January 2013 Analysis Prelim: An Exegesis

S. E. Payne

2 February 2013

1 Prolegomena

A prelim exam in real analysis was given on 18 January 2013 for Ph.D. students in the Department of Mathematical and Statistical Sciences of the University of Colorado Denver. In due course this exam along with brief sketches of solutions of the problems will be posted on the Department's web page. However, in discussing some of the solutions with students after the exam, it occurred to me that it would be helpful to use the recent exam as a guide to reviewing a significant bit of our curriculum in real analysis. So the present notes give a more nearly complete treatment of the ideas used in solving the problems on this exam.

2 Problem 1

.

Let $f_n(x) = (-1)^n \frac{x^n}{n}$.

(a) Show that $\sum_{n=1}^{\infty} f_n(x)$ converges uniformly on $[0, 1]$.

(b) Show that$\sum_{n=1}^{\infty} |f_n(x)|$ converges pointwise on $[0, 1)$.

(c) Show that$\sum_{n=1}^{\infty} |f_n(x)|$ does not converge uniformly on $[0, 1)$.

Solution: For part (a) the alternating series test provides a simple solution. But we need the complete theorem.

Theorem 2.1. *(Leibniz's Rule for Alternating Series) If $\{a_n\}$ is a monotonic decreasing sequence with limit 0, then the alternating series $\sum_{n=1}^{\infty}(-1)^{n+1}a_n$ converges. If S denotes its sum and s_n denotes its nth partial sum, we also have the inequalities*

$$0 < (-1)^n(S - s_n) < a_{n+1} \tag{1}$$

Note: The first inequality tells us that the error, $S - s_n$, has the sign $(-1)^n$, which is the same as the sign of the first neglected tern, $(-1)^n a_{n+1}$. The second inequality states that the absolute value of this error is less than that of the first neglected term.

Proof. Start with $s_n = \sum_{k=1}^{n}(-1)^{k+1}a_k$. Then $s_{2k+2} - s_{2k} = a_{2k+1} - a_{2k+2} > 0$. This says that the partial sums s_{2k} form an increasing subsequence. Similarly, $s_{2k+1} - s_{2k-1} = -(a_{2k} - a_{2k+1}) < 0$. This says that the partial sums s_{2k-1} form a decreasing subsequence. So the sequence $\{s_{2k+1}\}$ is clearly bounded above by s_1, but so is the sequence of even numbered partial sums: $s_{2r+2} = a_1 - \sum_{k=1}^{r}(a_{2k} - a_{2k+1}) - a_{2r+2}$. Similarly, both sequences $\{s_{2k}\}$ and $\{s_{2k-1}\}$ are bounded below by s_2. $(s_{2k+1} = a_1 - \sum_{t=1}^{k}(a_{2t} - a_{2t+1})$. Therefore, each sequence $\{a_{2k}\}$ and $\{a_{2k-1}\}$, being bounded and monotonic, converges to a limit. Say $a_{2k} \to S'$ and $a_{2k-1} \to S''$. But

$$S' - S'' = \lim_{n\to\infty} s_{2k} - \lim_{n\to\infty} s_{2k-1} = \lim_{n\to\infty}(s_{2k} - s_{2k-1}) = \lim_{n\to\infty}(-a_{2k} = 0.$$

Hence $S' = S''$, and we denote the common limit by S. Since s_{2k} is increasing and s_{2k-1} is decreasing, we have

$$s_{2k} < s_{2k+2} \leq S \text{ and } S \leq s_{2k+1} < s_{2k-1} \text{ for all } k \geq 1. \tag{2}$$

So $-S > -s_{2k}$ and we have

$$0 < S - s_{2k} \leq s_{2k+1} - s_{2k} = a_{2k+1} \implies |S - s_{2k}| \leq a_{2k+1}, \tag{3}$$

and

$$0 < s_{2k-1} - S \leq s_{2k-1} - s_{2k} = a_{2k} \implies |S - s_{2k-1}| \leq a_{2k}. \tag{4}$$

