1. Preliminaries: the contraction mapping theorem

When confronted with an equation that cannot be solved explicitly, one can often take a reasonable guess at a solution and then improve this guess according to some algorithm (e.g. Newton's method) or other. That is, given an initial guess x_0 at a solution, there is a function T such that $x_1 = T(x_0)$ is a better guess at a solution, and $x_2 = T(x_1)$ is a still better guess, etc. Repeating this proceedure one obtains an entire sequence (x_j) of better and better approximate solutions of the given equation, and if life is good one can show that there is a limit $x_\infty = \lim x_n$ that actually solves the given equation.

One of the most common features of 'guess-improvement functions' is that they tend to shrink distances between points.

Definition 1.1. *Let $X \subset \mathbf{R}^n$ be a closed set. A contraction mapping on X is a function $T : X \to X$ with the following property: there exists a constant $C < 1$ such that $\|T(x) - T(y)\| \leq C\|x - y\|$ for all $x, y \in X$.*

It should be pointed out that the definition implies rather directly that contraction mappings are continuous. That is, given $x \in X$ and $\epsilon > 0$, one can set $\delta = C^{-1}\epsilon$ and observe that if $\|x - y\| < \delta$, then

$$\|T(x) - T(y)\| \leq C\|x - y\| < C\delta < \epsilon.$$

Hence $\lim_{y \to x} T(y) = T(x)$, and T is continuous at x.

The condition that the source and target of a contraction mapping T be the same set allows us to iterate, composing T with itself as often as we like. The most important fact about contraction mappings concerns what happens when one does this.

Theorem 1.2 (Contraction Mapping Theorem)**.** *Suppose that $T : X \to X$ is a contraction mapping. Then there is a unique $x_{fix} \in X$ such that $T(x_{fix}) = x_{fix}$. Moreover, if $x_0 \in X$ is any point, and (x_j) is the sequence with initial term x_0 and other terms given inductively by $x_j = T(x_{j-1}) = T^j(x_0)$, then*

$$\lim x_j = x_{fix}.$$

The point x_{fix} is called a *fixed point* for T. So the theorem may be restated by saying that each contraction mapping has a unique fixed point and that when we appy T repeatedly to any other point $x \in X$, the images converges to x_{fix}.

Proof. We prove the theorem for $X \subset \mathbf{R}$ (i.e. when the dimension $n = 1$) and leave the case $n > 1$ as an exercise. Let $x_0 \in X$ be any point and $x_j = T^j(x_0)$ be as described in the theorem. Note that we can think of x_j as a partial sum for a telescoping series:

$$x_j = x_0 + \sum_{i=0}^{j-1}(x_{i+1} - x_i).$$

Thus $\lim x_j = x_0 + \sum_{i=1}^{\infty}(x_i - x_{i-1})$ provided we can show that the infinite series converges.

Now since T is a contraction mapping, we have the following convenient upper bound for the terms in the series:

$$|x_{i+1} - x_i| = |T(x_i) - T(x_{i-1}| \leq C|x_i - x_{i-1}| \leq \cdots \leq C^i|x_1 - x_0|.$$

That is, since $C < 1$, the terms of our series are dominated by the terms of a convergent geometric series. So by the comparison test for infinite series (see Apostol, volume 1), our series converges absolutely. In particular, the limit $\lim x_j$ exists and we call it x_{fix}. Since X is closed, we have $x_{fix} \in X$. That is, x_{fix} belongs to the domain of T.

Since T is continuous, we further have that

$$T(x_{fix}) = T(\lim x_j) = \lim T(x_j) = \lim x_{j+1} = x_{fix}.$$

That is, x_{fix} is a fixed point of T.

Finally, if $x \in X$ is some other fixed point of T, then we have

$$\|x - x_{fix}\| = \|T(x) - T(x_{fix})\| \leq C \|x - x_{fix}\|.$$

Since $C < 1$ and both sides are non-negative, both sides must be zero. That is, $x = x_{fix}$, so T has only one fixed point. □

In fact, the notion of a contraction mapping makes sense in much greater generality than we have defined it above. Any time one has a notion of distance between points (a 'metric') in a set X, one can define what it means for a mapping $T : X \to X$ to be a contraction. And then the contraction mapping theorem holds as stated above with the additional hypothesis that X is *complete*, a technical notion which we will not define here. In order to apply the contraction mapping theorem to produce solutions to ODEs, we will state and prove a version of the above the theorem to the situation where points in $\mathbf{R}^n$ are replaced by continuous functions. Keep in mind, however, that both the versions of the contraction mapping theorem given here are simply special cases of a single very general theorem.

