

CONVERGENCE ANALYSIS OF AN INTERIOR-POINT METHOD FOR MATHEMATICAL PROGRAMS WITH EQUILIBRIUM CONSTRAINTS

ARUN SEN

DEPT. OF ORFE, PRINCETON UNIVERSITY, PRINCETON, NJ 08544,
ASEN@PRINCETON.EDU

DAVID F. SHANNO

RUTCOR, RUTGERS UNIVERSITY, 640 BARTHOLOMEW RD.,
PISCATAWAY, NJ 0885,
SHANNO@RUTCOR.RUTGERS.EDU

ABSTRACT. We prove local and global convergence results for an interior-point method applied to mathematical programs with equilibrium constraints. The global result shows the algorithm minimizes infeasibility regardless of starting point, while one result proves local convergence when penalty functions are exact; another local result proves convergence when the solution is not even a KKT point.

1. INTRODUCTION

Equilibrium problems have applications in many areas, ranging from mechanics to problems in economics dealing with general equilibrium or game theory, to finance where certain options pricing problems may also be formulated as equilibrium models. What is of interest here is that an equilibrium problem can often be viewed in a natural way as a mathematical programming problem, what is commonly known as a *mathematical program with equilibrium constraints* (MPEC).

While MPEC's have many important applications, they present theoretical difficulties when viewed as nonlinear programs. To resolve these, an approach known as penalty methods has been proposed and much analyzed. The theoretical desirability of penalty methods for MPEC's is studied in, for instance, [1], while [2] shows that they work quite well in practice. The nonlinear solver used in [2] is LOQO, an interior-point nonlinear code. While LOQO has proved to be a rather robust code over the years, no formal convergence results have been

Date: November 23, 2004.

established for it, and indeed that task might be impossible. So considering convergence for MPEC's is even further out of the question. Indeed theoretical analysis of the application of interior-point methods to MPEC's is in its infancy. However, recently in [7] and [5] general convergence results are proved for interior-point methods. The intent of this paper is to extend the analysis to MPEC's. To date the only such paper we have seen in published form is [6], which uses a different approach than we do, and makes different assumptions.

We wish to emphasize that our results in this paper apply equally well to the algorithms in [7] and [5], but we will mostly refer to [5], as the assumptions in this paper are less restrictive while still sufficient for our purposes. The convergence results in [5] are somewhat unique in that they make *no* assumptions about the algorithm, only about the problem. Yet this is done in a way which does not reduce the generality of the problems which may be considered.

This paper is organized as follows: the IPM in question and its general convergence results are briefly described in section 2, then section 3 quickly reviews the main issues that arise when dealing with MPEC's. Section 4 describes our approach for applying IPM's to MPEC's- what we call an interior-point penalty method. Sections 5 and 6 get to the heart of the matter- a global convergence result for the IPM applied to MPEC's is presented in section 5 and local convergence is considered in section 6.

2. A GLOBALLY CONVERGENT INTERIOR-POINT METHOD

The following is only a brief description of the algorithm and its convergence. For details, see [5]. The analysis in [5] only covers problems with inequality constraints (Griva et. al. will extend this in the near future). Thus the problem considered is

$$(1) \quad \min \quad f(x)$$

$$\text{s.t.} \quad h_i(x) \geq 0, \quad i = 1, \dots, m,$$

where we assume that $f(x)$ and each of the $h_i(x)$ are twice continuously differentiable and $x \in R^n$. Adding nonnegative slacks, w , to the inequality constraints in (1) and putting these slacks in a logarithmic barrier term in the objective function, the problem becomes:

$$(2) \quad \min \quad f(x) - \mu \sum_{i=1}^m \log(w_i)$$

$$\text{s.t.} \quad h(x) - w = 0.$$

At this point we need to mention the assumptions that are made in [5] and which we will continue to observe for the rest of this paper:

Assumption 1: The objective function $f(x)$ is bounded below for all $x \in R^n$.

Assumption 2: The functions $h_i(x)$ satisfy the following conditions:

$$\lim_{\|x\| \rightarrow \infty} \min_{1 \leq i \leq m} h_i(x) \rightarrow -\infty,$$

and

$$\sqrt{\log(|\max_{1 \leq i \leq m} h_i(x)| + 1)} \leq -\min_{1 \leq i \leq m} h_i(x) + C,$$

where C is a positive constant depending only on the problem's data.