Putting these last two equations together gives the result

$$|S - s_n| \leq a_{n+1} \text{ for all } n \in \mathbb{N}.$$

□

For part (a), note that for each $x \in S = [0,1]$ the series $\sum_{n=1}^{\infty}(-1)^n \frac{x^n}{n}$ converges pointwise since it is an alternating series of decreasing positive terms approaching 0. So let $F(x)$ be the function to which the series converges on $S = [0,1]$. For the series to converge uniformly to F, we need to see that the uniform norm

$$\|\sum_{j=1}^{n}(-1)^j \frac{x^j}{j} - F(x)\|_S$$

approaches 0 as $n \to \infty$. This is clearly true, since by Eq. 1

$$|\sum_{n=m}^{\infty} f_n(x) < |f_m(x)| = \frac{1}{m} \to 0 \text{ as } m \to \infty.$$

For part (b), use the ratio test:

$$\left|\frac{f_{n+1}(x)}{f_n(x)}\right| = \frac{nx}{n+1} \to x \text{ as } n \to \infty,$$

so for $x \in [0,1)$, convergence follows from the ratio test.

For part (c) we offer two slightly different approaches. One way to deny uniform convergence is to show the existence of an $\epsilon_0 > 0$ such that no matter how large $m \in \mathbb{N}$ is, there is some $x \in [0,1)$ such that $\sum_{n=m+1}^{\infty} \frac{x^n}{n} \geq \epsilon_0$. So let m be given and put $\epsilon_0 = 1/4$. Choose x close enough to 1 so that $x^{2m} > 1/2$ (so $x^j > 1/2$ for $m+1 \leq j \leq m$). Then

$$\sum_{n=m+1}^{\infty} \frac{x^n}{n} > \sum_{n=m+1}^{2m} \frac{x^n}{n} > \frac{1}{2}\sum_{n=m+1}^{2m} \frac{1}{n} > \frac{1}{2}\sum_{n=m+1}^{2m} \frac{1}{2m} = \frac{1}{4}.$$

This completes our first proof.

For the second approach, denote the partial sums by

$$s_m(x) = \sum_{n=1}^{m} |f_n(x)| = \sum_{n=1}^{m} \frac{x^n}{n},$$

and put $s(x) = \lim_{m\to\infty} s_m$. Suppose that s_m converges to s uniformly on $[0,1)$. Then there must exist an m_0 such that

$$x \in [0,1) \implies |s_{m_0}(x) - s(x)| < 1.$$

Because s_{m_0} is bounded on $[0, 1)$, it follows that s is bounded on $[0, 1)$. But for any $m \in \mathbb{N}$,

$$\lim_{x \to 1^-} s(x) \geq \lim_{x \to 1^-} s_m(x) = \sum_{n=1}^{m} \frac{1}{n} \to \infty \text{ as } m \to \infty.$$

Hence s cannot be bounded on $[0, 1)$, a contradiction.

3 Problem 2

Let $\{a_n\}$ and $\{b_n\}$ be sequences of real numbers, and let $c_n = a_n + b_n$ for $n \in \mathbb{N}$.

(a) Prove or find a counterexample: If a is a limit point of $\{a_n\}$ and b is a limit point of $\{b_n\}$, then $a + b$ is a limit point of $\{c_n\}$.

(b) Prove or find a counterexample: If $a = \lim_{n \to \infty} a_n$ and b is a limit point of $\{b_n\}$, then $a + b$ is a limit point of $\{c_n\}$.

Solution: For part (a) there is a simple counterexample: Put $a_n = (-1)^{n+1}$ and $b_n = (-1)^n$ for $n \in \mathbb{N}$. Then $c_n = 0$ for all n and $a = b = 1$ is a limit point of both $\{a_n\}$ and $\{b_n\}$, but $a + b = 2$ is not a limit point of $\{c_n\}$.

For part (b), since b is a limit point of $\{b_n\}$, there is a subsequence $\{b_{n_k}\}_k^\infty$ that converges to b. Then the corresponding subsequence $\{a_{n_k}\}$ of $\{a_n\}$ must converge to the limit a, implying that the subsequence $\{c_{n_k}\}$ must converge to $a + b$, so $a + b$ is indeed a limit point of $\{c_n\}$.

4 Problem 3

Suppose that $f(x)$ is continuous and unbounded on $[a, b)$. Prove that $\lim_{x \to b^-} f(x)$ does not exist.

