*Electronic Journal of Differential Equations*, Vol. 2011 (2011), No. 44, pp. 1–16.
ISSN: 1072-6691. URL: http://ejde.math.txstate.edu or http://ejde.math.unt.edu
ftp ejde.math.txstate.edu

# OPTIMIZING SECOND-ORDER DIFFERENTIAL EQUATION SYSTEMS

TAMÁS HAJBA

Abstract. In this article we study some continuous versions of the Fletcher-Reeves iteration for minimization described by a system of second-order differential equations. This problem has been studied in earlier papers [19, 20] under the assumption that the minimizing function is strongly convex. Now instead of the strong convexity, only the convexity of the minimizing function will be required. We will use the Tikhonov regularization [28, 29] to obtain the minimal norm solution as the asymptotically stable limit point of the trajectories.

## 1. Introduction

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a convex, continuously differentiable function. Let us consider the minimization problem

$$\min_{\mathbf{x}\in\mathbb{R}^n} f(\mathbf{x}), \tag{1.1}$$

where the function $f(\mathbf{x})$ satisfies the following conditions:

$$f_* = \inf f(\mathbf{x}) > -\infty, \quad X_* = \{\mathbf{x} \in \mathbb{R}^n : f(\mathbf{x}) = f_*\} \neq \emptyset. \tag{1.2}$$

Several methods have been developed for the solution of this problem. The methods generated with an iterative process can be modelled with differential equations. These differential equations are usually called the continuous version of the method.

Modelling the iterative numerical methods of optimization with differential equations has been investigated in several papers. Some of them deal with either the gradient or the Newton's method and model the given method by a system of first order differential equations (e.g. [2, 3, 5, 8, 10, 14, 15, 16, 21, 9, 22, 23, 33, 34] etc.). In this article we investigate two models of the continuous version of the Fletcher-Reeves iteraiton. Both of them lead to the analysis of second-order differential equation systems (shortly SODE system). One of these models has not been studied earlier.

There is another approach to the study of second order differential equations with the optimization that arise in physical problems such as the heavy ball with friction. Results concerning such type of second-order differential equation models can be found in [1, 4, 7, 12, 13, 18, 31, 30]. There are also some papers discussing

---

2000 *Mathematics Subject Classification.* 90C25, 65K05, 34D05.
*Key words and phrases.* Fletcher-Reeves iteration; second-order differential equation; minimizing trajectory; stationary point in limit; Lyapunov-type methods.

Submitted October 25, 2010. Published March 31, 2011.

higher order methods; e.g. [32, 30, 26]. However, the mentioned papers deal with SODE systems that are linear in $\dot{\mathbf{x}}$. Since the Fletcher-Reeves iteration uses the new state point in the construction of the new direction, our system of second-order differential equations will not necessary be linear in the first derivative vector $\dot{\mathbf{x}}$. In connection with the optimization such type of second-order differential equation has been investigated in [19] assuming the minimizing function being strongly convex. The minimizing property of such type of second-order differential equation has not been investigated yet when the function is convex but necessary strongly convex. Since in this case the uniqueness of the minimum point can not be guaranteed the Tikhonov regularization will be used to obtain the so called minimal norm solution.

In this paper we consider the SODE system describing the so called heavy ball with friction as a simplification of the continuous version of the Fletcher-Reeves iteration using the old state point in the construction of the new direction. Since the regularized version of this type of differential equation is known only if the coefficient of $\dot{\mathbf{x}}$ is constant (see [12, 13]) we will show that the convergence of the trajectories to the minimal norm solution is valid with function-coefficient, too.

## 2. Second-order differential equation models of minimization

As it was pointed out in [23] the minimization models modelled by first order differential systems can be divided into two classes. Those models described by a system of first order differential equations for which the point $\mathbf{x}_*$ is a stationary point of the system belong to the first class. In this case the convergence of the trajectories to $\mathbf{x}_*$ is equivalent with the asymptotic stability of $\mathbf{x}_*$, therefore the Lyapunov function methods (see e.g. in [27]) are useful to prove the convergence with an appropriately chosen Lyapunov function (see e.g. [15, 16, 33]). To the second class of the models belong those continuous first order models, for which the minimum point is not stationary, but along the trajectories the right hand side vector of the differential equation system tends to the null-vector if $t \to \infty$. Following [23] we say in this case, that $\mathbf{x}_*$ is *stationary in limit*.

We extend this definition for the SODE systems, too. We will say, that a point is *stationary point* or *stationary in limit point* of a SODE system if it is stationary or stationary in limit point respectively for the equivalent first order system.

We will say, that a SODE system is a *minimizing model* for the minimization problem (1.1)-(1.2) if along its trajectories $\lim_{t\to\infty} f(\mathbf{x}(t)) = f_*$. It is *convergent* if any trajectory converges in norm to some $\mathbf{x}_* \in X_*$; i.e., $\|\mathbf{x}(t) - \mathbf{x}_*\| \to 0$. The trajectories of a convergent minimizing model are called *minimizing trajectories*.

It will be seen that the continuous version of the regularized Fletcher-Reeves iteration belongs to the class of methods stationary in limit both in the general and in the simplified cases.

As it was shown in [22, 23] the Lyapunov-type methods are also applicable to prove the convergence of the trajectories to a point stationary in limit. Namely, it has been proved, if the chosen Lyapunov function along the trajectory of the differential equality systems satisfies certain differential inequality on $[t_0, \infty)$, then it tends to zero if $t \to \infty$. This technique will be used in our proofs, too.

Here we describe one of the appropriate lemmas from [22] which will be fundamental in our investigation to prove the convergence of the trajectories to a stationary in limit minimum point.

**Lemma 2.1.** *Suppose that there exists $T_0 \geq 0$ such that*

(1) *for every fixed* $\tau \geq T_0$ *the scalar function* $g(t,\tau)$ *is defined and non-negative for all* $T_0 \leq t < \tau$ *and* $g(T_0,\tau) \leq K$ *uniformly in* $\tau$*, furthermore, it is continuously differentiable in* $t$*;*
(2) $g(t,\tau)$ *satisfies the following differential inequality:*

$$\frac{\mathrm{d}}{\mathrm{d}t} g(t,\tau) \leq -a(t)g(t,\tau) + b(t)(\tau - t)^s \tag{2.1}$$

*for* $T_0 \leq t < \tau$ *where* $s$ *is nonnegative integer and the functions* $a(t) > 0$ *and* $b(t)$ *are defined for all* $t \geq T_0$ *and integrable on any finite interval of* $[T_0, \infty)$ *and they are endowed with the following properties:*
(a) $\int_{T_0}^{\infty} a(t)dt = \infty$,
(b) $\lim_{t\to\infty} \frac{b(t)}{a^{s+1}(t)} = 0$,
(c) *In the case* $s \geq 1$ *the function* $a(t)$ *is differentiable and*

$$\lim_{t\to\infty} \frac{\dot{a}(t)}{a^2(t)} = 0.$$

*Then* $\lim_{\tau\to\infty} g(\tau,\tau) = 0$*.*

*Proof.* From (2.1) we have that

$$0 \leq g(\tau,\tau) \leq g(T_0,\tau)\mathrm{e}^{-\int_{T_0}^{\tau} a(\nu)\mathrm{d}\nu} + \int_{T_0}^{\tau} b(\theta)(\tau-\theta)^s \mathrm{e}^{\int_{\tau}^{\theta} a(\nu)\mathrm{d}\nu}\mathrm{d}\theta.$$

The convergence of the first term to zero follows from the condition 2(a).

