GLOBAL DYNAMICS OF GRADIENT SYSTEMS

1. Basic facts on gradient systems

Let $V : \mathbb{R}^n \longrightarrow \mathbb{R}$ a smooth(C^∞) function, and consider the following first order system of ODEs:

$$\dot{\mathbf{y}}(t) = -\nabla V\big(\mathbf{y}(t)\big),\ \mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{pmatrix},\ \dot{\mathbf{y}}(t) = \frac{d}{dt}\mathbf{y}(t), \tag{1.1}$$

subject to an initial condition $\mathbf{y}(t_0) = \mathbf{y_0} \in \mathbb{R}^n$, say. We use the convention $\nabla V = \begin{pmatrix} \partial_{y_1} V \\ \partial_{y_2} V \\ \dots \\ \partial_{y_n} V \end{pmatrix}$. Then since all components of $\nabla V(\mathbf{y})$ are locally Lipschitz (why?), we can apply the local Cauchy-Lipschitz theorem to conclude existence and uniqueness of local solutions for (1.1). Assume that we know a priori that the trajectory $\mathbf{y}(t)$ exists for all $t \geq t_0$. Then a very natural question to ask is:

Large time asymptotics: *Can we describe what a typical solution $\mathbf{y}(t)$ does for large times t?*

Remarkably, the special structure of (1.1) lets one make rather general statements about the asymptotic behaviour of solutions which exist globally, as long as these solutions stay in a bounded set. This would be rather hopeless for general first order systems. We observe right away that (1.1) is an example of an *autonomous system*, in the sense that the right hand side $-\nabla V(\mathbf{y})$ is a function of $\mathbf{y}$ only, and not involving the independent variable t. In particular, this implies that if $\mathbf{y}(t)$ is a trajectory, then so is any translate $\mathbf{y}(t + a)$, but of course the initial conditions change under such a time translation.

2. The case of isolated critical points

Recall that a critical point $\mathbf{y}_* \in \mathbb{R}^n$ for V is a point where $\nabla V(\mathbf{y}_*) = \mathbf{0}$. These play an obvious role for the dynamics of (1.1) since they are automatically *fixed points* or *equilibria*, i. e. the *constant function* $\mathbf{y}(t) = \mathbf{y}_*$ is in fact a solution. To visualise the orbits of (1.1), we note the following immediate

Lemma 2.1. *Assume that the level set $\mathcal{C} := \{V(\mathbf{y}) = c\}$ is such that $\nabla V(\mathbf{y}) \neq \mathbf{0}$ for all $\mathbf{y} \in \mathcal{C}$; in particular, $\mathcal{C}$ is a C^1 hypersurface in $\mathbb{R}^n$. Then the trajectories of* (1.1) *intersect $\mathcal{C}$ at right angles.*

Proof. This follows since the normal vectors to $\mathcal{C}$ in $\mathbf{y} \in \mathcal{C}$ are given by $\pm\frac{\nabla V(\mathbf{y})}{\|\nabla V(\mathbf{y})\|}$. □

Moreover, we observe that the trajectories move from higher to lower values of V:

Lemma 2.2. *If $\mathbf{y}(t)$ is a trajectory of* (1.1)*, then $t \to V(\mathbf{y}(\mathbf{t}))$ is decreasing.*

Proof. This follows from the fact that

$$\frac{d}{dt}\big(V(\mathbf{y}(\mathbf{t}))\big) = \dot{\mathbf{y}}(\mathbf{t}) \cdot \nabla\mathbf{V}\big(\mathbf{y}(\mathbf{t})\big) = -\big|\nabla\mathbf{V}(\mathbf{y}(\mathbf{t}))\big|^{\mathbf{2}}, \tag{2.1}$$

□

Observe that if $\mathbf{y}(t)$, $t \geq t_0$ is a bounded trajectory of (1.1), then by Bolzano-Weierstrass we can select a sequence of times $\{t_n\}_{n\geq 1}$ with $\lim t_n = +\infty$ and such that

$$\lim_{n\to\infty} \mathbf{y}(t_n) = \mathbf{y}_*$$

exists. The following theorem then gives powerful conclusions, especially in case the critical points of V are isolated:

Theorem 2.3. *Assume V is defined and smooth on $\mathbf{R}^n$. Let $\mathbf{y}_* \in \mathbf{R}^n$ be an ω-limit point or an α-limit point of a trajectory $\mathbf{y}(t)$ of (1.1). This means that either*

$$\lim_{t_n \to +\infty} \mathbf{y}(t_n) = \mathbf{y}_*$$

or

$$\lim_{t_n \to -\infty} \mathbf{y}(t_n) = \mathbf{y}_*$$

for some sequence $\{t_n\}_{n\geq 1} \subset \mathbf{R}$ (in particular, it is assumed that $\mathbf{y}(t)$ exists for $t \geq 0$ in the case of the ω-limit, and analogously for the α-limit). Then $\mathbf{y}_*$ is a critical point, $\nabla V(\mathbf{y}_*) = \mathbf{0}$. If the critical points of V are isolated, then the trajectory $\mathbf{y}(t) \to \mathbf{y}_*$ as $t \to \pm\infty$, respectively.

Proof. We consider the case of an ω-limit, the case of an α-limit being treated analogously. The first part of the theorem is equivalent to $\mathbf{y}_*$ being an equilibrium. To see this, consider the solution to

$$\widetilde{\mathbf{y}}' = -\nabla V(\widetilde{\mathbf{y}}),\ \widetilde{\mathbf{y}}(0) = \mathbf{y}_*, \tag{2.2}$$

which exists at least locally around $t = 0$. If $\nabla V(\mathbf{y}_*) \neq \mathbf{0}$, we get

$$\frac{d}{dt}\big(V(\widetilde{\mathbf{y}}(\mathbf{t}))\big)|_{\mathbf{t=0}} = -|\nabla \mathbf{V}(\mathbf{y}_*)|^{\mathbf{2}} < \mathbf{0} \tag{2.3}$$

We can now deduce a contradiction from this as follows: by assumption

$$\mathbf{y}_* = \lim_{n\to\infty} \mathbf{y}(t_n),$$

with t_n increasing to $+\infty$. By continuous dependence of solutions of (1.1) on initial data, given $\delta_0 > 0$, we can find $\delta t > 0$, $\delta_1 > 0$, such that for all $\mathbf{y_1} \in \mathbf{R}^n$ with

$$|\mathbf{y_1} - \mathbf{y}_*| < \delta_1,$$

the solution $\mathbf{y_1}(t)$ of

$$\mathbf{y_1}' = -\nabla V(\mathbf{y_1}),\ \mathbf{y_1}(0) = \mathbf{y_1}, \tag{2.4}$$

exists on $[0, \delta t]$, and satisfies

$$\max_{t\in[0,\delta t]} |\widetilde{\mathbf{y}}(t) - \mathbf{y_1}(t)| < \delta_0$$

It then follows from (2.3) that for $\delta t, \delta_0$ small enough, we get

$$V(\mathbf{y_1}(\delta t)) < V(\mathbf{y}_*) - \delta_0$$

Now pick $n_0 \in \mathbf{N}$ large enough, such that for $n \geq n_0$, we have

$$|\mathbf{y}(t_n) - \mathbf{y}_*| < \delta_1$$

Then we have

$$V(\mathbf{y}(t_n + \delta t)) < V(\mathbf{y}_*) - \delta_0,$$

and by monotonicity of V along trajectories, we conclude that

$$V(\mathbf{y}(t_m)) < V(\mathbf{y}_*) - \delta_0$$

for m sufficiently large. But this contradicts

$$\lim_{n\to\infty} V(\mathbf{y}(t_n) = V(\mathbf{y}_*)$$

in turn a consequence of $\mathbf{y}_* = \lim_{n\to\infty} \mathbf{y}(t_n)$.

Now we assume in addition that the critical points of V are isolated, i. e. that whenever $\mathbf{y}_*$ is a critical point, there exists $\delta > 0$ such that $B_\delta(\mathbf{y}_*)\backslash\{\mathbf{y}_*\}$ contains no critical point. In the situation from above, assume that we do not have $\lim_{t\to+\infty} \mathbf{y}(t) = \mathbf{y}_*$ (in the case of an ω-limit). This means that that there exists (check!) some $\delta_2 > 0$ and a sequence of times $\{\widetilde{t}_m\}_{m\geq 1}$ converging toward $+\infty$ and such that

$$|\mathbf{y}(\widetilde{t}_m) - \mathbf{y}_*| = \delta_2$$

We may as well assume by passing to a subsequence that $\mathbf{y}(\widetilde{t}_m) \to \widetilde{\mathbf{y}}_* \in \mathbf{R}^n$, with

$$|\mathbf{y}_* - \widetilde{\mathbf{y}}_*| = \delta_2$$

By the above, $\widetilde{\mathbf{y}}_*$ is a critical point. If we choose δ_2 sufficiently small, this contradicts our assumption of $\mathbf{y}_*$ being isolated. □

3. Real analytic V via gradient Lojasiewicz inequality

Two things may be criticised about the preceding theorem: the restriction on isolated critical points is somewhat limiting. For example, the function

$$V(x, y) = \cos(xy)$$

does not satisfy this requirement.
Second, the result is completely non-quantitative, in that we cannot say at all *how fast* a given trajectory is going to converge to the limiting critical point, in case of isolated critical points.

