D E P A R T M E N T   O F   M A T H E M A T I C S

Numerical Approximation of the Legendre Transformation

M.J. Baines

Numerical Analysis Report 15/88

U N I V E R S I T Y   O F   M A T H E M A T I C S

Numerical Approximation of the Legendre Transformation

M.J. Baines

Numerical Analysis Report 15/88

Department of Mathematics
University of Reading
P O Box 220
Whiteknights
Reading, U.K.

Abstract

It is proved that an approximate residual minimisation provides approximations to the dual functions appearing in the Legendre transformation. A technique is described for the construction of such approximations using the Moving Finite Element method.

§0. Introduction

The result contained in this report does two things. First, it provides a mechanism for constructing numerical (particularly finite element) approximations to the dual functions involved in the Legendre transformation in which there has been much recent interest [1,2,3]. Secondly, it reveals that the numerical method known as the Moving Finite Element (MFE) method [4,5,6,7], which is an example of an approximation to an envelope construction, is also a Legendre transformation.

The central idea is to seek the best approximation in a finite subspace to the Legendre dual of a given function by a double minimisation of the residual of the symmetric relation between these functions, in a least squares sense. The structure of the Legendre transformation, together with its envelope property (see[1]), is approximately reproduced. The procedure generates numerical approximations to the dual functions occuring in any Legendre transformation.

The result offers the opportunity of exploiting numerically the wealth of qualitative detail available on the Legendre transformation in reference [2]. Moreover, it demonstrates that the solution processes already developed for the MFE method may be used for generating the approximations.

The theorem and proof are first given here for the one-dimensional case but may easily be extended to any finite number of dimensions.

<u>Theorem</u> Let X, Y, x, y be scalar C^2 functions with Y a given function of y. Let S_n, T_n be (ordered) finite dimensional spaces capable of approximating C^2 functions arbitrarily closely pointwise as $n \to \infty$.

Define the sequences $X_o^{(n)}$, $x_o^{(n)} \in S_n$ of approximations to X, x by the problem:

Given $y^{(n)} \in T_n$, find the extremals $X_o^{(n)}$, $x_o^{(n)} \in S_n$ of the functional

$$I^{(n)} = ||X^{(n)}+Y(y^{(n)}) - x^{(n)}y^{(n)}||_{L_2}^2 . \qquad (0.1)$$

Then, if $Y(y^{(n)}) \to Y(y)$ as $n \to \infty$, the sequence $x_o^{(n)}$ approaches $Y'(y)$ and $X_o^{(n)}$ approaches $y\,Y'(y) - Y(y)$ provided that $T_n \subseteq S_n \cup (T_n \otimes S_n)$. The two sequences determine an arbitrarily close approximation to the envelope of the family of straight lines

$$X + Y(y) - xy = 0 \qquad (0.2)$$

as y varies.

<u>Proof</u>

The proof is in five parts. In the first part the continuous problem is studied from the variational point of view, while in the second part this description is repeated using finite subspaces. Limits of the solution as the spaces become richer are discussed in part 3. Part 4 is the main part connecting these limits with the envelope construction. In part 5 the results are summarised.

Part 6-8 are concerned with the numerical construction of approximations and the Moving Finite Element method.

§1. Proof (1) Consider stationary values of the L_2 norm (squared)

$$I = ||X + Y - xy||_2^2 \tag{1.1}$$

with first variation

$$\delta I = 2 \langle X + Y - xy, \delta X + \delta Y - x\delta y - y\delta x \rangle \tag{1.2}$$

Constrain Y to be a given function of y : then

$$\delta I = 2 \langle X + Y - xy, \delta X - y\delta x + (Y' - x)\delta y \rangle \tag{1.3}$$

For each y the conditions for a stationary value of I are

$$\langle X + Y - xy, \delta X \rangle = 0 \tag{1.4}$$

$$\langle X + Y - xy, y\delta x \rangle = 0 \tag{1.5}$$

for arbitrary $\delta X, \delta x$; i.e.

$$X + Y - xy = 0 \tag{1.6}$$

$$-y(X + Y - xy) = 0 . \tag{1.7}$$

These equations are satisfied by any pair x_o, X_o lying on the straight line (1.6) in x,X space , i.e.

$$X_o + Y - x_o y = 0 \ . \tag{1.8}$$

The stationary value I_o of I is evidently zero and, since $I \geq 0$, I_o must be a minimum. Moreover I_o is zero for any y so that δI is also zero with respect to variations in y at $I = I_o$. From (1.3)

$$\langle X_o + Y - x_o y,\ (Y' - x)\delta y \rangle = 0 \tag{1.9}$$

for arbitrary δy ; i.e.

$$(X_o + Y - x_o y)\ (Y - x_o) = 0 \ . \tag{1.10}$$

§2. Proof (2)

Now restrict the variations of x, X to a finite subspace S_n with the Weierstrass property of approximating continuous (and hence C^2) functions as closely as desired by taking n large enough. For example S_n may be the set of all polynomials of degree n, all piecewise polynomials on a partition of n points, or all finite Fourier expansions of degree n .

