

THE WEIERSTRASS EXCESS FUNCTION

RUSSELL D. RUPP

This paper presents properties of the Weierstrass excess function in a general setting. The results contained herein are appropriate to applications of Hestenes' method of multipliers.

1. **Introduction.** This paper presents properties of the Weierstrass excess function in a general setting. These results are appropriate to applications of Hestenes' method of multipliers (Hestenes [3, pp. 305-310], Rupp [10, Theorem 3.1]). Particular cases of most of the results here presented are due to or are similar to those of Hestenes [2, pp. 151-152; 4, pp. 518-521; 5] or Reid [6, pp. 951-952; 7, pp. 680-682; 8, pp. 667-671] who applied them to specific problems in the calculus of variations. In fact, Hestenes [1, pp. 263-264; 4, p. 514] obtained sufficiency theorems for constrained problems by adding a penalty term to the original functional and applying the unconstrained theory.

The method of multipliers was suggested by Hestenes in order to circumvent the numerical error to which the method of penalty functions is susceptible (Hestenes [3, pp. 305-308]). These two methods arise from the observation that a series of unconstrained minimization problems is often easier to work with than a single constrained problem. Thus instead of solving a single constrained problem, one might solve a series of unconstrained problems whose solutions tend to the minimum of the constrained problem.

Consider, for example, the problem of Lagrange: Minimize

$$I(x) = \int_{T^1}^{T^2} F(t, x(t), \dot{x}(t)) dt$$

over the subclass \mathfrak{D} of the class \mathfrak{A} of terminally admissible piecewise smooth arcs

$$x: x^i(t) \quad T^1 \leq t \leq T^2 \quad i = 1, 2, \dots, n$$

which is determined by the constraints

$$\phi_\alpha(t, x(t), \dot{x}(t)) = 0 \quad \alpha = 1, 2, \dots, m.$$

Let

$$\Psi(x) = \phi_\alpha(t, x(t), \dot{x}(t)) \phi_\alpha(t, x(t), \dot{x}(t))$$

where here and later the repeated index α is summed. The method

of penalty functions is to minimize

$$I(x) + \frac{1}{2\xi}\Psi(x)$$

over \mathfrak{X} as ξ tends to infinity. Hestenes' method of multipliers is to minimize

$$(1.1) \quad J(x, \lambda) = \int_{r^1}^{r^2} \left[F + \lambda_\alpha \phi_\alpha + \frac{1}{2} C \phi_\alpha \phi_\alpha \right] dt$$

over \mathfrak{X} where C is a fixed function and the $\lambda_\alpha(t)$ are approximations to the Lagrange multipliers. Thus an initial multiplier estimate fixes an initial minimization problem. A second multiplier estimate is obtained from the solution to the first minimization problem and so on.

The function C in (1.1) must be suitable, that is, if the constrained problem satisfies the usual sufficiency criteria, then the new family of problems must satisfy the sufficiency criteria in Rupp [9, Section 2]. In particular, a suitably strengthened Weierstrass condition must hold. The results here presented ensure that C may be chosen so that a suitably strengthened Weierstrass condition holds.

Let \mathfrak{Z} be a metric space with metric $d(z, z^*)$. Let $u = (u^1, u^2, \dots, u^n)$ be a vector in n -dimensional real space R^n . Repeated indices denote summation with respect to the index. For example, the norm of u is

$$|u| = (u^i u^i)^{1/2}.$$

Here and later, the range of the index i is $1, 2, \dots, n$. Given variables like z in \mathfrak{Z} and u in R^n , z -space or zu -space is the set of points z or (z, u) , that is, \mathfrak{Z} or $\mathfrak{Z} \times R^n$. Similarly given a set $\mathfrak{S} \subseteq \mathfrak{Z} \times R^n$, $\text{proj}_u(\mathfrak{S})$ or $\text{proj}_z(\mathfrak{S})$ is the projection of \mathfrak{S} into u or z -space respectively. Given z in $\text{proj}_z(\mathfrak{S})$, we define \mathfrak{S}_z to be the set of all u in R^n such that (z, u) is in \mathfrak{S} . Let $f(z, u)$ be a real-valued function defined on \mathfrak{S} . Suppose that \mathfrak{S}_z is open for each z in $\text{proj}_z(\mathfrak{S})$. Assume that $f(z, u)$ has at least one derivative with respect to each u^i . We define the Weierstrass excess or E -function E_f of f with respect to the u variables to be

$$E_f(z, u, v) = f(z, v) - f(z, u) - (v^i - u^i) f_{u^i}(z, u).$$

2. Minor E -dominance properties. This section contains lemmas which are both useful in applications of Hestenes' method of multipliers and in proving some of the results in the next section. These lemmas could be generalized to the type of E -dominance relations developed by Hestenes [5] in the parametric case. Most of the results in Hestenes [5] are stated so as to be equally valid in the nonparametric case.