Assumption 3: All solutions to (1) satisfy the standard second-order optimality conditions.

Assumption 4: All solutions to (2) for any fixed $\mu > 0$ satisfy standard second-order optimality conditions.

Assumption 5: The Hessians $\nabla^2 f$ and $\nabla^2 h_i(x)$ for all i satisfy Lipschitz conditions.

Assumption 1 is not really restrictive. Minimizing $f(x)$ is always equivalent to minimizing $\log(e^{f(x)} + 1)$. The first part of assumption 2 implies that the feasible set is bounded. In assumptions 3 and 4 the optimality conditions include regularity of the solution (i.e. linear independence of active constraint gradients), strict complementarity (that is, multipliers of active constraints are positive) as well as second order sufficiency. See [5] for an explanation of the rest of these assumptions. We should remark that in [5] another assumption is made that the constraints satisfy the Slater condition (that $\exists \bar{x} \in R^n$ s.t. $h_i(\bar{x}) > 0$ for all i). But in fact, as assumption 2 ensures the feasible set is bounded, a solution to (1) is guaranteed to exist, and by assumption 3 must be regular. It can be shown that this implies the existence of a point which satisfies the Slater condition. This can be seen as follows.

Suppose x^* is a regular solution to (1). Let $I = \{i : h_i(x) = 0\}$. We may assume that I is not empty, since otherwise x itself would satisfy the Slater condition. So consider the problem

$$\min \zeta$$

$$s.t. \ h_i(x) \geq -\zeta.$$

The Lagrangian for this problem is $L(x, \zeta, \lambda) = \zeta - \sum_{i \in I} \lambda_i(\zeta + h_i(x))$. The first order conditions include

$$\nabla_x L = \sum_{i \in I} \lambda_i \nabla_x h_i(x) = 0,$$

$$\nabla_\zeta L = 1 - \sum_{i \in I} \lambda_i = 0.$$

The second line implies all the λ_i 's cannot be 0, which means that by the first line the $\nabla_x h_i$'s are not linearly independent, which contradicts the regularity of x^* for the original problem. Thus $(x^*, 0)$ cannot be a solution to the above problem, hence there must be a point in a neighborhood of x^* which satisfies the Slater condition. Thus we do not need to state the Slater condition as a separate assumption.

Letting the vector y represent the Lagrange multipliers for (2) the first-order optimality conditions for the problem are

$$(3) \quad \nabla f(x) - A(x)^T y = 0,$$

$$\mu W^{-1} e - Y = 0,$$

$$w - h(x) = 0.$$

Here e is the vector of all ones, $A(x)$ is the Jacobian of the vector $h(x)$, and W and Y are the diagonal matrices with the coordinates of the vectors w and y on the diagonals, respectively.

Newton's method applied to (3) yields the linear system

$$\begin{bmatrix} H(x, y) & 0 & -A(x)^T \\ 0 & Y & W \\ A(x) & -I & 0 \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta w \\ \Delta y \end{bmatrix} = \begin{bmatrix} -\nabla f(x) + A(x)^T y \\ \mu e - WY e \\ -h(x) + w \end{bmatrix}.$$

Up to this point, the algorithm has followed the standard IPM paradigm, as used in LOQO. The standard practice is to now solve one iteration of the Newton system, choose an appropriate steplength to move to the next point while keeping y and w positive, and reduce μ to 0 in some fashion to approach a solution. The wrinkle that [5] adds here is to modify this system to become

$$\begin{bmatrix} H(x, y) & 0 & -A(x)^T \\ 0 & Y & W \\ A(x) & -I & \epsilon I \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta w \\ \Delta y \end{bmatrix} = \begin{bmatrix} -\nabla f(x) + A(x)^T y \\ \mu e - WY e \\ -h(x) + w \end{bmatrix},$$

where $\epsilon > 0$ is a regularizing parameter. This system is known as the *dual regularization* of the problem. This modification is crucial to making the IPM provably convergent, though many other modifications are required as well. The reader is encouraged to see [5] for further details. What matters for our purposes here is the two convergence results that are proved: Theorem 1 proves that the IPM will always converge to a point which locally minimizes the l_2 norm of the infeasibility, and if feasible, is a local minimum for the problem. Lemma 9 proves that if the IPM is started at a point sufficiently close to a solution, then it will converge to it at a quadratic rate. In this paper we will prove similar results for the IPM applied to MPEC's via a penalty method.

