5 Fixed Points and Perturbation

Problems Related to Fundamental Definitions

11/10: **1.** Are the following real functions Lipschitz continuous near 0? If yes, find a Lipschitz constant for some interval containing 0.

(1) $1/(1-x^2)$.

Proof. $\boxed{\text{Yes.}}$ Consider the interval $[-0.5, 0.5]$. Then we may take

$$\boxed{L = \frac{16}{9}}$$

□

(2) $x \log |x|$.

Proof. $\boxed{\text{No.}}$ □

(3) $x^2 \sin(1/x)$.

Proof. If we take the piecewise function consisting of the above expression on $\mathbb{R} \setminus \{0\}$ and 0 at 0, then $\boxed{\text{yes.}}$ Consider the interval $[-1, 1]$. Then we may take

$$\boxed{L = 2}$$

□

2. Find the first two elements $y_1(t), y_2(t)$ for the Picard iteration sequence of the following initial value problems, and estimate the error between $y_2(t)$ and the actual solution. Since they are all of separable form, the actual solutions can be explicitly found.

(1) $y' = 1 + y^2$, $y(0) = 0$.

Proof. We take $y_0(t) = 0$. Then

$$\begin{aligned} y_1(t) &= y_0(0) + \int_0^t [1 + y_0(t)^2] \, \mathrm{d}t \\ &= \int_0^t [1 + 0] \, \mathrm{d}t \\ \boxed{y_1(t) = t} & \end{aligned}$$

and

$$\begin{aligned} y_2(t) &= y_0(0) + \int_0^t [1 + y_1(t)^2] \, \mathrm{d}t \\ &= \int_0^t [1 + t^2] \, \mathrm{d}t \\ \boxed{y_2(t) = t + \frac{t^3}{3}} & \end{aligned}$$

The error is between y_2 and the actual solution $y(t) = \tan(t)$ is given by

$$\boxed{\varepsilon = \tan(t) - t - \frac{t^3}{3}}$$

□

(2) $y' = 2ty$, $y(0) = 1$.

Proof. We take $y_0(t) = 1$. Then

$$\begin{aligned} y_1(t) &= y_0(0) + \int_0^t 2ty_0(t)\,\mathrm{d}t \\ &= 1 + \int_0^t 2t\,\mathrm{d}t \\ \boxed{y_1(t) = 1 + t^2} \end{aligned}$$

and

$$\begin{aligned} y_2(t) &= y_0(0) + \int_0^t 2ty_1(t)\,\mathrm{d}t \\ &= 1 + \int_0^t [2t + 2t^3]\,\mathrm{d}t \\ \boxed{y_2(t) = 1 + t^2 + \frac{t^4}{2}} \end{aligned}$$

The error is between y_2 and the actual solution $y(t) = \mathrm{e}^{t^2}$ is given by

$$\boxed{\varepsilon = \mathrm{e}^{t^2} - 1 - t^2 - \frac{t^4}{2}}$$

□

(3) $y' = y/(1-t)$, $y(0) = 1$.

Proof. We take $y_0(t) = 1$. Then

$$\begin{aligned} y_1(t) &= y_0(0) + \int_0^t \frac{y_0(t)}{1-t}\,\mathrm{d}t \\ &= 1 + \int_0^t \frac{1}{1-t}\,\mathrm{d}t \\ \boxed{y_1(t) = 1 - \ln|1-t|} \end{aligned}$$

and

$$\begin{aligned} y_2(t) &= y_0(0) + \int_0^t \frac{y_1(t)}{1-t}\,\mathrm{d}t \\ &= 1 + \int_0^t \frac{1 - \ln|1-t|}{1-t}\,\mathrm{d}t \\ \boxed{y_2(t) = 1 - \ln|1-t| + \frac{1}{2}(\ln|1-t|)^2} \end{aligned}$$

The error between y_2 and the actual solution $y(t) = \mathrm{e}^{-\ln|1-t|}$ is given by

$$\boxed{\varepsilon = \mathrm{e}^{-\ln|1-t|} - 1 + \ln|1-t| - \frac{1}{2}(\ln|1-t|)^2}$$

□

3. Check whether the implicit equation $F(x, y) = 0$ uniquely determines an explicit function $y = f(x)$ around the given point (x_0, y_0). If it does, compute $f'(x_0)$.

