Communications in Applied Analysis **12** (2008), no. 2, 221–234

DYNAMICS OF FLOW IN POROUS MEDIA AND ITS CONTROL APPLIED TO OPTIMAL MANAGEMENT OF UNDERGROUND RESOURCES

N. U. AHMED

School of Information Technology and Engineering, University of Ottawa
Ottawa, Ontario, Canada
E-mail: ahmed@site.uottawa.ca

ABSTRACT. We consider the control of a porous media equation described by a class of degenerate nonlinear partial differential equations. We formulate control problems for aquifers (under ground water reservoirs) and present necessary conditions of optimality. The necessary conditions involve simultaneous optimization of location of wells in aquifers and ground water extraction rate from each of the wells. The method presented also applies to optimal extraction of other underground resources.

Keywords. Porous media equations, nonlinear partial differential equations, aquifers, optimal control

AMS (MOS) Subject Classification. 34A36, 34A60, 49J24, 93C10

1. INTRODUCTION

In the study of hydrological systems, in particular reservoir engineering, the dynamics of flow through porous media plays an important role in the design of extraction programs of water resources from aquifers. In particular, it is necessary to design and locate extraction wells and operate the program so as to maximize exploitation of resources while avoiding or reducing contamination from nearby landfills and sea water. Other applications include hi-tech methods of oil and gas recovery from underground storage. The methodology developed here is generally applicable to mining of underground resources like gas, oil, water etc.

The rest of the paper is organized as follows. In Section 2, the basic mathematical model of flow through porous media described by a class of nonlinear partial differential equations is presented along with the formulation of a control problem considered in this paper. Existence and regularity properties of solutions are discussed in Section 3. In Section 4, necessary conditions of optimality are developed. In the concluding remarks, the results of this paper are compared with those of our

Received June 6, 2008

previous paper on the same topic [1]. The paper is concluded with an indication to some open problems.

2. BASIC SYSTEM MODEL FOR EXPLOITED AQUIFERS

Let Σ denote a porous media filled with liquid or gas. The fluid can diffuse through the media from locations of higher pressure to those of lower ones. Let $\rho(t)(\cdot) \equiv \rho(t,\cdot)$ denote the spatial distribution of density of the fluid in Σ at time $t \geq 0$. In other words, ρ represents the amount of fluid contained in an unit volume of the porous medium. The temporal and spatial evolution of the density ρ is governed by a nonlinear partial differential equation of the form

$$(\partial/\partial t)(c\rho) - \triangle\Phi(\rho) = f \text{ on } I \times \Sigma, \tag{2.1}$$

$$\Phi(\rho) = 0 \text{ on } I \times \partial\Sigma, \tag{2.2}$$

$$\rho(0) = \rho_0,\ \xi \in \Sigma, \tag{2.3}$$

where $c(\xi)$ is a scalar valued function defined on Σ and taking values $0 \leq c(\xi) \leq 1$, representing porosity of the medium. This is an initial-boundary value problem where the operator $\triangle$ is the Laplacian in R^n $(n \leq 3)$. The function Φ is related to the flux of vector flow rate $J \equiv \nabla\Phi$. In general, Φ is a measurable function of $\xi \in \Sigma$ and a monotone increasing function of the density ρ. Without loss of generality, we may assume that Φ is independent of the spatial variable $\xi \in \Sigma$. For porous media, the typical form of Φ is $\Phi(\rho) \equiv \beta\rho^\gamma$, $\beta > 0$, $\gamma = 1 + (1/\alpha)$, $0 < \alpha < \infty$, and the pressure $P \equiv F(\rho) = b\rho^{\gamma-1}$, $b > 0$. The exact expression for Φ is dependent on Darcy's law [1, 4]. The function f represents the natural source term giving the rate at which resources are replenished. Detailed construction of the model based on physical arguments can be found in [4, p. 253] and also [5].

Remark 2.1 Since $\Phi'(0) = 0$, the system is degenerate parabolic (not strictly parabolic). Further, if the set $\Sigma^o \equiv \{\xi \in \Sigma : c(\xi) = 0\}$ is nonempty and has positive Lebesgue measure, the system (2.1)–(2.3) may change type. It is elliptic on Σ^o and parabolic on $\Sigma \setminus \Sigma^o$. [See also Remark 3.3.]

Aquifers are natural underground sources of water. Water can be extracted by drilling wells into the aquifer and pumping it on to the surface. For exploitation of water or other resources, a number of wells are drilled at strategically important locations $\{Z_i\} \in \Sigma$. Suppose there are N wells occupying spatial domains $\cup_{i=1}^N G_i(Z_i)$ in Σ with $G_i(Z_i) \cap G_j(Z_j) = \emptyset$ for $i \neq j$. Let u_i, $i = 1, 2, \ldots, N$, denote the pumping or extraction rate of the i-th well. Thus, the total extraction rate is given by the sum of the rates of individual wells. The system equation, representing the controlled aquifer, is then given by

$$(\partial/\partial t)(c\rho) - \triangle\Phi(\rho) = f - B(Z, u) \quad \text{for } (t, \xi) \in I \times \Sigma, \tag{2.4}$$

$$\Phi(\rho) = 0, \quad \text{for } (t,\xi) \in I \times \partial\Sigma, \tag{2.5}$$

$$\rho(0,x) = \rho_0(\xi), \quad \text{for } \xi \in \Sigma, \tag{2.6}$$

where the operator B is given by

$$B(Z,u)(t,\xi) \equiv \sum_{i=1}^{N} \Upsilon_{G_i(Z_i)}(\xi)u_i(t), \qquad \xi \in \Sigma, \quad t \in I, \tag{2.7}$$

with Υ_{G_i} denoting the characteristic function of the set G_i.

