

# ODE Final exam - Solutions

May 3, 2018

## 1 Computational questions (30)

For all the following ODE's with given initial condition, find the expression of the solution as a function of the time variable  $t$ . You do not have to justify existence, uniqueness, or to worry about the time of existence of the solutions, but you need to explain your computations.

1.

$$x' = \frac{t}{3+x}, \quad x(0) = 1.$$

2.

$$x'' - x' + x = 0, \quad x(0) = 1, \quad x'(0) = 0.$$

3.

$$X' = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} X, \quad X(0) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

4.

$$x' = t + tx^2, \quad x(0) = 0.$$

5.

$$x' - x = e^t, \quad x(0) = 1.$$

6.

$$tx' = x + te^{x/t}, \quad x(1) = 1.$$

1. As long as  $x(s) \neq 3$  we may write

$$x'(3+x(s)) = s.$$

Integrating between  $s = 0$  and  $s = t$  and using the initial condition  $x(0) = 1$ , we obtain

$$3x(t) + \frac{1}{2}x^2(t) - \left(3 + \frac{1}{2}\right) = \frac{1}{2}t^2.$$

Hence we get

$$x^2(t) + 6x(t) - (7 + t^2) = 0,$$

so we must have

$$x(t) = \frac{-6 \pm \sqrt{36 + 4(7 + t^2)}}{2}.$$

Since  $x(0) = 1$  and  $x$  cannot cross  $-3$ , the correct solution is

$$x(t) = \frac{-6 + \sqrt{36 + 4(7 + t^2)}}{2}.$$

2. Let us consider the associated characteristic polynomial.

$$x^2 - x + 1 = 0$$

its roots are  $\lambda_{\pm} = \frac{1 \pm i\sqrt{3}}{2}$ . Thus we know that the general solution takes the form

$$x(t) = Ae^{t\lambda_+} + Be^{t\lambda_-},$$

where  $A, B$  are coefficients to be determined. We have

$$x(0) = 1 = A + B, \quad x'(0) = 0 = \lambda_+ A + \lambda_- B.$$

We obtain

$$A + B = 1, \quad A - B = \frac{i}{\sqrt{3}},$$

thus

$$A = \frac{1}{2} + \frac{i}{2\sqrt{3}}, \quad B = \frac{1}{2} - \frac{i}{2\sqrt{3}}.$$

3. The eigenvalues of the matrix are easily seen to be  $\pm 1$ , with eigenvectors

$$\begin{pmatrix} 1 \\ \pm 1 \end{pmatrix}$$

We do the standard thing: changing basis, computing the exponential of a diagonal matrix, etc.

4. We write

$$\frac{x'}{1+x^2} = t,$$

and integrate to get

$$\tan^{-1}(x(t)) - \tan^{-1}(x(0)) = \frac{t^2}{2},$$

thus (since  $x(0) = 0$ ) we have

$$x(t) = \tan\left(\frac{t^2}{2}\right).$$

5. The associated homogeneous equation has general solution  $t \mapsto Ae^t$  for  $A$  arbitrary. Since the right-hand side is also of this form, we look for a particular solution as

$$t \mapsto ate^t.$$

We find  $a = 1$ . Hence the general solution of the ODE is  $(t + C)e^t$  with  $C$  arbitrary. The initial condition implies that  $C = 1$ , thus we obtain

$$x(t) = (t + 1)e^t.$$

6. This is an homogeneous equation (with the other meaning of *homogeneous*). For  $t \neq 0$  we write it as

$$x' - \frac{x}{t} - e^{x/t} = 0.$$

Introducing  $z = \frac{x}{t}$  we have  $x' = z + tz'$  and thus

$$z + tz' - z - e^z = 0,$$

so

$$z'e^{-z} = \frac{1}{t},$$

which gives, after integrating between 1 and  $t$

$$-e^{z(t)} + e^{z(1)} = \ln(t).$$

We have  $x(1) = 1$  hence  $z(1) = 1$  and we get

$$e^{z(t)} = e - \ln(t),$$

thus  $z(t) = \ln(e - \ln(t))$  and finally

$$\boxed{x(t) = t \ln(e - \ln(t))}$$

## 2 Time of existence (20)

We consider the ODE

$$x' = t^2x + (1 + \cos^2(t))x^2.$$

We denote by  $\gamma$  the maximal solution of this ODE with initial condition  $\gamma(0) = 1$ , defined on some interval  $(\alpha, \beta)$ .

