5 Applications of the Contraction Mapping Theorem

5.1 An algorithm to compute $\sqrt{x}$

The Contraction Mapping Theorem (Theorem 4.2) not only proves the existence of a unique solution of certain equations of the form:

$$Tx = x$$

but also provides us with an algorithm to obtain arbitrarily good approximations of this solution (see the proof of Theorem 4.2). Here we apply the Contraction Mapping Theorem to a method of computing the square root of a real number.

Define a map $T : (0, \infty) \to (0, \infty)$ by

$$Ty = \frac{1}{2}(y + \frac{x}{y}) \tag{5.1}$$

where $x \in (0, \infty)$ is the number we want to compute the square root of. Clearly, $y = \sqrt{x}$ is a fixed point of T. And T has the following structure: let $0 < y_0$ be a first guess for $\sqrt{x}$, and suppose that $y_0^2 \leq x$, i.e., we guessed too small. Then,

$$\frac{x}{y_0} \geq \sqrt{x}$$

and, therefore, Ty, defined by (5.1), is the average of two approximations of $\sqrt{x}$, one that is too big and one that is too small. That should bring us closer to the correct value, and, hopefully, by iterating the process we obtain better and better approximations of $\sqrt{x}$. Before we start using this algorithm we would like to be sure that $T^n y_0 \to \sqrt{x}$. A way to to this is to set up T as a contraction on some metric space. Of course we are considering only real numbers here, and we are interested in approximations in the usual distance on the real numbers. The only remaining freedom is then the choice of a subset $I \subset \mathbb{R}$.

First, let us see whether T contracts at all: for $y_1, y_2 > 0$

$$\begin{aligned} |Ty_1 - Ty_2| &= \left| \frac{1}{2}\left(y_1 + \frac{x}{y_1}\right) - \frac{1}{2}\left(y_2 + \frac{x}{y_2}\right) \right| \\ &= \frac{1}{2}|y_1 - y_2| \left| 1 - \frac{x}{y_1 y_2} \right| \end{aligned}$$

So, T is a contraction if one necessarily has $x < 3y_1y_2$, which means we have to exclude arguments y that are too small. Therefore, we try a metric space of the form $([a,\infty), d)$, with $a > 0$. This is a complete metric space because $[a,\infty) \subset \mathbb{R}$ is closed, and $(\mathbb{R}, d)$ is complete (see Assignment 6, Problem 1).

In order to make a good choice for a we first observe that

$$Ty = \sqrt{x} + \frac{(y-\sqrt{x})^2}{2y} \geq \sqrt{x} \tag{5.2}$$

for all $y > 0$. And for $y_1, y_2 \geq \sqrt{x}$ we have $|1 - x/(y_1, y_2)| < 1$ and hence

$$|Ty_1 - Ty_2| \leq \frac{1}{2}|y_1 - y_2|$$

Therefore, the restriction of T to $[\sqrt{x}, \infty)$ is well-defined (meaning that $([\sqrt{x}, \infty)) \subset [\sqrt{x}, \infty))$, and is a nice contraction ($c = 1/2$). We conclude that for any $y_0 > 0$, $y_i = Ty_0 \in [\sqrt{x}, \infty)$ and $y_n := T^n y_0 \to \sqrt{x}$. Moreover, the convergence is exponentially fast (see the proof of Theorem 4.2):

$$\begin{aligned}
|T^n y_0 - \sqrt{x}| &\leq |T^n y_0 - T^m y_0| \\
&\leq \frac{c^n}{1-c}|Ty_0 - y_0| \\
&\leq \frac{1}{2^n}|\frac{x}{y_0} - y_0|
\end{aligned}$$

Exercise 5.1 *a) Show that for any $y > 0$, the convergence of $T^n y$ to $\sqrt{x}$ is in fact faster than exponential. Start from (5.2) to obtain a good estimate for $|Ty - \sqrt{x}|$.*
b) Take $y = 1$ and $x = 2$. Find a good estimate of how large n needs to be for $T^n 1$ and $\sqrt{2}$ to have identical first d digits. E.g., how many iterations of the map T are sufficient to compute the first 63 digits of $\sqrt{2}$?

5.2 Discrete Dynamical Systems

The term dynamical system refers to models for the evolution in time of the state of some "real life" system, as they arise in many disciplines: physics, chemistry, engineering. biology, physiology, population dynamics, economics and finance, ... In fact, they can also be an auxiliary tool for solving other problems in mathematics, then sometimes referred to as "symbolic" dynamics. A dynamical system is then defined by a state space X (often a metric space), whose elements describe the different states the system can be in, and a prescription relating the state $x_t \in X$ at time t to the state at previous (and/or subsequent times). E.g, this could be given as a differential equation, leading to a function $t \mapsto x_t$ for a suitable initial condition x_0.