3. PENALTY METHODS AND MPEC'S

We briefly review the particular issues involving MPEC's.

A *mathematical program with equilibrium constraints* (MPEC) may be written as an NLP of the form

$$\begin{aligned} (4) \quad & \min f(x) \\ & \text{s.t.} \quad x^T F(x) \leq 0, x \geq 0, F(x) \geq 0, \\ & h(x, z) = 0, g(x, z) \leq 0, x \in R^n, z \in R^l \end{aligned}$$

$$F : R^n \rightarrow R^n, g : R^{n+l} \rightarrow R^m, h : R^{n+l} \rightarrow R^k$$

As shown, MPEC's commonly involve equality constraints, while the analysis for the IPM in section 2 at the moment only handles inequalities. We will address this matter in the next section. For now, we focus on the aspects of (4) that attract so much attention. Specifically, it is a well-known fact that the set of optimal Lagrange multiplier solutions for an MPEC is always unbounded (see, for instance, [8]), as a standard constraint qualification, the Managarian-Fromovitz Constraint Qualification (MFCQ) cannot hold at a solution. This implies trouble for many standard nonlinear solvers. What is most germane for this paper, however, in terms of the theoretical results for the IPM described in section 2 (again, ignoring the equality constraints for now), is that

its convergence proof fails to apply to an MPEC because a solution cannot be regular, violating assumption 3.

It is shown in [2] that these theoretical difficulties can indeed cause problems in practice for IPM's, as well as other types of solvers. That paper and many others analyze an approach to get around the difficulty of unbounded multipliers, namely penalty methods. These we now discuss.

3.1. Penalty Methods. Consider again our NLP in the form

$$\begin{aligned} & \min f(x) \\ \text{s.t.} \quad & x^T F(x) \leq 0, x \geq 0, F(x) \geq 0, \\ & h(x, z) = 0, g(x, z) \leq 0. \end{aligned}$$

The idea of a penalty method is to reduce a problem with complicated constraints to an unconstrained problem of the form $f(x) + \rho P(x)$ where $P(x)$ is nonnegative for all x , and is 0 if and only if x feasible. Then by increasing ρ , one hopes to get a solution which is feasible for the original problem and hence a true solution.

For instance, the ℓ_∞ penalty function, when applied to the complementarity constraints in the above problem results in

$$\begin{aligned} & \min f(x) + \rho |x^T F(x)| \\ \text{s.t.} \quad & x, F(x) \geq 0, h(x, z) = 0, g(x, z) \leq 0, \end{aligned}$$

which can be rewritten (so as to have differentiable functions) as

$$\begin{aligned} & \min f(x) + \rho \zeta \\ \text{s.t.} \quad & x^T F(x) \leq \zeta, \end{aligned}$$

$$x, F(x) \geq 0, h(x, z) = 0, g(x, z) \leq 0.$$

It is shown in, among other places, [1] and [4] that the l_∞ (and the related l_1) penalty function is *exact*, in the sense that it guarantees that every solution of the true problem is a solution to the penalized problem for sufficiently large (but finite) penalty ρ , under very general conditions. Also, importantly, the multipliers at a solution are now bounded. Thus the primary theoretical concern involving MPEC's is resolved. The application of penalty functions to MPEC's in practice is substantially investigated in [2], and proves to be quite successful. As

mentioned before, this was done using an interior-point solver, LOQO. The main purpose of the remainder of this paper is to analyze the theoretical convergence properties of the IPM from section 2 (itself a theoretical algorithm, though inspired by LOQO) when applied to MPEC's using penalty functions.