Solution: Suppose that $\lim_{x \to b^-} = L \in \mathbb{R}$. This means (putting $\epsilon = 1$) there is a $\delta > 0$ such that if $0 < b - x < \delta$ then $|f(x) - L| < 1$. Hence $|f(x)| \leq 1 + |L|$ for $x \in (b - \delta, b)$. Since f is continuous on $[a, b - \delta]$, which is a closed bounded interval, $f(x)$ is bounded on $[a, b - \delta]$. (Usually this is proved long before a student studies compact sets, but one quick argument is that, $[a, b - \delta]$ is a compact connected set, and the continuous image of a compact connected set is compact and connected. And the only compact and connected sets of real numbers are the closed bounded intervals.) Hence $f(x)$ is bounded on both $[a, b-\delta]$ and $(b-\delta, b)$, implying that it is bounded on $[a, b)$. Hence if $f(x)$ is unbounded on $[a, b)$ and continuous there, it must be that $\lim_{x\to b^-} f(x)$ does not exist.

We make an additional comment at this point. Suppose that $f\colon [a, b) \to \mathbb{R}$ is a given function. Consider the following three statements about f.

(i) $f(x)$ is continuous on $[a, b)$.

(ii) $\lim_{x\to b^-} f(x)$ exists as a real number.

(iii) $f(x)$ is unbounded on $[a, b)$.

The proof given above amounts to a proof that not all three statements (i), (ii) and (iii) can be true about the same function f. This means that if we take as hypothesis the truth of any two of the three statements, the proof given above shows that the third statement must be false. For example, if we are given that $f(x)$ is unbounded and $\lim_{x\to b^-} f(x)$ exists as a real number, then we may conclude that $f(x)$ is not continuous on $[a, b)$.

5 Problem 4

Suppose that $f(x)$ is continuous on the interval $[a, b]$. Prove each of the following statements:

(a) There is a $c \in (a, b)$ such that $f(c) = \frac{1}{b-a}\int_a^b f(x)dx$.

(b) There is a $c \in (a, b)$ such that $f(c) = \frac{1}{b-a}\int_a^b f(x)dx$.

Proof: (a) We offer two proofs. For $x \in [a, b]$ put $F(x) = \int_a^x f(t)dt$. Then F is continuous on $[a, b]$ and differentiable on (a, b) with $F'(x) = f(x)$. By the Mean Value Theorem there is a $c \in (a, b)$ for which

$$F'(c) = \frac{F(b) - F(a)}{b-a}, \text{ i.e., } f(c) = \frac{1}{b-a}\int_a^b f(t)dt.$$

This is a simple argument, but it does not immediately suggest a way to deal with part (b). So we offer a second proof that does suggest a way to deal with part (b).

If f is constant, the result is trivial. Suppose f is not constant. Since $[a, b]$ is compact, and f is contintuous, f attains its maximum at some $M \in [a, b]$ and minimum at some $m \in [a, b]$. Thus,

$$\int_a^b f(m)dx < \int_a^b f(x)dx < \int_a^b f(M)dx,$$

which implies

$$f(m) < \frac{1}{b-a}\int_a^b f(x)dx < f(M).$$

By the intermediate value thm, there is a c between m and M such that $f(c) = \frac{1}{b-a}\int_a^b f(x)dx$.

(b) This problem was incorrectly posed! There is a simple counterexample (discovered by one of the students while taking the exam): Let $f(x) = x$ and let $(a, b) = (0, 1)$ Then for all $c \in (a, b)$ we have

$$\frac{1}{c-a}\int_a^c f(x)dx = \frac{1}{c}\int_0^c xdx = \frac{1}{c}[\frac{x^2}{2}]_0^c = \frac{c}{2} < c = f(c).$$

The problem should have specified that $f(a)$ is not the minimum or maximum of $f(x)$, $x \in [a, b]$. In that case the function takes its minimum at some $m \in (a, b]$ and maximum at some $M \in (a, b]$, so

$$f(m) < \frac{1}{m-a}\int_a^m f(x)dx \qquad \text{and} \qquad f(M) > \frac{1}{M-a}\int_a^M f(x)dx.$$

Let $g(z) = f(z) - \frac{1}{z-a}\int_a^z f(x)dx$. Then $g(m) < 0$ and $g(M) > 0$. Since f is continuous, so is g. By the Intermediate Value theorem, there is c between m and M such that $g(z) = 0$, which proves the result.

