Volume 29, N. 2, pp. 125–138, 2010
ISSN 0101-8205
www.scielo.br/cam

On the global convergence of interior-point nonlinear programming algorithms

GABRIEL HAESER*

Department of Applied Mathematics
Institute of Mathematics, Statistics and Scientific Computing
University of Campinas, Campinas, SP, Brazil

E-mail: ghaeser@gmail.com

Abstract. Carathéodory's lemma states that if we have a linear combination of vectors in $\mathbb{R}^n$, we can rewrite this combination using a linearly independent subset. This lemma has been successfully applied in nonlinear optimization in many contexts. In this work we present a new version of this celebrated result, in which we obtained new bounds for the size of the coefficients in the linear combination and we provide examples where these bounds are useful. We show how these new bounds can be used to prove that the internal penalty method converges to KKT points, and we prove that the hypothesis to obtain this result cannot be weakened. The new bounds also provides us some new results of convergence for the quasi feasible interior point ℓ_2-penalty method of Chen and Goldfarb [7].

Mathematical subject classification: 90C30, 49K99, 65K05.

Key words: nonlinear programming, constraint qualifications, interior point methods.

1 Introduction

In 1911 Carathéodory proved that if a point $x \in \mathbb{R}^n$ lies on the convex hull of a compact set P, then x lies on the convex hull of a subset P' of P with no more than $n + 1$ points [6]. In 1914 Steinitz generalized this result for a general set P [18].

#CAM-111/09. Received: 30/VI/09. Accepted: 17/III/10.
*This work was supported by FAPESP Grant 05/02163-8.

Here we will see a different version of Carathéodory's result, which appears in [5] as "Carathéodory's theorem for cones", but is better known as "Carathéodory's lemma". We will provide bounds on the size of the multipliers given by the Carathéodory's lemma and we will apply this result to internal penalty methods. We address the following nonlinear optimization problem:

$$\text{Minimize } f(x) \quad \text{subject to} \quad h(x) = 0, \quad g(x) \leq 0, \tag{1}$$

where $f\colon \mathbb{R}^n \to \mathbb{R}, h\colon \mathbb{R}^n \to \mathbb{R}^m$ and $g\colon \mathbb{R}^n \to \mathbb{R}^p$ are continuously differentiable functions. Under a given constraint qualification, the solution x^* satisfies the KKT condition, that is, x^* is feasible with respect to equality and inequality constraints and there exist $\lambda \in \mathbb{R}^m$ and $\mu_j \geq 0$ for every $j \in A(x^*) = \{i \in \{1, \dots, p\} | g_i(x^*) = 0\}$ such that

$$\nabla f(x^*) + \sum_{i=1}^{m} \lambda_i \nabla h_i(x^*) + \sum_{j \in A(x^*)} \mu_j \nabla g_j(x^*) = 0.$$

A common constraint qualification usually employed is the Linear Independence constraint qualification, which states that

$$\left\{\nabla h_i(x^*)\right\}_{i=1}^{m} \cup \left\{\nabla g_j(x^*)\right\}_{j \in A(x^*)}$$

is linearly independent. We refer to this multi-set as the *active set of gradients at* x^*. The weaker Mangasarian-Fromovitz constraint qualification (MFCQ) [14, 16] states that the active set of gradients is positive-linearly independent, which means that there are no $\alpha \in \mathbb{R}^m$, $\beta_j \geq 0$ for every $j \in A(x^*)$ such that

$$\sum_{i=1}^{m} \alpha_i \nabla h_i(x^*) + \sum_{j \in A(x^*)} \beta_j \nabla g_j(x^*) = 0,$$

except if we take all α_i and β_j equal to zero.

Recently, a weaker constraint qualification appeared in the literature: the Constant Positive Linear Dependence constraint qualification (CPLD) [15, 4], which has been successfully applied to obtain new practical algorithms [1, 2, 10]. We say that the CPLD condition holds for a feasible x^* if for every $I \subset \{1, \dots, m\}$, $J \subset A(x^*)$ such that the set of gradients $\{\nabla h_i(x^*)\}_{i \in I} \cup \{\nabla g_j(x^*)\}_{j \in J}$ is positive-linearly dependent, there exists a neighborhood $V(x^*)$

of x^* such that the set of gradients $\{\nabla h_i(y)\}_{i \in I} \cup \{\nabla g_j(y)\}_{j \in J}$ remains positive-linearly dependent for every $y \in V(x^*)$. The CPLD condition is a natural generalization of the Constant Rank constraint qualification of Janin [13], which states the same as above, replacing "positive-linearly dependent" by "linearly dependent". The CPLD condition is weaker than the Constant Rank condition [17].

In practical algorithms, weaker constraint qualifications are preferred, since convergence results are stronger.