4. AN INTERIOR-POINT PENALTY METHOD FOR MPEC'S

We now describe how the IPM from section 2 may be applied to MPEC's written in the form of (4). To do so we must now address the issue that the general convergence proof we currently have for the IPM only deals with inequality constraints, while of course many MPEC's include equality constraints. This is not a bar to progress, however, since any equality constraint $h(x) = 0$ may be written as two inequalities, $h(x) \leq 0, -h(x) \leq 0$. Away from a solution, this does not present an issue. Thus when we are far from a solution, we will solve the following problem at each step:

$$(5) \quad \min f(x) + \rho\zeta$$

$$\text{s.t.} \quad x^T F(x) \leq \zeta, \quad x, F(x) \geq -\zeta,$$

$$-\zeta \leq h(x, z) \leq \zeta, g(x, z) \leq \zeta, \zeta \geq 0.$$

But this guarantees that at the solution there will be linearly dependent gradients for constraints which are tight, which violates assumption 3. So near a solution, we transform our problem: for a constraint $h(x) = 0$ we drop the relaxation on the lower bound, that is, we write it as

$$h(x) \geq 0, h(x) \leq \zeta.$$

There is the further complication that even without equality constraints, an MPEC will always have dependent constraint gradients at a solution if MPEC-LICQ holds. Thus near a solution to (4), instead of (5) we will solve the problem

$$(6) \quad \min f(x) + \rho(\zeta_1 + \zeta_2)$$

$$\text{s.t.} \quad x^T F(x) \leq \zeta_1,$$

$$x, F(x) \geq 0, 0 \leq h(x) \leq \zeta_2, g(x, z) \leq 0.$$

Assuming MPEC-LICQ, this problem has nice local properties, which we shall discuss in detail in section 6. For fixed ρ we shall refer to (5) or (6), whichever we are currently solving, as P_ρ . Now we are ready to describe the *Interior Point Penalty Method (IPPM)*:

The algorithm will have two phases: in *global mode*, it will always solve (5), while in *local mode* it will always solve (6). The method always starts in global mode. Assume there is a given ρ_o and corresponding solution $(x_{\rho_o}, \zeta_{\rho_o}, w_{\rho_o}, y_{\rho_o})$ to P_{ρ_o} , with $\zeta_{\rho_o} > 0$ (where of course, (w_{ρ_o}, y_{ρ_o}) are the starting slack and dual solutions). For each ρ_k we apply the IPM to P_{ρ_k} to get a solution $(x_{\rho_k}, \zeta_{\rho_k}, w_{\rho_k}, y_{\rho_k})$. ρ_k is then increased by some $\tau_k > 0$. The IPM is then applied to $P_{\rho_{k+1}}$ using $(x_{\rho_k}, \zeta_{\rho_k}, w_{\rho_k}, y_{\rho_k})$ as the starting point, except that we increase those (w_{ρ_k}, y_{ρ_k}) which are at 0 by some $\tau'_k > 0$.

When ζ_ρ becomes small, the algorithm enters *local mode*, and tries to solve (6). The current primal solution will be kept as the starting point, but the dual and slack variables will be re-estimated, as the problem to be solved has changed. And the penalty may have to be adjusted: the algorithm tries to solve (6) for successively higher values of ρ , starting with an estimate for the penalty based on the current approximation to the solution (recall that in local mode we are assumed to be close to a solution). If no solution is obtained, the algorithm goes back to global mode and tries to reduce ζ_ρ further.

5. GLOBAL CONVERGENCE OF THE IPPM FOR MPEC'S

In this section we provide a convergence proof for the IPPM in global mode. Without loss of generality we may assume that in global mode we are trying to solve a problem of the form

$$(7) \quad \min f(x) + \rho\zeta$$

$$s.t. \quad h(x) \leq \zeta, \zeta \geq 0$$

We make all the assumptions about this problem for fixed ρ as described in section 2, for solutions with $\zeta > 0$, except assumption 2. Clearly the feasible set for (7) is not bounded; instead we make this assumption about the original unrelaxed problem corresponding to (7). Also, the objective of (7) is of course not bounded below, but as stated earlier, in the objective ζ can be replaced by $\log(e^\zeta + 1)$. We now need the following definition.

Definition 1. *The l_∞ norm of the infeasibility of the unrelaxed problem corresponding to (7) at x is $\nu(x) = \max_i \{\max(h_i(x), 0)\}$.*

Then assumption 2 implies that ν cannot have a local minimum at ∞ . This will be crucial for the argument to follow.

