Inequalities and Approximation

WILLIAM W. HAGER

In this survey, we discuss a class of inequalities related to the following topics:

(1) Error estimates for variational inequality and optimal control approximation

(2) Stability for mathematical programs

(3) Solution regularity for optimal control problems

The material discussed in Sections 4–8 will be developed more fully in forthcoming papers [6–8].

1. ERROR ESTIMATES FOR QUADRATIC COST

Consider the problem

$$\text{minimize} \quad \{J(v) = a(v, v) + l(v) : v \in K\}, \tag{1.1}$$

where K is a convex subset of the Banach space $\mathscr{U}$, $l(\cdot)$ is a bounded linear functional on $\mathscr{U}$, and $a(\cdot, \cdot)$ is a symmetric, bounded bilinear form on $\mathscr{U}$. Suppose that there exists a solution $u \in K$ to (1.1), and let $K^h \subset \mathscr{U}$ be an approximation to K. No assumptions are made regarding K^h; in particular, it need not be convex. If $u^h \in K^h$ solves the problem

$$\text{minimize} \quad \{J(v) : v \in K^h\}, \tag{1.2}$$

we shall estimate the error $u - u^h$ in terms of energy $a(u - u^h, u - u^h)$.

Since K is convex and $J(\cdot)$ is differentiable, we have the standard variational inequality [12].

$$DJ[u](v - u) \geqq 0 \qquad \text{for all} \quad v \in K, \tag{1.3}$$

where

$$DJ[u](v) = 2a(u, v) + l(v). \tag{1.4}$$

Expanding $J(\cdot)$ about u gives us

$$J(u^h) = J(u) + DJ[u](u^h - u) + a(u^h - u, u^h - u). \tag{1.5}$$

Moreover, (1.3) implies that

$$\begin{aligned} DJ[u](u^h - u) &= DJ[u](u^h - v) + DJ[u](v - u) \\ &\geqq DJ[u](u^h - v) \end{aligned} \tag{1.6}$$

ISBN 0-12-178150-X

for all $v \in K$. On the other hand, since u^h minimizes $J(\cdot)$ over K^h, we have

$$J(u^h) \leqq J(v^h) = J(u) + DJ[u](v^h - u) + a(v^h - u, v^h - u) \tag{1.7}$$

for all $v^h \in K^h$. Finally combining (1.5)–(1.7), we get

$$a(u^h - u, u^h - u) \leqq DJ[u](v - u^h) + DJ[u](v^h - u) + a(v^h - u, v^h - u) \tag{1.8}$$

for all $v \in K$ and $v^h \in K^h$.

Special cases of (1.8) are the following:

(i) $K^h \subset K$. Choosing $v = u^h$ in (1.8) gives

$$a(u^h - u, u^h - u) \leqq DJ[u](v^h - u) + a(v^h - u, v^h - u) \tag{1.9}$$

for all $v^h \in K^h$.

(ii) $K = \mathscr{U}$. Hence (1.3) implies that $DJ[u](v) = 0$ for all $v \in K$ and (1.8) yields

$$a(u^h - u, u^h - u) \leqq a(v^h - u, v^h - u) \tag{1.10}$$

for all $v^h \in K^h$.

A classical application of (1.8) is the obstacle problem [3, 4] where we have

$$\begin{aligned} \mathscr{U} &= H_0^1(\Omega), \\ K &= \{v \in \mathscr{U} : v \geqq \psi \text{ on } \Omega\}, \\ J(v) &= \int_\Omega [|\nabla v|^2 - 2fv]. \end{aligned} \tag{1.11}$$

Here $\Omega \subset R^2$ is a bounded open set, $f \in L^2(\Omega)$, $H^m(\Omega)$ is the standard Sobolev space consisting of functions whose derivatives through order m are square integrable on Ω, $H_0^1(\Omega) \subset H^1(\Omega)$ is the subspace consisting of functions vanishing on $\partial\Omega$, and $\psi \in H^2(\Omega)$ is the given obstacle. If $\partial\Omega$ is sufficiently regular, there exists a solution $u \in H^2(\Omega)$ for problem (1.1). To simplify the exposition we assume that Ω is a polygon although this restriction is easily removed [3].