6 Problem 5

Suppose $f : \mathbb{R} \to \mathbb{R}$ is continuous with $\lim_{x\to-\infty} f(x) = \alpha$ and $\lim_{x\to\infty} f(x) = \beta$, where α, β are finite. Prove that f is uniformly continuous.

Proof. Let $\epsilon > 0$ be given. There is an m with $-\infty < m < 0$ such that $x < m$ implies $|f(x) - \alpha| < \epsilon/2$, and there is an M with $0 < M < \infty$ such that $x > M$ implies $|f(x) - \beta| < \epsilon/2$.

Since f is continuous on $\mathbb{R}$, it is uniformly continuous on $[m-1, M+1]$, so there is a δ with $0 < \delta < 1$ such that if $x, y \in [m-1, M+1]$ and $|x-y| < \delta$ then $|f(x) - f(y)| < \epsilon$. Now, choose any $x, y \in \mathbb{R}$ with $|x - y| < \delta$. Then x and y are either both in $[m-1, M+1]$ or both in $(-\infty, m)$ or both in (M, ∞). For example, if x and y are both in $(-\infty, m)$, then $|f(x) - f(y)| \leq |f(x) - \alpha| + |\alpha - f(y)| < \epsilon$. In all cases we have $|f(x) - f(y)| < \epsilon$, so f is uniformly continuous. □

One student offered a valid argument that under the hypotheses of the problem $f(x)$ is bounded on $(-\infty, +\infty)$, and then claimed that if $f(x)$ is continuous and bounded on $(-\infty, +\infty)$ it must be uniformly continuous there. So we give an interesting counterexample to that claim.

To prepare for the example, we first prove a useful lemma.

Lemma 6.1. *Let c be a fixed positive number. Then*

$$\lim_{x\to\infty} [\sqrt{x+c} - \sqrt{x}] = 0.$$

Proof of lemma:

$$|\sqrt{x+c} - \sqrt{x}| = \frac{(x+c) - x}{\sqrt{x+c} + \sqrt{x}} < \frac{c}{2\sqrt{x}} \to 0 \text{ as } x \to \infty.$$

Now consider the function $f(x) = \sin(x^2)$ for $x \in \mathbb{R}$. Put $c = \pi$, $x_k = \sqrt{k\pi + \pi/2}$ and $y_k = \sqrt{(k+1)\pi + \pi/2}$. Then $|x_k - y_k| \to 0$ as $k \to \infty$, but $|f(x_k) - f(y_k)| = |\sin((k+1)\pi + \pi/2) - \sin((k\pi) + \pi/2)| = |\pm 1 - \mp 1| = 2$. So we can find distinct x and y arbitrarily close together with $|f(x) - f(y)| = 2$. This shows that $f(x)$ is not uniformly continuous on $(-\infty, +\infty)$.

7 Problem 6

For which $\alpha > 0$ does

$$\sum_{k=1}^{\infty} \frac{\alpha^{k \ln k}}{k!}$$

converge?

Solution: The ratio of the $(k+1)^{st}$ to k^{th} term is

$$\frac{1}{k+1}\alpha^{(k+1)\ln(k+1) - k\ln(k)} = \frac{1}{k+1} e^{\ln(1+\frac{1}{k})^k \ln(\alpha)} e^{\ln(k+1)\ln(\alpha)}$$

$$= (k+1)^{\ln(\alpha)-1}\left((1+\frac{1}{k})^k\right)^{\ln(\alpha)}.$$

The second factor converges to α, since $\lim_{k\to\infty}(1+\frac{1}{k})^k = e$. The first factor converges to 0 if $\alpha < e$, to 1 if $\alpha = e$, and to ∞ if $\alpha > e$. So the ratio tends to a limit less than 1 if $\alpha < e$ and to a limit greater than 1 if $\alpha \geq e$. Thus the series converges iff $\alpha < e$.