Equation (2.4) describes the flow dynamics subject to controlled withdrawal of resources, equation (2.5) specifies the required boundary condition, and equation (2.6) gives the initial distribution of the density. It is important to protect coastal aquifers from contamination by sea (salt) water. One way of achieving this is to maintain the internal pressure close to a desired level which is determined by the hydrostatic pressure on the outer boundary of the reservoir. Salt water intrusion is expected if the external boundary pressure is higher than that of the internal pressure. Thus, an appropriate objective functional can be chosen as follows:

$$J(u) = \int_I \left\{ -\sum_{i=1}^{N} \gamma_i u_i(t) + (\lambda/2) \int_\Sigma \big(F(\rho) - F(\rho_d)\big)^2 d\xi \right\} dt, \tag{2.8}$$

for $\gamma_i \geq 0$ and $\lambda > 0$, where ρ is the solution of the system (2.4)–(2.6) corresponding to the control policy u, and ρ_d is the desirable density giving the pressure $P_d \equiv F(\rho_d)$. This is the desirable internal pressure level that is expected to prevent intrusion of salt water, or in general, contamination by external toxic sources. The objective is to find the best locations $\{Z_i^o\}$ of the wells and an withdrawal rate (water extraction program) $u^o \equiv \{u_i^o\}$ with values $u^o(t) \in U \subset R_+^N \equiv \{v \in R^N : 0 \leq v_i < \infty,\ t \in I,\ i = 1, 2, \ldots, N\}$ that minimizes the cost functional J. This problem was partially solved in a recent paper [1] of the author where the locations of wells were assumed to be given (fixed) and U assumed to be compact and convex. In this note, we include also the location of wells as decision variables like the extraction rates. Examining the objective functional, it is clear that minimizing this functional is equivalent to finding a balance between extraction of resources and avoidance of contamination. This is due to the fact that excessive withdrawal of water reduces the water table and the internal pressure leading to penetration of salt water from the surrounding region.

3. EXISTENCE AND REGULARITY OF SOLUTIONS

Define the operator:

$$E \equiv (-\triangle)^{-1} \tag{3.1}$$

subject to homogeneous Dirichlet boundary condition. Using this operator, system (2.4)–(2.6) can be written as an abstract differential equation

$$(d/dt)(E(c\rho)) + \Phi(\rho) = E\tilde{f} \equiv E(f - B(Z,u)), \qquad \rho(0) = \rho_0, \tag{3.2}$$

on suitable Banach spaces where we have set $\tilde{f} \equiv (f - B(Z,u))$. For this problem, suitable Banach spaces are the Gelfand triple $V \hookrightarrow H \hookrightarrow V^*$ where $H = L_2(\Sigma)$ and it is identified with its dual, $V \equiv W_0^{1,p}(\Sigma)$ and $V^* = W^{-1,q}(\Sigma)$, with q being the conjugate of p, that is, $(1/p\ + 1/q) = 1)$, and $1 < q \leq 2 \leq p < \infty$. Since E is a positive self adjoint operator in $H \equiv L_2(\Sigma)$, its square root is well defined and hence the operator $E_c^{1/2}(\varphi) \equiv E^{1/2}(c\varphi)$ is well defined. We introduce the vector space W as follows

$$W \equiv \{\rho : E_c^{1/2}\rho \in L_p(I,V)\ \&\ E_c^{1/2}\dot{\rho} \in L_q(I,V^*)\}.$$

Furnished with the norm topology given by

$$\| \rho \|_W \equiv \| E_c^{1/2}\rho \|_{L_p(I,V)} + \| E_c^{1/2}\dot{\rho} \|_{L_q(I,V^*)},$$

it is a Banach space, and the embedding $E_c^{1/2}W \hookrightarrow C(I,H)$ is continuous. The proof of this embedding is similar to that given in [3, Theorem 1.2.15, p. 27].

Following similar approach as given in [2, Theorem 1, p. 90], [3, Theorem 2.5.1, p. 107], it may be possible to give an operator theoretic proof for existence of solution for the system (3.2). In Showalter [4], considering $c(\xi) \equiv 1$, a simple and elegant proof is given on page 142, Example 6.6. Here we present a different and constructive proof based on finite dimensional projections and limiting arguments without assuming $c(\xi) \equiv 1$. This classical approach is also useful both for approximation and computation as required for real physical applications.

Theorem 3.1 Consider the system (3.2) and suppose

(A1): there exists a constant $c_1 \in (0,1]$ such that $\inf_{\xi\in\Sigma} c(\xi) = c_1$.

(A2): The function $\Phi : R \longrightarrow R$ is continuous and maximal monotone.

(A3): There exists a number $p \geq 2$ and constants $c_2 > 0$, $c_3 > 0$, such that

$$(1) : \Phi(r)r \geq c_2|r|^p \quad \text{and} \quad (2) : |\Phi(r)| \leq c_3|r|^{p-1}\ \forall\ r \in R.$$

Then, for each $\rho_0 \in V^*$ satisfying $E_c^{1/2}\rho_0 \in H$ and $\tilde{f} \in L_q(I,V^*)$ satisfying $E^{1/2}\tilde{f} \in L_2(I,H)$, system (3.2) has a unique weak solution $\rho \in L_p(I, L_p(\Sigma_c))$, in the sense that the following identity

$$-\int_I \langle \rho, (d/dt)E_c\phi\rangle dt + \int_I \langle \Phi(\rho), \phi\rangle_{V^*,V} dt = (E_c^{1/2}\rho(0), E_c^{1/2}\phi(0))_H + \int_I \langle E^{1/2}\tilde{f}, \phi\rangle_{V^*,V} dt \tag{3.3}$$

holds for all $\phi \in L_p(I,V)$ satisfying $\phi(T) = 0$. Further, the solution has the regularity properties $E_c^{1/2}\dot{\rho} \in L_q(I,V^*)$ and $E_c^{1/2}\rho \in L_\infty(I,H) \cap C(I,H)$.