Let $S^h \subset H_0^1(\Omega)$ denote a piecewise linear subspace that satisfies the standard interpolation bound

$$\|g - g^{\mathrm{I}}\|_{H^k} \leqq ch^{2-k} \tag{1.12}$$

for all $g \in H^2(\Omega)$ and $k = 0, 1$, where h denotes the diameter of the biggest triangle in the triangulation of Ω and c denotes a generic constant that is independent of h. Finally, we define the set

$$K^h = \{v^h \in S^h : v^h \geqq \psi^{\mathrm{I}} \quad \text{on} \quad \Omega\}. \tag{1.13}$$

Integrating by parts $DJ[u](v)$ given by (1.4) gives us

$$DJ[u](v) = \langle w, v\rangle, \qquad w = -2(\Delta u + f), \tag{1.14}$$

where $\langle\cdot,\cdot\rangle$ denotes the $L^2(\Omega)$ inner product. Furthermore, using the variational inequality, it can be shown that

$$w \geqq 0, \qquad w(u - \psi) = 0 \qquad \text{almost everywhere on } \Omega. \tag{1.15}$$

We now substitute $v = u$ and $v^h = u^{\mathrm{I}}$ into (1.8); applying (1.14) leads to

$$\begin{aligned} &DJ[u](u - u^h) + DJ[u](u^{\mathrm{I}} - u) \\ &= \langle w, u^{\mathrm{I}} - u^h\rangle \\ &= \underbrace{\langle w, \psi^{\mathrm{I}} - u^h\rangle}_{\leqq 0} + \underbrace{\langle w, \psi - \psi^{\mathrm{I}}\rangle}_{=O(h^2)} + \underbrace{\langle w, u - \psi\rangle}_{=0} + \underbrace{\langle w, u^{\mathrm{I}} - u\rangle}_{=O(h^2)} \\ &\leqq ch^2 \end{aligned} \tag{1.16}$$

since $w \geqq 0$ by (1.15), $\psi^{\mathrm{I}} - u^h \leqq 0$ by (1.13), $\|\psi - \psi^{\mathrm{I}}\|_{H^0} = O(h^2) = \|u - u^{\mathrm{I}}\|_{H^0}$ by (1.12), and $\langle w, u - \psi\rangle = 0$ by (1.15). Similarly, we have

$$a(u - u^{\mathrm{I}}, u - u^{\mathrm{I}}) \leqq \|u - u^{\mathrm{I}}\|^2_{H^1} = O(h^2). \tag{1.17}$$

Combining (1.8), (1.16), and (1.17), we obtain the estimate

$$a(u - u^h, u - u^h) \leqq O(h^2). \tag{1.18}$$

Using quadratic elements and sharper regularity results established by Brézis [1, 2], it can be shown that

$$\|u - u^h\|_{H^1} = O(h^{1\cdot 5 - \varepsilon}) \qquad \text{for any} \quad \varepsilon > 0 \quad [3].$$

2. ERROR ESTIMATES FOR DIFFERENTIABLE COST

Now let us consider the equation: Find $u \in H_0^1(\Omega)$ such that

$$u''(x) = e^{u(x)} \qquad \text{for all} \quad x \in \Omega \tag{2.1}$$

where $\Omega = (0, 1)$ and let $u \in H^2(\Omega)$ denote the solution. Defining the functional

$$J(v) = \int_\Omega [(v')^2 + 2e^v], \tag{2.2}$$

(2.1) is equivalent to the variational problem

$$\text{minimize } \{J(v) : v \in K\} \tag{2.3}$$

where $K \equiv H_0^1(\Omega)$. Letting $S^h \subset H_0^1(\Omega)$ denote the space of continuous, piecewise linear polynomials with h = maximum grid interval, we select $K^h = S^h$ and consider the approximation (1.2).

The estimate (1.8) no longer applies due to the e^v term included in $J(v)$. To generalize our earlier results, suppose that $J:\mathscr{U} \to R$ is differentiable; hence, (1.3) holds [12]. Moreover, suppose that there exists $\alpha > 0$ such that

$$J(v) - J(w) - DJ[w](v - w) \geqq \alpha\|v - w\|^2 \tag{2.4}$$

for all $v, w \in \mathscr{U}$ where $\|\cdot\|$ denotes the norm on $\mathscr{U}$. Consequently, (1.5) can be replaced by

$$J(u^h) \geqq J(u) + DJ[u](u^h - u) + \alpha\|u - u^h\|^2. \tag{2.5}$$

In addition, define the parameter

$$c(v, w) = \frac{J(v) - J(w) - DJ[w](v - w)}{\|v - w\|^2} \tag{2.6}$$

for all $v \neq w$. Hence (1.7) can be replaced by

$$J(u^h) \leqq J(u) + DJ[u](v^h - u) + c(v^h, u)\|v^h - u\|^2 \tag{2.7}$$

for all $v^h \in K^h$. Combining (2.5), (1.6), and (2.7), we get

$$\alpha\|u^h - u\|^2 \leqq DJ[u](v - u^h) + DJ[u](v^h - u) + c(v^h, u)\|v^h - u\|^2 \tag{2.8}$$

for all $v \in K$ and $v^h \in K^h$.