By induction on $s$ it can be proved that for all nonnegative integer $s$ the limit $\lim_{\tau\to\infty} \exp\big(\int_{T_0}^{\tau} a(\nu)\mathrm{d}\nu\big)a^s(\tau) = \infty$ holds true and hence we can estimate the second term by applying $(s+1)$ times the L'Hospital rule and the conditions 2(b) and 2(c):

$$\begin{aligned} &\lim_{\tau\to\infty} \frac{\int_{T_0}^{\tau} b(\theta)(\tau-\theta)^s \exp\big(\int_{T_0}^{\theta} a(\nu)\mathrm{d}\nu\big)\mathrm{d}\theta}{\exp\big(\int_{T_0}^{\tau} a(\nu)\mathrm{d}\nu\big)\mathrm{d}\theta} \\ &= \lim_{\tau\to\infty} \frac{b(\tau)s!}{[a(\tau)]^{s+1}} \lim_{\tau\to\infty} \prod_{j=0}^{s} \frac{1}{1 + \frac{j\,\dot{a}(\tau)}{[a(\tau)]^2}} = 0. \end{aligned}$$

□

The function $g(t,\tau)$ in the lemma constructed for a SODE problem will be called *Lyapunov-like function* of the model.

The focus of our interest is to formulate such SODE systems which are convergent and minimizing and for which the minimum point with the minimal norm is a stationary or stationary in limit point. Our motivation to construct such models of minimization was the following:

The Fletcher-Reeves iteration to minimize a function of $n$ variables starting from $\mathbf{x}_0$ and $\mathbf{p}_0 = -f'(\mathbf{x}_0)$ computes the pair of points

$$\begin{gathered} \mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{p}_k \\ \mathbf{p}_{k+1} = -f'(\mathbf{x}_{k+1}) + \delta_k \mathbf{p}_k \quad k = 1, 2, \dots. \end{gathered}$$

To obtain a convergent process we have to use well defined (here not detailed) changing rules for the sequences $\alpha_k$ and $\delta_k$.

Taking into consideration that the Fletcher-Reeves iteration uses the new state point in the construction of the new direction it is easy to see that this iteration can be considered as the Euler discretization with step size 1 of the non-autonomous first-order differential equation system of $2n$ variables

$$\dot{\mathbf{x}} = \alpha(t)\mathbf{p} \tag{2.2}$$

$$\dot{\mathbf{p}} = -\nabla f(\mathbf{x} + \alpha(t)\mathbf{p}) + \beta(t)\mathbf{p}, \tag{2.3}$$

$$\mathbf{x}(t_0) = \mathbf{x}_0, \quad \mathbf{p}(t_0) = \mathbf{p}_0, \tag{2.4}$$

where the changing rule of the parameters are described by continuous functions.

We will refer to the model (2.2)-(2.3) as *general model* (shortly Model G-FR) of the continuous version of the Fletcher-Reeves iteration. This model is equivalent with the SODE system of $n$ variable

$$\ddot{\mathbf{x}} + \gamma(t))\dot{\mathbf{x}} + \alpha(t)\nabla f(\mathbf{x} + \dot{\mathbf{x}}) = 0, \tag{2.5}$$

$$\mathbf{x}(t_0) = \mathbf{x}_0, \quad \dot{\mathbf{x}}(t_0) = \alpha(t_0)\mathbf{p}_0, \tag{2.6}$$

where

$$\gamma(t) = -\beta(t) - \frac{\dot{\alpha}(t)}{\alpha(t)}. \tag{2.7}$$

If we approximate $\nabla f(\mathbf{x}(t) + \alpha(t)\mathbf{p}(t))$ with $\nabla f(\mathbf{x}(t))$ then we obtain a much more simple model, namely

$$\dot{\mathbf{x}} = \alpha(t)\mathbf{p} \tag{2.8}$$

$$\dot{\mathbf{p}} = -\nabla f(\mathbf{x}) + \beta(t)\mathbf{p} \tag{2.9}$$

with the initial values (2.4). This model will be called *simplified model* (shortly Model S-FR) of the continuous version of the Fletcher-Reeves iteration. This model is equivalent with the SODE system

$$\ddot{\mathbf{x}} + \gamma(t)\dot{\mathbf{x}} + \alpha(t)\nabla f(\mathbf{x}) = 0 \tag{2.10}$$

with the initial values (2.6).

In [19] the asymptotic behavior of the trajectories of the Model G-FR and Model S-FR have been analyzed. It has been proved that under the assumption of the strong convexity of the function $f(\mathbf{x})$ there are such harmonizing conditions between the parameter functions $\alpha(t)$ and $\beta(t)$ which ensure that the differential equation system (2.8)-(2.9) or (2.2)-(2.3) is minimizing and the minimum point $\mathbf{x}_*$ is an asymptotically stable stationary point to which any trajectory tends if $t \to \infty$. Furthermore, several class of pairs of the functions $\alpha(t)$ and $\beta(t)$ satisfying the harmonization conditions has been given in [20].

The behavior of the trajectories of the second-order differential equation (2.10) has been investigated in several papers assuming that $\gamma(t)$ is a positive constant function and $\alpha(t) \equiv 1$ (e.g. [1, 4, 5, 7, 18]). This is the so called *heavy ball with friction* model. A detailed discussion of the minimizing properties of the trajectories of (2.10) with positive $\gamma(t)$ and $\alpha(t) \equiv 1$ functions have been given in the papers [12, 13].

## 3. Convergence theorems of the regularized SODE models

The strong convexity is too strict condition for most of the practical optimization problems.

In this paper we will require only the convexity of the minimizing function. But under this weaker assumption we can not expect that the set of minimum points consists of only one point. Therefore, as it will be shown in a numerical example in Section 4.2, it can happen that either the discrete Fletcher-Reeves method or its continuous versions stop in different minimum points starting from different initial points.

To avoid these problems a regularization technique is generally used. The regularization means that the minimizing function will be approximated with a bundle of strong convex functions depending on a damping parameter. Choosing the appropriate damping parameter one can expect that the sequence of the unique minimum points of the auxiliary functions tends to one of the well defined minimum point of the original minimizing function independently from the starting point. The possibility of this type of regularization is based on the following lemma due to Tikhonov [28, 29].

**Lemma 3.1.** *Let* $f : \mathbb{R}^n \to \mathbb{R}$ *be a convex function satisfying* (1.2) *and* $\lambda_k$, $k = 1, 2, \dots$ *be a positive monotone decreasing sequence for which* $\lim_{k\to\infty} \lambda_k = 0$. *Let the auxiliary strong convex function bundle defined by the sequence*

$$F_k(\mathbf{x}) = f(\mathbf{x}) + \frac{1}{2}\lambda_k\|\mathbf{x}\|^2, \quad k = 1, 2, \dots$$

*and let* $\mathbf{x}_k$ *denote the unique minimum point of* $F_k(\mathbf{x})$. *Then*

$$\lim_{k\to\infty} \|\mathbf{x}_k - \mathbf{x}_*\| = 0,$$

*where* $\mathbf{x}_*$ *is the minimal norm solution of* (1.1)*; i.e.,*

$$f(\mathbf{x}_*) = \inf_{\mathbf{x}\in\mathbb{R}^n} f(\mathbf{x}) \quad \textit{and} \quad \inf_{\mathbf{x}\in X_*} \|\mathbf{x}\| = \|\mathbf{x}_*\|,$$

*where* $X_*$ *is given in* (1.2).

In lots of minimization methods the damping parameter can be synchronized with the parameters of the used method modifying it step by step. Such regularized method is the Levenberg-Marquard algorithm [24, 25] for the Newton's method which was developed independently from the Tikhonov-regularization.

The regularization of the minimization methods modelled by differential equation systems means that instead of the function $f(\mathbf{x})$ and its first and higher order partial derivatives the auxiliary function

$$F(\mathbf{x}, t) = f(\mathbf{x}) + \frac{1}{2}\lambda(t)\|\mathbf{x}\|^2. \tag{3.1}$$

and its partial derivatives are used where the damping parameter $\lambda(t)$ continuously changes in time.

For the continuous gradient method which is modelled by first-order system of differential equations the regularization technique was applied in [22]. Other approaches can be found in [5] and [14]. Regularized second and higher order models have been examined e.g. in [6, 11, 31, 30]. Since second order dynamics are generally not descent methods hence they allow to overcome some drawbacks of the steepest descent method.