For any sets $A \subset \mathbf{R}^n, B \subset \mathbf{R}^m$, we let $C(A, B)$ denote the collection of all bounded, continuous mappings $f : A \to B$. We can measure the 'size' of any such f according to

$$\|f\| := \sup_{x \in A} \|f(x)\|.$$

and then declare the distance between $f, g \in C(A, B)$ to be $\|f - g\|$. Observe that by definition, $\|f - g\| \leq \epsilon$ if and only if $\|f(x) - g(x)\| \leq \epsilon$ for all $x \in A$. In particular, $\|f - g\| = 0$ if and only if $f \equiv g$ are the same function. Finally, observe that if $(f_j) \subset C(A, B)$ is a sequence of functions, then $f_j \to f \in C(A, B)$ uniformly on A if and only if $\lim_{j \to \infty} \|f_j - f\| = 0$.

Definition 1.3. *A transformation $T : C(A, B) \to C(A, B)$ is a* contraction mapping *if there exists $C > 0$ such that $\|T(f) - T(g)\| \leq C \|f - g\|$ for all $f, g \in C(A, B)$.*

Theorem 1.4 (Contraction Mapping Theorem for continous mappings)**.** *Suppose that $T : C(A, B) \to C(A, B)$ is a contraction mapping. If B is closed, then there is a unique function $f_{fix} \in C(A, B)$ such that $T(f_{fix}) = f_{fix}$. Moreover, if $f_0 \in C(A, B)$ is any function, then the sequence of functions $f_j := T^j(f)$ converges uniformly on A to f_{fix}.*

Proof. The proof is quite similar to the contraction mapping theorem for points. With $f_j = T^j(f)$ as in the statement of the theorem, we have

$$f_j = f_0 + \sum_{i=0}^{j-1} (f_{i+1} - f_i).$$

Hence (f_j) converges uniformly if and only if the infinite series $f_0 + \sum_{i=0}^{\infty} (f_{i+1} - f_i)$ does. The hypothesis that T is a contraction gives us as before that

$$\|f_{i+1} - f_i\| \leq C^i \|f_1 - f_0\|.$$

That is the terms of our series of functions are bounded above by the terms of a convergent geometric series. So the Weierstrass M-test tells us that $\sum(f_{i+1} - f_i)$ converges uniformly and absolutely on A, and we conclude that

$$\lim f_j = f_{fix} := f_0 + \sum_{i=0}^{\infty}(f_{i+1} - f_i)$$

uniformly on A.

To see that $T(f_{fix}) = f_{fix}$, let ϵ be any positive number. Then there exists $J \in \mathbf{N}$ such that $j \geq J$ implies $\|f_j - f_{fix}\| < \epsilon/2$. Hence by the triangle inequality and $f_{J+1} = T(f_J)$, we have

$$\begin{aligned} \|T(f_{fix}) - f_{fix}\| &= \|T(f_{fix}) - T(f_J) + f_{J+1} - f_{fix}\| \\ &\leq \|T(f_{fix}) - T(f_J)\| + \|f_{J+1} - f_{fix}\| \\ &\leq C\,\|f_{fix} - f_J\| + \|f_{J+1} - f_{fix}\| < C\epsilon/2 + \epsilon/2 < \epsilon. \end{aligned}$$

As $\epsilon > 0$ was arbitrary, we conclude $\|T(f_{fix}) - f_{fix}\| = 0$, i.e. $T(f_{fix}) = f_{fix}$. The argument that f_{fix} is the *unique* fixed point of T is the same as before. □

2. The existence and uniqueness theorem for first order ODEs

The fundamental fact about ordinary differential equations is that, under suitably nice circumstances and subject to appropriate initial conditions, one gets unique solutions. Here we will discuss this fact in the particular case of first order ODEs. The case of first order *systems* of ODEs is quite similar and essentially contains all other possible cases.

Let us set up the problem before stating any results. We begin with an open set $U \subset \mathbf{R}^2$ and a function $F : U \to \mathbf{R}$. Given any point $(t_0, y_0) \in U$ we seek solutions to the initial value problem

$$y'(t) = F(t, y(t)), \quad y(t_0) = y_0. \tag{1}$$

The domain of the function y is not so important here, so we allow ourselves to consider any differentiable function $y : I \to \mathbf{R}$ defined on an open interval I containing t_0. If such a y satisfies (1), then we refer to $y : I \to \mathbf{R}$ as a *(local) solution* of (1).