We are now ready to state and prove the main result of this section. The analysis here will let ρ approach ∞ . The issue of exactness and the finiteness of the penalty will be dealt with in the next section when we prove local convergence.

Theorem 1. *As $\rho_k \rightarrow \infty$, for sufficiently small $\{\tau_k, \tau'_k\}$ the IPPM in global mode gets arbitrarily close to a point which locally minimizes ν if it does not approach a solution to (4)*

Proof. Assume the algorithm does not approach a solution to (4) (so ζ never approaches 0). Then under our assumptions, by lemma 9 of [5] the IPM always converges to a solution of $P_{\rho_{k+1}}$ when the starting point is $(x_{\rho_k}, \zeta_{\rho_k}, w_{\rho_k}, y_{\rho_k})$ perturbed by a sufficiently small τ'_k , and $\rho_{k+1} - \rho_k = \tau_k$ is small enough. Next we show that the sequence $\{x_{\rho_k}\}$ for sufficiently small (τ_k, τ'_k) remains bounded. To do so, we first observe that

$$f(x_{\rho_k}) + \rho_k \zeta_{\rho_k} \leq f(x_{\rho_{k+1}}) + \rho_k \zeta_{\rho_{k+1}},$$

$$f(x_{\rho_{k+1}}) + \rho_{k+1} \zeta_{\rho_{k+1}} \leq f(x_{\rho_k}) + \rho_{k+1} \zeta_{\rho_k},$$

hold true for any k for (τ_k, τ'_k) sufficiently small. So combining the two inequalities we then have

$$(\rho_{k+1} - \rho_k) \zeta_{\rho_k} \geq (\rho_{k+1} - \rho_k) \zeta_{\rho_{k+1}},$$

which implies

$$\zeta_{\rho_k} \geq \zeta_{\rho_{k+1}}.$$

In other words, ν is nonincreasing. Now we make use of assumption 2 which gives us that the unbounded growth of $\{x_{\rho_k}\}$ would eventually increase ν . The result must be that $\{x_{\rho_k}\}$ remains bounded.

We should mention that there is no way for the algorithm to stall because of τ_k going to 0 prematurely. That is, even though it might be necessary to increase the penalty very slowly to ensure the monotonic decrease in ν , the penalty can always be increased from its current level. So now we study what happens to the sequence $\{x_{\rho_k}, \zeta_{\rho_k}\}$ as $\rho_k \rightarrow \infty$.

First, for any point $\{x, \zeta\}$ with $\zeta = \max(h_1(x), \dots, h_m(x), 0)$ we define the set

$$J(x) = \{j \mid h_j(x) = \zeta, j = 1, \dots, m\}.$$

Then it is known that $\{x_{\rho_k}, \zeta_{\rho_k}\}$ is a local minimum of P_{ρ_k} if and only if the following condition holds for all directions $d \in R^n$ such that

$\|d\| = 1$:

$$(8) \quad \nabla f(x_{\rho_k})^T d + \rho_k * \max_{j \in J(x)} \nabla h_j(x_{\rho_k})^T d \geq 0.$$

Similarly, ν is minimized at x_{ρ_k} if and only if

$$(9) \quad \max_{j \in J(x)} \nabla h_j(x_{\rho_k})^T d \geq 0 \quad \forall d \in R^n, \|d\| = 1$$

Now assume that ν is not minimized at x_{ρ_k} for any finite k (if it is, clearly $x_{\rho_l} = x_{\rho_k} \quad \forall l > k$, so we are done). Then by (9) we must have

$$(10) \quad \forall \rho_k \exists d \in R^n, \|d\| = 1 \text{ s.t. } \nabla h_j(x_{\rho_k})^T d < 0 \quad \forall j \in J(x_{\rho_k}),$$

in other words there must be a descent direction for infeasibility, if it is not minimized. Now suppose there exists an $\epsilon > 0$ such that the following condition holds:

$$(11) \quad \forall l \exists k > l \text{ s.t. } \exists d \in R^n, \|d\| = 1 \text{ s.t.}$$

$$\nabla h_j(x_{\rho_k})^T d < -\epsilon \quad \forall j \in J(x_{\rho_k}).$$

Consider what happens if (11) holds. Since $\{x_{\rho_k}\}$ is bounded, and hence f being continuously differentiable $\nabla f(x_{\rho_k})$ is also bounded, (8) obviously cannot hold for large enough ρ_k which satisfies (11). This contradicts the optimality of $\{x_{\rho_k}, \zeta_{\rho_k}\}$ for P_{ρ_k} , thus (11) cannot hold. It therefore follows that

$$\lim_{k \rightarrow \infty} \max_{j \in J(x)} \nabla h_j(x_{\rho_k})^T d \geq 0 \quad \forall d \in R^n, \|d\| = 1.$$