It turns out that in general, the theorem becomes false if one allows non-isolated critical points. However, a deep result, whose proof we can only partially give in the context of this course, gives a more powerful conclusion for *real analytic potentials functions* $V(\mathbf{y})$. Incidentally, this result also gives a quantitative bound on how fast the trajectory converges. For the sequel, recall that we say that V is real analytic on $\mathbb{R}^n$, provided for each $\mathbf{y}_0$ admits a $r > 0$ such that on the ball $B_r(\mathbf{y_0})$ we can represent V as a convergent power series

$$V(\mathbf{y}) = \sum_{\alpha \in \mathbf{N}^n_{\geq 0}} a_\alpha (\mathbf{y} - \mathbf{y}_0)^\alpha,\ \mathbf{y}^\alpha = \prod_{j=1}^n y_j^{\alpha_j}.$$

Then the following deep theorem will be assumed without proof:

Theorem 3.1. *(Lojasiewicz, 1960) Let f be a real analytic function on $\mathbb{R}^n$ and $x_* \in \mathbb{R}^n$. Then there exist $\beta \in (0, 1)$, $c_{1,2} > 0$, such that for each $x \in \mathbb{R}^n$ with $|x - x_*| < c_2$, we have*

$$\big|f(x) - f(x_*)\big|^\beta \leq c_1 \big|\nabla f(x)\big|.$$

Now consider the system

$$\dot{x}(t) = -\nabla f(x(t)),$$

with f real analytic on $\mathbb{R}^n$, and assume that $x(t)$ is a bounded trajectory forward in time (i. e. toward $t = +\infty$). In particular, there is a sequence $t_n \to \infty$ such that

$$x(t_n) \longrightarrow x_* \in \mathbb{R}^n.$$

Then we have the following

Proposition 3.2. *The trajectory $x(t)$ actually converges toward x_* as $t \to +\infty$, and there is $C > 0, \delta > 0$ such that*

$$\big|x(t) - x_*\big| \leq C t^{-\delta}.$$

Proof. We may assume that $f(x_*) = 0$. Recall from the preceding theorem that for suitable $\beta \in (0, 1), c_1 > 0, c_2 > 0$, we have

$$\big|f(x)\big|^\beta = \big|f(x) - f(x*)\big|^\beta \leq c_1 \big|\nabla f(x)\big|, \tag{3.1}$$

provided $|x - x_*| < c_2$. Here we may assume $\beta > \frac{1}{2}$. Then we claim that

$$\int_t^\infty |\dot{x}(t)| \leq C t^{-\delta} \tag{3.2}$$

for suitable $\delta > 0, C > 0$, which easily implies the proposition. To prove (3.2), let c_2 be as above, and pick t_m large enough such that $\big|x(t_m) - x_*\big| < \frac{c_2}{2}$. We will show that, increasing m if necessary, provided $t \geq t_m$, we have $\big|x(t) - x_*\big| < c_2$, and moreover (3.2) holds. For this, we may assume that $f(x(t)) \neq 0$ for all t. In fact, if not, then since $f(x(t))$ decreases towards its asymptotic value $\lim_{t_m \to +\infty} f(x(t_m)) = 0$, we necessarily have $\nabla f(x(t_*)) = 0$ for some t_* where $f(x(t_*)) = 0$, and this means that $x(t_*)$ is an equilibrium, whence the entire trajectory would have to be constant, and the conclusion of the proposition is trivial. Next, if $x(t) \in B_{c_2}(x_*)$, then (3.1) implies

$$\frac{d}{dt}\big[f(x(t))\big] = -\big|\nabla f(x(t))\big|^2 \leq -c_1^{-2} f(x(t))^{2\beta}.$$