(1) For $(x, y) \in \mathbb{R}^2$, $F(x, y) = x^2 + y^2 - 1$, $(x_0, y_0) = (\sqrt{2}/2, -\sqrt{2}/2)$.

Proof. From the implicit equation, we have that

$$\begin{aligned} 0 &= x^2 + y^2 - 1 \\ y &= \pm\sqrt{1 - x^2} \end{aligned}$$

Since

$$\begin{aligned} -\frac{\sqrt{2}}{2} &= -\sqrt{1 - \left(\frac{\sqrt{2}}{2}\right)^2} \\ y_0 &= -\sqrt{1 - x_0^2} \end{aligned}$$

our explicit function $\boxed{\text{is uniquely determined around } (x_0, y_0).}$

Moreover, we can compute that

$$\begin{aligned} f'(x_0) &= \frac{2x_0}{2\sqrt{1 - x_0^2}} \\ &\boxed{f'(x_0) = 1} \end{aligned}$$

□

(2) For $(x, y) \in \mathbb{R}^2$, $F(x, y) = x^2 - y^2 - 1$, $(x_0, y_0) = (1, 0)$.

Proof. From the implicit equation, we have that

$$\begin{aligned} 0 &= x^2 - y^2 - 1 \\ y &= \pm\sqrt{x^2 - 1} \end{aligned}$$

Since

$$y_0 = \sqrt{x_0^2 - 1} \qquad\qquad y_0 = -\sqrt{x_0^2 - 1}$$

our explicit function $\boxed{\text{is not uniquely determined around } (x_0, y_0).}$ □

(3) For $(x, y) \in \mathbb{R}^2$, $F(x, y) = x\mathrm{e}^y + y$, $(x_0, y_0) = (0, 0)$.

Proof. We apply the implicit function theorem.

F is defined on a subset of $\mathbb{R}^2$, as desired.

We have that

$$\frac{\partial F}{\partial x} = \mathrm{e}^y \qquad\qquad \frac{\partial F}{\partial y} = x\mathrm{e}^y + 1$$

Since both of the above partial derivatives are continuous, F is continuously differentiable on its domain, as desired.

$(x_0, y_0) = (0, 0) \in \mathbb{R}^2$, which is the domain of F, as desired.

$F(x_0, y_0) = 0\mathrm{e}^0 + 0 = 0$, as desired.

The truncated Jacobian matrix is 1×1 and contains a nonzero element at (x_0, y_0) — in particular, it contains $\partial F/\partial x$ — as desired.

Therefore, our explicit function $\boxed{\text{is uniquely determined around } (x_0, y_0).}$

Moreover, we can compute that

$$\begin{aligned}
f'(x_0) &= -\left(\frac{\partial F}{\partial y}\right)^{-1} \cdot \frac{\partial F}{\partial x} \\
&= -\left(0\mathrm{e}^0 + 1\right)^{-1} \cdot \mathrm{e}^0 \\
\boxed{f'(x_0) = -1}
\end{aligned}$$

□

Problems Involving the Banach Fixed Point Theorem

1. (1) Show that the condition "constant $q < 1$" in the statement of the Banach fixed point theorem is not redundant. You may give an example of a function $f : \mathbb{R} \to \mathbb{R}$ which satisfies the strict inequality $|f(x) - f(y)| < |x - y|$ but does not have a fixed point.