Therefore $\{x_{\rho_k}, \zeta_{\rho_k}\}$ approaches arbitrarily closely to a point which locally minimizes ν . \square

6. LOCAL CONVERGENCE OF THE IPPM FOR MPEC'S

In this section we consider the question of local convergence to a strict local minimum x^* of an MPEC for the IPPM. We shall provide two different results, one when exactness holds at x^* , one which doesn't require that assumption. First we deal with the exact case.

Until now we have made no particular assumptions about the solutions of the MPEC. We have made assumptions about the functions appearing in that problem and about the solutions of P_ρ in global mode for finite ρ . Now we add an assumption about the solution x^* of the MPEC, which allows us to raise the question of exactness.

It is shown in both [1] and [4] that for any MPEC, if a solution x^* is a KKT point and a weak second-order condition holds there, then the l_1 and l_∞ penalty functions are exact. Moreover, the second-order condition continues to hold for the relaxed problem. In general, if a

problem has all inequality constraints of the form $h(x) \leq 0$, the 2nd-order qualification known as the *quadratic growth condition* (QGC) holds at a solution x^* if

$$\max\{f(x) - f(x^*), h_1(x), \dots, h_m(x)\} \geq \|x - x^*\|^2$$

for all x in a neighborhood of x^* .

The important thing to realize about this condition is that it is shown in [3] that if a solution is a KKT point and MFCQ (equivalent to multipliers being bounded) holds, then the QGC holding is equivalent to the standard second-order sufficiency conditions holding.

Recall that in local mode we want to solve (6):

$$\min f(x) + \rho(\zeta_1 + \zeta_2)$$

$$\text{s.t. } x^T F(x) \leq \zeta_1,$$

$$x, F(x) \geq 0, 0 \leq h(x, z) \leq \zeta_2, g(x, z) \leq 0.$$

We may thus observe that the relaxed constraints must have gradients which are linearly independent of all the other constraint gradients, due to the presence of the auxiliary variables, which are not present in the other constraints. We make here one further assumption about (4): that it satisfies a condition known as the “MPEC linearly independent constraint qualification” or MPEC-LICQ. MPEC-LICQ simply says that solutions to an MPEC would be regular if the complementarity constraints were dropped from the problem (see [8] for details). Thus assumption 3 will hold for (6) if MPEC-LICQ and the QGC hold for (4). Furthermore, it is not hard to see that if a multiplier solution exists to (4) which is strictly complementary, then a strictly complementary solution exists for (6) as well for large enough (but finite) ρ .

We have the following result for the IPPM in local mode:

Theorem 2. *If MPEC-LICQ and the QGC and strict complementarity hold for (4) at a solution x^* , then for sufficiently large but finite ρ^* a solution $(x^*, 0, w^*, y_{\rho^*}^*)$ exists for (6) which satisfies the standard second-order optimality conditions. If the IPPM enters local mode at a point (x_o, ζ_o, w_o, y_o) sufficiently close to the solution $(x^*, 0, w^*, y_{\rho^*}^*)$, with the penalty parameter set to ρ^* , then under all previous assumptions it converges to x^* in one iteration (i.e. no further update of the penalty parameter is needed), and the rate of convergence is quadratic.*

Proof. The preceding discussion guarantees the existence of a solution as described. We may then conclude the result by lemma 9 in [5]. \square

This theorem thus complements the result from theorem 1: theorem 1 shows that in global mode the algorithm gets arbitrarily close to a point which minimizes the infeasibility. If such a point is in fact an isolated minimum for the original problem satisfying some standard conditions, theorem 2 proves that the algorithm in local mode will in fact converge to it. We do not want to enter local mode far from a solution, but rather approach it first in global mode. Note that if the solution does not satisfy MPEC-LICQ, the algorithm in global mode may still be able to approach it by gradually increasing the penalty. Even if the solution is not a KKT point at all, this may still be possible, as we now discuss.