In the following we will discuss the convergence of the regularized methods modelled by (2.8)-(2.9), (resp. by (2.10)) and by (2.2)-(2.3), (resp. by (2.5)).

### 3.1. Regularized general model.

The *regularized general model* (shortly RG-FR model) to solve the problem (1.1) can be given by the following first order system of differential equations of $2n$ variables:

$$\dot{\mathbf{x}} = \alpha(t)\mathbf{p} \tag{3.2}$$

$$\dot{\mathbf{p}} = -\nabla_{\mathbf{x}} F(\mathbf{x} + \alpha(t)\mathbf{p}, t) + \beta(t)\mathbf{p} \tag{3.3}$$

with the initial values (2.4), where $F(\mathbf{x}, t)$ is given by (3.1) and the function $\lambda(t)$ is a monotone decreasing positive function. This system is equivalent with the SODE system of $n$ variables

$$\ddot{\mathbf{x}} + \gamma(t)\dot{\mathbf{x}} + \alpha(t)\nabla_{\mathbf{x}} F(\mathbf{x} + \dot{\mathbf{x}}, t) = 0. \tag{3.4}$$

with the initial values (2.6), where $\gamma(t)$ is given by (2.7).

It can be seen that the difference between the RG-FR model and the G-FR model is that instead of the partial derivatives of the function $f(\mathbf{x})$ the partial derivatives of the auxiliary function $F(\mathbf{x}, t)$ are used.

**Proposition 3.2.** *Let us assume that the following hypotheses are satisfied:*

(1) *In the minimization problem* (1.1) $f$ *is defined and continuously differentiable convex function on* $\mathbb{R}^n$ *and its gradient* $\nabla f$ *is local Lipschitz continuous; i.e., it is Lipschitz continuous on all bounded subsets of* $\mathbb{R}^n$ *and the conditions given in* (1.2) *on page 1 hold;*
(2) *The parameter functions* $\alpha(t)$, $\beta(t)$ *and* $\gamma(t)$ *of the systems* (3.2)-(3.3) *and* (3.4) *fulfills the following conditions:*
   (a) $\alpha(t)$ *is a positive, upper bounded and continuously differentiable and* $\beta(t)$ *is a negative lower bounded function on* $[t_0, \infty)$ *;*
   (b) $\gamma(t)$ *is a monotone non-increasing, continuously differentiable function on* $[t_0, \infty)$ *and* $\inf_{t \geq t_0} \gamma(t) > 1$ *;*
(3) *For the damping parameter* $\lambda(t)$ *the following assumptions hold:*
   (a) $\lambda(t)$ *is a positive continuously differentiable monotone decreasing function on* $[t_0, \infty)$ *and convex for all* $t \geq t_1$*;*
   (b) $\alpha(t)\lambda(t)$ *is a monotone non-increasing function;*
   (c) $\lim_{t\to\infty} \lambda(t) = \lim_{t\to\infty} \dot{\lambda}(t) = 0$,

$$\lim_{t\to\infty} \frac{\dot{\alpha}(t)}{\alpha^2(t)\lambda(t)} = \lim_{t\to\infty} \frac{\dot{\lambda}(t)}{\alpha(t)\lambda^2(t)} = \lim_{t\to\infty} \frac{\dot{\lambda}(t)}{\alpha^2(t)\lambda(t)} = 0;$$

   (d) $\int_{t_0}^{\infty} \alpha(t)\lambda(t) = \infty$.

*Then*

(1) *the trajectories of* (3.2)-(3.3), *respectively of* (3.4) *exist and unique on the whole half-line* $[t_0, \infty)$ *with any initial point* (2.4)*;*
(2) *the RS-FR model given by* (3.2)-(3.3) *(or* (3.4)*) is minimizing; i.e.,*

$$\lim_{t\to\infty} f(\mathbf{x}(t)) = f(\mathbf{x}_*) = \inf_{\mathbf{x}\in\mathbb{R}^n} f(\mathbf{x});$$

(3) *the trajectories converge to the minimal norm solution; i.e., if* $\mathbf{x}_*$ *satisfies the condition* $\inf_{\mathbf{x}\in X_*} \|\mathbf{x}\| = \|\mathbf{x}_*\|$, *then* $\lim_{t\to\infty} \|\mathbf{x}(t) - \mathbf{x}_*\| = 0$*;*
(4) $\lim_{t\to\infty} \|\alpha(t)\mathbf{p}(t)\| = \lim_{t\to\infty} \|\dot{\mathbf{x}}(t)\| = 0$*;*
(5) *the minimal norm solution* $\mathbf{x}_*$ *is a stationary in limit minimum point; i.e.,* $\lim_{t\to\infty} \|\ddot{\mathbf{x}}(t)\| = 0$.

*Proof.* The existence and uniqueness of the trajectories on the whole $[t_0, \infty)$ follows from the convexity of the function $f(x)$ and the local Lipschitz continuity of the gradient $\nabla f(\mathbf{x})$.

For every fixed $t_0 < \tau < \infty$ the function $F(\mathbf{x}, \tau)$ defined by (3.1) is a strongly convex function, therefore it has a unique minimum point $\mathbf{x}_\tau^*$. Let $\mathbf{x}_*$ be the optimum point of the function $f$ with minimal norm on $X_*$. It follows from the Lemma 3.1 that $\lim_{\tau\to\infty} \|\mathbf{x}_\tau^* - \mathbf{x}_*\| = 0$.

We will show that $\lim_{\tau\to\infty} \|\mathbf{x}(\tau) - \mathbf{x}_*\| = 0$. To do this it is sufficient to prove that $\lim_{\tau\to\infty} \|\mathbf{x}(\tau) - \mathbf{x}_\tau^*\| = 0$ since $\|\mathbf{x}(\tau) - \mathbf{x}_*\| \le \|\mathbf{x}(\tau) - \mathbf{x}_\tau^*\| + \|\mathbf{x}_\tau^* - \mathbf{x}_*\|$.

Let us introduce the parametric function

$$
\begin{aligned}
g(t,\tau) = \frac{1}{2}\|\mathbf{x}(t) - \mathbf{x}_\tau^* + \alpha(t)\mathbf{p}(t)\|^2 + \\
+ \frac{1}{4}\alpha(t)\lambda(t)\|\mathbf{x}(t) - \mathbf{x}_\tau^*\|^2 + \frac{1}{2}(\gamma(t)-1)\|\mathbf{x}(t) - \mathbf{x}_\tau^*\|^2
\end{aligned}
$$

for fixed $\tau \ge t_0$.

It follows from the conditions 2(a), 2(b) and 3(a) that $g(t,\tau) \ge 0$ for all $t_0 \le t \le \tau$. For the derivative of $g(t,\tau)$ we have

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t} g(t,\tau) &= -\alpha(t)\big\langle \nabla_{\mathbf{x}} F(\mathbf{x}(t) + \alpha(t)\mathbf{p}(t), t), \mathbf{x}(t) - \mathbf{x}_\tau^* + \alpha(t)\mathbf{p}(t)\big\rangle + \\
&\quad + \frac{1}{4}\frac{\mathrm{d}}{\mathrm{d}t}(\alpha(t)\lambda(t))\|\mathbf{x}(t) - \mathbf{x}_\tau^*\|^2 + \frac{1}{2}\alpha(t)\lambda(t)\big\langle \mathbf{x}(t) - \mathbf{x}_\tau^*, \alpha(t)\mathbf{p}(t)\big\rangle \\
&\quad + (1-\gamma(t))\|\alpha(t)\mathbf{p}(t)\|^2 + \frac{1}{2}\dot{\gamma}(t)\|\mathbf{x}(t) - \mathbf{x}_\tau^*\|^2
\end{aligned}
$$

for all $t_0 \le t \le \tau$.