Now let us apply (2.8) to our particular equation (2.1). Observe that

$$\begin{aligned} DJ[w](v) &= 2\int_\Omega [w'v' + e^w v] \\ &= 2\langle -w'' + e^w, v\rangle \end{aligned} \tag{2.9}$$

and

$$J(v) - J(w) - DJ[w](v - w) = \langle (v - w)', (v - w)'\rangle + \langle e^\gamma (v - w), v - w\rangle \tag{2.10}$$

by Taylor's theorem where $\gamma(x)$ lies between $v(x)$ and $w(x)$. Hence we have

$$(2.10)\begin{cases} \geqq \langle (v - w)', (v - w)'\rangle, \\ \leqq \langle (v - w)', (v - w)'\rangle[\exp\{\|v'\|_{L^2} + \|w'\|_{L^2}\} + 1], \end{cases} \tag{2.11}$$

since

$$\begin{aligned} &\|v\|_{L^\infty} \leqq \|v'\|_{L^2} && \text{for all} \quad v \in H_0^1(\Omega), \\ &\gamma(x) \leqq \|v\|_{L^\infty} + \|w\|_{L^\infty} && \text{for all} \quad x \in \Omega. \end{aligned} \tag{2.12}$$

Finally (2.1) implies that $DJ[u](v) = 0$ for all $v \in H_0^1(\Omega)$ and (2.8)–(2.11) yield for $v = u^h$ and $v^h = u^{\mathrm{I}}$

$$\langle (u - u^h)', (u - u^h)'\rangle \leqq ch^2. \tag{2.13}$$

3. ERROR ESTIMATES FOR NONDIFFERENTIABLE COST

Consider the Bingham fluid problem that is given by (1.1) with the choices

$$\begin{aligned} \mathscr{U} &= H_0^1(\Omega) = K, \\ J(v) &= \int_\Omega [|\nabla v|^2 + |\nabla v| - 2fv] \end{aligned} \tag{3.1}$$

with $f \in L^2(\Omega)$. Letting $S^h \subset \mathscr{U}$ denote the piecewise linear subspace of Section 2, we again take $K^h = S^h$ and study the approximation (1.2).

Observe that (2.8) cannot be utilized since $J(\cdot)$ is nondifferentiable. To generalize (1.8) or (2.8), suppose that

$$J(v) = J_n(v) + J_d(v), \tag{3.2}$$

where $J_n(\cdot)$ is convex but possibly nondifferentiable, and

$$J_d(v) = a(v, v) + l(v). \tag{3.3}$$

If $u \in K$ solves (1.1), then the following variational inequality holds [12]:

$$DJ_d[u](v - u) + J_n(v) \geqq J_n(u) \tag{3.4}$$

for all $v \in K$ where

$$DJ_d[u](v) = 2a(u, v) + l(v). \tag{3.5}$$

Hence we have

$$J(u^h) = J_n(u^h) + J_d(u) + DJ_d[u](u^h - u) + a(u^h - u, u^h - u). \tag{3.6}$$

Moreover, by (3.4), we find that

$$DJ_d[u](u^h - u) \geqq DJ_d[u](u^h - v) + J_n(u) - J_n(v) \tag{3.7}$$

for all $v \in K$. On the other hand, we observe that

$$J(u^h) \leqq J(v^h) = J_n(v^h) + J_d(u) + DJ_d[u](v^h - u) + a(v^h - u, v^h - u) \tag{3.8}$$

for all $v^h \in K^h$. Finally the combination (3.6)–(3.8) yields

$$\begin{aligned} a(u^h - u, u^h - u) \leqq\ & J_n(v^h) - J_n(u) + J_n(v) - J_n(u^h) \\ & + DJ_d[u](v^h - u) + DJ_d[u](v - u^h) \\ & + a(v^h - u, v^h - u) \end{aligned} \tag{3.9}$$

for all $v \in K$ and $v^h \in K^h$.

Applying (3.9) to the Bingham fluid problem using $v = u^h$ and $v^h = u^{\mathrm{I}}$, we obtain

$$\begin{aligned} \|\nabla(u^h - u)\|_{H^0}^2 &\leqq \langle |(u^{\mathrm{I}} - u)|, 1\rangle - 2\langle \Delta u + f, u^{\mathrm{I}} - u\rangle + \|\nabla(u^{\mathrm{I}} - u)\|_{H^0}^2 \\ &\leqq ch. \end{aligned} \tag{3.10}$$

With a more careful analysis, one can establish the estimate

$$\|\nabla(u^h - u)\|_{H^0}^2 \leqq ch^{2-\varepsilon} \tag{3.11}$$

for any $\varepsilon > 0$. See Glowinski [5] for the details of (3.11).

4. PERTURBATIONS IN THE COST

In the previous sections, we studied the effect of replacing the constraint set K in (1.1) by an approximation K^h. Now let us consider the case where the constraint set is fixed, but the cost functional is permitted to depend on a parameter.

For example, consider the quadratic programs

$$\text{minimize} \quad \{u^T R_j u + 2r_j^T u : u \in K\} \tag{4.1}$$

for $j = 1, 2$ where $K \subset R^n$ is convex. Suppose that there exist solutions (u_1, u_2) to (4.1) associated with $j = 1, 2$, respectively. Hence the following variational inequality holds:

$$(R_j u_j + r_j)^T(v - u_j) \geqq 0 \qquad \text{for all} \quad v \in K \tag{4.2}$$

and $j = 1, 2$.