1. Show that  $\gamma(t)$  is always positive for  $t$  in  $(\alpha, \beta)$ .
2. Show that  $\gamma$  is increasing on  $(\alpha, \beta)$ .
3. Justify that  $\alpha = -\infty$ .
4. Justify that  $\beta$  is finite. *You may use the ODE  $x' = x^2$  for comparison.*

1. We observe that  $x(t) \equiv 0$  is solution to the ODE. By the uniqueness part of Cauchy-Lipschitz theorem, any solution that vanishes for a certain time must thus be the zero solution. In other words, any non-zero solution has a constant sign. Since  $\gamma(0) = 1$ , we deduce that  $\gamma$  stays positive for all times.
2. Since  $\gamma(t)$  is always positive, the right-hand side of the ODE is positive and thus  $\gamma'(t)$  is always positive, hence  $\gamma$  is increasing.
3. The function  $t \mapsto \gamma(t)$  is continuous, increasing, and bounded below by 0. Thus for any  $t \in (\alpha, 0)$  we have  $0 \leq \gamma(t) \leq 1$  and this negates the “blow up in finite time” criterion, thus  $\alpha = -\infty$ .
4. Since  $\gamma$  is positive, we have, for  $t \in (\alpha, \beta)$

$$\gamma'(t) = t^2\gamma(t) + (1 + \cos^2(t))\gamma^2(t) \geq \gamma^2(t).$$

In particular, integrating between 0 and  $t$  for  $t \in (\alpha, \beta)$  we obtain

$$1 - \frac{1}{\gamma(t)} \geq t,$$

so  $\boxed{\gamma(t) \geq \frac{1}{1-t}}$ , which proves that  $\gamma$  blows-up in finite time.

### 3 Qualitative study (30)

We consider the ODE

$$x'' - \sin(x) = 0. \tag{1}$$

1. Re-write (1) as a first-order ODE with unknown function  $X = (x, x')$ .
2. Find a (non-trivial) conserved quantity.
3. Sketch the allure of the orbits in  $\mathbb{R}^2$  with the system of coordinates  $(x, x')$ .
  - (a) Near the point  $(\pi/2, 0)$ .
  - (b) Near the point  $(0, 2)$ .(Briefly justify your drawing.)
4. Explain why we can find a change of variables that would transform the two sketches drawn in the previous question onto one another.

CANCELLED Explain why, near the point  $(0, 0)$ , the flow of this ODE looks like

$$e^{tA}, \quad A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

5. Sketch the allure of the orbits near  $(0, 0)$ .

1.

$$\boxed{X' = (x', x'') = (x', \sin(x))}$$

2. We apply the usual trick, multiplying by  $x'$  yields

$$x''x' - \sin(x)x' = 0,$$

and thus  $\boxed{Q(x, x') := \frac{(x')^2}{2} + \cos(x)}$  is conserved.

3. The orbits are contained in the level sets of  $Q$ . We can use the fact that

$$\cos(x) \approx (x - \pi/2) \text{ near } \pi/2, \quad \cos(x) \approx 1 - x^2/2 \text{ near } 0$$

to sketch these level sets.

4. We are considering two points that are not stationary points. By the “straightening of vector fields” theorem, the flow near both points can be mapped onto the flow of a constant vector field (i.e. an ODE with constant speed). By transitivity, the two flows can be mapped onto each other.
5. That would have been an application of Grobman-Hartman.
6. We use again the conserved quantity and the approximation  $\cos(x) \approx 1 - x^2/2$  near 0.

## 4 Around the gradient descent

Let  $d \geq 1$  be the dimension. In this problem,  $E$  is a function from  $\mathbb{R}^d$  to  $\mathbb{R}$  of class  $C^1$  such that:

- $E$  is **Lipschitz**, with a Lipschitz constant denoted by  $L$ . By definition, it means

$$\forall x, y \in \mathbb{R}^d, \quad |E(x) - E(y)| \leq L\|x - y\|.$$

You may use the following consequence:  $\boxed{\forall x \in \mathbb{R}^d, \quad \|\nabla E(x)\| \leq L.}$

- The gradient  $\nabla E$  is **Lipschitz**, with a Lipschitz constant denoted by  $M$ . By definition, it means

$$\boxed{\forall x, y \in \mathbb{R}^d, \quad \|\nabla E(x) - \nabla E(y)\| \leq M\|x - y\|.}$$

- $E$  is  $\alpha$ -**convex** for some  $\alpha$ . By definition, it means (we denote by  $\langle a, b \rangle$  the scalar product of two vectors).

$$\boxed{\forall x, y \in \mathbb{R}^d, \quad \langle \nabla E(x) - \nabla E(y), x - y \rangle \geq \alpha\|x - y\|^2.}$$

### Preliminary question

1. Show that, because  $E$  is  $\alpha$ -convex, then  $E$  has at most one critical point.
1. By contradiction, if there were two critical points  $x$  and  $y$ , we would have  $\nabla E(x) = \nabla E(y) = 0$  but  $\alpha$ -convexity would give us

$$\alpha\|x - y\|^2 \leq 0,$$

thus  $x = y$ .