Again, for fixed y as in (1.3), I is stationary if the (finite-dimensional) extremals $x_o^{(n)}$, $X_o^{(n)}$ satisfy

$$\langle X_o^{(n)} - yx_o^{(n)} + Y(y),\ \delta X^{(n)} \rangle = 0 \tag{2.1}$$

$$\langle X_o^{(n)} - yx_o^{(n)} + Y(y),\ - y\ \delta x^{(n)} \rangle = 0 \tag{2.2}$$

$\forall\ \delta x^{(n)},\ \delta X^{(n)} \epsilon S_n$. For a given finite subspace these equations determine unique $x_o^{(n)}$, $X_o^{(n)}$, in general. Although the corresponding stationary value $I_o^{(n)}$ is not (always) zero we note from (2.1) that

$$I_o^{(n)} = ||X_o^{(n)} - yx_o^{(n)} + Y(y)||_2^2 \tag{2.3}$$

$$= \langle X_o^{(n)} - yx_o^{(n)} + Y(y),\ [X_o^{(n)} - yx_o^{(n)} + Y(y)]_\perp \rangle \tag{2.4}$$

where the square bracket lies in $S_n^\perp$, the complement of S_n, and $\to 0$ as $n \to \infty$. It follows that $I_0^{(n)} \to 0$ as $n \to \infty$, as expected, and, from (2.3), that

$$||X_o^{(n)} - yx_o^{(n)} + Y(y)||_2 \to 0 \quad \text{as} \quad n \to \infty \tag{2.5}$$

$\forall y$: i.e., given $\epsilon > 0$, $\exists N(\epsilon)$ such that, for $n > N$

$$||X_o^{(n)} - yx_o^{(n)} + Y(y)||_2 < \epsilon \qquad \forall y \tag{2.6}$$

§3. Proof (3)

We now discuss the limits of $X_o^{(n)}$, $x_o^{(n)}$ as $n \to \infty$. Writing down (2.1), (2.2) for different integers n, m, with $S_n \subset S_m$ and subtracting, we obtain

$$\langle X_o^{(n)} - X_o^{(m)} - y(x_o^{(n)} - x_o^{(m)}),\ X^{(n)} \rangle = 0 \tag{3.1}$$

$$\langle -y(X_o^{(n)} - X_o^{(m)}) + y^2\ (x_o^{(n)} - x_o^{(m)}),\ X^{(n)} \rangle = 0\ . \tag{3.2}$$

It follows that the first factor in each case lies in $S_n^{\perp}$ and $\to 0$ as $n \to \infty$: moreover in general, given $\epsilon_1 > 0$,$\exists\ N_1(\epsilon_1)$ such that (provided $y \neq 0$) if $n > N_1$,

$$||X_o^{(n)} - X_o^{(m)}||_2 < \epsilon_1 \qquad ||x_o^{(n)} - x_o^{(m)}||_2 < \epsilon_1 \tag{3.3}$$

and hence both sequences $X_o^{(n)}$, $x_o^{(n)}$ converge in the Cauchy sense. There exist limits $X_o^{(\infty)}$, $x_o^{(\infty)}$ (not necessarily unique) with the property that, given $\epsilon_2 > 0$, $\exists\ N_2(\epsilon_2)$ such that, for $n > N_2$,

$$||X_o^{(n)} - X_o^{(\infty)}||_2 < \epsilon_2 \qquad ||x_o^{(n)} - x_o^{(\infty)}||_2 < \epsilon_2 \tag{3.4}$$

Since $I_o^{(n)} \to I_o = 0$ as $n \to \infty$, we have

$$X_o^{(\infty)} - yx_o^{(\infty)} + Y(y) = 0 . \tag{3.5}$$

§4. Proof (4)

Now consider variations of the residual

$$R_o^{(n)}(y) = X_o^{(n)} - yx_o^{(n)} + Y(y) \tag{4.1}$$

with y, regarding $X_o^{(n)}$, $x_o^{(n)}$ as functions of y . We have

$$R_o^{(n)}(y+\delta y) - R_o^{(n)}(y) = X_o^{(n)}(y+\delta y) - X_o^{(n)}(y)$$

$$- y\{x_o^{(n)}(y+\delta y) - x_o^{(n)}(y)\} + Y(y+\delta y) - Y(y) \tag{4.2}$$

$$= \left[\frac{\partial X_o^{(n)}}{\partial y} - y \frac{\partial x_o^{(n)}}{\partial y} - x_o^{(n)} + Y'(y) \right]_{y+\theta\delta y} \delta y \tag{4.3}$$

by a Mean Value Theorem, where $0 < \theta < 1$.