Proof. Our proof is based on a-priori bounds, Galerkin approximation, and limiting arguments. We start with the following a-priori bounds. Let

$$L_p(\Sigma_c) \equiv \left\{ \varphi : \int_\Sigma c(\xi)|\varphi(\xi)|^p d\xi \equiv |\varphi|^p_{L_p(\Sigma_c} < \infty \right\}.$$

This is a Banach space with the dual given by $L_q(\Sigma_c)$, $(1/p) + (1/q) = 1$, $\infty > p \geq q > 1$. Scalar multiplying equation (3.2) by $c\rho$ and integrating by parts, it is easy to derive the following estimates: there exist positive constants c_4, c_5 dependent possibly on the parameters $\{c_1, c_2, c_3, p, q, T\}$ so that

$$|E_c^{1/2}\rho(t)|^2_H \leq c_4 \left\{ |E_c^{1/2}\rho_0|^2_H + \int_I |E^{1/2}\tilde{f}|^2_H dt \right\} < \infty, \ \forall\, t \in I, \tag{3.4}$$

$$|E_c^{1/2}\rho(t)|^2_H + 2c_2 \int_0^t |\rho(s)|^p_{L_p(\Sigma_c)} ds \leq c_5 \left\{ |E_c^{1/2}\rho_0|^2_H + \int_I |E^{1/2}\tilde{f}|^2_H dt \right\} < \infty, \ \forall\, t \in I. \tag{3.5}$$

Straight computation shows that one can take $c_4 = e^T$ and $c_5 = (1 + Te^T)$. Further, there exists a positive number c_6, dependent possibly on the parameters mentioned above, so that

$$\| E_c^{1/2}\dot{\rho} \|_{L_q(I,V^*)} \leq c_6 \left\{ \| \rho \|^{p/q}_{L_p(I,L_p(\Sigma))} + \| E^{1/2}\tilde{f} \|_{L_q(I,V^*)} \right\}. \tag{3.6}$$

By virtue of assumption (A1), and the fact that $0 \leq c(\xi) \leq 1$ for $\xi \in \Sigma$, $L_p(\Sigma_c)$ is equivalent to $L_p(\Sigma)$ and hence it follows from the above estimates that if ρ is any solution of equation (3.2), then

$$E_c^{1/2}\rho \in L_\infty(I, H), \qquad \rho \in L_p(I, L_p(\Sigma_c)), \qquad E_c^{1/2}\dot{\rho} \in L_q(I, V^*).$$

Let ρ^1, ρ^2 be two solutions corresponding, respectively to the data ρ_0^1, ρ_0^2 with $E^{1/2}\rho_0^i \in H (i = 1, 2)$, and $\tilde{f}_i, (i = 1, 2) \in L_q(I, V^*)$ with $E^{1/2}\tilde{f}_i \in L_2(I, H), i = 1, 2$. Then using the assumption that Φ is monotone, it follows from similar computations that there exists a positive constant c_7 such that

$$\begin{aligned} &\sup_{t\in I}\{|E_c^{1/2}\rho^1(t) - E_c^{1/2}\rho^2(t)]|^2_H\} \\ &\leq c_7 \left\{ |E_c^{1/2}\rho_0^1 - E_c^{1/2}\rho_0^2]|^2_H + \int_I |E^{1/2}\tilde{f}_1 - E^{1/2}\tilde{f}_2]|^2_H dt \right\}. \end{aligned} \tag{3.7}$$

From this follows Lipschitz continuity of the solution map with respect to input data and hence the uniqueness. Now we prove existence. Since the embedding $V \hookrightarrow H$ is compact, there exists a sequence $\{v_i\} \in V$ which is a complete basis for the triple $\{V, H, V^*\}$. Using this basis and projecting the infinite dimensional system (3.2) to the sequence of finite dimensional systems,

$$\begin{aligned} &\left\langle (d/dt)E_c^{1/2}\left(\sum_{i=1}^n x_i^n v_i\right), E_c^{1/2}v_j \right\rangle + \left\langle \Phi\left(\sum_{i=1}^n x_i^n v_i\right), cv_j \right\rangle \\ &\qquad = \langle \tilde{f}, cv_j \rangle, \quad 1 \leq j \leq n, \quad n \in N, \end{aligned} \tag{3.8}$$

we obtain the following ordinary differential equation in R^n, $n \in N$, given by

$$\Gamma\dot{x} = -\tilde{\Phi}(x) + g, \qquad t \in I, \quad x(0) = x_0 \tag{3.9}$$

where $x(t) \in R^n$ is given by the vector $x(t) \equiv (x_i^n(t),\, i = 1, 2, \ldots, n)$,

$$\Gamma \equiv \{\gamma_{j,i} \equiv (E_c^{1/2} v_j, E_c^{1/2} v_i), 1 \le j, i \le n\}$$

is a symmetric positive matrix of dimension n, $\tilde{\Phi}$ is a n-vector valued function given by

$$\tilde{\Phi}(x) \equiv \left\{ \left\langle \Phi\left(\sum_{i=1}^{n} x_i^n v_i\right), cv_j \right\rangle, 1 \le j \le n \right\},$$

and

$$g = \{\langle \tilde{f}, cv_j \rangle, j = 1, 2, \ldots, n\} \quad \text{and} \quad x_0 = \{(\rho_0, v_j), 1 \le j \le n\}.$$