*In the following, we will denote by  $X_{\min}$  the unique critical point, we assume that it exists and is the unique global minimizer of  $E$ .*

## 4.1 Gradient descent in continuous time (40)

In this section, we fix  $X_0$  in  $\mathbb{R}^d$  and we study the ODE

$$\boxed{X'(t) = -\nabla E(X(t)), \quad X(0) = X_0} \quad (2)$$

where  $X$  is an unknown function with values in  $\mathbb{R}^d$ . This is known as a “gradient descent”.

### 4.1.1 Convergence to the minimizer

1. Explain why the maximal solution to the ODE (2) exists, is unique, and is defined for all times  $t$  in  $(-\infty, +\infty)$ .
2. Are there constant solutions to (2)? If yes, how many?
3. Show that either the solution is constant, or  $E(X(t))$  is (strictly) decreasing in  $t$ .
4. Is the equilibrium solution  $X(t) \equiv X_{\min}$  stable?
5. Prove that  $\lim_{t \rightarrow +\infty} X(t) = X_{\min}$  (for an initial condition close enough to  $X_{\min}$ , or, more difficult, for any choice of initial condition).

1. The function  $X \mapsto -\nabla E(X)$  is Lipschitz, by assumption. The Cauchy-Lipschitz theorem thus ensures existence and uniqueness of maximal solutions. Moreover over, it is **globally Lipschitz**, still by assumption, thus the maximal solutions are global (defined for all times - this is a result from class).
2. Constant solutions correspond to zeros of  $\nabla E$ , thus to critical points, and we know that there is only one such point. So there is exactly one constant solution, equal to  $X_{\min}$ .
3. If the solution is not equal to  $X_{\min}$ , we have

$$\frac{d}{dt} E(X(t)) = \langle \nabla E(X(t)), X'(t) \rangle = \boxed{-\|\nabla E(X(t))\|^2},$$

and the right-hand side is always negative, so the quantity  $E(X(t))$  is decreasing.

4. The quantity  $E(X(t))$  is a Liapounov function for the equilibrium (as shown in the previous question), thus we know the equilibrium is *asymptotically stable* and, in particular, it is stable.
5. Asymptotic stability implies that  $\lim_{t \rightarrow +\infty} X(t) = X_{\min}$  for an initial condition close enough to  $X_{\min}$ . In fact this is true for all initial conditions - we could prove it here by elementary methods, but the following questions will also provide a proof.

### 4.1.2 Speed of convergence

In this paragraph, we want to quantify the speed at which  $X(t)$  tends to  $X_{\min}$ . We suppose that the initial condition (at time 0)  $X_0$  is not equal to  $X_{\min}$ . For  $t \geq 0$ , we introduce the quantity

$$\mathcal{D}(t) := \|X(t) - X_{\min}\|^2.$$

1. Compute  $\mathcal{D}'(t)$ . You may use one of the auxiliary results.

2. Show that for  $t \geq 0$  we have

$$D'(t) \leq -\alpha D(t)$$

3. Prove that

$$D(t) \leq \|X_0 - X_{\min}\|^2 e^{-\alpha t}$$

and conclude about the speed of convergence of  $X(t)$  to  $X_{\min}$ .

1. We have, using the formula given to us and the fact that  $X$  satisfies the ODE

$$\boxed{D'(t) = 2\langle X'(t), X(t) - X_{\min} \rangle = -2\langle \nabla E(X(t)), X(t) - X_{\min} \rangle}.$$

2. We may also write, since  $\nabla E(X_{\min}) = 0$ ,

$$-2\langle \nabla E(X(t)), X(t) - X_{\min} \rangle = -2\langle \nabla E(X(t)) - \nabla E(X_{\min}), X(t) - X_{\min} \rangle,$$

and by the  $\alpha$ -convexity assumption we get

$$D'(t) \leq -2\alpha \|X(t) - X_{\min}\|^2 = -2\alpha D(t)$$

(the question is correct but not sharp, by a factor 2).