It is possible to use the root test instead of the ratio test if one remembers the appropriate form of the Stirling - De Moivre formula for $n!$.

$$\lim_{k\to\infty}\left(\frac{\alpha^{k\ln k}}{k!}\right)^{\frac{1}{k}} = \lim_{k\to\infty}\frac{\alpha^{\ln k}}{(k!)^{\frac{1}{k}}}.$$

First we write $c = \alpha/e$, i.e., $\alpha = ce$. So $\alpha^{\ln k} = c^{\ln k}\cdot k$. Recall Stirling's formula for $n!$:

$$n! = n^n e^{-n}(2n\pi)^{1/2}e^{\frac{\theta}{12n}},\ \text{where } 0 < \theta < 1.$$

(I can provide a proof of this if necessary. Just send me an email message asking for a pdf file giving a proof of Stirling's formula.)

So the k^{th} root of the k^{th} term is:

$$\frac{c^{\ln k}\cdot k}{(k!)^{1/k}} = c^{\ln k}\cdot\frac{1}{e^{-1}(2\pi)^{1/(2k)}(k^{1/k})^{1/2}e^{\frac{\theta}{12k^2}}}.$$

The factors of the denominator (except for the first one) approach 1 as $k\to\infty$. So the entire right hand factor approaches e. If $c = 1$, $c^{\ln k}\to 1$, so the entire limit is greater than 1 and the original series diverges. But if $0 < c < 1$, (i.e., if $0 < \alpha < e$), then the limit of the k^{th} root of the k^{th} term is $0 < 1$, and the series converges.

8 Problem 7

Suppose $f_n \to f$ uniformly on $A \subset X$, where (X, d) is a metric space, and let $x \in \bar{A}$, i.e., x is a limit point of A. Also assume that for $n = 1, 2, 3, \ldots$ the limits $\lim_{t\to x} f_n(t) = a_n \in \mathbb{R}$ exist.

(a) Prove $\{a_n\}$ converges. Hint: Show the sequence is Cauchy.

(b) Prove $\lim_{t\to x} f(t) = a$, where $a = \lim a_n$.

Proof: (a) Choose $\epsilon > 0$, and N large enough so that $m, n > N$ implies $|f_n(t) - f_m(t)| < \epsilon$ for every $t \in X$. Fix $n, m > N$. Choose $\delta > 0$ so that $d(t, x) < \delta$ imples $|f_n(t) - a_n| < \epsilon$ and $|f_m(t) - a_m| < \epsilon$. Then for $d(t, x) < \delta$, $|a_n - a_m| < |a_n - f_n(t)| + |f_n(t) - f_m(t)| + |f_m(t) - a_m| < 3\epsilon$. This shows that the sequence is Cauchy, and therefore converges.

(b) Since $a_n \to a$, there is M so that $n > M$ implies $|a_n - a| < \epsilon$. Choose $n > \max(N, M)$ and t so that $d(t, x) < \delta$. Then

$$|f(t) - a| < |f(t) - f_n(t)| + |f_n(t) - a_n| + |a_n - a| < 3\epsilon.$$

Hence $\lim_{t \to x} f(t) = a$.

9 Problem 8

Prove that the system of equations

$$\begin{aligned} 3x &= y + \sin x \\ 3y &= x + \cos y \end{aligned}$$

has a unique solution.

Solution We are going to apply a general mean value theorem for functions from $\mathbb{R}^n$ to $\mathbb{R}^n$ along with the Contraction Mapping theorem to show that certain systems of equation have a unique solution. First recall the Contraction Mapping theorem.

Theorem 9.1. *(The Contraction Mapping Theorem) Let (X, d) be a metric space. A function $f\colon X \to X$ is a* contraction *provided there is a number c with $0 < c < 1$ for which $d(f(x), f(y)) \le c \cdot d(x, y)$ for all $x, y \in X$.*

If (X, d) is a complete metric space and $f\colon X \to X$ is a contraction, then there is a unique $x_0 \in X$ for which $f(x_0) = x_0$.

Since the proof of this theorem is given in essentially all analysis courses, we do not give a proof here. However, there is not universal agreement on what theorem should be called "the" mean value theorem for continuous functions from $\mathbb{R}^n$ to $\mathbb{R}^m$, so we state such a theorem and give a proof.