When an arc $x(t)$ is considered, we make the correspondence $u = \dot{x}(t)$. Thus the function

$$L(u) = (1 + |u|^2)^{1/2}$$

is the integrand of the arc length integral. Integration of the E -function E_L of $L(u)$

$$\begin{aligned} (2.1) \quad E_L(u, v) &= L(v) - L(u) - (v^i - u^i)L_{u^i}(u) \\ &= L(v)[1 - \cos((1, u), (1, v))] \end{aligned}$$

or the function

$$(2.2) \quad L(u - v) - 1$$

provides a measure of the distance between two arcs. In order to prove some sufficiency theorems by expansion, Reid [6, pp. 951-952; 7, pp. 680-682] considered the function

$$|u - v|^2(1 + |u - v|)^{-1}$$

rather than (2.1) or (2.2). This function has the same qualitative growth as (2.1) or (2.2). However (2.2) is more useful in an indirect sufficiency proof because of its relation to directional convergence (Hestenes [2, pp. 152-159]). Results obtained by Reid [6, pp. 951-952; 7, pp. 680-682; 8, pp. 667-671] are contained in Lemmas 2.1, 2.2, 2.3 and Theorems 3.1, 3.2.

LEMMA 2.1. *Let positive constants M and m be given. There are positive constants $M_1, M_2, m_1,$ and m_2 such that*

$$M_1|v - u|^2 \geq E_L(u, v) \geq M_2|v - u|^2 \quad \text{when} \quad |u|, |v| \leq M$$

and

$$m_1|v - u| \geq E_L(u, v) \geq m_2(1 + |v - u|) \quad \text{when} \quad |u|, |v| \geq m.$$

The proof of the first inequality follows from the intermediate value theorem,

$$E_L(u, v) = |\pi|^2 \int_0^1 (1 - \theta)L(w)^{-3}[1 + |w|^2 \sin^2 \phi]d\theta$$

where $\pi = v - u$, $w(\theta) = u + \theta(v - u)$, and $\phi(\theta)$ is the angle between π and w . The second inequality follows from the identity

$$E_L(u, v) = L(v)(1 - \cos \phi)$$

where ϕ is the angle between $(1, u)$ and $(1, v)$. This proves the lemma.

LEMMA 2.2. *There are positive constants τ, τ^* such that*

$$\begin{aligned} \tau E_L(0, v - u) &\geq E_L(u, v) \\ &\geq \tau^* E_L(0, v - u) = \tau^* [L(v - u) - 1] = \tau^* \frac{|v - u|^2}{L(v - u) + 1} \end{aligned}$$

where u is in a bounded subset \mathfrak{S} of E^n and v is in E^n .

In order to prove this result, first note that if we assume $|v|$ is bounded, then Lemma 2.1 implies the existence of the required constants (which depend on the bound). Now observe that $E_L(0, v - u)$ is asymptotic to $|v|$. Hence Lemma 2.1 implies that there is another pair of positive constants such that the inequalities hold for large $|v|$. This proves the lemma.

In the next two lemmas we assume that \mathfrak{R} and \mathfrak{D} are subsets of zu -space and that \mathfrak{R} is compact. Let $f(z, u)$ be a continuous, real-valued function defined on a neighborhood of \mathfrak{D} union \mathfrak{R} and having two continuous u -derivatives on a neighborhood of \mathfrak{R} . Given (z_0, u_0) in \mathfrak{R} , we define the cone $\mathfrak{C}(z_0, u_0)$ of (z_0, u_0) with respect to \mathfrak{D} and \mathfrak{R} to consist of zero and all positive multiples of vectors π for which there are sequences (z_n, u_n) in \mathfrak{R} and (z_n, v_n) in \mathfrak{D} such that