Proof. Choose

$$\boxed{f(x) = \begin{cases} 1 & x \leq 0 \\ x + \mathrm{e}^{-x} & x > 0 \end{cases}}$$

The fact that

$$\frac{\mathrm{d}f}{\mathrm{d}x} = \begin{cases} 0 & x \leq 0 \\ 1 - \mathrm{e}^{-x} & x > 0 \end{cases}$$

implies that $|\,\mathrm{d}f/\mathrm{d}x\,| < 1$ for all x. Hence, f satisfies the desired strict inequality. Additionally, since the graph of $f(x) > x$ for all x (as can be readily verified from its definition), it has no fixed point, as desired. □

(2) Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be a Lipschitz mapping with uniform Lipschitz constant $q < 1$, that is,

$$|f(x) - f(y)| \leq q|x - y|$$

for all $x, y \in \mathbb{R}^n$. Prove that the mapping $x \mapsto x + f(x)$ is invertible with Lipschitz continuous inverse.

Proof. Let $g : \mathbb{R}^n \to \mathbb{R}^n$ be defined by $g(x) = x + f(x)$. To prove that g is invertible, it will suffice to show that g is one-to-one, that is, for every $b \in \mathbb{R}^n$, there exists a unique $a \in \mathbb{R}^n$ such that $g(a) = b$. Let $b \in \mathbb{R}^n$ be arbitrary. Define $h : \mathbb{R}^n \to \mathbb{R}^n$ by $h(x) = b - f(x)$. Then since

$$\begin{aligned}
|h(x) - h(y)| &= |[b - f(x)] - [b - f(y)]| \\
&= |f(y) - f(x)| \\
&= |f(x) - f(y)| \\
&\leq q|x - y|
\end{aligned}$$

we have by the Banach fixed point theorem that there exists a unique $a \in \mathbb{R}^n$ such that $a = h(a)$. It follows that

$$\begin{aligned}
a &= b - f(a) \\
a + f(a) &= b \\
g(a) &= b
\end{aligned}$$

as desired.

To prove that g^{-1} is Lipschitz continuous, it will suffice to show that

$$|g^{-1}(x) - g^{-1}(y)| \leq \frac{1}{1-q}|x-y|$$

for all $x, y \in \mathbb{R}^n$. Let $x, y \in \mathbb{R}^n$ be arbitrary. Define $a = g^{-1}(x)$ and $b = g^{-1}(y)$. Then since the first term below is nonnegative (as the product of two nonnegative numbers), we have that

$$\begin{aligned}
(1-q)|a-b| &= |a-b| - q|a-b| \\
&\leq |a-b| - |f(a) - f(b)| \\
&= |a-b| - |f(b) - f(a)| \\
&= \big||a-b| - |f(b) - f(a)|\big| \\
&\leq |[a-b] - [f(b) - f(a)]| \\
&= |[a + f(a)] - [b + f(b)]| \\
&= |g(a) - g(b)|
\end{aligned}$$

It follows by returning the substitution that

$$\begin{aligned}
(1-q)|g^{-1}(x) - g^{-1}(y)| &\leq |x-y| \\
|g^{-1}(x) - g^{-1}(y)| &\leq \frac{1}{1-q}|x-y|
\end{aligned}$$

as desired. □

2. Consider the following iterative algorithm to compute the square root of a given $a > 1$.

$$x_{n+1} = \frac{1}{2}\left(x_n + \frac{a}{x_n}\right)$$

(1) Show that the function

$$F(x) = \frac{1}{2}\left(x + \frac{a}{x}\right)$$

meets the requirements of the contraction mapping principle on the closed interval $[\sqrt{a/2}, a]$. Prove that $x_n \to \sqrt{a}$.

Proof. We want to show that

$$|F(x) - F(y)| \leq q|x-y|$$

for some $q \in (0,1)$ and all $x, y \in [\sqrt{a/2}, a]$.