Omitting the negative terms and taking into consideration that $F(\mathbf{x}(t), t)$ is strongly convex in its first variable with the convexity modulus $\frac{1}{2}\lambda(t)$ for every $t \ge t_0$, monotone decreasing in the second variable and $\mathbf{x}_\tau^*$ is the minimum point of $F(\mathbf{x}, \tau)$ we have

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t} g(t,\tau) &\le \alpha(t)\big(F(\mathbf{x}_\tau^*, t) - F(\mathbf{x}(t) + \alpha(t)\mathbf{p}(t), t)\big) - \\
&\quad - \frac{1}{2}\alpha(t)\lambda(t)\|\mathbf{x}(t) - \mathbf{x}_\tau^* + \alpha(t)\mathbf{p}(t)\|^2 + \frac{1}{2}\alpha(t)\lambda(t)\big\langle \mathbf{x}(t) - \mathbf{x}_\tau^*, \alpha(t)\mathbf{p}(t)\big\rangle \\
&= \alpha(t)\Big(\underbrace{F(\mathbf{x}_\tau^*, t) - F(\mathbf{x}_\tau^*, \tau)}_{= -\frac{1}{2}(\lambda(\tau)-\lambda(t))\|\mathbf{x}_\tau^*\|^2} + \underbrace{F(\mathbf{x}_\tau^*, \tau) - F(\mathbf{x}(t) + \alpha(t)\mathbf{p}(t), \tau)}_{\le 0} \\
&\quad + \underbrace{F(\mathbf{x}(t) + \alpha(t)\mathbf{p}(t), \tau) - F(\mathbf{x}(t) + \alpha(t)\mathbf{p}(t), t)}_{= \frac{1}{2}(\lambda(\tau)-\lambda(t))\|\mathbf{x}(t)+\alpha(t)\mathbf{p}(t)\|^2 \le 0}\Big) \\
&\quad - \frac{1}{2}\alpha(t)\lambda(t)\|\mathbf{x}(t) - \mathbf{x}_\tau^* + \alpha(t)\mathbf{p}(t)\|^2 + \frac{1}{2}\alpha(t)\lambda(t)\big\langle \mathbf{x}(t) - \mathbf{x}_\tau^*, \alpha(t)\mathbf{p}(t)\big\rangle \\
&\le -\frac{1}{2}\alpha(t)\lambda(t)\|\mathbf{x}(t) - \mathbf{x}_\tau^* + \alpha(t)\mathbf{p}(t)\|^2 + \frac{1}{2}\alpha(t)\lambda(t)\big\langle \mathbf{x}(t) - \mathbf{x}_\tau^*, \alpha(t)\mathbf{p}(t)\big\rangle \\
&\quad - \frac{1}{2}\alpha(t)\big(\lambda(\tau) - \lambda(t)\big)\|\mathbf{x}_\tau^*\|^2.
\end{aligned}
$$

Under the assumption 3(a) the inequalities

$$
\lambda(\tau) - \lambda(t) \ge \dot{\lambda}(t)(\tau - t), \quad \dot{\lambda}(t) < 0
$$

hold for all $t_0 \le t \le \tau$. Moreover, let us observe that $\|\mathbf{x}_\tau^*\|$ is uniformly bounded since

$$
f(\mathbf{x}_*) + \frac{1}{2}\lambda(\tau)\|\mathbf{x}_*\|^2 \ge f(\mathbf{x}_\tau^*) + \frac{1}{2}\lambda(\tau)\|\mathbf{x}_\tau^*\|^2 \ge f(\mathbf{x}_*) + \frac{1}{2}\lambda(\tau)\|\mathbf{x}_\tau^*\|^2,
$$

from where $\|\mathbf{x}_\tau^*\| \le \|\mathbf{x}_*\| = K$.

Decomposing $-\frac{1}{2}\alpha(t)\lambda(t)\|\mathbf{x}(t)+\alpha(t)\mathbf{p}(t)-\mathbf{x}_\tau^*\|^2$ into two equal terms and omitting the negative term $-\frac{1}{4}\alpha(t)\lambda(t)\|\alpha(t)\mathbf{p}(t)\|^2$ we have that

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}g(t,\tau) &\le -\frac{1}{4}\alpha(t)\lambda(t)\|\mathbf{x}(t)+\alpha(t)\mathbf{p}(t)-\mathbf{x}_\tau^*\|^2 \\ &\quad -\frac{1}{4}\alpha(t)\lambda(t)\|\mathbf{x}-\mathbf{x}_\tau^*\|^2 - \frac{1}{2}\alpha(t)\big(\lambda(t)-\lambda(\tau)\big)\|\mathbf{x}_\tau^*\|^2 \\ &= -A(t)\frac{1}{2}\|\mathbf{x}(t)+\alpha(t)\mathbf{p}(t)-\mathbf{x}_\tau^*\|^2 - B(t)\frac{1}{4}\alpha(t)\lambda(t)\|\mathbf{x}-\mathbf{x}_\tau^*\|^2 \\ &\quad - C(t)\frac{1}{2}(\gamma(t)-1)\|\mathbf{x}-\mathbf{x}_\tau^*\|^2 - \frac{1}{2}\alpha(t)\dot\lambda(t)K^2(\tau-t),\end{aligned}$$

where $A(t) = \frac{1}{2}\alpha(t)\lambda(t)$, $B(t) = \frac{1}{2}$ and $C(t) = \frac{1}{4(\gamma(t)-1)}\alpha(t)\lambda(t)$ Since $\gamma(t)$ is monotone nonincreasing, therefore $C(t) \ge \frac{\alpha(t)\lambda(t)}{4(\gamma(t_0)-1)} = C_1\alpha(t)\lambda(t)$. Otherwise, $\alpha(t)\lambda(t)$ is decreasing and tends to zero, so there exists $T \ge t_0$ such that $A(t) \le \frac{1}{2}$ and $C_1(t) \le \frac{1}{2}$ for every $t \ge T$. Consequently, there exists $K_1 > 0$, depending only on $\gamma(t_0)$ such that

$$\frac{\mathrm{d}}{\mathrm{d}t}g(t,\tau) \le -K_1\alpha(t)\lambda(t)g(t,\tau) - \frac{1}{2}\alpha(t)\dot\lambda(t)K^2(\tau-t).$$

Conditions 3(c) and 3(d) ensure that $g(t,\tau)$ satisfies the conditions of Lemma 2.1 and hence $\lim_{\tau\to\infty} g(\tau,\tau) = 0$.

Since $g(t,\tau)$ is a sum of non-negative functions every member of the sum tends to 0. This together with condition 2(b) proves the validity of

$$\lim_{\tau\to\infty}\|\mathbf{x}(\tau)-\mathbf{x}_\tau^*\| = 0 \quad\text{and}\quad \lim_{\tau\to\infty}\|\mathbf{x}(\tau)-\mathbf{x}_\tau^*+\alpha(\tau)\mathbf{p}(\tau)\| = 0.$$

It follows from the triangle inequality that

$$\|\alpha(\tau)\mathbf{p}(\tau)\| \le \|\mathbf{x}(\tau)-\mathbf{x}_\tau^*+\alpha(\tau)\mathbf{p}(\tau)\| + \|\mathbf{x}(\tau)-\mathbf{x}_\tau^*\| \to 0$$

which proves the limit

$$\lim_{\tau\to\infty}\|\dot{\mathbf{x}}(\tau)\| = \lim_{\tau\to\infty}\|\alpha(\tau)\mathbf{p}(\tau)\| = 0.$$

Since

$$0 \le \|\ddot{\mathbf{x}}(t)\| \le \alpha(t)\big(\|\nabla f(\mathbf{x}(t)+\alpha(t)\mathbf{p}(t))\| + \lambda(t)\|\mathbf{x}(t)+\alpha(t)\mathbf{p}(t)\|\big) + \gamma(t)\cdot\|\dot{\mathbf{x}}(t)\|,$$

the gradient $\nabla f(\mathbf{x})$ is continuous and the conditions 2(a), 2(b) and 3(c) hold, therefore $\|\ddot{\mathbf{x}}(t)\| \to 0$.