$$\lim (z_n, u_n) = \lim (z_n, v_n) = (z_0, u_0)$$

and

$$\frac{\pi}{|\pi|} = \lim \frac{v_n - u_n}{|v_n - u_n|} .$$

LEMMA 2.3. *If there is a positive constant τ such that (z, u) in \mathfrak{R} and (z, v) in \mathfrak{D} imply*

$$E_f(z, u, v) \geq \tau E_L(u, v) ,$$

then for each (z, u) in \mathfrak{R} ,

$$f_{u^i u^j}(z, u) \pi^i \pi^j \geq \tau [L(u)]^{-3/2} |\pi|^2$$

whenever π is in $\mathfrak{C}(z, u)$.

In order to prove this lemma, let (z, u) be in \mathfrak{R} and π be in $\mathfrak{C}(z, u)$. There are sequences (z_n, u_n) in \mathfrak{R} and (z_n, v_n) in \mathfrak{D} such that

$$\lim (z_n, u_n) = \lim (z_n, v_n) = (z, u)$$

and

$$\lim \frac{v_n - u_n}{|v_n - u_n|} = \frac{\pi}{|\pi|} .$$

Applying the intermediate value theorem to the hypothesis, we obtain

$$\int_0^1 (1 - \theta) f_{u^i u^j}(z_n, w_n) \pi_n^i \pi_n^j d\theta \geq \tau \int_0^1 (1 - \theta) L(w_n)^{-3} |\pi_n|^2 [1 + |w_n|^2 \sin^2 \phi_n] d\theta$$

where $\pi_n = v_n - u_n$, $w_n(\theta) = u_n + \theta\pi$, and $\phi(\theta)$ is the angle between π_n and w_n . Dividing n by $|\pi_n|^2$ and letting n tend to infinity, we complete the proof of the lemma.

LEMMA 2.4. *Suppose there is a positive constant τ^* such that (z, u) in \mathfrak{R} and (z, v) in \mathfrak{D} imply*

$$L(v) \geq \tau^* |f(z, v)| .$$

There is a positive constant τ such that (z, u) in \mathfrak{R} and (z, v) in \mathfrak{D} imply

$$E_L(u, v) \geq \tau |E_f(z, u, v)| .$$

This lemma follows from Lemma 2.1 and continuity considerations.

3. Major E -dominance properties. Let \mathfrak{R} be a compact subset of an open set \mathfrak{R} in zu -space, and let \mathfrak{D} be an arbitrary subset of \mathfrak{R} . We suppose $f(z, u)$ to be a continuous, real-valued function defined on \mathfrak{R} . In a neighborhood of \mathfrak{R} let $f(z, u)$ have two continuous partial derivatives with respect to the u -variables. We also suppose that the z -variables are of the form (x, y) where x and y are elements in metric spaces \mathfrak{D}' and \mathfrak{D}'' with metrics $d'(x, x^*)$ and $d''(y, y^*)$. Given a positive constant δ , \mathfrak{N}_δ is defined to be the set of all $(x, y, u) = (z, u)$ such that for some \tilde{y} , (x, \tilde{y}, u) is in \mathfrak{R} and $d'(y, \tilde{y}) < \delta$. In applications we sometimes consider \mathfrak{D}' or \mathfrak{D}'' to be the trivial space of one element.

This next definition is used to analyze situations in which some form of a Weierstrass condition occurs. The set \mathfrak{R} satisfies property \mathfrak{C} if there is a set \mathfrak{D}^* containing \mathfrak{R} with the two properties:

- (i) Given a positive constant ε , there is a positive constant δ such that for every (z, u) in \mathfrak{N}_δ , there is (z, u^*) in \mathfrak{D}^* with $|u - u^*| < \varepsilon$.
- (ii) There is a positive constant τ such that for (z, u^*) in \mathfrak{D}^* and (z, v) in \mathfrak{D} ,

$$(3.1) \quad E_f(z, u^*, v) \geq \tau E_L(u^*, v) .$$

In particular we note that property \mathfrak{C} is satisfied with $\mathfrak{D}^* = \mathfrak{N}_\delta$, if there are positive constants δ and τ such that (z, u) in \mathfrak{N}_δ and (z, v) in \mathfrak{D} imply

$$E_f(z, u, v) \geq \tau E_L(u, v) .$$

THEOREM 3.1. *Assume property \mathfrak{C} and that the matrix of second partials of $f(z, u)$ with respect to u is positive definite. There are positive constants δ, τ , and ε such that (z, u) in \mathfrak{N}_δ , (z, v) in \mathfrak{D} , and $|\pi| \leq \varepsilon$ imply*

$$E_f(z, u + \pi, v) \geq \tau E_L(u, v) .$$

This theorem follows by an argument similar to that of Hestenes [4, p. 519].