We have that

$$\begin{aligned}
|F(x) - F(y)| &= \left|\frac{1}{2}\left(x + \frac{a}{x}\right) - \frac{1}{2}\left(y + \frac{a}{y}\right)\right| \\
&= \frac{1}{2}\left|(x-y) + \left(\frac{a}{x} - \frac{a}{y}\right)\right| \\
&= \frac{1}{2}\left|(x-y) + a \cdot \frac{y-x}{xy}\right| \\
&= \frac{1}{2}\left|\left(1 - \frac{a}{xy}\right)(x-y)\right| \\
&= \frac{1}{2}\left|1 - \frac{a}{xy}\right||x-y|
\end{aligned}$$

□

(2) For $a = 2$, start the iteration $x_{n+1} = F(x_n)$ with $x_0 = 1$. Use a calculator to compute the first 10 values of this iteration, up to 11 digits after the decimal point. Compare it with the exponentially converging sequence $1.4, 1.41, 1.414, 1.4142, \ldots$. Which of the two algorithms is better?

Proof. We have that

$$\boxed{\begin{aligned} x_0 &= 1 \\ x_1 &= 1.5 \\ x_2 &= 1.41666666667 \\ x_3 &= 1.41421568627 \\ x_4 &= 1.41421356237 \\ x_5 &= 1.41421356237 \\ x_6 &= 1.41421356237 \\ x_7 &= 1.41421356237 \\ x_8 &= 1.41421356237 \\ x_9 &= 1.41421356237 \\ x_{10} &= 1.41421356237 \end{aligned}}$$

$\boxed{\text{The algorithm from part (1) is better.}}$ □

(3) Try to estimate the error $|x_n - \sqrt{a}|$ as well as possible. *Hint.* There should be something related to an iterative sequence $\{b_n\}$ satisfying

$$b_{n+1} \leq M b_n^2$$

You should prove that the sequence converges to zero faster than any geometric progression.

Context: This algorithm is referred to as **Newton's method**. It is a rapidly converging algorithm to find zeros/fixed points of functions, capable of giving very precise approximations within very few steps. A variation of it, called the **Nash-Moser technique**, is a very powerful tool for proving the existence of solutions to nonlinear differential equations.

3. In this question, we aim to prove that a certain differential equation admits a unique periodic solution. Let $f(t)$ be a smooth, real-valued function of period 1, that is, $f(t+1) \equiv f(t)$.

(1) Fix $a > 0$. Prove that there exists only one $x \in \mathbb{R}$ such that the solution of the initial value problem

$$y' + ay = f(t), \quad y(0) = x$$

has period 1, and express x in terms of the integral of f.

(2) Now let $g(w)$ be some smooth function such that $g'(0) = 0$. Prove the following nonlinear perturbative result: For any period 1 continuous function f close to 0, there is only one $x \in \mathbb{R}$ with small magnitude such that the solution of the perturbed initial value problem

$$y' + ay + g(y) = f(t), \quad y(0) = x$$

has period 1. *Hint*: Write the solution as $y(t; x)$ and consider the function $x \to y(1; x)$. Try to apply the continuous dependence result discussed in class to conclude that this function is a contraction near $0 \in \mathbb{R}$.

Problems Related to the Calculation of Perturbation

1. Determine the approximate solution of the following initial value problems up to order μ, where μ is the small perturbative parameter.

(1)

$$y'' + (1+\mu)y = 0, \quad y(0) = a, \quad y'(0) = b$$

This is in fact the harmonic oscillator equation. Compare your result with the actual solution, and estimate the time beyond which the approximate solution fails to match it well.

(2)

$$y' = ry - \mu y^2, \quad y(0) = a$$

This is in fact the logistic growing model with $\mu = r/M \ll 1$, M being the capacity. Compare your result with the actual solution of the logistic differential equation, and estimate the time beyond which the approximate solution fails to match it well.

It should be stressed again that these approximate solutions only work for a fixed time interval, and usually fail to approximate the actual solution well for large t.

2. Let h be a small real number. Consider the function

$$f(x) = \cos x + hx \sin x$$

We know that when $h = 0$, the point $x = 2\pi$ is a maximum point of $f(x)$. If $h \neq 0$, the extrema of f cannot be explicitly determined. However, determine the first positive maximum point $x_1(h)$ of f as a function of h, up to order h. *Hint*: $x_1(h)$ is the zero of $f'(x)$ with $x_1(0) = 2\pi$, and you may apply the implicit function theorem.