Taking the inner product with $R_o^{(n)}(y+\theta\delta y)$, we obtain

$$< R_o^{(n)}(y+\theta\delta y), \{R_o^{(n)}(y+\delta y) - R_o^{(n)}(y)\} >$$

$$= < R_o^{(n)}(y+\theta\delta y), \left[\frac{\partial X_o^{(n)}}{\partial y} - y\frac{\partial x_o^{(n)}}{\partial y} - x_o^{(n)} + Y'(y)\right]_{y+\theta\delta y} \delta y > \quad (4.4)$$

Returning to (2.1), (2.2), we make the choices

$$\delta X^{(n)} = \frac{\partial X_o^{(n)}}{\partial y}\delta y^{(n)} \qquad \delta x^{(n)} = \frac{\partial X_o^{(n)}}{\partial y}\delta y^{(n)} \quad (4.5)$$

assumed to lie in S_n . For this purpose we must restrict y to lie in an associated space T_n (see below). It then follows that the first two terms in the square bracket on the right hand side of (4.4) vanish, so that (if $y = y^{(n)} \epsilon T_n$)

$$< R_o^{(n)}(y^{(n)}+\theta^{(n)}\delta y^{(n)}), \{R_o^{(n)}(y^{(n)}+\delta y^{(n)}) - R_o^{(n)}(y^{(n)})\}$$

$$+\left[x_o^{(n)} - Y'(y^{(n)})\right]_{y^{(n)}+\theta^{(n)}\delta y^{(n)}} \delta y^{(n)} = 0 \quad (4.6)$$

Finally choose $\delta y^{(n)} = 0$ except in a neighbourhood $\eta(\xi)$ of a particular point ξ, where $\eta(\xi)$ is chosen sufficiently small that $R_o^{(n)}(y^{(n)}+\theta^{(n)}\delta y^{(n)}) \neq 0$. (If this is not possible, $R_o^{(n)} = 0$ in $\eta(\xi)$, see below). Then the inner products in (4.4) reduce to integrals over $\eta(\xi)$ and, using a Mean Value Theorem for integrals,

$$\left[\left\{R_o^{(n)}(y^{(n)}+\delta y^{(n)}) - R_o^{(n)}(y^{(n)}) + x_o^{(n)} - Y'(y^{(n)})\right\}_{y^{(n)}+\theta^{(n)}\delta y^{(n)}} \delta y^{(n)}\right]_{\xi^*} .$$

$$. \int R_o^{(n)}(y^{(n)}+\theta\delta y^{(n)})d\xi = 0 \quad (4.7)$$

for some $\xi^* \epsilon\ \eta(\xi)$. Without loss of generality we may take $\eta(\xi)$ to shrink to zero with $\delta y(n)$, i.e. $\exists$ a variable $\phi^{(n)}$ such that $\xi^* = \xi + \phi^{(n)}\delta y^{(n)}$. Hence, if $\int_{\eta(\xi)} R_o^{(n)} d\xi \neq 0$,

$$\left[x_o^{(n)} - Y(y^{(n)})\right]_{\substack{y^{(n)}+\theta^{(n)}\delta y^{(n)} \\ \xi+\phi^{(n)}\delta y^{(n)}}} = -\frac{1}{\delta y^{(n)}}\left[R_o^{(n)}\left[y^{(n)} + \delta y^{(n)}\right] - R_o^{(n)}(y)\right]_{\xi+\phi^{(n)}\delta y^{(n)}} \tag{4.8}$$

and, from (2.6) and (4.1), if $n > N$,

$$\left|\left|x_o^{(n)} - Y'(y^{(n)})\right|\right|_{\substack{y^{(n)}+\theta^{(n)}\delta y^{(n)} \\ \xi+\phi^{(n)}\delta y^{(n)}}} < \frac{2\epsilon}{|\delta y^{(n)}|} . \tag{4.9}$$

Therefore, given $\epsilon^* > 0$, $\exists$ constants γ^*, $\delta^*(\epsilon^*)$ and k such that, for $0 < \gamma^* \leq |\delta y^{(n)}| < \delta^*$ and $n > N$ chosen such that $\epsilon = \frac{1}{2}\epsilon^*\gamma^*$,

$$\left|\left|x_o^{(n)} - Y'(y^{(n)})\right|\right|_{\substack{y^{(n)}+\theta^{(n)}\delta y^{(n)} \\ \xi+\phi^{(n)}\delta y^{(n)}}} < \epsilon^* \tag{4.10}$$

for some $\theta^{(n)}$, $\phi^{(n)}$.