Choosing ($j = 1, v = u_2$) and ($j = 2, v = u_1$) and adding the resulting relations yields:

$$(u_2 - u_1)^T R_2 (u_2 - u_1) \leqq (r_1 - r_2)^T(u_2 - u_1) + u_1^T(R_1 - R_2)(u_2 - u_1). \tag{4.3}$$

If the smallest eigenvalue, α, of R_2 is positive, (4.3) implies that

$$\alpha |u_2 - u_1| \leqq |r_1 - r_2| + |u_1||R_1 - R_2|, \tag{4.4}$$

where $|\cdot|$ denotes the Euclidean norm. If K is compact or R_1 is positive definite, then an a priori bound can be given for $|u_1|$; hence (4.4) implies that the solution to (4.1) depends Lipschitz continuously on the perturbation.

We introduce the following notation: If $g : R^{m_1} \times R^{m_2} \times \cdots \times R^{m_l} \to R$, we let $\nabla_j g$ denote the gradient of $g(y_1, \ldots, y_l)$ with respect to y_j where $y_k \in R^{m_k}$ for $k = 1, \ldots, l$.

The following generalization of (4.4) is easily established [7]:

Theorem 4.1. *Let $K \subset R^n$ be nonempty, closed, and convex, $E \subset R^m$, $f: R^n \times R^m \to R$ be differentiable in its first n arguments on $K \times R^m$, and assume that there exists $\alpha > 0$ such that*

$$f(y, \xi) \geqq f(x, \xi) + \nabla_1 f(x, \xi)(y - x) + \alpha |x - y|^2 \tag{4.5}$$

for all $x, y \in K$ *and* $\xi \in E$. *Then for all* $\xi \in E$, *there exists a unique* $x(\xi) \in K$ *satisfying*

$$f(x(\xi), \xi) = \min\{f(y, \xi) : y \in K\}; \tag{4.6}$$

and given $\bar{x} \in K$, *we have*

$$|x(\xi) - \bar{x}| \leqq |\nabla_1 f(\bar{x}, \xi)|/\alpha. \tag{4.7}$$

Also, if $\nabla_1 f(z, \cdot)$ *is continuous for all* $z \in K$, *then* $z(\cdot)$ *is continuous on* E, *and if, moreover,* $\nabla_2 \nabla_1 f(\cdot, \cdot)$ *is continuous, then*

$$|x(\xi_1) - x(\xi_2)| \leqq \frac{|\xi_1 - \xi_2|}{2\alpha} \max_{0 \le s \le 1} |\nabla_2 \nabla_1 f(x(\xi_2), \xi_1 + s(\xi_2 - \xi_1))|. \tag{4.8}$$

5. DUAL APPROXIMATIONS

One case where the perturbation parameter appears linearly in the cost of a program arises in the study of dual methods for mathematical programs. Consider the program

$$\text{minimize} \quad \{f(z) : z \in R^n, g(z) \leqq 0\} \tag{5.1}$$

where $g : R^n \to R^m$ and $f : R^n \to R$. The Lagrange dual function is given by

$$\mathscr{L}(\eta) = \inf\{f(z) + \eta^{\mathrm{T}} g(z) : z \in R^n\} \tag{5.2}$$

and the associated dual problem becomes

$$\sup\{\mathscr{L}(\eta) : \eta \in R^m, \eta \geqq 0\}. \tag{5.3}$$

Under suitable assumptions, there exist solutions η^* to (5.3) and z^* to (5.1). Moreover, z^* achieves the minimum in (5.2) for $\eta = \eta^*$.

Now define the set

$$K = \{\eta \in R^m : \eta \geqq 0\}, \tag{5.4}$$

and let K^h be an approximation to K. The following approximation to (5.3) is considered:

$$\sup\{\mathscr{L}(\eta) : \eta \in K^h\}. \tag{5.5}$$

If η^h solves (5.5) and z^h achieves the minimum in (5.2) for $\eta = \eta^h$, let us estimate $z^* - z^h$.

Referring to our development in Section 1–3, we see the need for an inequality of the form (2.4) and an estimate of the parameter c in (2.6). To attack these estimates, define the Lagrangian

$$\mathscr{L}(z, \eta) = f(z) + \eta^{\mathrm{T}} g(z) \tag{5.6}$$

and assume the following:

(i) $f, g \in C^2$;
(ii) there exists $\alpha > 0$ such that

$$\nabla_1^2 \mathscr{L}(z, \eta) > \alpha I \qquad \text{for all} \quad z \in R^n \quad \text{and} \quad \eta \in E.$$

where $E \subset R^m$ and ∇_1^2 denotes the Hessian of $\mathscr{L}(z, \eta)$ with respect to z.