In Section 2 we will state Carathéodory's lemma and obtain new bounds on the size of the multipliers. Examples of possible applications of the new result will be given. In Section 3 we will illustrate the usefulness of the new bounds by proving that the internal penalty method converges to KKT points under the CPLD constraint qualification and the sufficient interior property. We conclude this section by proving that, in fact, convergence of the pure internal penalty method under MFCQ cannot be weakened in some sense. In Section 4 we address the interior point method of Chen and Goldfarb [7]. Using the new bounds for Carathéodory's lemma, we obtain stronger convergence results.

2 Generalized Carathéodory's lemma

The main tool which enables us to prove convergence results under the CPLD condition is Carathéodory's lemma. A simple modification of the classical proof provides us new bounds given by item (4) in Theorem 2.1, which can be very useful in applications of this result.

Theorem 2.1. *If $x = \sum_{i=1}^m \alpha_i v_i$ with $v_i \in \mathbb{R}^n$ and $\alpha_i \neq 0$ for every i, then there exist $I \subset \{1, \dots, m\}$ and scalars $\bar{\alpha}_i$ for every $i \in I$ such that*

(1) $x = \sum_{i \in I} \bar{\alpha}_i v_i$;
(2) $\alpha_i \bar{\alpha}_i > 0$ *for every* $i \in I$;
(3) $\{v_i\}_{i \in I}$ *is linearly independent;*
(4) $|\bar{\alpha}_i| \leq 2^{m-1} |\alpha_i|$ *for every* $i \in I$.

Proof. We assume that $\{v_i\}_{i=1}^m$ is linearly dependent, otherwise the result follows trivially. Then, there exists $\beta \in \mathbb{R}^m$, $\beta \neq 0$ such that $\sum_{i=1}^m \beta_i v_i = 0$. Thus, we may write

$$x = \sum_{i=1}^{m} (\alpha_i - \gamma\beta_i) v_i,$$

for every $\gamma \in \mathbb{R}$. Let $i^* = \text{argmin}_i \left|\frac{\alpha_i}{\beta_i}\right|$ and $\bar{\gamma} = \frac{\alpha_{i^*}}{\beta_{i^*}}$, then $\bar{\gamma}$ is the least modulus coefficient $\frac{\alpha_i}{\beta_i}$. Note that $\bar{\gamma}$ is such that $\alpha_i - \bar{\gamma}\beta_i = 0$ for at least one index $i = i^*$. If $\alpha_i(\alpha_i - \bar{\gamma}\beta_i) < 0$, then $|\alpha_i|^2 = \alpha_i^2 < \alpha_i\bar{\gamma}\beta_i = |\alpha_i||\bar{\gamma}||\beta_i|$, with $\alpha_i \neq 0$, $\beta_i \neq 0$, thus $|\bar{\gamma}| > \left|\frac{\alpha_i}{\beta_i}\right|$ which contradicts the definition of $\bar{\gamma}$. Therefore we conclude that $\alpha_i(\alpha_i - \bar{\gamma}\beta_i) \geq 0$. Also, $|\alpha_i - \bar{\gamma}\beta_i| \leq |\alpha_i| + |\bar{\gamma}||\beta_i| \leq 2|\alpha_i|$, since $|\bar{\gamma}| \leq \left|\frac{\alpha_i}{\beta_i}\right|$ for every i. Including in the sum only the indexes such that $\bar{\alpha}_i = \alpha_i - \bar{\gamma}\beta_i \neq 0$ we are able to write the linear combination x with at least one less vector. We can repeat this procedure until $\{v_i\}_{i \in I}$ is linearly independent with $\alpha_i\bar{\alpha}_i > 0$ and $|\bar{\alpha}_i| \leq 2^{m-1}|\alpha_i|$ for every $i \in I$. □

The new bounds $|\bar{\lambda}_i^k| \leq 2^{m+p-1}|\lambda_i^k|$ for every $i \in I$ and $|\bar{\mu}_j^k| \leq 2^{m+p-1}|\mu_j^k|$ for every $j \in J$ may be useful in many ways. For example, if we have that $\{(\lambda^k, \mu^k)\}$ is bounded, then the same is true for the sequence of new multipliers $\{(\bar{\lambda}^k, \bar{\mu}^k)\}$. The converse is not always true. Consider for instance $x^k = \alpha_1^k v_1^k + \alpha_2^k v_2^k$, $v_1^k \neq 0$ with $\beta_1^k v_1^k + \beta_2^k v_2^k = 0$ for $\beta_1^k = \beta_2^k = 1$, $\alpha_1^k = 1 + 10^k$, $\alpha_2^k = 10^k$. We have $\left|\frac{\alpha_1^k}{\beta_1^k}\right| > \left|\frac{\alpha_2^k}{\beta_2^k}\right|$ for every k, then $\bar{\alpha}_1^k = \alpha_1^k - \left(\frac{\alpha_2^k}{\beta_2^k}\right)\beta_1^k = 1$ and $x^k = \bar{\alpha}_1^k v_1^k$ for every k.