3. We may write

$$\frac{D'(t)}{D(t)} \leq -2\alpha$$

Integrating this inequality between 0 and  $t$ , we obtain

$$D(t) \leq D(0)e^{-2\alpha t} = \|X(0) - X_{\min}\|^2 e^{-2\alpha t}.$$

(again, the question is correct but not sharp, by a factor 2). We thus have

$$\|X(t) - X_{\min}\| \leq e^{-\alpha t} \|X(0) - X_{\min}\|,$$

thus  $X(t)$  converges to  $X_{\min}$  as  $t \rightarrow +\infty$ , exponentially fast.



## 4.2 Numerical study (40)

In this paragraph, we are interested in a numerical approach to gradient descent. It can be described as a sequence  $\{X_n\}_{n \geq 0}$  defined as follows:

- We start at some point  $X_0$ .
- At each step  $n \geq 0$ , we chose a step-size  $s_n \geq 0$  and we compute  $X_{n+1}$  in terms of  $X_n$  by

$$\boxed{X_{n+1} := X_n - s_n \nabla \mathbf{E}(X_n)}. \quad (3)$$

### 4.2.1 A model case

In this paragraph only, we take  $\mathbf{E}(X) = \|X\|^2$ . For the two following choices of step-sizes, show that the numerical scheme defined above does **not** converge to the minimizer of  $\mathbf{E}$  (here  $X_{\min} = 0$  of course), unless we start at this point.

1. For a constant step-size  $s_n = 1$ .
2. For a step-size  $s_n = n^{-100}$ .

It is thus important to chose the step-size carefully.

1. In this case, we have

$$X_{n+1} = X_n - 2X_n = -X_n,$$

and thus we have  $X_n = (-1)^n X_0$  for all  $n$ . This does not converge to 0, unless  $X_0$  is 0.

2. In this case, we have

$$X_{n+1} = X_n - \frac{1}{n^{100}} X_n = X_n (1 - n^{-100}).$$

We deduce that

$$X_n = X_0 \prod_{k=1}^n (1 - k^{-100}) = X_0 \exp \left( \sum_{k=1}^n \ln(1 - k^{-100}) \right).$$

The series  $\sum_{k=1}^+ \ln(1 - k^{-100})$  converges, thus  $X_n$  converges to  $CX_0$  for some constant  $C \neq 0$ , and in particular  $X_n$  does not converge to  $X_0$ .

In applied maths classes, the usual heuristics for step-sizes is to chose  $s_n$  such that

$$\begin{cases} \sum_n s_n & \text{diverges} \\ \sum_n s_n^2 & \text{converges.} \end{cases}$$

We will try to justify this heuristics.

#### 4.2.2 Convergence to the minimizer

- We let  $\{X_n\}_n$  be the sequence of points defined as above.
- We let  $t \mapsto X(t)$  be the solution to the “gradient descent” ODE (2) with initial condition  $X(0) = X_0$ .
- We let  $t_0 = 0$  and we let

$$\tilde{X}_0 := X(t_0) = X(0) = X_0.$$

- For any  $n \geq 0$ , we define

$$t_{n+1} = t_n + s_n, \quad \tilde{X}_{n+1} = X(t_{n+1}).$$

*In other words:  $t_n$  is the time after  $n$  steps,  $\tilde{X}_n$  is the value of the “real solution” at time  $t_n$  while  $X_n$  is the value of the numerical solution after  $n$  steps.*

1. Explain why, if we assume that the series  $\sum_n s_n$  diverges, then  $\tilde{X}_n$  tends to the minimizer  $X_{\min}$  as  $n \rightarrow +\infty$ .

*In the following, we will always assume that  $\sum_n s_n$  diverges.*

2. Show that  $\tilde{X}_n$  satisfies

$$\tilde{X}_{n+1} = \tilde{X}_n - \int_{t_n}^{t_{n+1}} \nabla E(X(s)) ds.$$

*The next questions are devoted to the analysis of this numerical scheme, and are thus of “real analysis” spirit.*

3. Show that we have

$$\tilde{X}_{n+1} = \tilde{X}_n - s_n \nabla E(\tilde{X}_n) + \varepsilon_n,$$

with an error term  $\varepsilon_n$  bounded by

$$\|\varepsilon_n\| \leq \frac{MLs_n^2}{2},$$

where  $L, M$  are the Lipschitz constants defined in the introduction.