The term discrete refers to the fact that time is taken to discrete, i.e., integer multiples of some time unit, i.e., we can take $t = n \in \mathbb{N}$. The dynamics is then usually defined by a map T relating the state x_{n+1} at time $t = n+1$ to the state x_n at time $t = n$:

$$x_{n+1} = Tx_n$$

If T is invertible, the dynamics can be defined for all $t = n \in Ir$.

One of the simplest and at the same time most famous dynamical systems is the *logistic map* of population dynamics. The state space is $[0, \infty) \subset \mathbb{R}$, representing the size of the population, and the dynamics is given by

$$Tx = 4\mu x(1-x)$$

where $\mu \geq 0$ is a parameter. The non-linear dependence of x_{n+1} on x_n expresses the fact that large populations lead to large off-spring, but too large populations have an adverse effect on the reproduction process.

The importance of this model is primarily that it convincingly illustrates the fact that very simple maps can have amazingly complex behavior. Analyzing the full variety of behavior of the logistic map is beyond the scope of this modest application of the Contraction Mapping Theorem. You can read more about it in : M. Schroeder, *Fractals, Chaos, Power Laws*, or R. Devaney, *An Introduction to Chaotic Dynamical Systems*.

Fixed points of the map T describe steady states (equilibrium states) of the system. If T is a contraction on a complete metric space the Contraction Mapping Theorem implies that there is a unique steady state, and that starting from any initial state, the system will evolve toward this steady state as time goes on. In the case of the logistic map we have the following.

Obviously, 0 is a fixed point of T. It is easy to show that when $\mu < 1/4$, T is a contraction on $[0, 1]$. The Contraction Mapping Theorem then tells us that for such values of μ, 0 is the unique fixed point, and that for an arbitrary initial population $x_0 \in [0, 1]$, $T^n x_0 \to 0$, i.e., the population dies out. For $\mu > 1/4$, but not too large, there is a second, non-zero, fixed point. As μ increases further, even more interesting things happen.

5.3 Integral Equations

Consider the (linear) Volterra equation for an unknown function $f : [a, b] \to \mathbb{R}$:

$$f(x) = g(x) + \int_a^b k(x,y) f(y)\, dy \tag{5.3}$$

where $g \in C([a, b])$ and $k : [0, 1] \times [0, 1] \to \mathbb{R}$, are given functions. The integral equation (5.3) is a fixed point equation, i.e., of the form $Tf = f$, for the map T defined by

$$Tf(x) = g(x) + \int_a^b k(x,y) f(y)\, dy \tag{5.4}$$

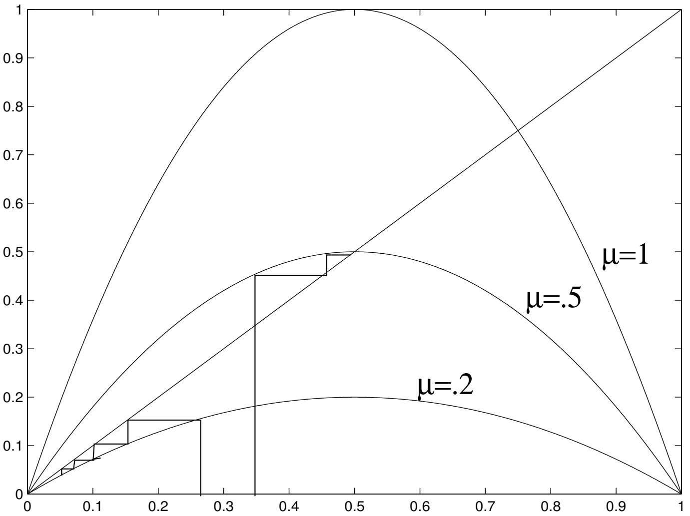

Figure 5.1: The logistic map for three different values of the parameter μ. For $\mu = .2$, 0 is the only fixed point and it is stable. For $\mu = .5$, 0 is an unstable fixed point and there is a non-zero, stable, fixed point. For $\mu = 1$, the two fixed points are unstable and more complex (and more interesting) asymptotic behavior of the dynamical system occurs.

Theorem 5.1 *If there exists a constant c satisfying*

$$\sup_{a \le x \le b} \int_a^b |k(x,y)|\, dy \le c < 1 \tag{5.5}$$

then there is a unique continuous function f satisfying (5.3).