Here $x_o^{(n)} \epsilon S_n$ and $y^{(n)} \epsilon T_n$. From (4.5) we can relate the spaces S_n, T_n . Recall that we used (2.1), (2.2) to eliminate two of the terms of (4.4). In fact it is easier to see the required relationship from (4.2), where the corresponding terms vanish by (2.1), (2.2), at least to

$O(\delta y)^2$, if

$$T_n \subseteq S_n \cup (T_n \otimes S_n) \quad (4.11)$$

Thus $T_n \equiv S_n$ is sufficient, but so for example is $T_n = \mathcal{P}_k$, $S_n = \mathcal{P}_\ell$ (k, ℓ any integers) where $\mathcal{P}_n$ is the space of polynomials of degree n, so that T_n can have more or less continuity than S_n. An important case is $S_n \equiv$ piecewise linear functions, $T_n \equiv$ piecewise constant functions, on a partition of the ξ interval, as we shall see below. This corresponds, however, to a special case of the proof which we now discuss.

If it is not possible to choose $\eta(\xi)$ sufficiently small that $R_o^{(n)} \neq 0$ or, what comes to the same thing,

$$\int_{\eta(\xi)} R_o^{(n)} \, d\xi = 0 \quad (4.12)$$

just before (4.8), where the argument $y^{(n)}+\theta^{(n)}\delta y^{(n)}$ of $R_o^{(n)}$ is understood, it follows that

$$R_o^{(n)} = X_o^{(n)} - yx_o^{(n)} + Y(y^{(n)}) \equiv 0 \quad (4.13)$$

in a neighbourhood of ξ. (We thus apparently have an exact Legendre transformation but in this case

$$x^{(n)} \neq Y'(y^{(n)}) \quad (4.14)$$

in general since they do not belong to the same space.)

Variations of $R_o^{(n)}$ within the neighbourhood are also zero. In particular, considering variations with respect to $X_o^{(n)}$ and $x_o^{(n)}$, we have from (4.1) that

$$\delta X_o^{(n)} = 0 = \delta x_o^{(n)} . \tag{4.15}$$

Now take variations of $R_o^{(n)}$ with respect to y at $y^{(n)} + \theta^{(n)}\delta y^{(n)}$ regarding $X_o^{(n)}$, $x_o^{(n)}$ as functions of y. We have (c.f. (4.3)) that in the neighbourhood $\eta(\xi)$

$$\left[\frac{\partial X_o^{(n)}}{\partial y} - \frac{\partial x_o^{(n)}}{\partial y} - x_o^{(n)} + Y'(y)\right]_{y=y^{(n)}+\theta^{(n)}\delta y^{(n)}} \equiv 0 \tag{4.16}$$

Finally, suppose that (4.10) does not hold, i.e.

$$\left|\left|x_o^{(n)} - Y'(y^{(n)})\right|\right|_{y^{(n)}+\theta^{(n)}\delta y^{(n)}} \tag{4.17}$$

is bounded away from zero for some $\xi \in \eta(\xi)$ and some $n > N$. Then, from (4.16),

$$\left|\left|\frac{\partial X_o^{(n)}}{\partial y} - \frac{\partial x_o^{(n)}}{\partial y}\right|\right|_{y^{(n)}+\theta^{(n)}\delta y^{(n)}} . \tag{4.18}$$

is also bounded away from zero for this ξ and n. Thus with $X_o^{(n)}$, $x_o^{(n)} \in S_n$, $y^{(n)} \in T_n$ so that variations

$$0 = \delta X_o^{(n)} = \frac{\partial X_o^{(n)}}{\partial y}\,\delta y , \quad 0 = \delta x_o^{(n)} = \frac{\partial x_o^{(n)}}{\partial y}\,\delta y \tag{4.19}$$

are permitted in (4.15), it is seen that (4.18) gives a contradiction. Hence (4.10) holds in this case also.

§5. Proof (5)

We have shown in (3.4) and (4.10) that $x_o^{(n)}$ tends to a limit as $n \to \infty$ and that $x_o^{(n)}$ is arbitrarily close to $Y'(y^{(n)})$. Suppose that the space T_n has the Weierstrass property of infinitely close approximation to continuous (and therefore C^2) functions. Then

$$||x_o^{(\infty)} - Y'(y)|| \le ||x_o^{(\infty)} - x_o^{(n)}|| + ||x_o^{(\infty)} - Y'(y^{(n)})|| + ||Y'(y^{(n)}) - Y'(y)|| \quad (5.1)$$

and, by (3.4), (4.10) and continuity of $Y'(y)$, is less than any given small positive number ϵ_2 , provided that $\delta y^{(n)}$ is sufficiently small and n is sufficiently large, where $x_o^{(n)}$ and $Y'(y^{(n)})$ (or $Y'(y)$ are evaluated at $y^{(n)} + \theta^{(n)}\delta y^{(n)}$ (or $y + \theta\delta y$) and $\xi + \phi^{(n)}\delta y^{(n)}$. Here $0 < \theta < 1$ and $\delta y^{(n)}$ is a small number bounded away from zero.