This in turn implies, recalling the hypothesis $\beta > \frac{1}{2}$, that

$$\frac{d}{dt}\big[f^{1-2\beta}(x(t))\big] = (1-2\beta)\cdot f^{-2\beta}\cdot \frac{d}{dt}\big[f(x(t))\big] \geq c_1^{-2}(2\beta-1),$$

which implies that as long as $x(s) \in B_{c_2}(x_*)$, $t \geq s \geq t_m$, we have (we have $c = c_1^{-1}$)

$$[f(x(t))]^{1-2\beta} \geq c^2(2\beta-1)(t-t_m) + D,$$

whence

$$0 \leq f(x(t)) \leq \frac{1}{[c^2(2\beta-1)(t-t_m)+D]^{\frac{1}{2\beta-1}}},$$

where the key is that $\frac{1}{2\beta-1} = 1+\varepsilon$ with $\varepsilon > 0$ since $\frac{1}{2} < \beta < 1$. The constant $D = [f(x(t_m))]^{1-2\beta}$, which can be made arbitrarily large by picking m large enough (why?). We can now show that if we increase m if necessary and let $t \geq t_m$, then we have $x(t) \in B_{c_2}(x_*)$, and the inequality (3.2) is valid. Indeed, observe that using the Cauchy-Schwarz inequality we have

$$\begin{aligned}\int_{t_m}^{t} \big|\dot{x}(s)\big|\,ds &\leq \int_{t_m}^{t}\big|\nabla f(x(s))\big|\,ds\\ &= \int_{t_m}^{t}\sqrt{-\frac{d}{ds}[f(x(s))]}\,ds\\ &\leq \Big(\int_{t_m}^{t} -\frac{d}{ds}[f(x(s))]\cdot(s-t_m+1)^{1+\frac{\varepsilon}{2}}\,ds\Big)^{\frac{1}{2}}\cdot\Big(\int_{t_m}^{t}(s-t_m+1)^{-1-\frac{\varepsilon}{2}}\,ds\Big)^{\frac{1}{2}}.\end{aligned}$$

Here we can use integration by parts for the first integral:

$$\begin{aligned}&\int_{t_m}^{t} -\frac{d}{ds}[f(x(s))]\cdot(s-t_m+1)^{1+\frac{\varepsilon}{2}}\,ds\\ &= -f(x(s))\cdot(s-t_m+1)^{1+\frac{\varepsilon}{2}}\Big|_{t_m}^{t} + (1+\frac{\varepsilon}{2})\cdot\int_{t_m}^{t} f(x(s))\cdot(s-t_m+1)^{\frac{\varepsilon}{2}}\,ds\\ &< \min\{\varepsilon,1\}\cdot\frac{c_2^2}{10},\end{aligned}$$

for $D = D(\varepsilon, c_2)$ large enough above, as long as $x(s) \in B_{c_2}(x_*)$ for $t_m \leq s \leq t$. In turn, these imply that

$$\int_{t_m}^{t}\big|\dot{x}(s)\big|\,ds < \frac{c_2}{2},$$

say, as long as $x(s) \in B_{c_2}(x_*)$, $t_m \leq s \leq t$ A simple continuity argument[1] then implies that $x(t) \in B_{c_2}(x_*)$ for all $t \geq t_m$, and then the inequality (3.2) follows with $\delta = \frac{\varepsilon}{4}$, provided $t \geq t_m$, and then with suitably modified C for all $t \geq 0$. Indeed, we can bound for $t \geq t_m$ (with m as in the preceding sufficiently large)

$$\begin{aligned}\big|x(t)-x_*\big| &\leq \int_{t}^{\infty}\big|\dot{x}(s)\big|\,ds\\ &\leq \Big(\int_{t}^{\infty} -\frac{d}{ds}[f(x(s))]\cdot(s-t_m+1)^{1+\frac{\varepsilon}{2}}\,ds\Big)^{\frac{1}{2}}\cdot\Big(\int_{t}^{\infty}(s-t_m+1)^{-1-\frac{\varepsilon}{2}}\,ds\Big)^{\frac{1}{2}}\\ &\leq C_1\cdot t^{-\frac{\varepsilon}{4}}\end{aligned}$$

for suitable $C_1 = C_1(t_m)$, while we also have the bound

$$\big|x(t)-x_*\big| \leq C_2\cdot t^{-\frac{\varepsilon}{4}},\ t\in[0,t_m]$$

[1] Assume that there is a first time $t \geq t_m$such that $x(t) \notin B_{c_2}(x_*)$, but $x(s) \in B_{c_2}(x_*)$ for $t_m \leq s < t$, and derive a contradiction from this.

for suitable $C_2 = C_2(t_m)$ by continuity. We can then set $C = \max\{C_1, C_2\}$.

$\square$