Theorem 2.1 (Existence and Uniqueness Theorem)**.** *Suppose that $F = F(t, y)$ is continuous on U and furthermore continuously differentiable with respect to the second variable y. Then for any $(t_0, y_0) \in U$ there is a solution $y : I \to \mathbf{R}$ of the initial vale problem* (1)*. This solution is unique in the following sense: if $\tilde{y} : \tilde{I} \to \mathbf{R}$ is another solution, then $\tilde{y}(t) = y(t)$ for all $t \in I \cap \tilde{I}$.*

The solution y will appear as the fixed point of a certain contraction mapping T. In order to introduce T, we observe that if $y : I \to \mathbf{R}$ solves (1), then y is differentiable and therefore continous. By hypothesis, f is also continuous, so (1) says among other things that y' is equal to a composition of continous functions and therefore itself continuous. So for any $t \in I$, we may apply the fundamental theorem of calculus to obtain

$$y(t) - y(t_0) = \int_{t_0}^{t} y'(s)\,ds = \int_{t_0}^{t} f(s, y(s))\,ds.$$

That is y is a fixed point of the mapping $T : C(I, [y_0 - R, y_0 + R]) \to C(I, \mathbf{R})$ given by

$$Ty(t) = y(t_0) + \int_{t_0}^{t} f(s, y(s))\, ds. \tag{2}$$

Here we choose $R > 0$ so that $I \times [y_0 - R, y_0 + R] \subset U$. Otherwise the integrand in the definition of $Ty(t)$ might fail to be defined. Note also that at this point, the source and target of T are different, because it is not a priori clear that $|Ty(t) - y_0| \leq R$ for all t. Anyhow, by reversing the above computation, we find

Lemma 2.2. *Let I be an open interval containing t_0 and $y : I \to \mathbf{R}$ be continuous. Then y solves* (1) *if and only if $Ty = y$, where $T : C(I, [y_0 - R, y_0 + R]) \to C(I, \mathbf{R})$ is defined by* (2)*.*

Proof. We have already seen that if y solves (1), then $Ty = y$. If, on the other hand, $y : I \to J$ is any continuous function satisfying $Ty = y$, then we may employ the (other) fundamental theorem of calculus to differentiate $Ty(t)$, obtaining

$$y'(t) = (Ty)'(t) = f(t, y(t)).$$

Moreover, setting $t = t_0$ in (2) gives $y(t_0) = Ty(t_0) = y_0 + 0$. Hence y solves (1). □

For the remainder of the proof, we choose $\delta_0, R > 0$ so that $\overline{I(\delta_0)} \times K \subset U$, where $I(\delta_0) = (t_0 - \delta_0, t_0 + \delta)$ and $K = [y_0 - R, y_0 + R]$. Note that for any $\delta < \delta_0$, we have $\overline{I(\delta) \times K} \subset U$, too.

Lemma 2.3. *There exists $\delta < \delta_0$ such that $Ty(t) \in K$ for all $y \in C(I(\delta), K)$ and $t \in I(\delta)$. That is T maps $C(I(\delta), K)$ into itself.*

Proof. The Extreme Value Theorem tells us that there exists $A > 0$ such that $|f(t, y)| \leq A$ for all (t, y) in the compact set $\overline{I(\delta_0)} \times K$. Hence if $\delta < \delta_0$, we have for any continous $y : I(\delta) \to K$ and any $t \in I(\delta)$ that

$$|Ty(t) - y_0|| = \left| \int_{t_0}^{t} f(s, y(s))\, ds \right| \leq \int_{t_0}^{t} |f(s, y(s))|\, ds \leq A|t - t_0| < A\delta.$$

So if we take $\delta < R/A$, the conclusion of the lemma is satisfied. □

Lemma 2.4. *There exists $\delta < \delta_0$ such that $T : C(I(\delta), K) \to C(I(\delta), K)$ is a contraction mapping.*