Letting $z(\eta)$ denote the minimizing value of z in (5.2), we know that

$$\nabla_1 \mathscr{L}(z(\eta), \eta) = 0 \tag{5.7}$$

since the minimization is unconstrained. Moreover, by Theorem 4.1, $z(\cdot)$ is differentiable almost everywhere on E. Hence the chain rule and (5.7) give us

$$D\mathscr{L}[\mu](\eta) = \eta^{\mathrm{T}} g(z(\mu)) \tag{5.8}$$

for almost every $\mu \in E$. Motivated by relation (5.8) and our earlier development in Section 1–3, we study the quantity

$$\mathscr{L}(\eta_1) - \mathscr{L}(\eta_2) - g(z(\eta_2))^{\mathrm{T}}(\eta_1 - \eta_2). \tag{5.9}$$

The following result can be established [7]:

$$(5.9)\begin{cases} \leqq -(\alpha/2)|z(\eta_1) - z(\eta_2)|^2, \\ \geqq -|\nabla g(z(\eta_2))^{\mathrm{T}}(\eta_1 - \eta_2)|^2/\alpha. \end{cases} \tag{5.10}$$

Therefore, proceeding as in Section 2, we get the estimate:

$$\begin{aligned}(\alpha^2/2)|z(\eta^h) - z(\eta^*)|^2 \leqq{} & \alpha g(z(\eta^*))^{\mathrm{T}}(\eta^h - \mu + \eta^* - \mu^h) \\ & + |\nabla g(z(\eta^*))^{\mathrm{T}}|^2 |\mu^h - \eta^*|^2 \end{aligned} \tag{5.11}$$

for all $\mu^h \in K^h$ and $\mu \in K$.

6. RITZ–TREFFTZ FOR OPTIMAL CONTROL

As an application of the results in Section 5, consider the following control problem:

$$\text{minimize} \quad \left\{ C(x, u) = \int_0^1 f(x(t), u(t), t)\, dt \right\}$$

subject to

$$\begin{aligned} &\dot{x}(t) = Ax(t) + Bu(t) && \text{for almost every} \quad t \in [0, 1], \\ &x(0) = x_0, && \\ &G(u(t), t) \leqq 0 && \text{for all} \quad t \in [0, 1], \\ &x \in \mathscr{A}(R^n), \quad u \in L^\infty(R^m), && \end{aligned} \tag{6.1}$$

where $x: [0, 1] \to R^n$, $u: [0, 1] \to R^m$, and $\mathscr{A}$ denotes absolutely continuous functions. The Lagrange dual functional associated with (6.1) is given by

$$\mathscr{L}(p, \lambda) = \inf\{C(x, u) + \langle p, \dot{x} - Ax - Bu\rangle + \langle G(u), \lambda\rangle : x \in \mathscr{A}(R^n), u \in L^\infty(R^m), x(0) = x_0\}, \tag{6.2}$$

and the dual problem becomes [9]

$$\begin{gathered} \sup\{\mathscr{L}(p, \lambda) : (p, \lambda) \in K\}, \\ K \equiv \{(p, \lambda) : p \in BV, p(1) = 0, \lambda \in L^1, \lambda \geqq 0\}, \end{gathered} \tag{6.3}$$

where BV denotes the space of functions with bounded variation. Under suitable assumptions [7, 9], there exist solutions (x^*, u^*) to (6.1) and (p^*, λ^*) to (6.3). Moreover,

$$G(u^*(t), t)^{\mathrm{T}}\lambda^*(t) = 0 \qquad \text{for almost every} \quad t \in [0, 1] \tag{6.4}$$

and (x^*, u^*) achieve the minimum in (6.2) for $(p, \lambda) = (p^*, \lambda^*)$.

Now let $K^h \subset K$ denote the subset consisting of continuous, piecewise linear functions p such that $p(1) = 0$ and piecewise constant functions λ such that $\lambda \geqq 0$. As usual, the superscript h denotes the maximum grid interval. Consider the following approximation to (6.3):

$$\sup\{\mathscr{L}(p, \lambda) : (p, \lambda) \in K^h\}. \tag{6.5}$$

Suppose that (6.5) has the solution $(p^h, \lambda^h) \in K^h$ and that (x^h, u^h) achieve the minimum in (6.2) for $(p, \lambda) = (p^h, \lambda^h)$. Let us estimate the errors $(x^h - x^*)$ and $(u^h - u^*)$.

Assume that $f, G \in C^2$, the components of $G(\cdot, t)$ are convex for all $t \in [0, 1]$, and there exists $\alpha > 0$ such that

$$\nabla_1^2 f(z, t) > \alpha I \tag{6.6}$$

for all $z \in R^{n+m}$ and $t \in [0, 1]$. Integrating by parts in (6.2), it can be shown that [7]

$$\mathscr{L}(p, \lambda) = -p(0)^{\mathrm{T}}x_0 + \int_0^1 F(t)\, dt, \tag{6.7}$$

where

$$F(t) = \inf\{f(x, u, t) - \dot{p}(t)^{\mathrm{T}}x - p(t)^{\mathrm{T}}(Ax + Bu) + G(u, t)^{\mathrm{T}}\lambda(t) : x \in R^n, u \in R^m\}. \tag{6.8}$$

That is, $\mathscr{L}(p, \lambda)$ can be expressed as the integral of a pointwise minimum.