The next theorem provides a connection between two possible Weierstrass conditions which are stated below as (i) and (ii).

THEOREM 3.2. *Let \mathfrak{D} be a subset of \mathfrak{R} and suppose that the matrix of second partials of $f(z, u)$ with respect to u is positive definite. The following two conditions are equivalent:*

- (i) *There are positive constants τ and δ such that (z, u) in \mathfrak{R}_δ and (z, v) in \mathfrak{D} imply*

$$E_f(z, u, v) \geq \tau E_L(u, v) .$$

- (ii) *There are positive constants ε and δ such that (z, u) in \mathfrak{R}_δ , (z, v) in \mathfrak{D} , and $|\pi| \leq \varepsilon$ imply*

$$E_f(z, u + \pi, v) \geq 0 .$$

That (i) implies (ii) is trivial. The reverse implication follows in the manner of Hestenes [2, pp. 151-152], that is, from the intermediate value theorem and the identity

$$\begin{aligned} E_f(z, u, u + k\pi) &= E_f(z, u + \pi, u + k\pi) \\ &+ kE_f(z, u, u + \pi) + (k - 1)E_f(z, u + \pi, u) . \end{aligned}$$

The next theorem, which is important in applications to differential equation constraints, requires several new concepts. Let the real-valued, nonnegative function $h(z, u)$ be zero on \mathfrak{R} and have the same continuity and differentiability properties as $f(z, u)$. We say that $f(z, u)$ satisfies the Legendre condition with respect to $h(z, u)$ if there is a positive constant h_0 such that

$$f(z, u) + h_0 h(z, u)$$

has a positive definite matrix of second partial derivatives with respect to u . We also define a subset \mathfrak{S} of \mathfrak{R} to be Cartesian with respect to the x -variables if there is a subset \mathfrak{X} of x -space and \mathfrak{Y} of yu -space such that \mathfrak{S} is the Cartesian product of \mathfrak{X} and \mathfrak{Y} . The set \mathfrak{S} is locally Cartesian with respect to x if given (x^*, y^*, u^*) in \mathfrak{S} , there is a neighborhood \mathfrak{R} of x^* in x -space such that the Cartesian product of \mathfrak{R} and $\text{proj}_{yu}(\mathfrak{S})$ is contained in \mathfrak{S} . We remark that if $f(z, u)$ is continuous on the closure of \mathfrak{R} , then \mathfrak{R} may be considered to be Cartesian with respect to z .

THEOREM 3.3. *Let \mathfrak{D} be given by the set of zeros of $h(z, u)$, and*

assume the strong condition of Legendre with respect to $h(z, u)$. Suppose there is a positive constant τ^* such that (z, u) in \mathfrak{R} and (z, v) in \mathfrak{D} imply

$$E_f(z, u, v) \geq \tau^* E_L(u, v) .$$

There exists a continuous, nonnegative function $H(z, u)$ which is infinitely differentiable in its Euclidean arguments and a positive constant τ such that if we define

$$F(z, u) = f(z, u) + H(z, u)h(z, u) ,$$

then for (z, u) in \mathfrak{R} and (z, v) in \mathfrak{R} ,

$$E_F(z, u, v) \geq \tau E_L(u, v) .$$

Furthermore if \mathfrak{R} is locally Cartesian with respect to x , then $H(z, u) = H(x, y, u)$ may be taken to be a function of (y, u) alone.

Theorem 3.3 follows by an argument similar to that of Hestenes [4, p. 520].

This last theorem is a criterion for determining when the function $H(z, u)$ in Theorem 3.3 may be taken as a constant.

THEOREM 3.4. *Let \mathfrak{D} be defined by $h(z, u)$ as in Theorem 3.3. Suppose that $f(z, u)$ is defined on the closure of \mathfrak{R} and that there are constants $\alpha > 1, c > 0$, and d such that for large $|u|$,*

$$(3.18) \quad f(z, u) \geq d + c|u|^\alpha .$$

The conclusion of Theorem 3.3 holds with $H(z, u)$ a constant. Furthermore if

$$(3.19) \quad \sum_i |f_{u^i}| < c \quad \text{on } \mathfrak{R} ,$$

then the conclusion holds with $\alpha = 1$.