Proof. Choose $\delta < \delta_0$ so that the conclusion of the previous lemma holds. By hypothesis $\frac{\partial f}{\partial y}$ is continuous on U and therefore by the Extreme Value Theorem bounded on $\overline{I(\delta_0)} \times K$. That is, there is a constant B such that $\frac{\partial f}{\partial y} \leq B$ everywhere on the latter set. So if $y, \tilde{y} \in C(I(\delta), K)$, we can apply the mean value theorem to f (as a function of y only) to compute for any $t \in I(\delta)$ that

$$Ty(t) - T\tilde{y}(t) \leq \int_{t_0}^{t} f(s, y(s)) - f(s, \tilde{y}(s))\, ds \leq \int_{t_0}^{t} \frac{\partial f}{\partial y}(s, z(s)(y(s) - \tilde{y}(s))\, ds,$$

where $z(s)$ lies between $y(s)$ and $\tilde{y}(s)$. Thus we may estimate

$$|Ty(t) - T\tilde{y}(t)| \leq \int_{t_0}^{t} B|y(s) - \tilde{y}(s)|\, ds \leq B\delta\, \|y - \tilde{y}\|$$

for all $t \in I(\delta)$. That is, $\|Ty - T\tilde{y}\| \leq B\delta \|y - \tilde{y}\|$. Choosing δ so that $B\delta < 1$, we obtain that T is a contraction mapping on $C(I(\delta), K)$. □

Our final lemma is the main step in proving uniqueness of solutions to (1).

Lemma 2.5. *Suppose that $y_1 : I(\delta_1) \to \mathbf{R}$, $y_2 : I(\delta_2) \to \mathbf{R}$ both solve* (1)*. Then there exists $\delta > 0$ such that $y_1(t) \equiv y_2(t)$ for all $t \in I(\delta)$.*

Proof. Since y_1 and y_2 both solve (1), both functions are differentiable and therefore continuous at t_0. Thus there exists $\delta > 0$ such that $|t - t_0| < \delta$ implies that $|y_j(t) - y_0| < R$. That is, $y_j \in C(I(\delta), K)$. By Lemma 2.4 we may shrink δ if necessary and assume that $T : C(I(\delta), K) \to C(I(\delta), K)$ is a contraction mapping. But Lemma 2.2 tells us that $T(y_1) = y_1$ and $T(y_2) = y_2$. Since fixed points of contraction mappings are unique, we conclude that $y_1 \equiv y_2$ on $I(\delta)$. □

Proof of Theorem 2.1. Choose $0 < \delta < \delta_0$ small enough that Lemmas 2.3 and 2.4 hold. Then $T : C(I(\delta), K)$ is a contraction mapping, so Theorem 1.4 gives us $y \in C(I(\delta), K)$ such that $T(y) = y$. By Lemma 2.2, this y solves (1), and we have proven the existence part of Theorem 2.1.

For uniqueness, suppose that $y : I \to \mathbf{R}$ and $\tilde{y} : \tilde{I} \to \mathbf{R}$ both solve (1). Let $J \subset I \cap \tilde{I}$ denote the largest interval (open or closed) containing t_0 on which $y(t) = \tilde{y}(t)$. Lemma 2.5 shows that $J \neq \emptyset$. Suppose, to get a contradiction, that $J \neq I \cap \tilde{I}$. Then some endpoint t_1 of J belongs to $I \cap \tilde{I}$. By continuity $y(t_1) = \tilde{y}(t_1)$ so that $t_1 \in J$. Thus y and $\tilde{y}$ both solve the same initial value problem at t_1. But by Lemma 2.5 again, this implies that $y(t) = \tilde{y}(t)$ for all t in some small interval $(t_1 - \epsilon, t_1 + \epsilon)$ about t_1. Thus $y(t) \equiv \tilde{y}(t)$ on all of $J \cup (t_1 - \epsilon, t_1 + \epsilon)$, contradicting the fact that J is the largest open interval about t_0 on which the two solutions agree. Thus $J = I \cap \tilde{I}$, and the proof of uniqueness is complete. □

Observe that the uniqueness part of Theorem 2.1 ensures us that there is a solution $y : I \to \mathbf{R}$ for which the interval I is as large as possible. To see that this is so, let

$$I = \bigcup \{J \subset \mathbf{R} : J \text{ is an open interval about } t_0 \text{ on which (1) admits a solution}\}$$

be the union of all solution intervals. Then certainly, I is an open interval about t_0, and we can define our 'maximal' solution $y : I \to \mathbf{R}$ at any point $t \in I$ by setting $y(t) = \tilde{y}(t)$ where $\tilde{y} : J \to \mathbf{R}$ is a solution whose domain J contains t. Since any two solutions solutions agree on the intersection of their domains, it will not matter which solution $\tilde{y}$ we use to define $y(t)$. We will have moreover that $y \equiv \tilde{y}$ on all of J, so that in particular $y' = F(t, y)$ holds at t. That is, y satisfies (1) on all of I. We call y a *global* solution of the (1) and note that in this paragraph we have just established a stronger version of Theorem 2.1