Hence $x_o^{(\infty)}$ is unique and arbitrarily close to $Y'(y)$. From this result and (3.5) we deduce that

$$x_o^{(\infty)} = Y'(y) \quad , \quad X_o^{(\infty)} = y\, Y'(y) - Y(y) \; . \quad (5.2)$$

It follows that $x_o^{(\infty)}$, $X_o^{(\infty)}$ lie on the envelope of (3.5) as y varies and that there exists a Legendre transformation between the dual functions $Y(y)$ and $X_o^{(\infty)}$, as in the continuous case.

This completes the proof. The extension to many variables is simply stated by replacing x, y in the statement of the Theorem in §0 by x_i, y_i (and using the summation convention) and by replacing the words "straight lines" by "hyperplanes". The proof is then an extension of the above.

We go on now to discuss the numerical construction of the approximations.

§6. Construction of Approximations

We have shown that the extremals $X_o^{(n)}$, $x_o^{(n)} \epsilon S_n$ of the functional

$$||X^n + Y(y^{(n)}) - x^{(n)} y^{(n)}||_{L_2} \tag{6.1}$$

approximate the envelope, as y varies, of

$$X + Y(y) - xy = 0 . \tag{6.2}$$

Since the equation of the envelope

$$x = \frac{\partial Y}{\partial y} \tag{6.3}$$

induces a Legendre transformation between $Y(y)$ and $X(x)$ (see [1]) the extremals of (6.1) can be used to furnish approximations to the functions appearing in any Legendre transformation [2] through the conditions (2.1), (2.2).

A standard approach is that of finite elements. The functions $X^{(n)}$, $x^{(n)}$ are expanded in terms of a finite number of basis functions ϕ_j (spanning S_n) and the function $y^{(n)}$ is expanded in terms of a finite number of basis functions ψ_k (spanning $T^{(n)}$). Then (2.1), (2.2) imply that

$$< \Sigma X_j^{(n)} \phi_j - \Sigma y^{(n)} x_j^{(n)} \phi_j + Y(y^{(n)}) \ , \ \phi_i > \tag{6.4}$$

$$< \Sigma X_j^{(n)} \phi_j - \Sigma y^{(n)} x_j^{(n)} \phi_j + Y(y^{(n)}) \ , -y^{(n)} \phi_i > \tag{6.5}$$

(orthogonality with respect to $\delta X^{(n)}$ or $\delta x^{(n)}$ being equivalent to orthogonality with respect to ϕ_i). The suffix o indicating extremals has been dropped. When the basis functions are of compact support (6.4) and (6.5) lead to sparse matrix equations for the unknown coefficients X_j, x_j , for each $y^{(n)} \epsilon \, T_n$.

In the particular case when S_n is the set of piecewise linear functions and T_n is the space of piecewise constant functions on a partition of an interval, equation (6.2) can be exactly satisfied although (6.3) is only satisfied approximately. The solutions for $x^{(n)}$, $X^{(n)}$ in this case (see [6]) are the piecewise linear functions with nodal intersections

$$\frac{Y(y_R^{(n)})-Y(y_L^{(n)})}{y_R^{(n)}-y_L^{(n)}} , \qquad \frac{y_R^{(n)}Y(y_L^{(n)})-y_L^{(n)}Y(y_R^{(n)})}{y_R^{(n)}- y_L^{(n)}} , \tag{6.6}$$

respectively, where $y_L^{(n)}$ and $y_R^{(n)}$ are the piecewise constant values of y associated with elements to the left and right of each node.

§7. The Moving Finite Element Method

A finite element method with this structure is the Moving Finite

Element (MFE) method of K. Miller [4] in the form studied by Mueller and Carey [5] and by Baines and Wathen [6]. This is a method of approximately solving the time-dependent pde

$$v_t = Lv \tag{7.1}$$

where Lv is an operator incorporating space derivatives but not time, by writing it in a frame moving with speed $\dot{x}$ as

$$\dot{v} - v_x\dot{x} - Lv = 0 \tag{7.2}$$

and obtaining $\dot{v}$, $\dot{x}$ by minimising the functional

$$J = ||\dot{v} - v_x\dot{x} - Lv||_{L_2} \tag{7.3}$$

over $\dot{v}$, $\dot{x}$ in an approximation space. The unknown functions $\dot{v}$, $\dot{x}$ (or v, x) are expanded in piecewise linear basis functions while the function v_x is expanded (consistently) in piecewise constant basis functions.