4. Using  $\alpha$ -convexity, show that

$$\|X_n - s_n \nabla E(X_n) - \tilde{X}_n + s_n \nabla E(\tilde{X}_n)\|^2 \leq \|X_n - \tilde{X}_n\|^2 (1 - 2\alpha s_n + s_n^2 M^2),$$

where  $\alpha, M$  are the constants defined in the introduction.

5. For any  $n \geq 0$ , we let  $V_n$  be the difference  $V_n := X_n - \tilde{X}_n$ . Prove that

$$\|V_{n+1}\| \leq \|V_n\| \sqrt{1 - 2\alpha s_n + s_n^2 M^2} + \frac{MLs_n^2}{2}.$$

6. Using the discrete version of Grönwall’s lemma recalled in the “Auxiliary results” section, show that if  $\sum_n s_n^2$  converges, then  $X_n$  tends to  $X_{\min}$  as  $n \rightarrow +\infty$ .

1. Since  $t_n = \sum_{k=0}^n s_k$ , the divergence of the series is equivalent to the fact that  $\lim_{n \rightarrow \infty} t_n = +\infty$ , but we know from previous questions that

$$\lim_{t \rightarrow \infty} X(t) = X_{\min},$$

so, since  $\tilde{X}_n = X(t_n)$ , we have

$$\boxed{\lim_{n \rightarrow \infty} \tilde{X}_n = \lim_{n \rightarrow \infty} X(t_n) = X_{\min}.}$$

2. We use the fundamental theorem of calculus and the fact that  $X$  is a solution to the ODE:

$$tX_{n+1} = X(t_{n+1}) = X(t_n) + \int_{t_n}^{t_{n+1}} X'(s)ds = \tilde{X}_n - \int_{t_n}^{t_{n+1}} \nabla E(X(s))ds.$$

3. We need to show that

$$\left\| \int_{t_n}^{t_{n+1}} \nabla E(X(s))ds - s_n \nabla E(X(t_n)) \right\| \leq \frac{MLs_n^2}{2}.$$

We may write

$$\left| \int_{t_n}^{t_{n+1}} \nabla E(X(s))ds - s_n \nabla E(X(t_n)) \right| = \int_{t_n}^{t_{n+1}} \|\nabla E(X(s)) - \nabla E(X(t_n))\| ds.$$

Since  $\nabla E$  is Lipschitz, we have

$$\|\nabla E(X(s)) - \nabla E(X(t_n))\| \leq M\|X(s) - X(t_n)\|.$$

The mean value theorem gives

$$\|X(s) - X(t_n)\| \leq (s - t_n) \sup_{\|X'(t)\|}$$

and, since  $E$  is assumed to be Lipschitz, we have

$$\sup_t \|X'(t)\| = \sup_t \|\nabla E(X(t))\| \leq \sup_x \|\nabla E(x)\|L.$$

So we obtain

$$\left\| \int_{t_n}^{t_{n+1}} \nabla E(X(s))ds - s_n \nabla E(X(t_n)) \right\| \leq \int_{t_n}^{t_n + s_n} ML(s - t_n)ds = \frac{MLs_n^2}{2}.$$

4. Expand the square as

$$\begin{aligned} \|X_n - s_n \nabla E(X_n) - \tilde{X}_n + s_n \nabla E(\tilde{X}_n)\|^2 &= \|X_n - \tilde{X}_n\|^2 + s_n^2 \|\nabla E(\tilde{X}_n) - \nabla E(X_n)\|^2 \\ &\quad + 2s_n \langle X_n - \tilde{X}_n, \nabla E(\tilde{X}_n) - \nabla E(X_n) \rangle. \end{aligned}$$

The  $\alpha$ -convexity assumption implies that

$$\langle X_n - \tilde{X}_n, \nabla E(\tilde{X}_n) - \nabla E(X_n) \rangle \leq -\alpha \|X_n - \tilde{X}_n\|^2,$$

and the Lipschitz-ness of  $\nabla E$  gives

$$\|\nabla E(\tilde{X}_n) - \nabla E(X_n)\|^2 \leq M^2 \|\tilde{X}_n - X_n\|^2$$

and thus we get

$$\|X_n - s_n \nabla E(X_n) - \tilde{X}_n + s_n \nabla E(\tilde{X}_n)\|^2 \leq \|X_n - \tilde{X}_n\|^2 (1 - 2\alpha s_n + M^2 s_n^2).$$