Theorem 9.2. *(A Mean Value Theorem) Let D be open in R^n and suppose that $\mathbf{f} : D \to R^m$ is differentiable at each point of D. Let $\mathbf{x}$, $\mathbf{y} \in D$ be such* *that the closed line segment $L(\mathbf{x},\mathbf{y}) \subseteq D$. Then for each point $\mathbf{a} \in R^m$ there is a point $\mathbf{z} \in L(\mathbf{x},\mathbf{y})$ different from $\mathbf{x}$ or $\mathbf{y}$ such that*

$$\mathbf{a} \cdot (\mathbf{f}(\mathbf{y}) - \mathbf{f}(\mathbf{x})) = \mathbf{a} \cdot (\mathbf{f}'(\mathbf{z})(\mathbf{y} - \mathbf{x}))\,.$$

Proof. Let $\mathbf{u} = \mathbf{y} - \mathbf{x}$. Since D is open and $L(\mathbf{x},\mathbf{y}) \subseteq D$, an open line segment slightly longer than $L(\mathbf{x},\mathbf{y})$ is contained in D. So there is a $\delta > 0$ such that $\mathbf{x} + t\mathbf{u} \in D$ for all $t \in (-\delta, 1+\delta)$. This corresponds to the open line segment from $\mathbf{x} - \delta\mathbf{u} = (1+\delta)\mathbf{x} - \delta\mathbf{y}$ to the point $\mathbf{x} + (1+\delta)\mathbf{u} = -\delta\mathbf{x} + (1+\delta)\mathbf{y}$. Now let $\mathbf{a}$ be a fixed vector in R^m, and let F be the real-valued function defined on $(-\delta, 1+\delta)$ by $F(t) = \mathbf{a} \cdot \mathbf{f}(\mathbf{x} + t\mathbf{u})$. Then F is differentiable on $(-\delta, 1+\delta)$. By a routine argument we see that $\mathbf{G} : t \mapsto \mathbf{f}(\mathbf{x} + t\mathbf{u})$ has derivative $\mathbf{G}'(t) = \mathbf{f}'(\mathbf{x}+t\mathbf{u};\mathbf{u})$. (This is just another notation for $\frac{\partial \mathbf{f}}{\partial \mathbf{u}}(\mathbf{c}+t\mathbf{u})$.)

Here is the proof: If $\mathbf{G}(t) = \mathbf{f}(\mathbf{x} + t\mathbf{u})$, for $a < t < b$, then

$$\mathbf{G}'(t) = lim_{h\to 0}\frac{\mathbf{G}(t+h) - \mathbf{G}(t)}{h} =$$

$$= lim_{h\to 0}\frac{\mathbf{f}(\mathbf{x} + (t+h)\mathbf{u}) - \mathbf{f}(\mathbf{x} + t\mathbf{u})}{h}$$

$$= lim_{h\to 0}\frac{\mathbf{f}((\mathbf{x} + t\mathbf{u}) + h\mathbf{u}) - \mathbf{f}(\mathbf{x} + t\mathbf{u})}{h} =$$

$$= \frac{\partial \mathbf{f}}{\partial \mathbf{u}}(\mathbf{x} + t\mathbf{u}).$$

Moreover, the map

$$A \colon \mathbb{R}^n \to \mathbb{R} \colon \mathbf{v} \mapsto \mathbf{a} \cdot \mathbf{v}$$

is linear with $[A] = \mathbf{a} = [a_1, \ldots, a_n]$. So $A'_{\mathbf{c}}(\mathbf{v}) = \mathbf{a} \cdot \mathbf{v}$ (for any $\mathbf{c} \in D$). In particular this means that

$$A'(\mathbf{G}(t))(\mathbf{G}'(t)) = A'_{\mathbf{G}(t)}(\mathbf{G}'(t)) = \mathbf{a} \cdot \mathbf{G}'(t) = \mathbf{a} \cdot (\mathbf{f}'(\mathbf{x} + t\mathbf{u};\mathbf{u})).$$

So if $F(t) = \mathbf{a} \cdot \mathbf{f}(\mathbf{x} + t\mathbf{u}) = (A \circ \mathbf{G})(t)$, then by the Chain Rule $F'(t) = \mathbf{a} \cdot (\mathbf{f}'(\mathbf{x} + t\mathbf{u};\mathbf{u}))$. By the usual mean value theorem, $F(1) - F(0) = F'(\theta)$, where $0 < \theta < 1$. Put $\mathbf{z} = \mathbf{x} + \theta\mathbf{u} \in L(\mathbf{x},\mathbf{y})$. Then $F'(\theta) = \mathbf{a} \cdot (\mathbf{f}'(\mathbf{z})(\mathbf{u}))$ and $F(1) - F(0) = \mathbf{a} \cdot (\mathbf{f}(\mathbf{y}) - \mathbf{f}(\mathbf{x}))$. It is easy to see that we now have proved Theorem 9.2. □