We now go back to the situation of the previous section, that is we go back to global mode, drop all assumptions about the solution, but do assume a strict local minimum x^* exists for (4). Without loss of generality, we may again assume that the problem we want a solution for is of the form

$$(12) \quad \begin{aligned} & \min f(x) \\ & \text{s.t. } h(x) \leq 0 \end{aligned}$$

while the problem we actually solve at each step has all the constraints relaxed, (7):

$$\begin{aligned} & \min f(x) + \rho\zeta \\ & \text{s.t. } h(x) \leq \zeta, \zeta \geq 0 \end{aligned}$$

We want to be able to make the following claim: that there is a sequence of solutions $(x_{\rho_k}, \zeta_{\rho_k})$ which converges to $(x^*, 0)$ as $\rho_k \rightarrow \infty$. We proved a similar result in theorem 3 of [2], but that was done with some assumptions, which we now drop. To this end, we first establish that any open ball centered at x^* must contain x_ρ for all sufficiently large ρ .

Lemma 1. $\forall B_{r>0}(x^*) \exists \bar{\rho}$ s.t. $\forall \rho > \bar{\rho} x_\rho \in B_r$ for some solution (x_ρ, ζ_ρ) to P_ρ .

Proof. We choose r to be small enough such that $f(x^*) < f(x) \forall x \in \overline{B_r(x^*)}$ s.t. x feasible for (12). Suppose $B_r(x^*)$ does not contain x_ρ for

some ρ . What must be true is that the function $f(x) + \rho\zeta$ attains a global minimum over the set

$$h(x) \leq \zeta, \zeta \geq 0, x \in \overline{B_r(x^*)}$$

since x is now restricted to a closed and bounded set (and ζ will always be $\max(0, h_i(x), \dots, h_m(x))$). Call this global minimum $(\hat{x}_\rho, \hat{\zeta}_\rho)$. By assumption this global minimum must lie on the boundary of $\overline{B_r(x^*)}$, and must be better than any point in the interior:

$$(13) \quad f(\hat{x}_\rho) + \rho\hat{\zeta}_\rho < f(x) + \rho\zeta$$

$$\text{for } x \in B_r(x^*), \zeta = \max(0, h_i(x), \dots, h_m(x)).$$

Furthermore, it must be true that $\hat{\zeta}_\rho > 0$, since otherwise \hat{x}_ρ would be feasible for (12) with $f(\hat{x}_\rho) < f(x^*)$ which cannot happen. Now let S be the set of all such ρ just described (i.e. those for which P_ρ has no solutions in $B_r(x^*)$). We assume for the moment that S is unbounded. Say it were the case that

$$(14) \quad \exists \epsilon > 0 \text{ s.t. } \hat{\zeta}_\rho > \epsilon \forall \rho \in S.$$

Then it is not hard to see that, since f is bounded on $\overline{B_r(x^*)}$, we must have for sufficiently large ρ in S ,

$$f(\hat{x}_\rho) + \rho\hat{\zeta}_\rho > f(x^*),$$

which of course contradicts (13). So suppose (14) doesn't hold. Then we have that, for some sequence $\rho_1, \rho_2, \dots, \rho_k$ in S : $\lim_{k \rightarrow \infty} \hat{x}_{\rho_k} = \bar{x}$ which must be feasible for (12) and lie in $\overline{B_r(x^*)}$ (since a closed set must contain its limit points). It must be true that $f(\bar{x}) > f(x^*)$, and yet $\lim_{k \rightarrow \infty} f(\hat{x}_{\rho_k}) + \rho_k \hat{\zeta}_{\rho_k} = f(\bar{x})$, and $f(\hat{x}_{\rho_k}, \hat{\zeta}_{\rho_k}) + \rho_k \hat{\zeta}_{\rho_k} < f(x^*)$ for all k . Obviously this cannot happen since the limit of a sequence cannot be strictly greater than a number if all the elements of the sequence are strictly less than it. Therefore S must be bounded and the lemma is proved. \square

The proof of the claim now follows:

Theorem 3. *There is a sequence of solutions $(x_{\rho_k}, \zeta_{\rho_k})$ which converges to $(x^*, 0)$ as $\rho_k \rightarrow \infty$.*

Proof. We simply apply the lemma, and the result is obvious. In fact there are uncountably many such sequences. \square

Given the presence of such sequences, absent truly pathological behavior the IPPM will be able to track such a path if a starting point is chosen sufficiently close to x^* for sufficiently large ρ . This behavior has in fact been seen in practice (see [2]).