Another situation in which bounds may be useful is when $\mu_j^k \to 0$ for some j. This appears for example in the internal penalty method, in which quasi-KKT points are defined as

$$\nabla f(x^k) + \sum_{i=1}^{m} \lambda_i^k \nabla h_i(x^k) + \sum_{j=1}^{p} \mu_j^k \nabla g_j(x^k) = \varepsilon_k, \tag{2}$$

with $\mu_j^k \to 0$ when $g_j(x^*) < 0$. With the new bounds, we have that $\bar{\mu}_j^k \to 0$ whenever $\mu_j^k \to 0$ (we point out that the reciprocal is also not true, this can be observed by taking the previous counter-example with α_1^k and α_2^k divided by 10^k). This result is crucial to obtain the complementarity condition $\sum_{j=1}^{p} \mu_j g_j(x^*) = 0$ of the KKT condition. We will give the details in the next section, where we also show the impossibility to weaken the hypothesis that guarantee convergence of the pure internal penalty method to KKT points.

3 Internal penalty method

In this section we will consider problem (1) with only inequality constraints:

$$\text{Minimize } f(x) \quad \text{subject to} \quad g(x) \leq 0. \tag{3}$$

The internal penalty method consists of solving the following subproblem:

$$\text{Minimize } f(x) - r_k \sum_{j=1}^{p} \frac{1}{g_j(x)} \quad \text{subject to} \quad g(x) < 0, \tag{4}$$

for a sequence of positive scalars $r_k \to 0$. If there are additional constraints $x \in \Omega$, they are added to the constraints of the subproblems.

It is a well known fact that if x^* is a limit point of the sequence $\{x^k\}$ generated by the internal penalty method, such that x^* satisfies the sufficient interior property, that is, x^* can be approximated by a sequence of strictly feasible points $y^k \to x^*$ $(g(y^k) < 0)$, then x^* is a solution to problem (3) [8, 5, 11].

We assume that x^* is a local solution of problem (3) such that the sufficient interior property holds, and we apply the internal penalty method to:

$$\begin{aligned} &\text{Minimize} \quad f(x) + \tfrac{1}{2}\|x - x^*\|_2^2, \\ &\text{subject to} \quad \|x - x^*\|_2 \leq \delta, \quad g(x) \leq 0, \end{aligned} \tag{5}$$

for a sufficiently small δ (note that x^* is the unique global solution of this problem). The corresponding subproblem is:

$$\begin{aligned} &\text{Minimize} \quad \varphi(x) = f(x) + \tfrac{1}{2}\|x - x^*\|_2^2 - r_k \textstyle\sum_{j=1}^{p} \frac{1}{g_j(x)}, \\ &\text{subject to} \quad \|x - x^*\|_2 \leq \delta, \quad g(x) < 0. \end{aligned} \tag{6}$$

It's a classical result of internal penalty methods that the subproblems (6) admit a global solution x^k [8, 11]. Since every limit point of the sequence of solutions $\{x^k\}$ of (6) is a global solutions of (5), we have that $x^k \to x^*$, thus, for sufficiently large k, we have $\nabla\varphi(x^k) = 0$, that is,

$$\nabla f(x^k) + \sum_{j=1}^{p} \mu_j^k \nabla g_j(x^k) = x^* - x^k, \, \mu_j^k = \frac{r_k}{g_j(x^k)^2}.$$

We can then repeat standard arguments (see [15, 1, 2, 3, 11, 17]) to prove that under the CPLD constraint qualification, there exist $J \subset \{1, \ldots, p\}$ and new non-negative multipliers $\bar{\mu}_j^k$, $j \in J$, given by Carathéodory's lemma, such that we can take a subsequence in wich $\bar{\mu}_j^k$ converges to some non-negative μ_j for every j and

$$\nabla f(x^*) + \sum_{j \in J} \mu_j \nabla g_j(x^*) = 0.$$

To obtain that x^* is a KKT point, we note that if $g_j(x^*) < 0$, then $\mu_j^k \to 0$, thus, by the new bounds $\bar{\mu}_j^k \leq 2^{p-1}\mu_j^k$, we have $\bar{\mu}_j^k \to 0$, that is, $\mu_j = 0$, and thus complementarity holds. So, under the CPLD constraint qualification and the sufficient interior property, limit points of the internal penalty method are KKT points. We will prove next that these hypotheses are equivalent to the Mangasarian-Fromovitz condition when only inequality constraints are present.