If Lv is a function of v_x only, i.e.

$$\dot{v} - v_x\dot{x} - Lv = \dot{v} - v_x\dot{x} + f(v_x) = 0 \tag{7.4}$$

say, there is a precise correspondence between (7.2) and the form

$$X_o^{(n)} - y^{(n)}x_o^{(n)} + Y(y^{(n)}) = 0 \tag{7.5}$$

of previous sections, with

$$X_o = \dot{v} \ , \quad x_o = \dot{x} \ , \quad y = u_x \ , \quad Y = - f(v_x) \ . \qquad (7.6)$$

S_n, T_n are the spaces of piecewise linear, piecewise constant functions, respectively, so that (4.11) is satisfied.

In this case (7.2) can be satisfied exactly [6],[7] as in equation (6.6). Hence the MFE method is a Legendre transformation between the dual functions $Y(y^{(n)}) = f(v_x)$ and $X_o(x_o) = \dot{v}(\dot{x})$ and, by the earlier theorem, $\dot{v}(\dot{x})$ approximates the envelope of (7.4) as v_x varies, i.e. $\dot{x}$ approximates $\partial f/\partial v_x$.

Moreover, a numerical procedure has already been developed [6] in the MFE method for the solution of (6.4), (6.5), which is available for the construction of approximations to the Legendre transformation. Briefly, (6.4), (6.5) are written in the matrix form

$$A(y^{(n)}) \, \underline{x} = \underline{b} \qquad (7.7)$$

where $\underline{x}^T = (\ldots, X_j^{(n)}, x_j^{(n)}, \ldots)$, $\underline{b} = \{b_i\}$, $A = \{A_{ij}\}$.

$$b_{2k-1} = \langle Y(y^{(n)}) \ , \phi_{2k-1} \rangle$$

$$b_{2k} = \langle Y(y^{(n)}), \phi_{2k} \rangle$$

$$A_{ij} = \begin{bmatrix} \langle \phi_j, \phi_i \rangle & \langle -y^{(n)}\phi_j, \phi_i \rangle \\ \langle - y^{(n)}\phi_j, \phi_i \rangle & \langle (y^{(n)})^2\phi_j, \phi_i \rangle \end{bmatrix}$$

In the usual MFE method $S_n \equiv \text{span} \{\phi_j\}$ = piecewise linear functions and $T_n = \text{span} \{\psi_k\}$ = piecewise constant functions. In this case the matrix A has a simple decomposition (equivalent to a double assembly) which allows fast inversion by the pre-conditioned conjugate gradient technique [5],[6],[8]. The resulting approximations

$$X^{(n)} = \Sigma X_j^{(n)} \phi_j \qquad x^{(n)} = \Sigma x_j^{(n)} \phi_j \tag{7.8}$$

provide the Legendre dual of the function

$$Y(\Sigma y_k^{(n)} \psi_k) . \tag{7.9}$$

We observe that for Lv of the form $f(v_x)$ and for these spaces the MFE method is an example of a Legendre transformation. For other spaces the same approach may be used although the algebraic techniques will be generally less simple [6],[9].

For other forms of Lv we can say the following (c.f. [7]). In the case

$$Lv = - f(v,v_x) \tag{7.10}$$

it can be shown that the same structure holds and that again the MFE method may be regarded as a Legendre transformation but using a related variable. In the original variable, however, it is only an approximation since in this case equation (7.5) can no longer be satisfied exactly. The same applies to

$$Lv = - f(x, v_x) \quad \text{or} \quad Lv = - f(x,v,v_x) . \tag{7.11}$$

In addition ξ may be introduced as a passive variable [1].

Finally, for

$$Lv = v_{xx} \tag{7.12}$$

it can be argued that an envelope structure exists in a particular sense (see [7]) and hence also a Legendre structure. More specialised and detailed analysis is required in this case, however.

