1 \quad \delta\xi|_{\Gamma_1} = 0 \), then

\[
E(\xi + \delta\xi) - E(\xi) = 2 \int_{\Omega} \rho(|\nabla \xi|^2) \left(1 - \frac{1}{2} \|\nabla \xi\|^2 + \frac{1}{2} \|\nabla \phi\|^2 \right) \nabla \xi \cdot \nabla \delta \xi \, dx + o(\delta \xi)
\]

\( (\rho' = -\kappa(1-k|\nabla \xi|^2)^{a-1} ) \)

From (5.1) and (5.2)

\[
E(\xi + \delta\xi) - E(\xi) = 2 \int_{\Omega} \left[ 2\rho \nabla \xi \cdot \nabla \delta \xi \nabla (\phi - \xi) \nabla (\phi - \xi) - \rho \nabla \phi \delta \xi - \rho \nabla \phi \delta \xi \right] \nabla \xi \, dx + o(\delta \xi) + o(\delta \phi)
\]

where

\[
\rho' = -\kappa(1-k|\nabla \xi|^2)^{a-1}
\]

From (5.3)

\[
\int_{\Omega} \rho \nabla \phi \nabla w \, dx = -\int_{\Omega} 2\rho \nabla \phi \nabla \delta \xi \nabla \phi \nabla w + o(\delta \xi) \quad \forall w \in H^1_{01}(\Omega)
\]

and since \( \rho(|\nabla (\xi + \delta \xi)|^2) \) is bounded from below by a positive number, there exists \( K \) such that \( \|\nabla \phi\| < K \|\nabla \delta \xi\| \)

Therefore, by letting \( w = \phi - \xi \) in (5.8), (5.7) becomes

\[
\delta E = -2 \int_{\Omega} \left[ \rho \nabla (\phi - \xi) \cdot \nabla \delta \xi + \rho \left(|\nabla \phi|^2 - |\nabla \xi|^2\right) \nabla \xi \cdot \nabla \delta \xi \right] \, dx
\]

and from (5.2) the term \( \rho \nabla \phi \nabla \delta \xi \) disappears.
Corollary 1

If $\xi$, $\phi$ is a stationary point of $E$, it satisfies:

\begin{align}
(5.9) \quad \nabla \cdot \bar{\rho}(1 + \frac{\bar{M}^2}{2} (1 - |\nabla \bar{\xi}|^2 |\nabla \bar{\phi}|^{-2}) \nabla \bar{\xi}) &= 0 \text{ in } \Omega \\
(5.10) \quad \bar{\rho} (1 + \frac{\bar{M}^2}{2} (1 - |\nabla \bar{\xi}|^2 |\nabla \bar{\phi}|^{-2}) \frac{\partial \bar{\xi}}{\partial n} |_{\Gamma_2} &= 0 ; \bar{\xi} |_{\Gamma_1} = \phi_1
\end{align}

Remark: It should be noted that in most cases (5.3) has no other stationary point than the solutions of (3.4)-(3.7). Indeed let $(x_{\xi}, y_{\xi}, z_{\xi})$ be a curvilinear system of coordinate such that

$$\nabla \xi = (\frac{\partial \xi}{\partial x_{\xi}}, 0, 0)$$

Then, from (5.9), (5.10)

\begin{align}
(5.11) \quad \frac{\partial}{\partial x_{\xi}} \left[ \bar{\rho}(1 + \frac{\bar{M}^2}{2} (1 - |\nabla \bar{\xi}|^2 |\nabla \bar{\phi}|^{-2}) \frac{\partial \bar{\xi}}{\partial x_{\xi}} \right] &= 0, \frac{\partial \bar{\xi}}{\partial n} |_{\Gamma_2} = 0 \\
&\text{or } \bar{M}^2(1 - |\nabla \bar{\xi}|^2 |\nabla \bar{\phi}|^{-2}) |_{\Gamma_2} = -2, \bar{\xi} |_{\Gamma_1} = \phi_1
\end{align}

This system looks like the one dimensional transonic equation for a compressible fluid with density

$$\bar{\rho} (1 + \frac{\bar{M}^2}{2} (1 - |\nabla \bar{\xi}|^2 |\nabla \bar{\phi}|^{-2}))$$

Therefore, if the $\xi$-stream lines meet two boundaries and $\Delta \xi < +\infty$ at the shocks and

$$1 + \frac{\bar{M}^2}{2} (1 - |\nabla \bar{\xi}|^2 |\nabla \bar{\phi}|^{-2}) > 0$$

then $\bar{\phi} = \bar{\xi}$.