For this purpose we shall define the quasi-normality constraint qualification [12, 5].

Definition 3.1. We say that a feasible point x^* to problem (3) satisfies the quasi-normality constraint qualification if x^* satisfies MFCQ, or if there exist $\mu_j \geq 0$ for every $j \in A(x^*)$, not all zero, with $\sum_{j \in A(x^*)} \mu_j \nabla g_j(x^*) = 0$ then there does not exist a sequence $z^k \to x^*$, such that $\mu_j > 0 \Rightarrow g_j(z^k) > 0$ for every $j \in A(x^*)$.

We will use the result proved in [4] that CPLD implies quasi-normality.

Theorem 3.2. *A feasible point x^* satisfies CPLD and the sufficient interior property if, and only if, x^* satisfies MFCQ.*

Proof. Suppose a feasible point x^* satisfies the CPLD condition and the sufficient interior property. Then x^* satisfies the CPLD condition for the problem:

$$\text{Minimize } f(x) \quad \text{subject to} \quad -g_i(x) \leq 0, \quad \forall i \in A(x^*), \tag{7}$$

therefore x^* satisfies the quasi-normality condition for problem (7). If MFCQ does not hold, then there exist not all zero scalars $\mu_j \geq 0$ such that

$$\sum_{j \in A(x^*)} \mu_j \nabla g_j(x^*) = 0,$$

multiplying by -1 we get that MFCQ does not hold for problem (7). Thus, by the quasi-normality for this problem we get that there is no sequence $z^k \to x^*$ such that $\mu_j > 0 \Rightarrow -g_j(z^k) > 0$ for every $j \in A(x^*)$. Since there is at least one index $j \in A(x^*)$ such that $\mu_j > 0$, we conclude that there is no sequence $z^k \to x^*$ such that $g_j(z^k) < 0$, which contradicts the sufficient interior property.

The converse holds trivially since one can easily prove that the sufficient interior property holds using the direction given by the original MFCQ definition, see details in [9, 11]. Clearly, MFCQ also implies the CPLD condition. $\square$

This shows that the internal penalty method converges to a KKT point under MFCQ, and relaxing this condition to CPLD does not provide a stronger result. This is clear since we cannot expect convergence of the internal penalty method if the sufficient interior property does not hold.

We conclude this section with a counter-example showing that a stronger form of Theorem 3.2, in which CPLD is replaced by quasi-normality, does not hold. Consider the problem:

$$\text{Minimize } x \quad \text{subject to} \quad -x^2 \leq 0,$$

at the point $x^* = 0$. It is clear that MFCQ does not hold and the sufficient interior property holds. Also, the quasi-normality condition holds since there are no infeasible points.

In the next section we will use the new bounds obtained in Carathéodory's lemma to prove some stronger convergence results for Chen and Goldfarb's interior point method [7].

4 Chen and Goldfarb's interior point method

Consider the following nonlinear optimization problem:

$$\text{Minimize } f(x) \quad \text{subject to} \quad h(x) = 0, \quad c(x) \geq 0, \tag{8}$$

where $f\colon \mathbb{R}^n \to \mathbb{R}$, $h\colon \mathbb{R}^n \to \mathbb{R}^m$ and $c\colon \mathbb{R}^n \to \mathbb{R}^p$ are twice continuously differentiable functions and $\mathcal{F}^0 = \{x \in \mathbb{R}^n | c(x) > 0\}$ is non-empty.

Chen and Goldfarb's quasi-feasible interior point method consists in two parts: the first part is to apply the log-barrier method to problem (8), obtaining subproblems (FP_μ) below:

$$\begin{array}{ll} \text{Minimize} & f(x) - \mu \sum_{i=1}^{m} \log(c_i(x)) \\ \text{subject to} & h(x) = 0, \quad c(x) > 0, \end{array}$$

for a sequence of positive parameters $\mu \to 0$. The second part consists in applying, for every μ, an ℓ_2-penalty method to solve (FP_μ), yielding subproblems ($\ell_2\mathrm{FP}_\mu$) below:

$$\begin{array}{ll} \text{Minimize} & f(x) - \mu \sum_{i=1}^{m} \log(c_i(x)) + r\|h(x)\|_2 \\ \text{subject to} & c(x) > 0, \end{array}$$

for a sequence of parameters $r \to +\infty$. The idea of the method is to solve ($\ell_2\mathrm{FP}_\mu$) by a Newton-like approach. Here follows the details of the algorithm to solve (FP_μ), for a fixed $\mu > 0$, according to [7].