Moreover, with the identifications

$$\eta = \begin{bmatrix} \dot{p}(t) \\ p(t) \\ \lambda(t) \end{bmatrix}, \qquad z = \begin{bmatrix} x \\ u \end{bmatrix},$$

and

$$g(z) = \begin{bmatrix} -x \\ -Ax - Bu \\ G(u, t) \end{bmatrix}, \tag{6.9}$$

we can apply the results of Section 5 for each time t and integrate over $t \in [0, 1]$ to get

$$\begin{aligned}(\alpha/2)\{\|x^h - x\|_{L^2}^2 + \|u^h - u^*\|_{L^2}^2\} \\ \leqq \langle G(u^*), \lambda^* - \mu^h\rangle + C\{\|\dot{p}^* - \dot{q}^h\|_{L^2}^2 + \|p^* - q^h\|_{L^2}^2 + \|\lambda^* - \mu^h\|_{L^2}^2\}\end{aligned} \tag{6.10}$$

for all $(q^h, \mu^h) \in K^h$ satisfying $q^h(0) = p^*(0)$ where C depends on A, B, and $\nabla_1 G(u^*(\cdot), \cdot)$.

Now suppose that $(\dot{p}^*, \lambda^*)$ are Lipschitz continuous [6], select $q^h = p^{\mathrm{I}}$, and let $\mu^h = \lambda^{\mathrm{I}}$ = the piecewise constant function agreeing with the minimum value of λ^* on each grid interval. Since $G(u^*(t), t) \leqq 0 \leqq \lambda^*(t)$ for all $t \in [0, 1]$ and (6.4) holds, we conclude that

$$\langle G(u^*), \lambda^* - \lambda^{\mathrm{I}}\rangle = 0 \tag{6.11}$$

while the remaining terms on the right side of (6.10) are $O(h^2)$. To summarize,

$$\|x^h - x^*\|_{L^2}, \|u^h - u^*\|_{L^2} = O(h). \tag{6.12}$$

(For additional results in this area, see [7 and 11].

7. SEMIDUAL METHODS IN OPTIMAL CONTROL

In the previous section, we introduced a dual multiplier for both the differential equation and the control constraint. Now let us consider a semidual approach where a dual multiplier is only used for the differential equation. In particular, let us consider the problem

$$\text{minimize} \quad \left\{ C(x, u) = \frac{1}{2} \int_0^1 [x(t)^T Q x(t) + u(t)^T R u(t)] \, dt \right\}$$

subject to

$$\dot{x}(t) = Ax(t) + Bu(t) \qquad \text{for almost every} \quad t \in [0, 1], \tag{7.1}$$

$$x(0) = x_0,$$

$$G(u(t), t) \leqq 0 \qquad \text{for all} \quad t \in [0, 1],$$

$$x \in \mathscr{A}(R^n), \qquad u \in L^\infty(R^m).$$

Define the set

$$U = \{u \in L^\infty(R^m) : G(u(t), t) \leqq 0 \text{ for all } t \in [0, 1]\},$$

and the semidual functional

$$\mathscr{L}(p) = \inf\{C(x, u) + \langle p, \dot{x} - Ax - Bu \rangle : x \in \mathscr{A}(R^n), x(0) = x_0, u \in U\}. \tag{7.2}$$

The semidual problem becomes

$$\sup\{\mathscr{L}(p) : p \in K\} \qquad \text{where} \quad K \equiv \{p \in BV : p(1) = 0\}. \tag{7.3}$$

Suppose that Q and R are positive definite; hence, it follows from [7, 9] that there exist solutions (x^*, u^*) to (7.1) and p^* to (7.3). Moreover, (x^*, u^*) achieve the minimum in (7.2) for $p = p^*$ and the following adjoint equation and minimum principle hold:

$$\dot{p}^*(t) = -A^T p^*(t) + Q x^*(t) \qquad \text{for almost every} \quad t \in [0, 1], \tag{7.4}$$

$$M(u^*(t), p^*(t)) = \min\{M(u, p^*(t)) : u \in R^m, G(u, t) \leqq 0\}$$
$$\text{for almost every} \quad t \in [0, 1], \tag{7.5}$$

where

$$M(u, p(t)) \equiv \tfrac{1}{2} u^T R u - p(t) B u.$$

Let $K^h \subset K$ be the space of continuous, piecewise linear polynomials, and consider the following approximation to (7.3):

$$\sup\{\mathscr{L}(p) : p \in K^h\}. \tag{7.6}$$

Suppose that p^h solves (7.6) and that $(x, u) = (x^h, u^h)$ achieves the minimum in (7.2) for $p = p^h$. We study the errors $(x^* - x^h)$ and $(u^* - u^h)$.