§8. Examples

(i) Consider

$$- L_v = f(v_x) = \frac{v_x^2}{v_x^2+½(1-v_x)^2} \tag{8.1}$$

which corresponds to the equation

$$v_t + \frac{v_x^2}{v_x^2+½(1-v_x)^2} = 0 \tag{8.2}$$

whose x derivative is the Buckley-Leverett equation in the variable $u = v_x$, namely,

$$u_t + \left[\frac{u^2}{u^2+½(1-u)^2}\right]_x = 0 \tag{8.3}$$

In the notation of the earlier sections (8.2) corresponds to

$$Y(y) = \frac{y^2}{y^2+½(1-y)^2} \tag{8.4}$$

and the Legendre transformation yields

$$x = Y' = \frac{y(1-y)}{[y^2+\frac{1}{2}(1-y)^2]^2} \tag{8.5}$$

$$X = \frac{1}{2}\frac{y^2\ (1-3y^2)}{[y^2+\frac{1}{2}(1-y)^2]^2} \tag{8.6}$$

Take S_n, T_n to be the spaces of piecewise linear, constant functions, respectively. Then, according to the theorem of §1, minimisation of (0.1) over $X^{(n)}$, $x^{(n)}$ in the finite space S_n (or setting (0.1) to zero) provides the approximate envelope corresponding to (8.5). In the present example $y^{(n)}$ lies in T_n and equation (6.2) can be satisfied exactly (as in equation (6.6)). Given a partition, it is therefore straightforward to calculate the "approximation". (Note that the approximation is only to the envelope equations, because the Legendre transformation is exact.

Figure 1 shows the y, Y and x, X curves in the continuous (infinite dimensional) case and the corresponding finite dimensional approximations. The most distinctive feature is the cusp in the x, X plane. It shows that x, which is the velocity $\dot{x}$ in the original problem, increases to a maximum and then decreases. This is associated with the formation of a half-shock (a compression abutted with an expansion). The variable X is the time rate of change of v.

In the finite dimensional case the constant values of $Y(y^{(n)})$ in T_n are obtained by sampling points on the y, Y curve: as a result the dual in S_n samples the tangents enveloping the x, X curve. This is in fact a Legendre transformation in itself. Alternatively, we could have obtained the constant values of $Y(y^{(n)})$ by a least squares projection.

In that case we do not expect the projection points to lie on the y, Y curve nor the lines in the dual space to be tangents to the x, X curve.

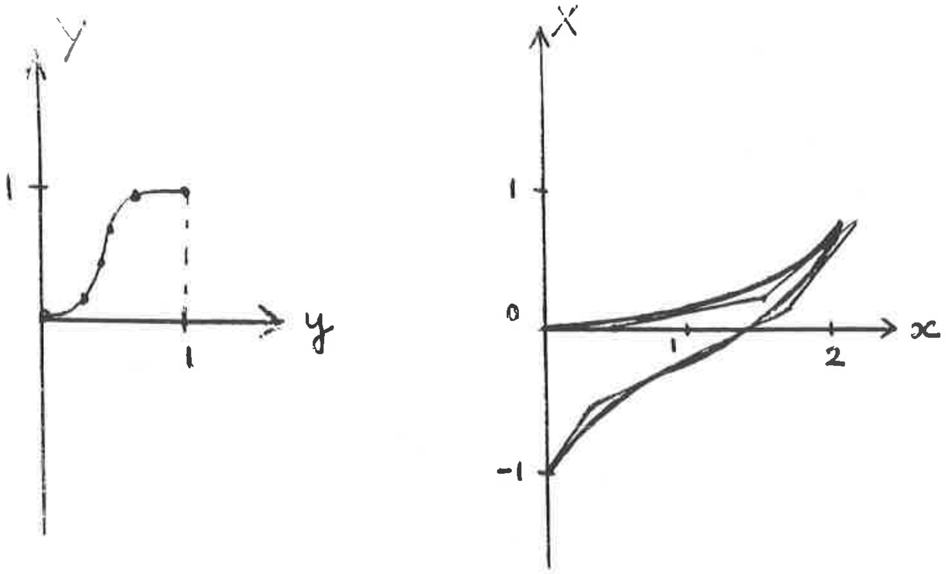

Figure 1.

(ii) The second example has more generality. It is the general form of (8.3), namely,

$$u_t + f(u)_x = 0 \; . \tag{8.7}$$

This time u is the object function while the approximation spaces S_n, T_n are as before. The transformed left hand side of (8.7),

$$\dot{u} - u_x\dot{x} + f(u)_x \tag{8.8}$$

has a non-zero residual when $\dot{u} \sim \dot{U} \in S_n$, $\dot{x} \sim \dot{X} \in S_n$, $u_x \sim U_x \in T_n$, and the least squares procedure comes into play. The weak forms are

(6.4) and (6.5) in the form

$$\left.\begin{array}{l} \langle \sum \dot{U}_j\, \phi_j + \sum \dot{X}_j (-U_x)_j\, \phi_j + f(U)_x \;,\; \phi_i \rangle = 0 \\ \langle \sum \dot{U}_j\, \phi_j + \sum \dot{X}_j (-U_x)_j\, \phi_j + f(U)_x \;,\; (-U_x)_i \phi_j \rangle = 0 \end{array}\right] \qquad (8.9)$$