Finally, using the continuity of the function $f$ the limit

$$\lim_{\tau\to\infty} f(\mathbf{x}(\tau)) = f(\mathbf{x}_*)$$

holds, too. The last statement is trivial from the definition. □

3.2. **Regularized simplified model.** Approximating $\nabla_{\mathbf{x}}F(x(t)+\alpha(t)p(t),t)$ by $\nabla_{\mathbf{x}}F(x(t),t)$ the *regularized simplified model* (shortly RS-FR model) to solve the problem (1.1) can be given by the following first order system of differential equations:

$$\dot{\mathbf{x}} = \alpha(t)\mathbf{p} \tag{3.5}$$

$$\dot{\mathbf{p}} = -\nabla_{\mathbf{x}}F(\mathbf{x},t) + \beta(t)\mathbf{p} \tag{3.6}$$

with the initial values (2.4), where $F(\mathbf{x},t)$ is given by (3.1) in which the damping parameter $\lambda(t)$ is a positive monotone decreasing function.

The equivalent SODE system is as follows:

$$\ddot{\mathbf{x}}+\gamma(t)\dot{\mathbf{x}}+\alpha(t)\nabla_{\mathbf{x}}F(\mathbf{x},t)=0. \tag{3.7}$$

with the initial values (2.6), where $\gamma(t)$ is given by (2.7).

The convergence of the trajectories of this SODE to a minimum point of the function $f(\mathbf{x})$ has been analyzed in detail in papers of [6] and [11] when both $\alpha(t)$ and $\gamma(t)$ are constant functions. Now we formulate a theorem on the convergence of its trajectories to the stationary in limit minimal norm solution with function parameters and prove it by constructing an appropriate Lyapunov-like function for the RS-FR model given by (3.5)-(3.6), respectively by (3.7).

**Proposition 3.3.** *Let the following assumptions hold:*

1. *In the minimization problem* (1.1) $f$ *is defined and continuously differentiable convex function on* $\mathbb{R}^n$ *and its gradient* $\nabla f$ *is local Lipschitz continuous and the conditions given in* (1.2) *on page 1 hold;*
2. *The parameter functions* $\alpha(t)$ *and* $\beta(t)$ *satisfy the following conditions*
   - (a) $\alpha(t)$ *is a positive upper bounded and* $\beta(t)$ *is a negative lower bounded continuously differentiable function on* $[t_0,\infty)$*; both* $\alpha(t)$ *and* $\beta(t)$ *are continuously differentiable on* $[t_0,\infty)$ *and* $\frac{\dot\alpha(t)}{\alpha(t)}$ *is bounded on* $[t_0,\infty)$*;*
   - (b) *there exists* $t_1\ge t_0$ *such that* $\alpha(t)+\beta(t)<0$ *and* $\frac{\beta(t)}{\alpha(t)}$ *is nondecreasing on* $[t_1,\infty)$*;*
3. *Let the damping parameter* $\lambda(t)$ *satisfy the following conditions*
   - (a) $\lambda(t)$ *is a positive, continuously differentiable monotone decreasing convex function on* $[t_0,\infty)$*;*
   - (b) $\lim_{t\to\infty}\lambda(t)=\lim_{t\to\infty}\dot\lambda(t)=\lim_{t\to\infty}\frac{\dot\lambda(t)}{\lambda^2(t)}=0$,
   - (c) $\int_{t_0}^{\infty}\lambda(t)\mathrm{d}t=\infty$*;*
   - (d) $\alpha(t)+\beta(t)\le-\frac12\lambda(t)$ *for every* $t_1\le t$*;*
   - (e) $-\frac{\dot\alpha(t)}{\alpha(t)}-\frac{\alpha(t)}{2}\le-\frac14\lambda(t)$ *for all* $t_1\le t$.

*Then*

1. *the trajectories of* (3.5)-(3.6)*, respectively of* (2.10) *exist and unique on the whole half-line* $[t_0,\infty)$ *with any initial point* (2.4) *(resp.* (2.6)*);*
2. *the RS-FR model given by* (3.5)-(3.6) *(or 3.7) is minimizing; i.e.,*
$$\lim_{t\to\infty}f(\mathbf{x}(t))=f(\mathbf{x}_*)=\inf_{\mathbf{x}\in\mathbb{R}^n}f(\mathbf{x});$$
3. *the trajectories converge to the minimal norm solution; i.e., if* $\mathbf{x}_*$ *satisfies the condition* $\inf_{\mathbf{x}\in X_*}\|\mathbf{x}\|=\|\mathbf{x}_*\|$*, then* $\lim_{t\to\infty}\|\mathbf{x}(t)-\mathbf{x}_*\|=0$*;*
4. $\lim_{t\to\infty}\|\mathbf{p}(t)\|=\lim_{t\to\infty}\|\dot{\mathbf{x}}(t)\|=0$*;*
5. *the minimal norm solution* $\mathbf{x}_*$ *is a stationary in limit minimum point; i.e.,* $\lim_{t\to\infty}\|\ddot{\mathbf{x}}(t)\|=0$.

*Proof.* Analogously to the proof of the Proposition 3.2 it is sufficient to prove that

$$\lim_{\tau\to\infty}\|\mathbf{x}(\tau)-\mathbf{x}_\tau^*\|=0.$$

Let us introduce the function

$$g(t,\tau)=\frac{2}{\alpha(t)}\Big(F(\mathbf{x}(t),t)-F(\mathbf{x}_\tau^*,\tau)\Big)+\frac12 h(t)\|\mathbf{x}(t)-\mathbf{x}_\tau^*\|^2$$

$$+\frac{1}{2}\|\mathbf{x}(t)-\mathbf{x}_\tau^*+\mathbf{p}(t)\|^2+\frac{1}{2}\|\mathbf{p}(t)\|^2,$$

where $h(t)=-1-\frac{\beta(t)}{\alpha(t)}>0$. This function is defined for all $t\in[t_0,\tau]$, for every fixed $\tau<\infty$ and $g(t,\tau)\geq 0$, in these intervals since $\lambda(t)$ is monotone decreasing.

For all $t_0\leq t\leq\tau$ the derivative of $g(t,\tau)$ by $t$ with a fixed $\tau$ is

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}g(t,\tau)=&\frac{-2\dot\alpha(t)}{\alpha^2(t)}\Big(F(\mathbf{x}(t),t)-F(\mathbf{x}_\tau^*,\tau)\Big)+\frac{\dot\lambda(t)}{\alpha(t)}\|\mathbf{x}(t)\|^2\\&+\frac{1}{2}\dot h(t)\|\mathbf{x}(t)-\mathbf{x}_\tau^*\|^2+\big(\alpha(t)+\beta(t)+h(t)\alpha(t)\big)\big\langle\mathbf{x}(t)-\mathbf{x}_\tau^*,\mathbf{p}(t)\big\rangle\\&+\big(\alpha(t)+2\beta(t)\big)\|\mathbf{p}(t)\|^2-\big\langle\nabla_{\mathbf{x}}F(\mathbf{x}(t),t),\mathbf{x}(t)-\mathbf{x}_\tau^*\big\rangle.\end{aligned}$$

Taking into consideration the conditions 2, 3(a) and 3(c), we obtain

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}g(t,\tau)\leq&\frac{-2\dot\alpha(t)}{\alpha^2(t)}\Big(F(\mathbf{x}(t),t)-F(\mathbf{x}_\tau^*,\tau)\Big)-\frac{1}{2}\lambda(t)\|\mathbf{p}(t)\|^2\\&-\big\langle\nabla_{\mathbf{x}}F(\mathbf{x}(t),t),\mathbf{x}(t)-\mathbf{x}_\tau^*\big\rangle,\end{aligned}$$

for all $t_1\leq t\leq\tau$.