Proof: We prove this result by showing that, when (5.5) is satisfied, the map T is a contraction on the complete metric space $(C([a,b]), |\cdot\|_\infty)$ (see Theorem 2.10). For any $f_1, f_2 \in C([a,b])$ we have

$$\begin{aligned}
\|Tf_1 - Tf_2\|_\infty &= \sup_{a \le x \le b} \left| \int_a^b k(x,y)(f_1(y) - f_2(y))\, dy \right| \\
&\le \sup_{a \le x \le b} \int_a^b |k(x,y)|\, |f_1(y) - f_2(y)|\, dy \\
&\le \|f_1 - f_2\|_\infty \sup_{a \le x \le b} \int_a^b |k(x,y)|\, dy \\
&\le c\|f_1 - f_2\|_\infty
\end{aligned}$$

■

From the proof of the Contraction Mapping Theorem we also know that, when $c < 1$, we can obtain the unique fixed point f as a limit:

$$f = \lim_{n\to\infty} T^n f_0 \tag{5.6}$$

for any $f_0 \in C([a,b])$. It is interesting to reinterprete this limit as a series. Define a map $K : C([a,b]) \to C([a,b])$ by

$$Kf = \int_a^b k(x,y)f(y)\,dy$$

One has that $\|K^n f_0\|_\infty \le c^n |f_0\|_\infty \to 0$. Then, (5.3) becomes $f = g + Kf$, or

$$(\mathbb{1} - K)f = g \tag{5.7}$$

where $\mathbb{1}$ is the identity map, i.e., $\mathbb{1}f = f$. Solving (5.7) for f, given an arbitrary g, is equivalent to inverting the map $\mathbb{1} - K$, and (5.6) can be reinterpreted as the solution of this problem:

$$f = (\mathbb{1} - K)^{-1}g = \sum_{n=0}^{\infty} K^n g \tag{5.8}$$

This series is called the *Neumann series*, and using its partial sums to approximate f is often called the *Born approximation*. Note the similarity of (5.8) with the geometric series $(1-x)^{-1} = \sum_{n=0}^\infty x^n$, for $|x| < 1$. In fact, we will see later that it *is* a geometric series that is absolutey convergent for a natural norm on the linear operators on a Banach space. This also explains why we only needed a condition on k, and none on g: (5.5) is the condition needed for $\mathbb{1} - K$ to be invertible, which does not involve g.

5.4 Differential Equations (Boundary Value Problems)

Consider a copper rod or pipe, say of length 1 in some units, and partly insulated. The ends of the rod are kept at constant temperatures T_0 and T_1. Denote by x the coordinate along the length of the pipe ($x \in [0,1]$), and let t be time. We are interested in the temperature $u(x,t)$ as it varies along the rod, and with time. To a reasonable degree of accuracy $u(x,t)$ satisfies

$$u_t = u_{xx} - q(x)u \tag{5.9}$$

where $q(x)$ is a given function describing the cooling rate at x, and

$$u_t(x,t) = \frac{\partial}{\partial t}u(x,t) \tag{5.10}$$

$$u_{xx}(x,t) = \frac{\partial^2}{\partial x^2}u(x,t) \tag{5.11}$$

is a standard notation for partial derivatives. Of course, this equation does not uniquely determine u by itself. It has to be supplemented with an inital condition $u_0(x)$, and boundary conditions, e.g., $u(0,t) = T_0, u(1,t) = T_1$, for all t. If one is interested in time independent solutions $u(x,t) = u(x)$, e.g., describing the situation as $t \to \infty$, one has to solve the following ODE (put $u_t = 0$ in the PDE (5.9)) + boundary conditions:

$$-u''(x) + q(x)u(x) = 0 \tag{5.12}$$
$$u(0) = T_0, \quad u(1) = T_1 \tag{5.13}$$

If $q(x)$ is a constant function, (5.12) is easy to solve explicitly. But if q is not constant explicit integration of this equation is in general impossible. We will show that, when $\|q\|_\infty$ is not too large, the Contraction Mapping Theorem implies the existence of a unique solution, and provides a method for calculating approximations of that solution.

First, it is convenient to get rid of the inhomogeneous boundary conditions (5.13), by deriving an equation for a function v related to u by

$$u(x) = T_0 + (T_1 - T_0)x + v(x)$$

v then satisfies

$$v''(x) = q(x)v(x) + (T_0 + (T_1 - T_0)x)q(x) \tag{5.14}$$
$$v(0) = 0, \quad v(1) = 0 \tag{5.15}$$

Next, we will turn this problem into a fixed point problem of the Volterra type with the aid of the following proposition.