Algorithm 4.1 (Chen and Goldfarb**).** *Parameters:* $\varepsilon_\mu > 0$, $\sigma \in \left(0, \frac{1}{2}\right)$, $\chi > 1$, $\kappa_1 \in (0, 1)$, $\kappa_2 > 1$, $\pi_\mu = \max\{\mu, 0.1\}$, $\nu > 0$, $0 < \gamma_{\min} < 1 < \gamma_{\max}$, $k := 0$. *Given initial interior points* $x^0 \in \mathcal{F}^0$, $u^0 > 0$, *an initial penalty parameter* $r_0 > 0$ *and an initial approximation* $\mathcal{H}^0 \in \mathbb{R}^{n\times n}$ *for the Hessian of the Lagrangian* $L(x, \lambda, y) = f(x) - \lambda^{\mathrm{T}} c(x) + y^{\mathrm{T}} h(x)$.

Step 1: **Search direction**

Modify $\mathcal{H}^k$*, if necessary, such that condition C-5 below, holds:*

$$\begin{cases} d^{\mathrm{T}} \tilde{\mathcal{H}}^k d \geq \nu \|d\|^2, \forall d \neq 0 & \text{if } \|h(x^k)\|_2 > 0 \\ d^{\mathrm{T}} \tilde{\mathcal{H}}^k d \geq \nu \|d\|^2, \forall d \neq 0, \nabla h(x^k)^{\mathrm{T}} d = 0 & \text{if } \|h(x^k)\|_2 = 0 \end{cases},$$

where

$$\tilde{\mathcal{H}}^k = \begin{cases} \hat{\mathcal{H}}^k & \text{if } \|h(x^k)\|_2 = 0 \\ \hat{\mathcal{H}}^k + \dfrac{1}{\delta_{x^k}} \nabla h(x^k)^{\mathrm{T}} \nabla h(x^k) & \text{if } \|h(x^k)\|_2 > 0 \end{cases},$$

with

$$\delta_{x^k} = \frac{\|h(x^k)\|_2}{r_k}, \ \hat{\mathcal{H}}^k = \mathcal{H}^k + \nabla c(x^k) C(x^k)^{-1} \mathcal{U}^k \nabla c(x^k)^{\mathrm{T}},$$
$$\mathcal{U}^k = diag(u^k), \ C(x^k) = diag(c(x^k)).$$

Calculate $(\Delta x^k, \lambda^k, y^k)$*, solution of the KKT system*

$$M^k \begin{pmatrix} \Delta x^k \\ \lambda^k \\ y^k \end{pmatrix} = \begin{pmatrix} -\nabla f(x^k) \\ \mu e \\ -h(x^k) \end{pmatrix}, \tag{9}$$

where

$$M^k = \begin{pmatrix} \mathcal{H}^k & -\nabla c(x^k) & \nabla h(x^k) \\ \mathcal{U}^k \nabla c(x^k)^{\mathrm{T}} & C(x^k) & 0 \\ \nabla h(x^k)^{\mathrm{T}} & 0 & -\delta_{x^k} \mathrm{I} \end{pmatrix},$$

I *is the identity matrix and* $e = (1, \ldots, 1)$ *of appropriate dimensions.*

Step 2: **Termination**

Stop if $\left\| \begin{pmatrix} \nabla_x L(x^k, \lambda^k, y^k) \\ C(x^k)\lambda^k - \mu e \\ h(x^k) \end{pmatrix} \right\|_2 \leq \varepsilon_\mu$ *and* $\lambda^k \geq -\varepsilon_\mu e$*.*

Step 3: **Penalty parameter update**

If the following conditions hold

(C-1): $\|h(x^k)\|_2 > 0$*,*

(C-2): $\|\Delta x^k\|_2 \leq \pi_\mu$*,*

(C-3): $\kappa_1 \mu e \leq C(x^k)\lambda^k \leq \kappa_2 \mu e$*,*

(C-4): $\|\bar{w}^k\| < \pi_\mu$, $\bar{w}^k = y^k - \frac{r_k}{\|h(x^k)\|_2} h(x^k)$*,*

then

$r_{k+1} := \chi r_k$*,*

$(x^{k+1}, u^{k+1}, \mathcal{H}^{k+1}) = (x^k, u^k, \mathcal{H}^k)$*,*

$k := k + 1$*,*

and go back to Step 1.

Step 4: **Line search**

Initialize $t_k = 1$ *and successively divide it by* 2*, if necessary, until the following conditions hold*

T-1: $c(x^k + t_k \Delta x^k) > 0$

T-2: $\Phi_{\mu,r_k}(x^k + t_k \Delta x^k) - \Phi_{\mu,r_k}(x^k) \leq -\sigma t_k (\Delta x^k)^{\mathrm{T}} \tilde{\mathcal{H}} \Delta x^k$*,*

where $\Phi_{\mu,r} = f(x) - \mu \sum_{i=1}^m \log(c_i(x)) + r\|h(x)\|_2$*.*

Step 5: **Update**
Define u_i^{k+1} *to be the projection of* λ_i^k *on the interval*

$$\left[\mu\frac{\gamma_{\min}}{c_i(x^k)}, \mu\frac{\gamma_{\max}}{c_i(x^k)}\right],$$

for each i.
$x^{k+1} = x^k + t_k \Delta x^k$,
$r_{k+1} = r_k$.
Calculate the new estimative $\mathcal{H}^{k+1}$ *for the Hessian of the Lagrangian.*
$k := k + 1$
go back to Step 1.