Since $F(\mathbf{x},t)$ is a strongly convex function in the variable $\mathbf{x}$ for all $t\geq t_0$ and its convexity modulus is $\frac{1}{2}\lambda(t)$ for every $t_0\leq t$, therefore for all $t\geq t_1$, we have the inequality

$$\begin{aligned}&-\big\langle\nabla_{\mathbf{x}}F(\mathbf{x}(t),t),\mathbf{x}(t)-\mathbf{x}_\tau^*\big\rangle\\&\leq-\big(F(\mathbf{x}(t),t)-F(\mathbf{x}_\tau^*,t)\big)-\frac{1}{2}\lambda(t)\|\mathbf{x}(t)-\mathbf{x}_\tau^*\|^2\\&=-\big(F(\mathbf{x}(t),t)-F(\mathbf{x}_\tau^*,\tau)\big)-\big(F(\mathbf{x}_\tau^*,\tau)-F(\mathbf{x}_\tau^*,t)\big)-\frac{1}{2}\lambda(t)\|\mathbf{x}(t)-\mathbf{x}_\tau^*\|^2\\&\leq-\big(F(\mathbf{x}(t),t)-F(\mathbf{x}_\tau^*,\tau)\big)-\frac{1}{2}\big(\lambda(\tau)-\lambda(t)\big)\|\mathbf{x}_\tau^*\|^2-\frac{1}{2}\lambda(t)\|\mathbf{x}(t)-\mathbf{x}_\tau^*\|^2.\end{aligned}$$

Substituting this inequality into the estimation of $\frac{\mathrm{d}}{\mathrm{d}t}g(t,\tau)$ we can obtain the inequality

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}g(t,\tau)\leq&\Big(-\frac{2\dot\alpha(t)}{\alpha^2(t)}-1\Big)\Big(F(\mathbf{x}(t),t)-F(\mathbf{x}_\tau^*,\tau)\Big)\\&-\frac{1}{2}\lambda(t)\|\mathbf{x}(t)-\mathbf{x}_\tau^*\|^2-\frac{1}{2}\lambda(t)\|\mathbf{p}(t)\|^2-\frac{1}{2}\big(\lambda(\tau)-\lambda(t)\big)\|\mathbf{x}_\tau^*\|^2.\end{aligned}$$

for all $t_1\leq t\leq\tau$. Since the inequality

$$-\|\mathbf{x}(t)-\mathbf{x}_\tau^*\|^2-\|\mathbf{p}\|^2\leq-\frac{1}{2}\|\mathbf{x}(t)-\mathbf{x}_\tau^*\|^2-\frac{1}{2}\|\mathbf{p}(t)\|^2-\frac{1}{4}\|\mathbf{x}(t)-\mathbf{x}_\tau^*+\mathbf{p}(t)\|^2$$

and the conditions 3(c)-3(d) of the proposition hold, with the coefficients

$$A(t)=\frac{\dot\alpha(t)}{\alpha(t)}+\frac{\alpha(t)}{2},\quad B(t)=\frac{\lambda(t)}{2h(t)},\quad C(t)=\frac{1}{2}\lambda(t),\quad D(t)=\frac{1}{4}\lambda(t)$$

we obtain, for all $t_1\leq t\leq\tau$,

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}g(t,\tau)\leq&-A(t)\cdot\frac{2}{\alpha(t)}\Big(F(\mathbf{x}(t),t)-F(\mathbf{x}_\tau^*,\tau)\Big)-B(t)\frac{1}{2}h(t)\|\mathbf{x}(t)-\mathbf{x}_\tau^*\|^2\\&-C(t)\frac{1}{2}\|\mathbf{p}(t)\|^2-D(t)\frac{1}{2}\|\mathbf{x}(t)-\mathbf{x}_\tau^*+\mathbf{p}(t)\|^2-\frac{1}{2}\big(\lambda(\tau)-\lambda(t)\big)\|\mathbf{x}_\tau^*\|^2.\end{aligned}$$

It is obvious that $-C(t) \le -D(t)$, and from the condition 3(e) we have that $-A(t) \le -D(t)$, too. After a short calculation we can obtain that

$$-B(t) \le -D(t) \quad \text{if } h(t) \le 2 \quad \text{and} \quad -B(t) \ge -D(t) \quad \text{if } h(t) \ge 2.$$

Since $h(t)$ is nonincreasing, there are two cases:

**Case 1.** $h(t) \ge 2$ (or equivalently $3\alpha(t) + \beta(t) \le 0$) for all $t \ge t_1$. In this case $-B(t) = \max(-A(t), -B(t), -C(t), -D(t)$ for all $t \ge t_1$. It means that

$$\frac{\mathrm{d}}{\mathrm{d}t} g(t,\tau) \le -B(t) g(t,\tau) - \frac{1}{2}\big(\lambda(\tau) - \lambda(t)\big)\|\mathbf{x}_\tau^*\|^2$$

for all $t_1 \le t \le \tau$. Using the definition of $B(t)$ and the the fact, that $h(t_1) \ge h(t)$ for all $t \ge t_1$ we can give the following upper bound:

$$-B(t) = -\frac{\lambda(t)}{2h(t)} \le -\frac{\lambda(t)}{2h(t_1)},$$

consequently,

$$\frac{\mathrm{d}}{\mathrm{d}t} g(t,\tau) \le -\frac{\lambda(t)}{2h(t_1)} g(t,\tau) - \frac{1}{2}\big(\lambda(\tau) - \lambda(t)\big)\|\mathbf{x}_\tau^*\|^2$$

for all $t_1 \le t \le \tau$.

**Case 2.** There exists $t_2 \ge t_1$ such that $h(t) \le 2$ (or equivalently $3\alpha(t) + \beta(t) \ge 0$) for all $t \ge t_2$. Then $-D(t) = \max(-A(t), -B(t), -C(t), -D(t)$ for all $t \ge t_2$, therefore

$$\frac{\mathrm{d}}{\mathrm{d}t} g(t,\tau) \le -\frac{1}{4}\lambda(t) g(t,\tau) - \frac{1}{2}\big(\lambda(\tau) - \lambda(t)\big)\|\mathbf{x}_\tau^*\|^2$$

for all $t_2 \le t \le \tau$.

The estimation of the last term in both cases can be done as in the proof of Proposition 3.2. So, in both cases there exists a positive constant $K_1$ and time $T \ge t_0$ such that the inequality

$$\frac{\mathrm{d}}{\mathrm{d}t} g(t,\tau) \le -K_1 \lambda(t) g(t,\tau) - \frac{1}{2} K^2 \dot{\lambda}(t)(\tau - t)$$

holds for all $T \le t \le \tau$. To complete the proof one can follow the proof of the Proposition 3.2. □

## 4. Analysis and comparison of the methods

4.1. **Existence of parameters.** For both models one can give the triplet of parameter functions $(\alpha(t), \beta(t), \lambda(t))$ such that conditions of the propositions are satisfied. Namely,

(A) for the RG-FR model
  (a) if

$$\alpha(t) = \alpha_0,$$
$$\gamma(t) = -\beta(t) = -\beta_0 - B(1+t)^{-b},$$
$$\lambda(t) = L(1+t)^{-\ell},$$

  then the conditions of the proposition 3.2 are fulfilled if either

$$b = 0, \quad \alpha_0 > 0, \quad \beta_0 + B < -1, \quad 0 < \ell < 1, \quad L > 0,$$

  or

$$b > 0, \quad \alpha_0 > 0, \quad \beta_0 < -1, \quad B < 0, \quad 0 < \ell < 1, \quad L > 0;$$

(b) if

$$\lambda(t) = \alpha(t) = \alpha_0(1+t)^{-a},$$
$$-\beta(t) = -\beta_0 - B(1+t)^{-1},$$

then the conditions of the proposition 3.2 are fulfilled if

$$\alpha_0 > 0, \quad \beta_0 < -1, \quad \frac{1}{2} > a \geq B > 0.$$

(B) for the RS-FR model

(a) if

$$\alpha(t) = \alpha_0,$$
$$\gamma(t) = -\beta(t) = -\beta_0 - B(1+t)^{-\ell},$$
$$\lambda(t) = L(1+t)^{-\ell},$$

then the conditions of the proposition 3.3 are fulfilled if

$$0 < \alpha_0, \quad \beta_0 < -\alpha_0, \quad B < 0, \quad 0 < \ell < 1, \quad L > 0,$$

(b) if

$$\alpha(t) = \alpha_0(1+t)^{-\ell},$$
$$\beta(t) = -\beta_0(1+t)^{-\ell},$$
$$\lambda(t) = L(1+t)^{-\ell}$$

then the conditions of the proposition 3.3 are fulfilled if

$$\alpha_0 > 0, \quad \beta_0 > 0, \quad L > 0, \quad 0 < \ell < 1, \quad 2(\alpha_0 - \beta_0) < -L, \quad L < 2\alpha_0.$$

More families of parameters satisfying the conditions of the proposition can be obtained by the technique given in [20].