Note that equation (3.9) is the finite dimensional approximation of the infinite dimensional system (3.2). The finite dimensional space here is given by the closure of the linear span of $\{v_i, 1 \le i \le n\}$ which is evidently isomorphic to R^n. Since Γ is a symmetric positive matrix and $\tilde{\Phi}$ is maximal monotone from R^n to R^n, for every $\beta > 0$, the operator $(\Gamma + \beta\tilde{\Phi})$ is a maximal monotone map in R^n and hence equation (3.9) has a unique solution $x \in C(I, R^n)$. This is easily justified as follows. Partition the interval I into a finite number, say m, of disjoint intervals $\{I_i \equiv (t_i, t_{i+1}], i = 0, 1, \ldots, m-1\}$ with (Lebesgue measure) $\ell(I_i) = \beta_m$ for all $i = 0, 1, \ldots, m-1$. Use the implicit scheme and interpolation to construct the approximate solution defined by the expression

$$x_m(t) \equiv ((t_{i+1} - t)/\beta_m)\nu_m(t_i) + ((t - t_i)/\beta_m)\nu_m(t_{i+1}), \quad t \in I_i, \quad i = 0, 1, \ldots, m-1,$$

where the nodes $\{\nu_m(t_i), i = 0, 1, \ldots, m-1\}$, with $\nu_m(t_0) = \nu_m(0) = x_0$, are given by

$$\nu_m(t_{i+1}) \equiv (\Gamma + \beta_m\tilde{\Phi})^{-1}\left(\Gamma\nu_m(t_i) + \int_{t_i}^{t_{i+1}} g(s)ds\right), \quad i = 0, 1, \ldots, m-1. \tag{3.10}$$

Since g is locally integrable, it is easy to verify that $\{x_m\} \in C(I, R^n)$ and that it is a Cauchy sequence, and so there exists a function $x \in C(I, R^n)$ to which the sequence converges uniformly on I as $m \to \infty$. The components of this vector valued function x are denoted by $x_i^n, i = 1, 2, \ldots n$. Thus the function given by $\rho^n \equiv \sum_{i=1}^n x_i^n(t) v_i$ is the unique solution of the system of finite dimensional equations (3.8) which is equivalent to

$$\langle (d/dt)E_c^{1/2}\rho^n, E_c^{1/2} v_j \rangle + \langle \Phi(\rho^n), cv_j \rangle = \langle \tilde{f}, cv_j \rangle, \quad 1 \le j \le n. \tag{3.11}$$

Now using the a-priori bounds (3.4)–(3.5), and the fact that the Banach spaces involved are reflexive, there exists a subsequence relabeled, as the original sequence, and an element $\rho^o \in L_p(I, L_p(\Sigma_c))$ satisfying $E_c^{1/2}\rho^o \in L_\infty(I, H)$, $E_c^{1/2}\dot{\rho}^o \in L_q(I, V^*)$ such that

$$E_c^{1/2}\rho^n \xrightarrow{w^*} E_c^{1/2}\rho^o \text{ in } L_\infty(I, H), \tag{3.12}$$

$$E_c^{1/2}\dot{\rho}^n \xrightarrow{w} E_c^{1/2}\dot{\rho}^o \text{ in } L_q(I,V^*), \tag{3.13}$$

$$\Phi(\rho^n) \xrightarrow{w} \Phi(\rho^o) \text{ in } L_q(I\times\Sigma_c). \tag{3.14}$$

Since Φ is a nonlinear function, the proof of the convergence result (3.14) is nontrivial. However, it follows from the assumptions (A2), (A3), continuity and monotonicity of the function Φ, and the well known Mazur's theorem that states that every weakly convergent sequence has a suitable convex combination that converges strongly. Indeed, from the a-priori bound (3.5) it is clear $\{\rho^n\}$ is contained in a bounded subset of $L_p(I\times\Sigma_c)$ and further it follows from assumption (A3)(2) that $\{\Phi(\rho^n)\}$ is contained in a bounded subset of $L_q(I\times\Sigma_c)$. Since $1<q\le 2\le p<\infty$, these are reflexive Banach spaces and therefore there exist $\rho^o\in L_p(I\times\Sigma_c)$ and $\eta\in L_q(I\times\Sigma_c)$ such that, along a subsequence if necessary, $\rho^n \xrightarrow{w} \rho^o$ and $\Phi(\rho^n)\xrightarrow{w}\eta$. By Mazur's theorem, there exists $\{\alpha_k^n, 1\le k\le n, \alpha_k^n\ge 0, \sum\alpha_k^n=1\}$ for all $n\in N$ such that $\zeta^n\equiv\sum_{k=1}^n\alpha_k^n\rho^k \xrightarrow{s} \rho^o$ in $L_p(I\times\Sigma_c)$. On the other hand, for any $\rho\in L_p(I\times\Sigma_c)$, it follows from monotonicity of Φ that

$$\int_I\langle\Phi(\rho)-\Phi(\zeta^n),\rho-\zeta^n\rangle dt\ge 0.$$

Letting $n\to\infty$, we obtain

$$\int_I\langle\Phi(\rho)-\eta,\rho-\rho^o\rangle dt\ge 0,\quad \forall\ \rho\in L_p(I\times\Sigma_c).$$