Theorem 2.6 (Existence and uniqueness of global solutions)**.** *Under the hypotheses of Theorem 2.1, there exists a global solution $y : I \to \mathbf{R}$ of* (1) *that is unique in the sense that if $\tilde{y} : J \to \mathbf{R}$ is any solution of* (1)*, then $J \subset I$ and $y \equiv \tilde{y}$ on J.*

An important feature of global solutions is that they persist until their graphs 'exit' the domain of existence U for the righthand side $F(t, y)$ of (1). More precisely, we have

Theorem 2.7. *Let $y : I \to \mathbf{R}$ be the global solutio of* (1)*. If $K \subset U$ is any compact (i.e. closed and bounded) set, then there is a closed interval $I_K \subset I$ such that $(t, y(t)) \notin K$ for any $t \notin I_K$.*

The conclusion of this theorem is often summarized by saying that the graph of $y : I \to \mathbf{R}$ is a curve that is 'properly embedded' in U. The proof of Theorem 2.7 depends on one aspect of another fundamental theorem concerning solutions to ODEs.

Theorem 2.8 (Stability Theorem). *Under the hypotheses of Theorem 2.1, we have that solutions to* (1) *depend continuously on the initial condition* (t_0, y_0). *Specifically, if* $K \subset U$ *is a compact subset, then there exists* $\epsilon > 0$ *such that for every* $(t_0, y_0) \in K$*, the global solution* $y = y_{t_0,y_0} : I \to \mathbf{R}$ *of* (1) *has domain* $I = I_{t_0,y_0}$ *containing* $(t_0 - \epsilon, t_0 + \epsilon)$. *Moreover, if* $\varphi : K \times (-\epsilon, \epsilon) \to \mathbf{R}$ *is defined by setting* $\varphi(t_0, y_0, t) = y_{t_0,y_0}(t + t_0)$*, then* φ *is continuous.*

This theorem follows from a slightly more careful version of the proof of existence used for Theorem 2.1. We omit the details here. For purposes of proving Theorem 2.7, the important part of the stability theorem is that it gives a positive lower bound ϵ on the lifespan of any solution that begins in K.

Proof of Theorem 2.7. Suppose that the theorem is false for some compact set $K \subset U$. Let $I = (a, b)$. Let $\epsilon > 0$ be the constant associated to K in the Stability Theorem. Then taking the closed interval $J = [a + \epsilon, b - \epsilon] \subset I$, we may choose a point $t_1 \in I - J$ such that $(t_1, y_1) \in K$, where $y_1 := y(t_1)$. Thus y is a solution of $y' = f(t, y)$ subject to the initial condition $y(t_1) = y_1$. The stability theorem guarantees us that there is another solution $\tilde{y} : (t_1 - \epsilon, t_1 + \epsilon) \to \mathbf{R}$ of the same initial value problem. By uniqueness, we therefore have $\tilde{y} \equiv y$ on $I \cap (t_0 - \epsilon, t_0 + \epsilon)$. By setting,

$$\hat{y}(t) = \begin{cases} y(t) & \text{if } t \in I \\ \tilde{y}(t) & \text{if } |t - t_1| < \epsilon, \end{cases}$$

we obtain that $\hat{y}$ is a solves (1) on $I \cup (t_1 - \epsilon, t_1 + \epsilon)$. Since this last interval is bigger than I, we have contradicted the fact that I is the domain of existence for the global solution of (1). So the theorem holds. □

3. Asymptotic Behavior of Solutions to Autonomous 1st Order Equations

In this section we consider solutions of the initial value problem

$$y' = f(y), y(t_0) = y_0 \tag{3}$$

where $f : \mathbf{R} \to \mathbf{R}$ is a C^1 function. Differential equations like the one here, in which the right side does not depend explicitly on t, are called *autonomous*. Such ODEs are both common in applications and important in theory[1]

The ODE in (3) is separable and therefore in principle solvable by integration. In practice, however, the integration can be unmanageable and will only give an implicit and fairly unenlightening formula for the solution. Here way take a more qualitative approach to analyzing the problem, and in particular, understanding what happens to the global solution as t tends toward the ends of the domain. This is, in pedestrian terms, somewhat akin to plugging information about today's weather into the equations of fluid mechanics to try and infer whether one ought to plan picnics in the year 10,000. In this light, it is somewhat remarkable that we will be able to say anything sensible at all.