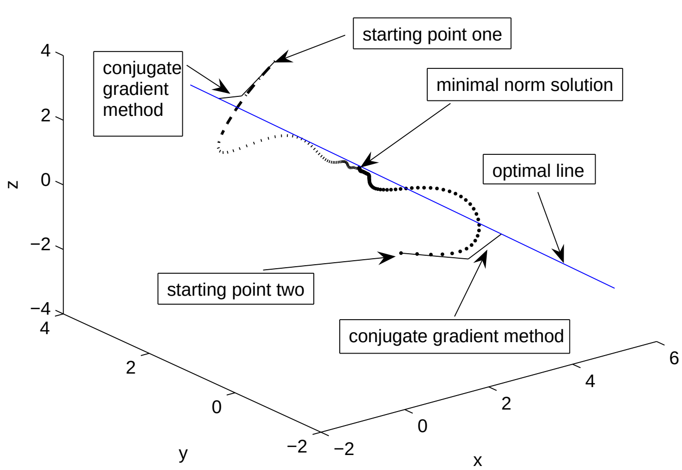


Figure 1. Trajectory of the continuous method for the RS-FR model

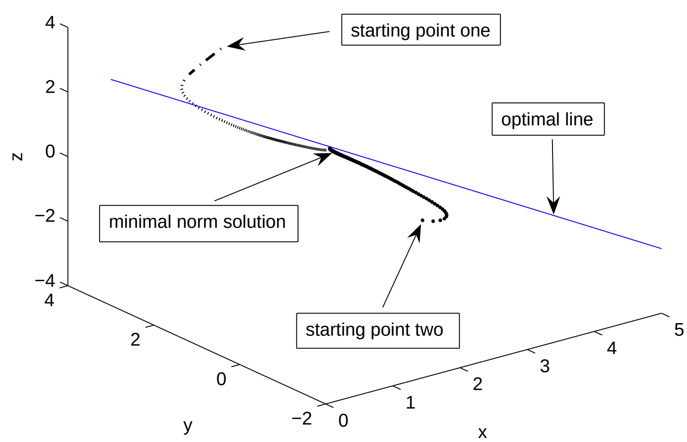


FIGURE 2. Trajectories of the continuous method for the RG-FR model

4.2. **Comparison of the generalized and simplified models.** Let us illustrate the behaviour of the trajectories of the given models on a numerical example. Let us minimize the the function

$$f(x, y, z) = (x + y - 3)^2 + (x + z - 3)^2.$$

where $f$ is a convex function and the minimum points of $f$ lie on the line

$$x = 3 - t, \quad y = t, \quad z = t.$$

The minimum point with the minimum norm is $(2, 1, 1)$. We have solved the RS-FR model and the RG-FR model with the third ordered Runge-Kutta method with different parameter functions and with two different initial points $x_0 = (1, 2, 4)$ and $x_0 = (4, 2, -3)$ and step size $h = 0.1$. The results can be seen in Figures 1-reffig2.

On Figure 1 we have drawn the minimizing lines obtained by the discrete Fletcher-Reeves algorithm which show that starting from different initial points the obtained minimum points could be different.

On the other hand both the generalized and simplified models converge to the unique minimal norm solution. However we can see, that the shape of the trajectories are quite different in the two models, especially the RG-FR model gives a "smoother" trajectory. Since in the RS-FR model $\nabla_{\mathbf{x}} F(x(t) + \alpha(t)p(t), t)$ is approximated by $\nabla_{\mathbf{x}} F(x(t), t)$ we can expect that the RG-FR model converges faster to the minimum point but the RS-FR model could be easier to solve numerically.

4.3. **Comparison of the heavy ball with friction and the simplified models.** Let us consider the system

$$\ddot{\mathbf{x}} + \gamma\dot{\mathbf{x}} + \nabla f(\mathbf{x}) + \lambda(t)\mathbf{x} = 0. \tag{4.1}$$

where $\gamma$ is a constant. This equation is known as the regularized version of the heavy ball system with friction model and has been studied in papers [6] and [11].

If we assume that $\alpha(t) \equiv 1$ and $\gamma(t) \equiv \gamma$ in the RS-FR-model (3.7), then the regularized heavy ball with friction model can be considered as a special case of it. In this special case our proposition turns into the following result.

**Corollary 4.1.** *Under assumption 1. of Proposition 3.3, the trajectories of the SODE system* (4.1) *exist and unique on the whole half-line* $[t_0, \infty)$ *with any initial point and the following limits hold*

$$\lim_{t\to\infty} f(\mathbf{x}(t)) = f(\mathbf{x}_*) = \inf_{\mathbf{x}\in\mathbb{R}^n} f(\mathbf{x}), \quad \lim_{t\to\infty} \|\mathbf{x}(t) - \mathbf{x}_*\| = 0,$$

*where* $\mathbf{x}_*$ *is the minimal norm solution of* (1.1) *and*

$$\lim_{t\to\infty} \|\dot{\mathbf{x}}(t)\| = 0, \quad \lim_{t\to\infty} \|\ddot{\mathbf{x}}(t)\| = 0,$$

*if the following four conditions hold:*

(1) $\gamma > 1$*;*
(2) $\lambda(t)$ *is positive monotone decreasing continuously differentiable convex function on* $[t_0, \infty)$*;*
(3) $\lim_{t\to\infty} \lambda(t) = \lim_{t\to\infty} \dot{\lambda}(t) = \lim_{t\to\infty} \frac{\dot{\lambda}(t)}{\lambda^2(t)} = 0$;
(4) $\int_{t_0}^{\infty} \lambda(t)\mathrm{d}t = \infty$.

According to the convergence conditions of the theorems in [6] and [11] the condition 1, the third term of the condition 3 and the convexity of $\lambda(t)$ can be omitted. However we wanted to give common conditions which guarantee the convergence of the trajectories without doing difference between the cases when the coefficient of $\dot{\mathbf{x}}$ is a positive constant or a function. So, on one hand our result is weaker and on the other hand it is stronger then the results of [6] and [11].

Otherwise in our models (not only in the simplified but in the generalized one, too) there is a function parameter $\alpha(t)$ in the coefficient of the gradient of the function. It is true, that applying a time-transformation $t = z(s)$ this function parameter turns into constant 1 if we get the transformation from the differential equation

$$\frac{\mathrm{d}z(s)}{\mathrm{d}s} = \frac{1}{\sqrt{\alpha(z(s))}},$$

but the transformed $\gamma(z(s))$ will be constant only for a special function of $\gamma(t)$. So, the heavy ball with friction model with constant $\gamma$ in general can not be obtained from our model by time-transformation.

The discrete Fletcher-Reeves iteration has two parameters. Therefore we have insisted on such models which has two corresponding function parameters, too.

The Fletcher-Reeves iteration has some very favorable properties which have not been investigated in this paper. It would be interesting to know which properties preserved in the proposed continuous GM-FR and SM-FR models. This is the subject of our further research.

## References

[1] Alvarez, F.; *On the minimizing property of a second order dissipative system in Hilbert space.* SIAM J. on Control and Optimization, 38(4)(2000):1102–1119.

[2] Antipin A. S.; *Minimization of convex functions on convex sets by means of differential equations.* Differential Equations 30(9)(1994):1365–1375. Translated from: Antipin A. S., *Minimizatsiya vypuklykh funktsiĭ na vypuklykh mnozhestvakh s pomosh'yu differentsialp'nykh uravneniĭ*, Differentsialp1nye uravneniya, 30(9)(1994):1475–1486.