In [7], the authors prove that if the primal iterate sequence $\{x^k\}$ lies in a bounded set and the modified Hessian sequence $\{\mathcal{H}^k\}$ is bounded, then, under MFCQ, the limit points of $\{x^k\}$ are stationary for an infeasibility measure problem, and, if the limit point is feasible, KKT condition holds for FP_μ.

We will prove, using the new bounds for Carathéodory's lemma, that if the penalty parameter $r_k \to +\infty$ and x^* is infeasible with respect to the equality constraints, then we can weaken the constraint qualification hypothesis and assume only the CPLD condition to obtain that x^* is stationary for an infeasibility measure problem.

Proposition 4.2. *If the penalty parameter* $r_k \to +\infty$ *and* x^* *is a limit point of the sequence* $\{x^k\}$ *generated by Algorithm* 4.1 *such that* $\|h(x^*)\|_2 > 0$, *and* x^* *satisfies the CPLD constraint qualification for problem*

$$\text{Minimize } \|h(x)\|_2^2 \quad \text{subject to} \quad c(x) \geq 0, \tag{10}$$

then x^* *is a KKT point for this problem.*

Proof. Let's consider a subsequence $\{x^k\}$ such that $x^k \to x^*$ and r_k is increased for every k, thus, conditions C-1 to C-4 are fulfilled. From (9), we can write

$$-\mathcal{H}^k \Delta x^k + \sum_{i=1}^{m} \lambda_i^k \nabla c_i(x^k) - \sum_{i=1}^{p} \left(r_k \frac{h_i(x^k)}{\|h(x^k)\|_2} + \bar{w}_i^k \right) \nabla h_i(x^k) = \nabla f(x^k).$$

By C-3, $0 < \kappa_1\mu \leq c_i(x^k)\lambda_i^k$, thus $\lambda_i^k > 0$, since $c_i(x^k) > 0$. By Carathéodory's lemma, there exist a subset $I_k \subset \{1, \dots, m\}$ and scalars $\bar{\lambda}_i^k > 0$ such that

$$\begin{aligned}\textstyle\sum_{i\in I_k} \bar{\lambda}_i^k \nabla c_i(x^k) - \sum_{i=1}^p \frac{r_k}{\|h(x^k)\|} h_i(x^k)\nabla h_i(x^k) = \\ = \nabla f(x^k) + \mathcal{H}^k \Delta x^k + \textstyle\sum_{i=1}^p \bar{w}_i^k \nabla h_i(x^k),\end{aligned} \tag{11}$$

and $\{\nabla c_i(x^k)\}_{i\in I_k}$ is linearly independent.

Let's take a subsequence such that $I_k = I$. Since $\lambda_i^k \leq \frac{\kappa_2\mu}{c_i(x^k)}$, from the new bounds on Carathéodory's lemma, we have $\bar{\lambda}_i^k \leq \frac{2^{m-1}\kappa_2\mu}{c_i(x^k)}$, hence $0 < c_i(x^k)\bar{\lambda}_i^k \leq 2^{m-1}\kappa_2\mu$.

If $\left\{\frac{\bar{\lambda}^k}{r_k}\right\}$ admits a limited subsequence, we may consider a subsequence such that $\frac{\bar{\lambda}^k}{r_k} \to \lambda'$. Dividing (11) by r_k, taking limits for k and observing that $\{\Delta x^k\}$ and $\{\bar{w}^k\}$ are limited sequences since C-2 and C-4 hold, we obtain

$$\sum_{i=1}^p \frac{h_i(x^*)}{\|h(x^*)\|_2}\nabla h_i(x^*) - \sum_{i\in I} \lambda_i' \nabla c_i(x^*) = 0,$$

and

$$0 < c_i(x^k)\frac{\bar{\lambda}_i^k}{r_k} \leq \frac{2^{m-1}\kappa_2\mu}{r_k} \Rightarrow c_i(x^*)\lambda_i' = 0,$$

thus x^* is a KKT point of problem (10).