Since, for these choices of S_n, T_n, the space spanned by the functions ϕ_i and $(-U_x)_i\, \phi_i$ is also spanned by the restriction of the ϕ_i's to each element (provided that $(U_x)_i$ differs from element to element), these equations can be written [6], [8]

$$\frac{1}{3} h_L(\dot{U}_j - m_L \dot{X}_j) + (f_j - \bar{f}^L_j) = 0 \qquad (8.10)$$

$$\frac{1}{3} h_R(\dot{U}_j - m_R \dot{X}_j) + (f_j - \bar{f}^R_j) = 0 \qquad (8.11)$$

where h and m are the element length and slope u_x in an element, L and R refer to elements to the left and right of node j,

$$f_j = f(U_j) \;,\; \bar{f}^L_j = \frac{1}{h_L} \int_{j-1}^{j} f(U)dx \;,\; \bar{f}^R_j = \frac{1}{h_R} \int_{j}^{j+1} f(U)dx \;, \qquad (8.12)$$

leading to

$$\dot{X}_j = \frac{\frac{3}{h_R}(-f_j + \bar{f}^R_j) - \frac{3}{h_L}(f_j - \bar{f}^L_j)}{m^R_j - m^L_j} \qquad (8.13)$$

approximating $\dot{x} = \partial f_x / \partial u_x$ (see[8]). By subtracting equation (8.10) from (8.11) with j replaced by $j-1$ in (8.11) we also obtain

$$\dot{m}_{j-\frac{1}{2}} = \frac{3}{h_{j-\frac{1}{2}}}(f_j + f_{j-1} - \bar{f}^L_j - f^R_{j-1}) \tag{8.14}$$

approximating $\dot{u}_x = -\partial f/\partial x$. Equations (8.13) and (8.14) may be solved for X_j, $m_{j-\frac{1}{2}}$.

It is the equivalence of (a) the two equations (8.9) and (b) the equations (8.13) together with the first of (8.9), at the level of general approximation spaces, which is the essence of this report. Equally, both are equivalent to (8.13) and (8.14).

§9. Conclusion

It has been proved that minimisation in a finite space of functions of the L_2 norm of the residual of the left hand side of the equation

$$X + Y(y) - xy = 0 \tag{9.1}$$

occurring in the Legendre transformation over both of the variables X, x, generates an approximation to the envelope of (9.1) as y varies, where Y is a prescribed function of y, and hence to the Legendre dual function $X(x)$ of $Y(y)$. The conditions for the minimum provide a mechanism for the construction of the approximation using finite element subspaces, in the manner already studied in the Moving Finite Element method.

In the particular case of piecewise linear/piecewise constant function spaces, the minimum is zero and (9.1) is satisfied exactly, i.e. the Legendre transformation exists within the approximation spaces as well as in the continuous problem. In this case, however, the envelope equation is still only approximately satisfied.

10. Acknowledgements

I should like to thank Simon Chynoweth, David Porter and Mike Sewell for the stimulus for this work and Susan Davis for her patient typing.

§10. References

[1] Sewell, M.J. (1982) Legendre Transformations and Extremum Principles. In Mechanics of Solids (H.G. Hopkins and M.J. Sewell (eds)). Pergamon Press.

[2] Sewell, M.J. (1987) Maximum & Minimum Principles. CUP.

[3] Sewell, M.J. & Porter, D. (1980) Constitutive Surfaces in Fluid Mechanics. Math. Proc. Camb. Phil. Soc. 88,517.

[4] Miller, K & Miller, R.N. (1981) Moving Finite Elements Part I. Miller K. (1981) Moving Finite Elements Part II. SIAM J. Numer. An. 18,1019-1057.

[5] Mueller, A.C. & Carey, G.F. (1985) Continuously Deforming Finite Elements for Transport Problems. Int. J. Num. Meths. in Eng. 21,2099.

[6] Wathen, A.J. & Baines, M.J. (1985) On the Structure of the Moving Finite Element Equations. IMA. Journal of Numer. An. 5, 161-182.

[7] Baines, M.J. Moving Finite Envelopes. Num. Anal. Rept. 12/87. Department of Mathematics, University of Reading.

[8] Baines, M.J. & Wathen, A.J., (1988) Moving Finite Elements for Evolutionary Problems I (Theory). Johnson, I.W., Wathen, A.J. & Baines, M.J. (1988) Moving Finite Elements for Evolutionary Problems II (Applications). J. Comp. Phys. (to appear).

[9] Jimack, P. (1988) High Order Moving Finite Elements. Report No. AM-88-03, School of Mathematics, University of Bristol.