[3] Archetti F., Szegő G. P.; *Global optimization algorithms*, in: Dixon L.C.W., Spedicato G.P. (eds), *Nonlinear Optimization: Theory and Algorithms*, Birkhauser, 1980, 429–469.
[4] Attouch H.; *An inertial proximal method for maximal monotone operators via discretization of a nonlinear oscillator with damping*. Set-Valued Analysis, 9(2001): 3–11.
[5] Attouch H., Cominetti R.; *A dynamical approach to convex minimization coupling approximation with the steepest descent method*. J. Differential Equations, 128(2)(1996):519–540.
[6] Attouch H., Czarnecki M-O.; *Asymptotic control and stabilization of nonlinear oscillators with non isolated equilibria*. J. Differential Equations, 179(1)(2002): 278–310.
[7] Attouch H., Goudou X., Redont P.; *The heavy ball with friction method. I. The continuous dynamical system*. Comm. Contemp. Math., 2(1)(2000):1–34.
[8] Botsaris C. A.; *Differential gradient method*, J. Math. Anal. Appl., 63(1978):177–198.
[9] Branin F. H., Hoo S. K.; *A method for finding multiple extrema of a function of N variables*, Proceedings of the Conference on Numerical Methods for Nonlinear Optimization, University of Dundee, Scotland, June 28-July 1, 1971, in: Lootsma, F. (ed.), *Numerical Methods of Nonlinear Optimization*, Academic Press, London, 1972.
[10] Brown A. A., Bartholomew-Biggs M. C.; *Some effective methods for unconstrained optimization based on the solution of systems of ordinary differential equations*, J. Optim. Theory and Appl., 62(2)(1988): 211–224.
[11] Cabot A.; *Inertial gradient-like dynamical system controlled by a stabilizing term*. J. Optim. Theory Appl., 120(2004): 275–303.
[12] Cabot A., Engler H., Gadat S.; *On the long time behavior of second order differential equations with asymptotically small dissipation*. Trans. Amer. Math. Soc., 361(11)(2009): 5983–6017.
[13] Cabot A., Engler H., Gadat S.; *Second-order differential equations with asymptotically small dissipation and piecewise flat potentials*. Conf. Seventh Mississippi State - UAB Conference on Differential Equations and Computational Simulations, Electronic Journal of Differential Equations, Conference 17(2009): 33–38.
[14] Cominetti R., Peypouquet J., Sorin S., *Strong asymptotic convergence of evolution equations governed by maximal monotone operators with Tikhonov regularization*. J. Differential Equations, 245(12)(2008):3753–3763.
[15] Evtushenko Yu. G., Zhadan V. G.; *Application of the method of Lyapunov functions to the study of convergence of numerical methods*. USSR Comput. Math. Math. Phys., 15(1)(1975):96–108. Translated from: Evtushenko Yu. G., Zhadan V.G., *Primeneniya metoda funktsiĭ Lyapunova dlya issledovaniya skhodimosti chislennykh metodov*, Zh. vychisl. mat. i mat. fiz., 15(1)(1975): 101-112.
[16] Flåm S. D.; *Solving convex programming by means of ordinary differential equations*. Mathematics of Operations Research 17(2)(1992):290–302
[17] Fletcher R., Reeves C. M.; *Function minimization by conjugate gradients*. Comput. J., 7(1964): 149–154.
[18] Goudou X., Munier J.; *The gradient and heavy ball with friction dynamical systems: the quasiconvex case*. Math. Program. Ser. B, 116(1-2)(2009): 173–191.
[19] Hajba T.; *Optimization methods modelled by second order differential equation*. Annales Univ. Sci. Budapest, Sect. Comp. 26(2006): 145–158.
[20] Hajba T.; *A másodrendű differenciálegyenlettel modellezett optimalizálók megengedett paramétereinek elemzése, (Parameter analysis of optimization methods modelled by second order differential equation)*. Alkalmazott Matematikai Lapok 25(2008): 61–79, (in Hungarian).
[21] Hauser R., Nedić J.; *The continuous Newton-Raphson method can look ahead*, SIAM J. Optim. 15(3)(2005): 915–925.
[22] Kovacs M.; *Continuous analog of gradient-type iterative regularization*. J. Mosc. Univ. Comput. Math. Cybern. 3(1979):37–44. Translated from: Kovach M., *Nepreryvnyĭ analog iterativnoĭ regulyarizatsii gradientnogo tipa*. Vestnik Mosk. un-ta, Ser. XV. Vychisl. mat. i kibern., 3 (1979): 36-42.
[23] Kovacs M.; *Some convergence theorems on nonstationary minimization precesses*. Math. Operationsforschung u Statist., 15(2)(1984): 203–210.
[24] Levenberg K.; *A method for the solution of certain non-linear problems in least squares*. Quaterly of Applied Mathematics, 2(1944): 164–168.

[25] Marquard D. V.; *An algorithm for least squares estimation of non-linear parameters*. SIAM Journal, 11(1963):431–441.

[26] Nedić A.; *The continuous projection-gradient method of the fourth order*. Yugosl. J. Oper. Res., 5(1)(1995):27–38.

[27] Rouche N., Habets P., Laloy M.; *Stability Theory by Liapunov's Direct Method*. Springer Verlag, 1977.

[28] Tikhonov A. N. *Solution of incorrectly formulated problems and the regularization method*. Soviet Math. Dokl. 4(1963):1035–1038. Translated from: Tikhonov A.N., *O reshenii nekorrektno postavlennykh zadach i metode regulyarizatsii*, Dokl. Akad. Nauk SSSR, 153.1(1963): 49–52.

[29] Tikhonov A. N., Vasilev F. P.; *Methods for the solution of ill-posed extremal problems*. In: Banach Center Publications, Vol. 3, PWN, Warszawa, 1978, 297–342, (in Russian: Tikhonov A. N., Vasilp1ev F. P.: Metody resheniya nekorrektnykh ekstremal'nykh zadach).

[30] Vasil'ev F. P., Nedich A.; *Regularized continuous third-order gradient projection method*. Differential equation 30(12)(1994): 1869–1877. Translated from: Vasil'ev F.P., Nedich A., Regulyarizovannyĭ nepreryvnyĭ metod tret'ego poryadka proektsii gradienta, Differentsialp1nye uravneniya, 30(12)(1994), 2033-2042.

[31] Vasil'ev F. P., Amochkina T. V., Nedić A.; *On a regularized variant of the second-order continuous gradient projection method*. J. Mosc. Univ. Comput. Math. Cybern., 3(1995):33–39. Tanslated from: Vasil'ev F. P., Amochkina T. V., Nedich A., *Ob odnom regulyarizovannom variente nepreryvnogo metoda proektsii gradienta vtorogo poryadka*, Vestnik Mosk. un-ta, Ser. XV. Vychisl. mat. i kibern., 3(1995), 39-46.

[32] Vasiljev F. P., Nedić A.; *A regularized continuous projection-gradient method of the fourth order*. Yugosl. J. Oper. Res., 5(2)(1995):195–209.

[33] Venec V. I., Rybashov M. V.; *The method of Lyapunov function in the study of continuous algorithms of mathematical programming*. USSR Comput. Math. Math. Phys 17(1977):64–73. Translated from: V.I. Venets, M.V. Rybashov, *Metod funktsiĭ Lyapunova v issledovanii nepreryvnykh algoritmov matematicheskogo programmirovaniya*. Zh. vychisl. mat. i mat. fiz., 17(3)(1977): 622-633.

[34] Zhang L.-H., Kelley C. T, Liao L.-Z.; *A continuous Newton-type method for unconstrained optimization*, Pacific Journal of Optimization, 4(2)(2008):259-277.



Tamás Hajba

Department of Mathematics and Computer Science Faculty of Engineering Sciences, Széchenyi István University, Egyetem tér 1., H-9026 Győr, Hungary

*E-mail address*: hajbat@sze.hu, Tel +36 (96) 503-400, Fax +36 (96) 613-657