Proposition 5.2 *Let $f : [0,1] \to \mathbb{R}$ be a continuous function, then the unique solution v of*

$$-v'' = f \tag{5.16}$$
$$v(0) = v(1) = 0 \tag{5.17}$$

can be given as

$$v(x) = \int_0^1 g(x,y)f(y)\,dy \tag{5.18}$$

where

$$g(x,y) = \begin{cases} x(1-y) & \text{if } 0 \le x \le y \le 1 \\ y(1-x) & \text{if } 0 \le y \le x \le 1 \end{cases} \tag{5.19}$$

We postpone the proof of this proposition to the end of this section and first see where it takes us. Using this proposition, (5.14) can be written in the form

$$v(x) = -\int_0^1 g(x,y)q(y)v(y)\,dy - \int_0^1 g(x,y)(T_0 + (T_1 - T_0)y)q(y)\,dy$$

which is of the form

$$(\mathbb{1} - K)v = h$$

with

$$Kv(x) = -\int_0^1 \underbrace{g(x,y)q(y)}_{:=k(x,y)} v(y)\,dy \tag{5.20}$$

$$h(x) = -\int_0^1 g(x,y)(T_0 + (T_1 - T_0)y)q(y)\,dy \tag{5.21}$$

We can now apply Theorem 5.1 to obtain the following result.

Theorem 5.3 *If*

$$\sup_{0\le x\le 1} \int_0^1 |g(x,y)q(y)|\,dy < 1 \tag{5.22}$$

the boundary value problem (5.12)-(5.13) has a unique continuous solution.

Note that it is not necessary that q be continuous, as long as h is and the condition (5.22) is satisfied.

We still have to prove Proposition 5.2.

Proof: (Proposition 5.2)
By integrating (5.16) twice we obtain:

$$v(x) = -\int_0^x \underbrace{\int_1^y f(s)\,ds}_{:=a(y)}\, dy + C_1 x + C_2$$

where C_1 and C_2 are two real constants to be determined later. Integration by parts of the integral over y in the RHS gives

$$\begin{aligned} v(x) &= -[ya(y)]_0^x + \int_0^x yf(y)\,dy + C_1x + C_2 \\ &= x\int_x^1 f(y)\,dy + \int_0^x yf(y)\,dy + C_1x + C_2 \end{aligned}$$

We can determine C_1 and C_2 by imposing (5.17). One finds

$$\begin{aligned} C_1 &= -\int_0^1 y f(y)\, dy \\ C_2 &= 0 \end{aligned}$$

It follows that

$$v(x) = \int_0^x y(1-x)f(y)\, dy + \int_x^1 x(1-y)f(y)\, dy$$

which is what we had to prove. ■

5.5 Differential Equations (Existence and Uniqueness Theorem)

Another important application of Contraction Mapping Theorem to differential equations is an Existence and Uniqueness Theorem for initial value problems of the type

$$\begin{cases} \dot{x}\ (t) = F(t, x(t)) \\ x(t_0) = x_0 \end{cases} \tag{5.23}$$

with F a Lipschitz continuous function on a suitable domain. Here, x takes values in $\mathbb{R}^n$, and so do its derivative and F. (5.23) is often presented as a set of coupled ODE's, e.g., describing a dynamical system. (2.24) can be reformulated as an integral equation:

$$x(t) = x_0 + \int_{t_0}^t F(t, x(t))\, dt \tag{5.24}$$

By the fundamental theorem of calculus the continuous solutions of (5.23) and (5.24) coincide. (5.24) is easily written as a fixed point equation for the map T defined by

$$Tx(t) = x(t) = x_0 + \int_{t_0}^t F(s, x(s))\, ds \tag{5.25}$$

We are then looking for solutions of

$$Tx = x \tag{5.26}$$

If we can find sufficient conditions that guarantee that T is a contraction on a suitable complete metric space of continuous functions, the Contraction Mapping Theorem immediately implies that (5.26) has a unique solution in that space.

Theorem 5.4 *Let $I \subset \mathbb{R}$ be a (possibly infinite) interval, and t_0 a point in the interior. Suppose $F : I \times \mathbb{R}^n \to \mathbb{R}^n$ is continuous in both arguments and uniformly Lipschitz continuous in its second argument, i.e., suppose there exists a constant $C > 0$ such that for all, $t \in I$, and $x, y \in \mathbb{R}^n$,*

$$\|F(t, x) - F(t, y)\| \leq C\|x - y\| \tag{5.27}$$

Then, there is a unique continuous function $x : I \to \mathbb{R}^n$ solving (5.23).