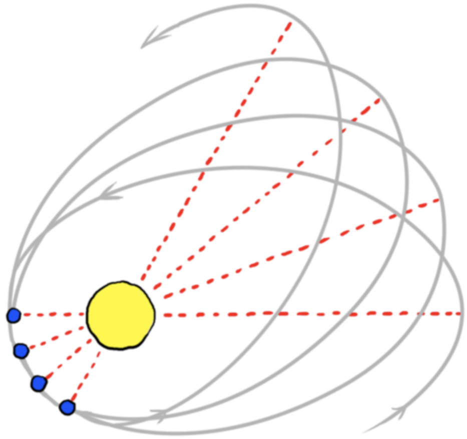
3. Let $A, B > 0$ be fixed constants, μ be a real number of small magnitude compared to A, B. Determine the approximate solution of the following differential equation up to order μ.

$$\frac{\mathrm{d}^2 u}{\mathrm{d}\varphi^2} + u = A + \mu u^2, \quad u(0) = A + B, \quad u'(0) = 0$$

Hint: For the zeroth order approximation, the solution can be easily guessed without using the Duhamel formula.

Bonus Problem

We have already determined that the orbits of the two-body problem under Newtonian gravity are conic sections in the center of mass frame. However, in actual physical situations, there are multiple effects that lead to correction of the orbit. For Mercury, the closest known planet to the Sun, this correction is much easier to observe than with other planets. These effects break the closedness of the ideally elliptic orbit of Mercury, resulting in very slow precession of the perihelion of the orbit (see the picture, which is very much exaggerated).



The total observed precession of the perihelion of Mercury is about 5600.73 arcsec per century. After deducing the effect of axial precession of the Earth itself, the causes of precession explainable by classical Newtonian gravity theory are shown below.

Amount, arcsec per century	Cause
532.3035	Gravity of other solar bodies
0.0286	Quadrupole moment of the sun
-0.0020	Lense-Thirring precession
532.3301	Total of classical effects

The observed total after deducing the axial precession of the Earth is, however, 574.10 ± 0.65 arcsec per century. There are about 42-43 arcsec per century that cannot be explained by any of the classical effects listed above. This anomalous precession was first observed by Urbain Le Verrier, who assumed that there should be another planet evolving inside the orbit of Mercury. However, his computation did not match with the observation of other planets. Celestial physicists came up with several other assumptions but none of them gave a result consistent with this case. It was only after Albert Einstein's discovery of general relativity that the correction-matching observation was finally obtained. Einstein computed the distortion of Kepler orbits under a special solution, called the **Schwarzchild solution**, of the relativistic gravitational equation, and his result matches very much for Mercury and for light rays from stars far away in space. These two cases are considered to be the initial powerful experimental supports for general relativity.

In this problem, you will be able to reproduce Einstein's result. The physics has been reduced to a minimum and you will find that Einstein's core mathematical argument is a pure perturbative problem for ODEs.

In Schwarzchild spacetime, we assume that there is only one static star (the Sun) whose gravitational field is present. The mass of the star is assumed to be M, and the spacetime is assumed to be **spherically symmetric**. The spacetime is parameterized by four coordinates: t, r, θ, φ. The t coordinate can be imagined as *time*, and r, θ, φ can be imagined as the spherical coordinates. We describe the "distance" between spacetime points using the Schwarzchild metric

$$\mathrm{d}s^2 = -\left(1 - \frac{r_s}{r}\right) c^2 \,\mathrm{d}t^2 + \left(1 - \frac{r_s}{r}\right)^{-1} \mathrm{d}r^2 + r^2 \left(\mathrm{d}\theta^2 + \sin^2\theta \,\mathrm{d}\varphi^2\right)$$

where c is the speed of light in a vacuum, G is the gravitational constant, and $r_s = 2GM/c^2$ is the **Schwarzchild radius**.