6. DISCRETIZATION AND NUMERICAL SOLUTIONS

Let $\mathcal{C}_h$ be a set of triangles or tetraedra of $\Omega$ where $h$ is the length of the greatest side. Suppose that

$$\bigcup_{T \in \mathcal{C}_h} T \subset \Omega , T_1 \cap T_2 = \varnothing \text{ or a vertex } \forall T_1, T_2 \in \mathcal{C}_h,$$

Let $\Omega_h = \bigcup_{T \in \mathcal{C}_h} T$ and $\Gamma_{1h}, \Gamma_{2h}$ parts of $\partial \Omega_h$ which approximate $\Gamma_1$ and $\Gamma_2$. 

Let \( H_h \) an approximation of \( H^1(\Omega) \):

\[
(6.1) \quad H_h = \{ w_h \in C^0(\Omega_h) | w_h \text{ linear on } T \forall T \in \mathcal{T}_h \}
\]

Note that any element of \( H_h \) is completely determined by the values that it takes at the nodes of \( \mathcal{T}_h \). Therefore if we assume that \( \mathcal{T}_h \) has \( N = n+p+m \) nodes \( P_i \) with \( P_i \in \Gamma_1 \) if \( i > n+p \), \( P_i \in \Gamma_2 \) if \( i \in [n,n+p] \), and if we define \( w_i \in H_h \) by

\[
(6.2) \quad w_i = 1 \text{ at node } i \text{ and zero at all other nodes}
\]

Then any function \( w \in H_h \) is written as

\[
(6.3) \quad \phi = \sum_{i=1}^{N} \xi^i w_i
\]

**Algorithms 1**

Let \( \xi_h = \sum_{i=1}^{N} \xi^i w_i \), then (5.2) becomes

\[
\int_{\Omega} (1-k|\nabla \xi_h|^2) \nabla \phi_h \cdot \nabla w_i \, dx = 0 \quad i=1,\ldots,n+p
\]

\[
(6.4) \quad \phi_h = \sum_{i=1}^{n+p} \phi^i w_i + \sum_{n+p+1}^{N} \phi^i w_i
\]

and (5.6) becomes

\[
(6.5) \quad \frac{1}{2} \delta E_h = \sum_{i=1}^{N} \delta \xi^i \delta E^i_h + o(\delta \xi^i)
\]

\[
(6.6) \quad \delta E^i_h = \int_{\Omega} [\rho - \rho'(|\nabla \phi_h|^2 - |\nabla \phi|^2)] \nabla \xi_h \cdot \nabla w_i \, dx
\]

Consider the following algorithm

**Step 0** Choose \( \mathcal{T}_h, \xi_{ho} \) set \( j=0 \)

**Step 1** Compute \( \phi_{hj} \) by solving (6.4) with \( \xi_h = \xi_{hj} \)

**Step 2** Compute \( \{ \delta E^i_{hj}, i = 1,\ldots,N \} \) by (6.6)

**Step 3** Compute \( \delta \xi_h = \sum_{i=1}^{N} \delta \xi^i w_i \) by solving
(6.7) \[ \int_{\Omega_h} \nabla \delta \xi_h \nabla \omega_h \, dx = \delta \xi_{hj}^i, \quad i=1, \ldots, n+p \]

**Step 4** Compute an approximation \( \tilde{\lambda}_j \) of the solution of

(6.8) \[ \min_{\lambda \in [0,1]} \int_{\Omega_h} \rho(\lambda) \left| \nabla (\xi_h(\lambda) - \phi_h(\lambda)) \right|^2 \, dx \]

where

\[ \xi_h(\lambda) = \sum_{i=1}^{N} (\xi_{hj}^i - \lambda \delta \xi_{hj}^i) \omega_i \]

**Step 5** Set

\[ \xi_{hj}^{j+1} = \xi_h(\tilde{\lambda}_j), \quad j=j+1 \]

and go to step 1.