Integrating by parts in (7.2), it can be shown that

$$\mathscr{L}(p) = \mathscr{L}_1(p) + \mathscr{L}_2(p), \tag{7.7}$$

where

$$\mathscr{L}_1(p) = -\frac{1}{2}\int_0^1 (\dot{p}(t) + A^{\mathrm{T}}p(t))^{\mathrm{T}}Q(\dot{p}(t) + A^{\mathrm{T}}p(t))\,dt, \tag{7.8}$$

$$\mathscr{L}_2(p) = -p(0)^{\mathrm{T}}x_0 + \int_0^1 \gamma(p(t))\,dt, \tag{7.9}$$

$$\gamma(p(t)) \equiv \inf\{M(u, p(t)) : u \in R^m,\ G(u, t) \leqq 0\}. \tag{7.10}$$

Furthermore, it can be shown that (x^h, u^h) satisfy relations (7.4)–(7.5) with superscript $*$ replaced by h.

Since $\mathscr{L}_1(p)$ is a quadratic functional and $\mathscr{L}_2(p)$ is possibly nondifferentiable, we can apply the bound (3.9) choosing $v = p^h$ and $v^h = p^{\mathrm{I}}$:

$$\begin{aligned} -\mathscr{L}_1(p^h - p^*) \leqq {} & -\mathscr{L}_1(p^{\mathrm{I}} - p^*) \\ & + \langle \dot{p}^* + A^{\mathrm{T}}p^*, Q[\dot{p}^{\mathrm{I}} - \dot{p}^* + A^{\mathrm{T}}(p^{\mathrm{I}} - p^*)] \rangle \\ & + \mathscr{L}_2(p^*) - \mathscr{L}_2(p^{\mathrm{I}}). \end{aligned} \tag{7.11}$$

If there exists a continuous control $\bar{u}$ such that $G(\bar{u}(t), t) \leqq 0$ for all $t \in [0, 1]$, it follows from (7.4)–(7.5) that $\ddot{p}^* \in L^\infty$. Integrating by parts the $\dot{p}^{\mathrm{I}} - \dot{p}^*$ term appearing in (7.11) and applying the interpolation estimate $\|p^* + p^{\mathrm{I}}\|_{L^2} = O(h^2)$, we get:

$$\begin{aligned} -\mathscr{L}_1(p^h - p^*) & \leqq ch^2 + \mathscr{L}_2(p^*) - \mathscr{L}_2(p^{\mathrm{I}}) \\ & \leqq ch^2 + \int_0^1 (p^{\mathrm{I}}(t) - p^*(t))^{\mathrm{T}}Bu^{\mathrm{I}}(t)\,dt, \end{aligned} \tag{7.12}$$

where $u = u^{\mathrm{I}}(t)$ achieves the minimum in (7.10) for $p(t) = p^{\mathrm{I}}(t)$. Applying (4.7), $\|u^{\mathrm{I}}\|_{L^\infty}$ is bounded uniformly in h, and (7.12) gives us:

$$-\mathscr{L}_1(p^h - p^*) \leqq ch^2. \tag{7.13}$$

Since $p^h(1) = p^*(1) = 0$, we show in [10] that there exists a constant $\beta > 0$ such that

$$\beta\|p^h - p^*\|_{H^1}^2 \leqq -\mathscr{L}_1(p^h - p^*). \tag{7.14}$$

Therefore, we obtain

$$\|p^h - p^*\|_{H^1} \leqq ch. \tag{7.15}$$

Since (p^h, x^h) satisfy (7.4) with superscript $*$ replaced by h, (7.14) also implies that

$$\|x^* - x^h\|_{L^2} \leqq ch. \tag{7.16}$$

Finally we combine the inequality (7.14), relation (7.5) with superscript $*$ replaced by h, and the quadratic program stability bound (4.4) to get:

$$\|u^* - u^h\|_{L^2} \leqq ch. \tag{7.17}$$

8. LIPSCHITZ CONTINUITY FOR CONSTRAINED PROCESSES

In the previous sections, we studied the effect on a program of replacing a constraint set K by an approximation K^h, and the dependence of the solution to a program on a parameter appearing in the cost. Now we study stability (or more specifically Lipschitz continuity) in the abstract setting of a "constrained process."

A constrained process can be described as follows: Let $\mathscr{S}$ be a Banach space, $\mathscr{D}$ be a convex subset of a Banach space, $z: \mathscr{D} \to \mathscr{S}$, and $c: \mathscr{D} \to$ (power set of $\{1, \ldots, n\}$). Two examples of constrained processes are the following:

(1) A control problem such as (6.1) with optimal solution (x^*, u^*); we choose $\mathscr{D} = [0, 1]$, the time interval, $z(d) = (x^*(d), u^*(d))$, and $c(d) =$ indices of the binding constraints associated with $u^*(d)$.