The trajectory of a planet in the spacetime is a curve parameterized by the parameter $\tau \in \mathbb{R}$ called the **proper time of the planet**. The orbit is a **geodesic** (which is a geometric object, but the definition does not really matter here) in four-dimensional spacetime. Some considerations in symmetry enable us to assume that $\theta \equiv \pi/2$ along the trajectory ("in the ecliptic plane"). Differential geometry guarantees that the spacetime inner product along the orbit should be a constant, i.e.,

$$-c^2 = -\left(1 - \frac{r_s}{r}\right) c^2 \left(\frac{\mathrm{d}t}{\mathrm{d}\tau}\right)^2 + \left(1 - \frac{r_s}{r}\right)^{-1} \left(\frac{\mathrm{d}r}{\mathrm{d}\tau}\right)^2 + r^2 \left(\frac{\mathrm{d}\varphi}{\mathrm{d}\tau}\right)^2 \tag{5.1}$$

Further considerations in symmetry guarantee that along the trajectory, there are two conserved quantities. These are the relativistic counterparts of total energy and angular momentum.

$$\left(1 - \frac{r_s}{r}\right) \frac{\mathrm{d}t}{\mathrm{d}\tau} = \frac{E}{mc^2} \qquad\qquad mr^2 \frac{\mathrm{d}\varphi}{\mathrm{d}\tau} = L \tag{5.2}$$

Here, m is the mass of the planet, which is very small compared to M.

1. We are now interested in the shape of the orbit that we observe as a curve in the ecliptic plane, which can be parameterized as a polar coordinate equation $r = r(\varphi)$. Equations 5.1-5.2 should imply a differential equation satisfied by $r = r(\varphi)$. Find that differential equation.

2. Introducing $u = 1/r$ and differentiating that equation with respect to φ again, you should arrive at an equation of identical form to that of Problem 3.3, i.e., of the form

$$\frac{\mathrm{d}^2 u}{\mathrm{d}\varphi^2} + u = A + \mu u^2$$

It can be considered to be a perturbation of the equation presented in HW2 that gives the Kepler orbit. Determine the A and μ here in terms of the physical constants G, M, m, c, E, L. *Hint*: You need to cancel the term $(\mathrm{d}t/\mathrm{d}\tau)^2$ in Equation 5.1 and try to express $(\mathrm{d}r/\mathrm{d}\varphi)^2$ in terms of r alone, just as in HW2. This will be the perturbative problem that we are interested in.

3. Assuming that μ is small, the equation in the previous problem is a perturbation of the Kepler orbit equation. Thus, it is natural to choose the ray joining the Sun and the perihelion as the polar axis so that we can fix the initial conditions as

$$u(0) = p, \quad u'(0) = 0$$

where p is the distance from the perihelion to the Sun. The perihelion of the planet corresponds to the maximum of u in the previous problem. Determine the first value of φ at which u has a local maximum, up to order μ. Its derivation with 2π is the **precession angle**, as the non-periodicity of the solution breaks the closedness of the orbit.

Hint: This is in fact Problem 3.2. Your expression should not involve p.

Substitute in the following physical constants: $G = 6.673 \times 10^{-11}\,\mathrm{kg\,m^3\,s^{-2}}$, $M = 1.989 \times 10^{30}\,\mathrm{kg}$, $c = 2.998 \times 10^{8}\,\mathrm{m\,s^{-1}}$. Also, the mass of Mercury is $m = 3.301 \times 10^{23}\,\mathrm{kg}$, the angular momentum of Mercury is $L = 8.983 \times 10^{38}\,\mathrm{kg\,m^2\,s^{-1}}$, and the revolution period of Mercury is 88 days. Use a calculator to compute the precession angle and convert its unit to arcsec per century (the expression you obtain is, of course, in radians per revolution; also recall that $1\,\mathrm{arcsec} = 1^\circ/3600$), and compare it with the unexplained precession.