In order to prove this theorem, it suffices to show $E_F \geq 0$ since by Lemma 2.1, subtracting a small positive constant times L from f does not change the hypotheses. Let \mathfrak{N} be a neighborhood of \mathfrak{R} which is bounded with respect to u and so large that (z, u) in \mathfrak{R} and (z, v) in the exterior of \mathfrak{N} imply by (3.19) (if $\alpha = 1$) and (3.18),

$$(3.20) \quad \begin{aligned} E_f(z, u, v) &= f(z, v) - f(z, u) - (v^i - u^i)f_{u^i}(z, u) \\ &\geq d + c|v|^\alpha - f(z, u) - (v^i - u^i)f_{u^i}(z, u) \geq 0 . \end{aligned}$$

On the other hand, by continuity there are constants $\varepsilon > 0$ and b such that for (z, u) in \mathfrak{Q} , (z, v) in \mathfrak{R} , and $h(z, v) \leq \varepsilon$,

$$(3.21) \quad E_f(z, u, v) > 0$$

and for (z, u) in \mathfrak{R} , (z, v) in \mathfrak{X} , and $h(z, v) > \varepsilon$,

$$(3.22) \quad E_f(z, u, v) > b .$$

Let H be a positive constant such that

$$(3.23) \quad H\varepsilon + b > 0 .$$

By (3.21) we have that for (z, u) in \mathfrak{R} , (z, v) in \mathfrak{X} , and $h(z, v) \leq \varepsilon$,

$$(3.24) \quad E_{f+Hh}(z, u, v) = E_f(z, u, v) + Hh(z, v) \geq 0 .$$

Also by (3.22) and (3.23) we have that for (z, u) in \mathfrak{R} , (z, v) in \mathfrak{X} , and $h(z, v) > \varepsilon$,

$$(3.25) \quad \begin{aligned} E_{f+Hh}(z, u, v) &= E_f(z, u, v) + Hh(z, v) \\ &\geq b + \varepsilon H \geq 0 . \end{aligned}$$

Combining (3.20), (3.24), and (3.25), we prove the theorem.

REFERENCES

1. M. R. Hestenes, *An alternate sufficiency proof for the normal problem of Bolza*, Trans. Amer. Math. Soc., **61** (1947), 256-264.
2. ———, *Calculus of Variations and Optimal Control Theory*, John Wiley and Sons, Inc., New York, 1966.
3. ———, *An indirect sufficiency proof for the problem of Bolza in nonparametric form*, Trans. Amer. Math. Soc., **62** (1947), 509-535.
4. ———, *Multipier and gradient methods*, J. Optimization Theory and Applications, **4** (1969), 303-320.
5. ———, *The Weirstrass E-function in the calculus of variations*, Trans. Amer. Math. Soc., **60** (1946), 51-71.
6. W. T. Reid, *Expansion methods for the isoperimetric problem of Bolza in non-parametric form*, Amer. J. Math., **71** (1949), 946-975.
7. ———, *Isoperimetric problems of Bolza in non-parametric form*, Duke Math. J., **5** (1939), 675-691.
8. ———, *Sufficient conditions by expansion methods for the problem of Bolza in the calculus of variations*, Annals of Math., **38** (1937), 662-678.
9. R. D. Rupp, *A new type of variational theory sufficiency theorem*, (to appear, Pacific Journal of Mathematics).
10. ———, *A nonlinear optimal control minimization technique*, (to appear, Trans. Amer. Math. Soc.).
11. ———, *A unified sufficiency theory in the Calculus of Variations*, Dissertation, University of California, Los Angeles, 1970.

Received December 4, 1970 and in revised form June 22, 1971. The preparation of this paper was sponsored in part by the U. S. Army Research Office under Grant DA-31-124-ARO (D)-355. Reproduction in whole or in part is permitted for any purpose of the United States Government. The preparation of this paper was also partially supported by National Science Foundation Grant GU 3171.

UNIVERSITY OF CALIFORNIA, LOS ANGELES

AND

STATE UNIVERSITY OF NEW YORK AT ALBANY