If $f(y_0) = 0$ then we call y_0 an *equilibrium point* of the ODE. In this case, one checks easily that $y : \mathbf{R} \to \mathbf{R}$ given by $y(t) \equiv y_0$ is the global solution of (3). In particular, the

[1] In somewhat the same way we reduced solving higher order ODEs to solving first order systems, one can always reduce a non-autonomous ODE to a first order autonomous system.

domain of the global solution is all of $\mathbf{R}$, and we have $\lim_{t\to\infty} y(t) = y_0$. If $f(y_0) \neq 0$, then things are certainly more complicated. Nevertheless, we have

Theorem 3.1. *Suppose that $f(y_0) > 0$ and that $y : I \to \mathbf{R}$ is the global solution of* (1)*. Then y is strictly increasing on all of I.*

- *If there is an equilibrium point of $y' = f(y)$ that is larger than y_0, then the domain I of y includes all $t > t_0$, and we have that*
$$\lim_{t\to\infty} y(t) = L$$
where L is the smallest equilibrium point of $y' = f(y)$ larger than y_0.
- *Otherwise $\lim_{t\to b} y(t) = \infty$, where $b \leq \infty$ is the right endpoint of I.*

We leave it to the reader to puzzle out the statement of this theorem in the case $f(y_0) < 0$ and to draw the appropriate conclusions about the asymptotic behavior of the global solution $y : I \to \mathbf{R}$ as t decreases toward the *left* endpoint of the domain I. In essence, Theorem 3.1 is telling us that solutions to first order autonomous ODEs will either drift off to infinity or settle down and become asymptotically constant. If the reader finds this unsurprising, then he or she should try to imagine what the analogous assertion should be for solutions of autonomous systems of 2 or 3 ODEs (Hint: don't even try when there are 3 or more equations involved.)

Proof. To justify the first assertion, it is sufficient to show that $y'(t) > 0$ for all $t \in I$. Suppose to the contrary that $y'(t) \leq 0$ for some $t \in I$. Then there exists t_1 between t and t_0 such that $y'(t_1) = f(y(t_1)) = 0$. Thus $y(t_1)$ is an equilibrium point of the ODE. Since the graph of y passes through $(t_1, y(t_1))$, it follows by uniqueness of solutions to initial value problems that $y(t) \equiv y(t_1)$ on I. However, $y(t_0) \neq y(t_1)$ since $f(y(t_0)) > 0$, whereas $f(y(t_1)) < 0$. This contradiction proves that y' is positive on all of I.

Since y is increasing it follows that if b is the left endpoint of I, then $\lim_{t\to b} y(t) \leq \infty$ exists and is equal to $\sup_{t\in I} y(t)$. If $y_0 < y_1$ for some equilibrium point y_1, then we set
$$L = \inf\{y > y_0 : f(y) = 0\}.$$
By continuity, we must have $f(L) = 0$, and by definition $f(y) > 0$ for all $y \in [y_0, L)$.

We observe also that $y(t) \in (y_0, L)$ for all $t > t_0$. That is, $y(t) > y(t_0)$ since y is strictly increasing, and if $y(t) \geq L$ for some t, then $y(t_1) = L$ for some t_1, and we again contradict uniqueness of solutions to initial value problems. Hence $\lim_{t\to b} y(t) \leq L$.

To see that $b = \infty$, we invoke Theorem 2.7. Indeed, given any $M > 0$, the theorem guarentees the existence of a closed interval $[c, d] \subset (a, b)$ such that $(t, y(t)) \notin [-M, M] \times [y_0, L]$. Since $y_0 \leq y(t) \leq L$ for all $t \geq t_0$, we see that this can happen only if $d \geq M$. In particular $b > M$. Since M was arbitrary, we conclude $b = \infty$.

Finally, we suppose to get one last contradiction that $\lim_{t\to\infty} y(t) = z < L$. Then $f(z) > 0$ and we can choose $M > 0$ large enough that $|f(y(t)) - f(z)| < f(z)/2$ for all $t > M$—in particular, $y'(t) = f(y(t)) > f(z)/2$. Hence $y(t) \geq (t - M)f(z)/2 + y(M)$, from which we obtain that $\lim_{t\to\infty} y(t) = \infty$. This contradicts $y(t) < L$ for all $t \geq t_0$, and we conclude that in fact $\lim_{t\to\infty} y(t) = L$. ☐