In the case $\frac{\bar{\lambda}^k}{r_k} \to +\infty$, dividing (11) by $\|\bar{\lambda}^k\|_\infty$ and taking limits for a subsequence such that $\frac{\bar{\lambda}^k}{\|\bar{\lambda}^k\|_\infty} \to \bar{\lambda} \geq 0, \bar{\lambda} \neq 0$ we have

$$\sum_{i\in \mathrm{I}} \bar{\lambda}_i \nabla c_i(x^*) = 0,$$

and

$$0 < c_i(x^k)\frac{\bar{\lambda}_i^k}{\|\bar{\lambda}^k\|_\infty} \leq \frac{2^{m-1}\kappa_2\mu}{\|\bar{\lambda}^k\|_\infty} \Rightarrow c_i(x^*)\bar{\lambda}_i = 0.$$

Excluding from the set I all indexes such that $\bar{\lambda}_i = 0$, we have $I \subset A(x^*)$ and CPLD is not fulfilled. □

Chen and Goldfarb's algorithm to solve (8) consists of defining positive sequences $\mu_k \to 0$, $\varepsilon_k \to 0$ and using Algorithm 4.1 to approximately solve

(FP_{μ_k}), that is, obtaining iterates satisfying the stopping criterium of Step 2. In this case, they prove that under MFCQ, limit points are stationary for an infeasibility measure problem, and in case the limit point is feasible, KKT condition holds for (8). We will prove that, under CPLD, if the limit point is feasible, then the KKT condition holds.

Proposition 4.3. *Assume x^* is a limit point of the sequence $\{x^k\}$ generated by Chen and Goldfarb's algorithm to solve* (8)*, such that x^* satisfies the CPLD constraint qualification for problem* (8). *Assume also that Algorithm* (4.1) *is well-defined, thus x^* is a KKT point of problem* (8).

Proof. Let's take a subsequence such that $x^k \to x^*$. By the stopping criterium of Step 2, we have

$$\nabla f(x^k) - \sum_{i=1}^{m} \lambda_i^k \nabla c_i(x^k) + \sum_{i=1}^{p} y_i^k \nabla h_i(x^k) = \delta_1^k \tag{12}$$

$$C(x^k)\lambda^k - \mu_k e = \delta_2^k \tag{13}$$

$$h(x^k) = \delta_3^k, \tag{14}$$

such that $\|(\delta_1^k, \delta_2^k, \delta_3^k)\|_2 \leq \varepsilon_k$ e $\lambda_i^k \geq -\varepsilon_k$.

By Carathéodory's lemma, there are scalars $\bar{\lambda}_i^k$, $\bar{y}_i^k$, and subsets

$$I_k \subset \{1, \ldots, m\}, J_k \subset \{1, \ldots, p\}$$

(we will take a subsequence that satisfies $I_k = I$ and $J_k = J$ for every k) such that

$$\nabla f(x^k) - \sum_{i \in I} \bar{\lambda}_i^k \nabla c_i(x^k) + \sum_{i \in J} \bar{y}_i^k \nabla h_i(x^k) = \delta_1^k, \tag{15}$$

$\{\nabla c_i(x^k)\}_{i\in \mathrm{I}} \cup \{\nabla h_i(x^k)\}_{i \in \mathrm{J}}$ is linearly independent and $|\bar{\lambda}_i^k| \leq 2^{m+p-1}|\lambda_i^k|$, $\bar{\lambda}_i^k \lambda_i^k > 0$, thus $\bar{\lambda}_i^k \geq -2^{m+p-1}\varepsilon_k$. Define $\alpha_k = \|(\bar{\lambda}^k, \bar{y}^k)\|_\infty$.

If $\{\alpha_k\}$ admits a limited subsequence, let's consider a subsequence such that $(\bar{\lambda}^k, \bar{y}^k) \to (\bar{\lambda}, \bar{y})$. Since $\varepsilon_k \to 0$, taking limits in (15) we obtain

$$\nabla f(x^*) - \sum_{i \in \mathrm{I}} \bar{\lambda}_i \nabla c_i(x^*) + \sum_{i \in \mathrm{J}} \bar{y}_i \nabla h_i(x^*) = 0.$$

Since $\bar{\lambda}_i^k \geq -2^{m+p-1}\varepsilon_k$, we have $\bar{\lambda} \geq 0$, and from (13) we get

$$|\bar{\lambda}_i^k| \leq 2^{m+p-1}|\lambda_i^k| \leq 2^{m+p-1}\frac{|[\delta_2^k]_i + \mu_k|}{c_i(x^k)},$$

which implies $\bar{\lambda}_i c_i(x^*) = 0$. By (14) we get $h(x^*) = 0$, thus x^* is a KKT point of problem (8).

If $\alpha_k \to +\infty$ consider a subsequence such that $\left(\frac{\bar{\lambda}_i^k}{\alpha_k}, \frac{\bar{y}_i^k}{\alpha_k}\right) \to (\hat{\lambda}, \hat{y}) \neq 0$, and since

$$\frac{\bar{\lambda}_i^k}{\alpha_k} \geq -\frac{2^{m+p-1}\varepsilon_k}{\alpha_k} \to 0,$$

we have $\hat{\lambda} \geq 0$.