Take $\rho=\rho^o+\varepsilon w$ for arbitrary $w\in L_p(I\times\Sigma_c)$ and $\varepsilon>0$. Using this ρ in the above inequality, we have

$$\int_I\langle\Phi(\rho^o+\varepsilon w)-\eta,w\rangle dt\ge 0,\quad \forall\ w\in L_p(I\times\Sigma_c),$$

and $\varepsilon>0$. Letting $\varepsilon\downarrow 0$ it follows from continuity of Φ that

$$\int_I<\Phi(\rho^o)-\eta,w>dt\ge 0,\ \forall\ w\in L_p(I\times\Sigma_c).$$

This implies that $\eta=\Phi(\rho^o)$ proving (3.14). Multiplying equation (3.8) by $z\in C^1(I)$ with $z(T)=0$, and integrating by parts over I, one has

$$\begin{aligned}-\int_I\langle E_c^{1/2}\rho^n,(d/dt)E_c^{1/2}z(t)cv_j\rangle dt+\int_I\langle\Phi(\rho^n),z(t)cv_j\rangle_{V^*,V}dt\\ =(E_c^{1/2}\rho^n(0),E_c^{1/2}z(0)v_j)_H+\int_I\langle\tilde{f},z(t)cv_j\rangle_{V^*,V}dt\end{aligned}\tag{3.15}$$

for $1\le j\le n$, where $\rho^n(0)=\sum_{i=1}^n x_i^n(0)v_i=\sum_{i=1}^n(\rho_0,v_i)v_i$. Letting $n\to\infty$ in (3.15), along a subsequence if necessary, it follows from the convergence results (3.12)–(3.14) that

$$\begin{aligned}-\int_I\langle E_c^{1/2}\rho^o,(d/dt)E_c^{1/2}z(t)v_j\rangle dt+\int_I\langle\Phi(\rho^o),z(t)cv_j\rangle_{V^*,V}dt\\ =(E_c^{1/2}\rho_0,E_c^{1/2}z(0)v_j)_H+\int_I\langle E^{1/2}\tilde{f},z(t)cv_j\rangle_{V^*,V}dt.\end{aligned}$$

This is true for all $j \in N$, and since $\{v_j\}$ is a basis, it follows from this identity that ρ^o is the unique weak solution of equation (3.2). Since $\rho^o \in L_p(I, L_p(\Sigma_c))$ and $L_p(\Sigma_c) \cong L_p(\Sigma) \subset L_2(\Sigma) \equiv H$ and $E^{1/2} : H \to V$, we have $E_c^{1/2}\rho^o \in L_p(I, V)$. We have already seen that $E_c^{1/2}\dot{\rho}^o \in L_q(I, V^*)$. Thus, it follows from the embedding theorem stated in the introduction of this section that $E_c^{1/2}\rho^o \in C(I, H)$. Hence, we have all the regularity properties as stated. This completes our proof. •

Remark 3.2 We have assumed that the porosity coefficient $c(\xi)$, $\xi \in \Sigma$, is bounded away from zero. This is used to ensure that $L_p(\Sigma_c) \cong L_p(\Sigma)$, which, in turn, is used to prove the regularity of solutions as stated in the theorem. It would be interesting to relax this assumption.

Remark 3.3 Recall the set Σ^o as introduced in Remark 2.1. On this set the system is elliptic. Since Φ is strictly monotone, the solution in the elliptic phase is given by

$$\rho(t, \xi) \equiv \Phi^{-1}(E\tilde{f})(t, \xi), \qquad (t, \xi) \in I \times \Sigma^o$$

provided the data satisfies the compatibility condition

$$\lim_{t \downarrow 0} \Phi^{-1}(E\tilde{f})(t, \xi) = \rho_0(\xi)$$

for $\xi \in \Sigma^o$. On the other hand, for physical reasons, it is evident that the fluid content of any part of the medium that has zero porosity must be zero. Hence the initial condition and the data $\tilde{f}$ must be identically zero on Σ^o.

4. NECESSARY CONDITIONS OF OPTIMALITY

As stated earlier, our objective is to find optimal locations $\{Z_i^o, i = 1, 2, \ldots N\}$ for the wells, and the corresponding optimal extraction rates $\{u_i^o, i = 1, 2, \ldots, N\}$ which minimize the cost functional (2.8) subject to dynamic constraints (2.4)–(2.6). Following a similar procedure as given in [1, Theorem 4.1], we can prove existence of optimal control for the problem (2.8). Here, we assume that the optimization problem has a solution and concentrate only on the necessary conditions of optimality. By use of these necessary conditions one can determine the locations of wells and formulate optimum extraction policies.

Without loss of generality, we may assume that the aquifer body represented by Σ is an open connected convex domain. It suffices if it is only given by a finite union of such convex bodies. Let the sets introduced in the introduction be given by $G_i(Z_i) \equiv C_r(Z_i)$ where C_r denotes a cylindrical well of radius $r > 0$ vertically installed at the point $Z_i \in \Sigma$ through which the central axis of the well passes. We let $C_r(Z_i)$ denote only the part inside the aquifer body. Let Σ_0 denote a closed bounded convex subset of Σ contained entirely in its interior with $d(\xi, \partial\Sigma) \geq r$, for all $\xi \in \Sigma_0$. Clearly, we may choose $Z_i \in \Sigma_0$ and take $r > 0$ sufficiently small, so that $\bigcup_{i=1}^N C_r(Z_i) \subset \Sigma$.

For convenience of notation, we write $Z \in \Sigma^N$ (cartesian product of N copies of Σ) for $\{Z_i, i = 1, 2, \ldots, N\}$ with each component $Z_i \in \Sigma \subset R^n$.