**Proposition 3**

Let \( \{\xi_{hj}^i\}_{j \geq 0} \) be a sequence generated by algorithm 1 such that

\[ |\nabla \xi_{hj}(x)| \leq k^{-1/2} \quad \forall x, \forall j. \]

Every accumulation point of \( \{\xi_{hj}^i\}_{j \geq 0} \) is a stationary point of the functional

(6.9) \[ E_h(\xi_h) = \int_{\Omega_h} \left| \nabla (\phi_h - \xi_h) \right|^2 \, dx \]

Where \( \phi_h = \phi_h(\xi_h) \) is the solution of (6.4), in

\[ \Xi_h = \{ \xi_h \in \mathcal{H}_h \mid |\nabla \xi_h(x)| \leq k^{-1/2} \quad \forall x \in \Omega_h \} \]

**Proof**

Algorithm 1 is the method of steepest descent applied to minimize (6.9) in \( \Xi_h \), with the norm

(6.10) \[ \|\xi_h\|_h^2 = \int_{\Omega_h} \nabla \xi_h \nabla \xi_h \, dx \]

Therefore \( \{E_h(\xi_{hj})\}_j \) decreases until \( \delta E_{hj} \) reaches zero.

**Remark 6.1** : (6.4) should be solved by a method of relaxation but (6.7) can be factorized once and for all by the method of Choleski.

**Remark 6.2** : Problem (6.8) is usually solved by a Golden section search or a Newton method.
Remark 6.3: Step 5 can be modified so as to obtain a conjugate gradient method.

Remark 6.4: The restriction: \( |u_{h_j}(x)| \leq k^{-1/2} \) in theorem 5.1 is not a problem if \( u \) is not too close to \( k^{-1/2} \) otherwise one must treat this restriction as a constraint in the algorithm. Also, even though theorem (5.1) ensures the computation of stationary points only, it is a common experience that global minima can be obtained by this procedure if there is a finite number of local minima.

Remark 6.5: The entropy condition \( \Delta \xi_h < +\infty \) can be taken into account numerically. Let \( M(x) \) be a real valued function then \( \Delta \xi_h \leq M(x) \) becomes, from (6.7)

\[
-\sum_{j} \lambda_j \delta E_{h_j}^i \leq M(x_i) \quad i=1,\ldots,n+p
\]

Therefore, to satisfy (6.11) at iteration \( j+1 \), it suffices to take \( \delta E_{h_j}^i = 0 \) in (6.7) for all \( i \) such that (6.11) at iteration \( j \) is an equality. This procedure amounts to control \( \omega = \Delta \xi \) instead of \( \xi \).

7. NUMERICAL RESULTS

The method was tested on a nozzle discretized as shown on figure 1, (300 triangular elements, 180 nodes). The Polak-Ribiere method of conjugate gradient was used with an initial control: \( \Delta \xi = 0 \) (incompressible flow). A mono-dimensional optimization subroutine based on a dichotomic search was given to us by Lemarechal. Several boundary conditions were tested

1°) Subsonic mach number \( M_{\infty} = 0.63 \) at the entrance, zero potential on exit, the method had already converged in 10 iterations (to be compared with the Gelder-Periaux method) giving a criterium \( E_{h10} = 2 \times 10^{-13} \) (\( E_{h0} = 10^{-4} \)).
2°) Entrance and exit potential specified.
   For a decrease of potential of $\phi_1 - \phi_2 = 0.7$ the method had converged in 20 iterations without including the entropy condition, giving a criterium of $E_{h20} = 5 \times 10^{-7}$, the results are shown on figure 2.

3°) Supersonic mach number $M_\infty = 1.25$ at the entrance.
   The method computes a solution that has a shock at the first section of discretization. An other boundary condition must be added. One iteration of the method takes 3" on an IBM 370/158 on this example.
   A three dimensional nozzle is being tested: the result will be shown at the conference. 20 to 40 iterations are usually sufficient for the algorithm to converge. The results are in good agreement with the tabulated data. Simple and multi-bodies airfoils are also being tested. For them it is necessary to include the entropy condition; 80 iterations are usually more than sufficient for the convergence.

8. CONCLUSIONS
   Thus this method seems very promising. It compares very well with the finite differences method available and it has the advantage of allowing complicate two and three dimensional geometries. This work illustrates the fact that optimal control theory is a powerful tool with unexpected applications sometimes.

ACKNOWLEDGEMENT
   We wish to thank M. Periaux, Perrier and Poirier for allowing us to use their data files and computer and for their valuable comments.

REFERENCES

   GELDER D. - Solution of the compressible flow equation. Int. J. on Num. Meth. in Eng., Vol. 3 pp. 35-43 (1971)