Proof: As mentioned above, our strategy is to show that T, defined in (5.25), is a contraction on a suitable metric space of continuous functions. Note that T depends on t_0 and t. It will be necessary to first consider solutions for sufficiently short times, i.e., $t \in [t_0, t_0 + \delta] \subset I$, with $\delta > 0$ to be determined later. Let $f, g : [t_0, t_0 + \delta] \to \mathbb{R}^n$, be two continuous functions. Then

$$\begin{aligned}
\|Tf - Tg\|_{\sup} &= \sup_{t_0 \leq t \leq t_0 + \delta} |Tf(t) - Tg(t)| \\
&= \sup_{t_0 \leq t \leq t_0 + \delta} \left| \int_{t_0}^{t} F(s, f(s)) - F(s, g(s))\, ds \right| \\
&\leq \sup_{t_0 \leq t \leq t_0 + \delta} \int_{t_0}^{t} |F(s, f(s)) - F(s, g(s))|\, ds \\
&\leq \sup_{t_0 \leq t \leq t_0 + \delta} \int_{t_0}^{t} C|f(s) - g(s)|\, ds \\
&\leq \delta C \|f - g\|_{\sup}
\end{aligned}$$

where we have used the definition of T and the Lipschitz condition (5.27). Now, it is clear that T is a contraction on $C([t_0, t_0 + \delta])$ with the supremum norm as long as $|t - t_0| \leq \delta < 1/C$. Therefore, there is a unique solution $x : [t_0, t_0 + \delta] \to \mathbb{R}^n$. The argument holds for *any* $t_0 \in I$. By covering I with overlapping intervals of length $< 1/C$, we see that (5.23) has a unique continuous solution defined on all of I. ■

Note that, due to (5.24), the solution is in fact C^1 (continously differentiable). In fact, F need not be continuous in t to have a unique solution, but one looses differentiability of the solution at the times where F is not continuous. The condition (5.27) can also be relaxed to hold with a time dependent Lipschitz constant $C(t)$, as long as it is bounded on compact intervals.

Examples:

1. Linear ODE with continuous coefficients have unique global solutions. It is sufficient to consider *first order* ODE's (Why?):

$$\begin{aligned}
\dot{x}(t) &= a(t)x(t) + b(t) \\
x(0) &= x_0
\end{aligned}$$

where $a : \mathbb{R} \to \mathbb{R}^{n\times n}$ is a continuous matrix-valued function (with respect to any norm on the $n\times n$ matrices with real entries), and $b : \mathbb{R} \to \mathbb{R}^n$ is continuous as well. Then, for any bounded interval $I \subset \mathbb{R}$, there exists a constant C such that

$$\|a(t)v\| \leq C\|v\|, \quad \text{for all } t \in I, v \in \mathbb{R}^n$$

and it is then obvious that

$$\|a(t)f(t) - a(t)g(t)\|_{\sup} \leq C\|f - g\|_{\sup}$$

Therefore, the hypotheses of Theorem 5.4 are satisfied and we have a unique continuous solution on I. As I was an arbitrary interval, we can conclude that there exists a unique continuous solution for all $t \in \mathbb{R}$.

2. The applications of Theorem 5.4 are by no means limited to linear ODE. E.g., consider the nonlinear scalar ODE given by

$$\begin{aligned} \dot{x}\ (t) &= \sqrt{a(t)^2 + x(t)^2} =: F(t, x(t)) \\ x(0) &= x_0 \end{aligned}$$

where $a : \mathbb{R} \to \mathbb{R}$ is a continuous function. It is not hard to prove that

$$|F(t, x) - F(t, y)| \leq |x - y|, \quad \text{for all } t, x, y, \in \mathbb{R}$$

This bound is obvious for $a^2 = 0$, and the LHS is decreasing as a function of a^2. Again we can conclude by Theorem 5.4 that there is a unique global solution of this ODE.

3. Nonlinear F can easily be continuous without being uniformly Lipschitz continuous. A simple example of this situation shows that global solutions may not exist in this case.

$$\begin{aligned} \dot{x} &= x^2 \\ x(0) &= 1 \end{aligned}$$

It is easy to check that

$$x(t) = \frac{1}{1-t}, \quad t \in [0, 1)$$

is a solution. The singularity at $t = 1$ is possible because the RHS is not uniformly Lipschitz continuous:

$$|x^2 - y^2| = |x + y|\ |x - y|$$

and $|x + y|$ can become arbitrarily large. In this simple example it is easy to understand that the quadratic RHS makes x grow faster the larger it becomes (positive, nonlinear feedback). This leads to blowup of the solution in finite time, at which point the equation is not well-defined.