Dividing (15) by α_k and taking limits we get

$$\sum_{i\in I}\hat{\lambda}_i \nabla c_i(x^*) - \sum_{i\in J}\hat{y}_i \nabla h_i(x^*) = 0,$$

with

$$|\bar{\lambda}_i^k| \leq 2^{m+p-1}|\lambda_i^k| \leq 2^{m+p-1}\frac{|[\delta_2^k]_i + \mu_k|}{c_i(x^k)},$$

thus, multiplying this inequality by $\frac{c_i(x^k)}{\alpha_k}$ and taking limits we get $c_i(x^*)\hat{\lambda}_i = 0$, therefore, removing from I all indexes such that $\hat{\lambda}_i = 0$, we get $I \subset A(x^*)$, which contradicts CPLD. □

We point out that since problem (8) includes also equality constraints, the result of Theorem 3.2 does not apply.

Acknowledgement. The author is indebted to an anonymous referee for insightful comments and suggestions.

REFERENCES

[1] R. Andreani, E.G. Birgin, J.M. Martínez and M.L. Schuverdt, *On augmented lagrangian methods with general lower-level constraints.* SIAM Journal on Optimization, **18** (2007), 1286–1309.

[2] R. Andreani, E.G. Birgin, J.M. Martínez and M.L. Schuverdt. *Augmented lagrangian methods under the constant positive linear dependence constraint qualification.* Mathematical Programming, **112** (2008), 5–32.

[3] R. Andreani, G. Haeser and J.M. Martínez, *On sequential optimality conditions for smooth constrained optimization.* To appear in Optimization, 2010. Available at Optimization Online http://www.optimization-online.org/DB_HTML/2009/06/2326.html.

[4] R. Andreani, J.M. Martínez and M.L. Schuverdt, *On the relation between constant positive linear dependence condition and quasinormality constraint qualification.* Journal of Optimization Theory and Applications, **125** (2005), 473–485.

[5] D.P. Bertsekas, *Nonlinear Programming: 2nd Edition*. Athena Scientific (1999).

[6] C. Carathéodory, *Über den Variabilitätsbereich der Fourierschen Konstanten von positiven harmonischen Funktionen.* Rend. Circ. Mat. Palermo, **32** (1911), 193–217.

[7] L. Chen and D. Goldfarb, *Interior-point l2-penalty methods for nonlinear programming with strong global convergence.* Mathematical Programming, **108**(1) (2006), 1–36.

[8] A.V. Fiacco and G.P. McCormick, *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*. Wiley, 1968 (reprinted, SIAM, 1990).

[9] A. Forsgren, P.E. Gill and M.H. Wright, *Interior methods for nonlinear optimization.* SIAM Review, **44**(4) (2002), 525–597.

[10] M.A. Gomes-Ruggiero, J.M. Martínez and S.A. Santos, *Spectral projected gradient method with inexact restoration for minimization with nonconvex constraints.* SIAM Journal on Scientific Computing, **31**(3) (2009), 1628–1652.

[11] G. Haeser, *Condições sequenciais de otimalidade.* PhD thesis, IMECC-UNICAMP, Departamento de Matemática Aplicada, Campinas-SP, Brazil, 2009. http://libdigi.unicamp.br/document/?code=000466897.

[12] M.R. Hestenes, *Optimization theory: the finite dimensional case*. John Wiley (1975).

[13] R. Janin, *Directional derivative of the marginal function in nonlinear programming.* Mathematical Programming Study, **21** (1984), 110–126.

[14] O.L. Mangasarian and S. Fromovitz, *The Fritz John optimality conditions in the presence of equality and inequality constraints.* Journal of Mathematical Analysis and Applications, **17** (1967), 37–47.

[15] L. Qi and Z. Wei, *On the constant positive linear dependence condition and its applications to* SQP *methods.* SIAM Journal on Optimization, **10**(4) (2000), 963–981.

[16] R.T. Rockafellar, *Lagrange multipliers and optimality.* SIAM Review, **35**(2) (1993), 183–238.

[17] M.L. Schuverdt, *Métodos de Lagrangiano aumentado com convergência utilizando a condição de dependência linear positiva constante.* PhD thesis, IMECC-UNICAMP, Departamento de Matemática Aplicada, Campinas-SP, Brazil (2006). http://libdigi.unicamp.br/document/?code=vtls000375801.

[18] E. Steinitz, *Bedingt konvergente Reihen und konvexe Systeme i-iv.* J. Reine Angew. Math., **143** (1913), 128–175.