Theorem 4.1 Consider the system given by equations (2.4)–(2.6) with the control (2.7) and the cost functional given by (2.8). Suppose the assumptions of Theorem 3.1 hold and $\Sigma_0 \subset \Sigma$ is a closed bounded convex set away from the boundary $\partial\Sigma$ by at least a distance $r > 0$, and $\bigcup C_r(Z_i) \subset \Sigma$. Suppose U is a compact convex subset of R_+^N and let $\mathcal{U}_{ad} \equiv L_\infty(I, U)$ denote the class of admissible controls and suppose $\{p, q\}$ is the conjugate pair satisfying $5/2 < p \le 4$. Then for the triple $(Z^o, u^o, \rho^o) \in \Sigma_0^N \times \mathcal{U}_{ad} \times L_p(I, V)$ to be optimal it is necessary that there exists a $\psi \in L_p(I, V)$ with $E_c^{1/2}\psi \in C(I, H)$ satisfying the following inequalities and equations:

$$dJ(Z^o, Z - Z^o; u^o, u - u^o) \equiv -\int_I \Big\{\sum \big(r_i + \int_{C_r(0)} \psi(t, Z_i^o + e)de\big)(u_i - u_i^o)\Big\}dt$$
$$+\int_I \Big\{\sum_{i=1}^{N} \big(\int_{C_r(0)} \nabla\psi(t, Z_i^o + e)de, Z_i - Z_i^o\big)u_i^o\Big\}dt \ge 0, \forall\ u \in \mathcal{U}_{ad}, Z_i \in \Sigma_0, \quad (4.1)$$

where ψ is the solution of the adjoint evolution equation given by

$$(\partial/\partial t)(c\psi) + \Phi^{'}(\rho^o)\triangle\psi = -\lambda F^{'}(\rho^o)\big(F(\rho^o) - F(\rho_d)\big),\ \text{on } I \times \Sigma, \quad (4.2)$$
$$\psi_{|\partial\Sigma} = 0,\ \text{on } I \times \partial\Sigma, \quad (4.3)$$
$$\psi(T) = 0,\ \text{on } \Sigma, \quad (4.4)$$

and ρ^o is the solution of the state equation corresponding to the optimal pair (Z^o, u^o)

$$(\partial/\partial t)(c\rho^o) - \triangle\Phi(\rho^o) = f - B(Z^o, u^o), (t, \xi) \in I \times \Sigma, \quad (4.5)$$
$$\Phi(\rho^o)|_{\partial\Sigma} = 0, (t, x) \in I \times \partial\Sigma, \quad (4.6)$$
$$\rho^o(0, \xi) = \rho_0, \xi \in \Sigma. \quad (4.7)$$

Proof. Let $(Z^o, u^o) \in \Sigma_0^N \times \mathcal{U}_{ad}$ be the optimal location-control pair and (Z, u) an arbitrary element of the admissible set $\in \Sigma_0^N \times \mathcal{U}_{ad}$. Clearly, by convexity, $Z^\varepsilon \equiv Z^o + \varepsilon(Z - Z^o) \in \Sigma_0^N$ and $u^\varepsilon \equiv u^o + \varepsilon(u - u^o) \in \mathcal{U}_{ad}$ for all $0 \le \varepsilon \le 1$. Let $\{\rho^\varepsilon, \rho^o\}$ denote the solutions of the system (2.4)–(2.6) corresponding to the pairs $(Z^\varepsilon, u^\varepsilon)$ and (Z^o, u^o) respectively. Since the pair (Z^o, u^o) is optimal, it is evident that $J(Z^\varepsilon, u^\varepsilon) - J(Z^o, u^o) \ge 0$. Dividing this by ε and letting $\varepsilon \downarrow 0$, it is easy to verify that the Gateaux differential dJ of J at (Z^o, u^o) in the direction $(Z - Z^o, u - u^o)$ must satisfy the following inequality,

$$dJ(Z^o, Z - Z^o; u^o, u - u^o)$$
$$= \int_I \Big\{\sum_{i=1}^{N} -\gamma_i(u_i - u_i^o) + \lambda \int_\Sigma \{(F(\rho^o) - F(\rho_d))\}(F^{'}(\rho^o)\varphi)d\xi\Big\}dt \ge 0 \quad (4.8)$$

for all $(Z, u) \in \Sigma_0^N \times \mathcal{U}_{ad}$, where φ is the weak solution of the variational equation given by

$$(\partial/\partial t)(c\varphi) - \triangle(\Phi^{'}(\rho^o)\varphi) = -g \text{ in } I \times \Sigma, \tag{4.9}$$

$$\Phi^{'}(\rho^o)\varphi|_{\partial\Sigma} = 0, \text{ in } I \times \partial\Sigma, \tag{4.10}$$

$$\varphi(0) = 0, \text{ in } \Sigma, \tag{4.11}$$

with the function g given by

$$\begin{aligned} g(t,\xi) &\equiv \sum_{i=1}^{N} u_i^o(t)(\nabla \Upsilon_{C_r(Z_i^o)}(\xi), Z_i - Z_i^o)_{R^n} + \sum_{i=1}^{N}(u_i(t) - u_i^o(t))\Upsilon_{C_r(Z_i^o)}(\xi) \\ &\equiv \mathcal{G}(Z - Z^o, u - u^o). \end{aligned} \tag{4.12}$$