5. We simply combine the results of question 3 and question 4.
6. Applying the lemma, we obtain

$$\|V_n\| \leq \exp\left(\sum_{k=0}^{n-1} \ln\left(\sqrt{1 - 2\alpha s_k + s_k^2 M^2}\right)\right) \left(\|V_0\| + \sum_{k=0}^{n-1} \frac{MLs_k^2}{2}\right).$$

We assumed that  $\sum_k s_k^2$  is finite (and in particular the step-sizes tend to 0). We may thus write

$$\left(\|V_0\| + \sum_{k=0}^{n-1} \frac{MLs_k^2}{2}\right) \leq C,$$

for some constant  $C$ , and on the other hand a first-order expansion of  $\ln(1 - x)$  yields

$$\ln\left(\sqrt{1 - 2\alpha s_k + s_k^2 M^2}\right) = \frac{1}{2}(-2\alpha s_k + O(s_k^2)).$$

Since  $\sum_k s_k$  diverges and  $\sum_k s_k^2$  converges a well-known theorem about comparison of series with positive terms implies

$$\lim_{n \rightarrow \infty} \sum_{k=0}^n \ln\left(\sqrt{1 - 2\alpha s_k + s_k^2 M^2}\right) = -\infty.$$

So since  $\lim_{x \rightarrow -\infty} \exp(x) = 0$ , we obtain  $\boxed{\lim_{n \rightarrow \infty} \|V_n\| = 0}$ .

### 4.3 A noisy version (20)

In this section, we fix the dimension  $d = 1$  and we consider the ODE

$$\boxed{x'_\varepsilon(t) = -\mathbf{E}'(x_\varepsilon(t)) + \varepsilon A(t), \quad x_\varepsilon(0) = x_0} \quad (4)$$

where  $t \mapsto x_\varepsilon(t)$  is an unknown function with real values,  $\mathbf{E}$  satisfies the same assumptions as before (but in addition, we assume  $\mathbf{E}$  to be of class  $C^2$ ),  $\varepsilon$  is some fixed real parameter and  $A$  is a continuous function such that

$$T \mapsto \left| \int_0^T A(t) dt \right| \text{ is bounded.}$$

Let  $\bar{x}$  be the solution to (4) when  $\varepsilon = 0$ . We look for an expression of  $x_\varepsilon$  as

$$x_\varepsilon = \bar{x} + \varepsilon \tilde{x} + O(\varepsilon^2).$$

1. Write down (without rigorous justification) the ODE satisfied by  $\tilde{x}$ .
2. Write down an expression for  $\tilde{x}$ .
3. Show that  $\tilde{x}(t)$  is bounded as  $t \rightarrow +\infty$ .

1. We have

$$\bar{x}'(t) + \varepsilon \tilde{x}'(t) + O(\varepsilon^2) = -\mathbf{E}'(\bar{x} + \varepsilon \tilde{x} + O(\varepsilon^2)) + \varepsilon A(t),$$

so

$$\bar{x}'(t) + \varepsilon \tilde{x}'(t) + O(\varepsilon^2) = -\mathbf{E}'(\bar{x}(t)) - \varepsilon \mathbf{E}''(\tilde{x}(t)) + \varepsilon A(t) + O(\varepsilon^2).$$

Formally,  $\tilde{x}$  must satisfy

$$\boxed{\tilde{x}'(t) = -\mathbf{E}''(\tilde{x}(t)) + A(t).}$$

2. The  $\alpha$ -convexity assumption translates into the fact that  $\mathbf{E}''$  is positive, and bounded below by  $\alpha$ . We may thus consider the map  $x \mapsto \frac{-1}{\mathbf{E}''(x)}$ , and we denote by  $G$  an antiderivative of this map. Since  $G'$  has a sign,  $G$  is a one-to-one continuous map and we denote by  $G^{-1}$  its inverse bijection (defined on the codomain).

We have

$$G'(\tilde{x}(t))\tilde{x}'(t) = A(t),$$

and thus, integrating between 0 and  $t$ , we obtain

$$G(\tilde{x}(t)) - G(\tilde{x}(0)) = \int_0^t A(s) ds,$$

so we obtain the formal expression

$$\boxed{\tilde{x}(t) = G^{-1} \left( \int_0^t A(s) ds + G(\tilde{x}(0)) \right)}.$$

3. Since  $\int_0^t A(s) ds$  stays bounded as  $t$  varies (by assumption), and since  $G^{-1}$  is continuous, we see that  $\tilde{x}(t)$  stays bounded as  $t$  varies.