(2) A mathematical program such as

$$\text{minimize} \quad \{f(x, \xi): x \in R^n, g(x, \xi) \leqq 0\}, \tag{8.1}$$

where $\xi \in R^m$ is a given parameter. Suppose that there exists a solution $x(\xi)$ to (8.1) for $\xi \in \mathscr{D} \subset R^m$, a convex subset. We then choose $z(d) = x(d)$ and $c(d) =$ indices of the binding constraints associated with $x(d)$.

Our goal is to estimate the Lipschitz constant for $z(\cdot)$. First consider the program (8.1), and let $\# c(d)$ denote the number of elements in the set $c(d)$. Recall that the Kuhn–Tucker conditions give us a system of $n + \# c(\xi)$ equations in the same number of unknowns: $x(\xi)$ and the dual multipliers associated with binding constraints. If $c(\xi_1) = c(\xi_2)$, then $|x(\xi_1) - x(\xi_2)|$ can often be estimated in terms of $|\xi_1 - \xi_2|$ using the implicit function theorem. Similar results apply to the control problem but with the Kuhn–Tucker conditions replaced by the Pontryagin minimum principle and the adjoint equation.

On the other hand, suppose that $c(\xi_1) \neq c(\xi_2)$. A key result on this subject is given in [6]; namely, a Lipschitz constant that is valid for compatible parameters (where the binding constraints agree) is also valid for noncompatible parameters. To be more precise, assume that $z(\cdot)$ is continuous, and $c(\cdot)$ has the following property: If $\{d_k\} \subset \mathscr{D}$, $d_k \to d \in \mathscr{D}$ as $k \to \infty$, and $I \subset c(d_k)$ for all k, then $I \subset c(d)$.

Given $d, e \in \mathscr{D}$, we define the segment

$$[d, e] = \{(1 - \lambda)d + \lambda e : 0 \leqq \lambda \leqq 1\}$$

and we say that (d, e) are *compatible* if $c(d) = c(e)$ and $c(\delta) \subset c(d)$ for all $\delta \in [d, e]$.

Theorem 8.1. *If γ satisfies*

$$\|z(d) - z(e)\|_{\mathscr{S}} \leqq \gamma \|d - e\|_{\mathscr{D}} \tag{8.2}$$

for all compatible data $(d, e) \in \mathscr{D} \times \mathscr{D}$, *then* γ *satisfies* (8.2) *for all data* $(d, e) \in \mathscr{D} \times \mathscr{D}$.

The application of Theorem 8.1 to derive Lipschitz continuity results for both mathematical programs and optimal control problems is given in [6].

REFERENCES

1. H. Brézis, Nouveaux théorèmes de régularité pour les problèmes unilatéraux, *Recontre Physiciens Théoriciens Mathématiciens, Strasbourg* **12** (1971).
2. H. Brézis, Seuil de régularité pour certain problèmes unilateraux, *C. R. Acad. Sci., Ser. A* **273** (1971), 35–37.
3. F. Brezzi, W. W. Hager, and P. A. Raviart, Error estimates for the finite element solution of variational inequalities, *Numer. Math.* **28** (1977), 431–443.
4. R. S. Falk, Error estimates for the approximation of a class of variational inequalities, *Math. Comp.* **28** (1974), 308–312.
5. R. Glowinski, Sur l'approximation d'une inéquation variationnelle elliptique, *RAIRO Anal. Numer.* **10**, No. 12 (1976), 13–30.
6. W. W. Hager, Lipschitz continuity for constrained processes, *SIAM J. Control Optim.* **17** (1979), 321–338.
7. W. W. Hager, Convex control and dual approximations, *Control Cybernet.* **8** (1979), 5–22.
8. W. W. Hager and G. Ianculescu, Semidual approximations in optimal control (to appear).
9. W. W. Hager and S. K. Mitter, Lagrange duality theory for convex control problems, *SIAM J. Control Optim.* **14** (1976), 843–856.
10. W. W. Hager, Rates of Convergence for Discrete Approximations to Problems in Control Theory, Ph.D. Thesis, Mass. Inst. of Tech., Cambridge, Massachusetts, 1974.
11. W. W. Hager, The Ritz–Trefftz method for state and control constrained optimal control problems, *SIAM J. Numer. Anal.* **12** (1975), 854–867.
12. J. L. Lions, "Optimal Control of Systems Governed by Partial Differential Equations" (transl. by S. K. Mitter), Springer-Verlag, Berlin and New York, 1971.

This research was supported in part by Office of Naval Research Grant N00014-76-C-0369, and the Naval Surface Weapons Center, Silver Springs, Maryland.

DEPARTMENT OF MATHEMATICS
CARNEGIE–MELLON UNIVERSITY
PITTSBURGH, PENNSYLVANIA