The derivative of the characteristic function is understood in the sense of distributions. It is a simple exercise to verify that for every $u \in \mathcal{U}_{ad}$ and $Z \in \Sigma$, $B(Z, u) \in L_\infty(I, L_\infty(\Sigma)) \equiv L_\infty(I \times \Sigma)$, and since the set Σ has finite Lebesgue measure, $L_\infty(\Sigma) \subset L_q(\Sigma) \subset W^{-1,q}(\Sigma) = V^*$, and therefore $g \in L_q(I, V^*)$. Indeed, the reader can easily verify that for any $h \in L_p(I, V)$, the pairing

$$\int_I \langle g(t,\cdot), h(t,\cdot)\rangle_{V^*,V} dt = \int_{I\times\Sigma} g(t,\xi)h(t,\xi)d\xi dt$$

is well defined. Note that $\mathcal{G}$ is the Gateaux derivative of the operator B at (Z^o, u^o) in the direction $(Z - Z^o, u - u^o)$ with value g which is an element of $L_q(I, V^*)$. The system (4.9)–(4.11) is a linear homogenous boundary value problem, a special case of the nonlinear system equation (3.2). Thus existence of a unique weak solution follows from Theorem 3.1 as a corollary. Clearly, the expression (4.12) defines a linear map from $\Sigma^N \times \mathcal{U}_{ad}$ to $L_q(I, V^*)$. Emphasizing this dependence we observe that the map

$$-\mathcal{G}(Z - Z^o, u - u^o) \longrightarrow \varphi$$

is a continuous linear map from the Banach space $L_q(I, V^*)$ to the Banach space W as introduced in section 3. Considering the expression (4.8), we may denote the second term by the functional

$$L(\varphi) \equiv \int_I \int_{\partial\Sigma} \lambda\{F(\rho^o) - F(\rho_d)\}\{F^{'}(\rho^o)\varphi\}dx\, dt. \tag{4.13}$$

Clearly, it is a linear functional in $\varphi \in L_p(I, V)$ and it follows from our assumption on p given by, $5/2 < p \leq 4$, that $F^{'}(\rho^o)(F(\rho^o) - F(\rho_d)) \in L_q(I, V^*)$. Thus, L is a bounded linear functional on $L_p(I, V)$. Then, the composition map

$$-\mathcal{G}(Z - Z^o, u - u^o) \longrightarrow \varphi \longrightarrow L(\varphi)$$

is a continuous linear functional on $L_q(I, V^*)$. Hence, there exists a ψ (not necessarily unique) in the dual $(L_q(I, V^*))^* = L_p(I, V)$ such that

$$L(\varphi) = \int_I \langle -\mathcal{G}(Z - Z^o, u - u^o), \psi\rangle_{V^*,V} dt. \tag{4.14}$$

Thus, the inequality (4.8) can be written as

$$dJ(Z^o, Z - Z^o; u^o, u - u^o) \\ = \int_I \Big\{ \sum -\gamma_i(u_i - u_i^o) + \langle -\mathcal{G}(Z - Z^o, u - u^o), \psi \rangle_{V^*,V} \Big\} dt \geq 0, \qquad (4.15)$$

for all $(Z, u) \in \Sigma_0^N \times \mathcal{U}_{ad}$. Now integrating by parts, the second term of the above expression yields

$$\langle -\mathcal{G}(Z - Z^o, u - u^o), \psi \rangle_{V^*,V} = -\sum_{i=1}^{N} (u_i(t) - u_i^o(t)) \int_{C_r(0)} \psi(t, Z_i^o + e) de \\ + \sum_{i+1}^{N} u_i^o(t) \left(Z_i - Z_i^o, \int_{C_r(0)} \nabla \psi(t, Z_i^o + e) de \right). \qquad (4.16)$$

Since $\psi \in L_p(I, V)$, both the integrals displayed in the above expression are well defined for almost all $t \in I$. For example, considering the second term, it follows from Holder's inequality that

$$| \int_{C_r(0)} \nabla \psi(t, Z_i^o + e) de |_{R^n} \leq \ell(C_r(0))^{1/q} \left(\int_{C_r(0)} |\nabla \psi(t, Z_i^o + e)|_{R^n}^p de \right)^{1/p},$$

where $\ell(C_r(0))$ denotes the Lebesgue measure (volume) of the cylinder $C_r(0)$. Since $V = W_0^{1,p}$ and $\psi \in L_p(I, V)$, it is clear that the righthand integral is in $L_p(I) \subset L_1(I)$. Substituting the expression (4.16) in equation (4.15), we obtain the necessary condition (4.1). Now we show that ψ is determined by the weak solution of the adjoint system(4.2)–(4.4). Using the variational equation (4.9)–(4.11) and carrying out the necessary integration by parts, one can easily arrive at the following expression for $L(\varphi)$, given by (4.14),

$$L(\varphi) = -\int_I \langle \mathcal{G}(Z^o, Z - Z^o; u^o, u - u^o), \psi \rangle_{V^*,V} dt = \langle c\varphi, \psi \rangle |_0^T \\ + \int_I \langle \varphi, -\{ (\partial/\partial t)(c\psi) + \Phi'(\rho^o) \triangle \psi \} \rangle_{V,V^*} dt \\ + \int_{I \times \partial\Sigma} (\Phi'(\rho^o)\varphi)(\partial/\partial\nu)\psi d\xi \, dt - \int_{I \times \partial\Sigma} (\partial/\partial\nu)(\Phi'(\rho^o)\varphi)\psi d\xi \, dt. \qquad (4.17)$$

For $\psi(T, \cdot) = 0$, it follows from (4.11) that the first term on the right hand side of equation (4.17) vanishes. The boundary condition (4.10) forces the third term to vanish. Setting the adjoint boundary condition, $\psi|_{I \times \partial\Sigma} = 0$, the last term also vanishes leaving only the second term. Finally, setting

$$(\partial/\partial t)(c\psi) + \Phi'(\rho^o) \triangle \psi = -\lambda F'(\rho^o)(F(\rho^o) - F(\rho_d))$$

it follows from the expression (4.17) that the functional $L(\varphi)$, so obtained, coincides with the expression given by (4.13) as required. In conclusion, the function ψ, whose existence is proved above, is given by the solution of the following initial (final) boundary value problem,

$$(\partial/\partial t)(c\psi) + \Phi^{'}(\rho^o)\triangle\psi = -\lambda F^{'}(\rho^o)(F(\rho^o) - F(\rho_d)), \text{ on } I \times \Sigma,$$
$$\psi|_{I\times\partial\Sigma} = 0 \text{ on } I \times \partial\Sigma,$$
$$\psi(T, \cdot) = 0, \text{ on } \Sigma.$$