7. CONCLUSION

In this paper we have extended the convergence results for interior-point methods applied to standard problems, shown in places such as [5] and [7], to MPEC's. We want to emphasize that as in [5], the results we have proved for MPEC's rely solely on *assumptions made on the problem*, not on the algorithm itself. For instance, unlike [6] we do *not* assume that the sequence (primal or dual) generated by our algorithm is bounded. This we believe sets our results apart from other papers on IPM's applied to MPEC's, which are starting to appear in the literature. Moreover, as far as we know, few global results have been proved (for IPM's or otherwise) for MPEC's, of the kind shown in theorem 1, and little or no analysis done when the solution is not a KKT point, as in theorem 3. Indeed, theorems 1 and 3 apply to *any* problem which satisfies some very loose assumptions, thus making a general point about penalty methods, not just when applied to MPEC's.

This paper combined with [2] begins to give a complete picture, theoretical and practical, of IPM's applied to MPEC's. This is highly desirable, as IPM's are in general superior in practice to other kinds of solvers for large-scale problems. The IPPM as described is of course theoretical (as is the IPM it depends upon). In practice one would prefer not to increase the penalty by only very tiny increments. We needed this to ensure monotonic decrease in ν , to make the theory rigorous. Similarly, in practice it is problematic to set slack and dual starting values too close to 0 for an IPM. Moreover, in reality we do not need to convert equality constraints to inequalities. And indeed, significant numerical success has been experienced (see [2]) without satisfying these theoretical requirements. Also note that in [2] the penalty parameter was dynamically adjusted during the algorithm's run, instead of solving a sequence of separate optimization problems. This is likely the preferable way to proceed in practice. Future work could involve analyzing this aspect theoretically, while also implementing in LOQO some of the changes suggested in [5] (this has already begun), and then applying the modified code to MPEC's. The goal being to constantly close the gap between theory and practice in a way that actually improves the chances of solving real problems.

8. ACKNOWLEDGEMENT

Research of the first author was supported by NSF grant DMS-9870317 and ONR grant N00014-98-1-0036. Research of the second author was supported by NSF grant DMS-0107450.

REFERENCES

- [1] Anitescu, Mihai. “Nonlinear Programs with Unbounded Lagrange Multiplier Sets.” Technical Report ANL/MCS-P793-0200.
- [2] Benson, Hande Y. and Arun Sen and David F. Shanno and Robert J. Vanderbei. “Interior-Point Algorithms, Penalty Methods and Equilibrium Problems.” Princeton University ORFE Technical Report 03-02.
- [3] Bonnans, F. and A. Ioffe. “Second-order Sufficiency and Quadratic Growth for Non-isolated Minima.” *Mathematics of Operations Research* (20) 4, 801-817, 1995.
- [4] Bonnans, F. and A. Shapiro. *Perturbation Analysis of Optimization Problems*. Springer-Verlag, New York, 2000.
- [5] Griva, I. and David F. Shanno and Robert J. Vanderbei. “Convergence Analysis of a Primal-Dual Interior-Point Method for Nonlinear Programming.” Working paper.
- [6] Liu, X. and J. Sun. “Generalized Stationary Points and an Interior-Point Method for Mathematical Programs with Equilibrium Constraints.” *Mathematical Programming B* (101) 1: 231-261, 2004.
- [7] Liu, X. and J. Sun. “A Robust Primal-Dual Interior-Point Algorithm for Nonlinear Programs.” *SIAM Journal on Optimization* (14) 4: 1163-1186, 2004.
- [8] Scholtes, Stefan. “Convergence Properties of a Regularization Scheme for Mathematical Programs with Complementarity Constraints.” *SIAM Journal on Optimization* 11, 918-936, 2001.