Thus, we have the adjoint system given by the equations (4.2)–(4.4) where ρ^o is the solution of the system equation (4.5)–(4.7) corresponding to the optimal pair (Z^o, u^o). For p as specified, the function $h \equiv \lambda F^{'}(\rho^o)(F(\rho^o) - F(\rho_d))$ belongs to $L_q(I, V^*)$. Hence, it follows from theorem 3.1, that it has a weak solution $\psi \in W$. Thus, we have demonstrated all the necessary conditions of optimality. This completes the proof. •

Remark 4.2 Taking $\lambda = 0$, it follows from the adjoint system (4.2)–(4.4) that $\psi \equiv 0$. In this situation the optimality condition (4.1) reduces to

$$dJ(u^o, u - u^o) = -\int_I \sum r_i(u_i - u_i^o)dt \geq 0, \quad \forall u \in \mathcal{U}_{ad}.$$

This means that the cost is minimal when the extraction rate is set to the maximum limit. Intuitively, this is what is expected when there are no (side effects) penalty for extraction.

Remark 4.3 If the locations are fixed a-priori at $\{Z_i^o\} \in \Sigma_0$, we expect to revert back to optimal policies obtained for given locations. If no variation in positions is permitted, the first term in the expression for g given by (4.12) vanishes leaving only the second term. As a consequence, the optimality condition (4.1) reduces to

$$dJ(u^o, u-u^o) \equiv -\int_I \Big\{\sum\Big(r_i+\int_{C_r(Z_i^o)} \psi(t,\xi)d\xi\Big)(u_i-u_i^o)\Big\}dt \geq 0, \quad \forall\, u \in \mathcal{U}_{ad}. \quad (4.18)$$

This is precisely the result given in [1, Theorem 5.1] where $G_i(Z_i)$ takes the place of $C_r(Z_i^o)$.

Remark 4.4 For the resource extraction problem considered here, point wise control does not make physical sense though it has significant theoretical interest because it poses some interesting mathematical difficulties. Reffering to equation (4.1) and dividing the spatial integrals by the volume of the set $C_r(0)$ and letting $r \downarrow 0$, we obtain the following necessary condition

$$dJ(Z^o, Z - Z^o; u^o, u - u^o) \equiv -\int_I \Big\{\sum\big(r_i + \psi(t, Z_i^o)\big)(u_i - u_i^o)\Big\}dt$$
$$+ \int_I \Big\{\sum_{i=1}^{N} \big(\nabla\psi(t, Z_i^o), Z_i - Z_i^o\big)u_i^o\Big\}dt \geq 0, \forall\, u \in \mathcal{U}_{ad}, Z_i \in \Sigma_0. \quad (4.19)$$

This derivation is formal and mathematically is not always justifiable. Consider the spatial dimension to be $n = 3$ and $p = 5/2$ and $V \equiv W_0^{1,p}(\Sigma)$. By the Sobolev embedding theorem, $W_0^{m,p} \subset C^{\alpha}(\Sigma)$ provided $m \geq (n/p) + \alpha$. In case $V = W_0^{1,p}$, this is not satisfied and hence for $\psi(t, \cdot) \in W_0^{1,p}$, the point values may not be finite let alone those of its gradient.

Some Open Problems. An interesting direction of future research is inclusion of stochastic terms to account for randomness in the natural source term, for example, the recharge process of aquifers. Additive noise is not admissible because the solution must be nonnegative. However, one can use any bounded Lipschitz function satisfying $\sigma(r) \geq 0$ for $r \geq 0$ and $\sigma(r) = 0$ for $r \leq 0$. Using σ as the volatility of the recharge process (source process), one can introduce the following model

$$\begin{aligned} d(c\rho) - \triangle\Phi(\rho)dt = f dt + \sigma(\rho)dW, (t, \xi) \in I \times \Sigma, \\ \Phi(\rho) = 0, (t, \xi) \in I \times \partial\Sigma, \\ \rho(0) = \rho_0, \xi \in \Sigma, \end{aligned} \tag{4.20}$$

where W is a space time Brownian motion.

Remark 4.5 We believe that the methodology developed here is also applicable to mining of underground resources such as gas, oil etc.

ACKNOWLEDGMENTS

This work was partially supported by the Science and Engineering Research Council of Canada under grant no A7109.

REFERENCES

[1] N. U. Ahmed, Porous Medium Dynamics with Optimal Control and Extraction of Underground Resources, *Dyn. Contin. Discrete Impuls. Syst.*, 2008, (to appear).

[2] F. Browder, Nonlinear Maxiaml Monotone Operators in Banach Spaces, Math. Ann. **175** (1968), 89–113.

[3] N. U. Ahmed and K. L.Teo, *Optimal Control of Distributed Parameter Systems*, Elsevier, North Holland, New York, 1981.

[4] R. E. Showalter, *Monotone Operators in Banach Space and Nonlinear Partial Differential Equations*, Mathematical Surveys and Monographs, Vol 49, American Mathematical Society, Providence, RI, 1997.

[5] J. L. Vazquez, *The Porous Media Equation*, Oxford University Press, Oxford 2007.