Let $A = (a_{ij})$ be an $m \times n$ matrix over $\mathbb{R}$, and put

$$M_A = \left[\sum_{\substack{1 \le i \le m \\ 1 \le j \le n}} a_{ij}^2 \right]^{\frac{1}{2}}.$$

If $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$, define the **norm** $\|\mathbf{x}\|$ of $\mathbf{x}$ by

$$\|\mathbf{x}\| = \sqrt{\sum_{i=1}^n x_i^2}.$$

Lemma 9.3. *If A is an $m \times n$ matrix over $\mathbb{R}$, then we have*

$$\|A\mathbf{x}\| \le M_A \cdot \|\mathbf{x}\| \text{ for all } \mathbf{x} \in \mathbb{R}^n.$$

Proof. Suppose that $\mathbf{x} = (x_1, \ldots, x_n)^T$, $\mathbf{y} = A\mathbf{x} = (y_1, \ldots, y_m)^T$. Then

$$\begin{aligned}
\|\mathbf{y}\|^2 &= \|A\mathbf{x}\|^2 = \sum_{i=1}^m y_i^2 = \sum_{i=1}^m \left| \sum_{j=1}^n a_{ij} x_j \right|^2 \\
&\le \sum_{i=1}^n \left\{ \sum_{j=1}^n |a_{ij}| \cdot |x_j| \right\}^2 \quad \text{(now use Cauchy-Schwarz)} \\
&\le \sum_{i=1}^m \left\{ \sum_{j=1}^n |a_{ij}|^2 \cdot \sum_{j=1}^n |x_j|^2 \right\} \\
&= \sum_{i=1}^m \left\{ \|\mathbf{x}\|^2 \cdot \sum_{j=1}^n |a_{ij}|^2 \right\} = M_A^2 \|\mathbf{x}\|^2.
\end{aligned}$$

So $\|A\mathbf{x}\| \le M_A \|\mathbf{x}\|$. □

We now put these ideas together!

Theorem 9.4. *Let D be open in R^n and suppose that $\mathbf{f} : D \to R^m$ is differentiable at each point of D. Let $\mathbf{x}$, $\mathbf{y} \in D$ be such that the closed line segment $L(\mathbf{x}, \mathbf{y}) \subseteq D$. Then for some $\mathbf{z} \in L(\mathbf{x}, \mathbf{y})$, $\mathbf{x} \ne \mathbf{z} \ne \mathbf{y}$, we have*

$$\|\mathbf{f}(\mathbf{y}) - \mathbf{f}(\mathbf{x})\| \le \|\mathbf{f}'(\mathbf{z}) \cdot (\mathbf{y} - \mathbf{x})\| \le M \cdot \|\mathbf{y} - \mathbf{x}\|.$$

Here M is the square root of the sum of the squares of the entries of $\mathbf{f}'(\mathbf{z})$.

Proof. If $\mathbf{f}(\mathbf{x}) = \mathbf{f}(\mathbf{y})$, the result is trivial. For $\mathbf{f}(\mathbf{x}) \neq \mathbf{f}(\mathbf{y})$, put $\mathbf{a} = \frac{\mathbf{f}(\mathbf{y})-\mathbf{f}(\mathbf{x})}{\|\mathbf{f}(\mathbf{y})-\mathbf{f}(\mathbf{x})\|}$. Then $\|\mathbf{a}\| = 1$ and

$$\mathbf{a} \cdot (\mathbf{f}(\mathbf{y}) - \mathbf{f}(\mathbf{x})) = \|\mathbf{f}(\mathbf{y}) - \mathbf{f}(\mathbf{x})\|.$$

Now apply the Mean Value theorem (and use the Cauchy - Schwarz inequality) to get a point $\mathbf{z} \in L(\mathbf{x}, \mathbf{y})$ different from $\mathbf{x}$ or $\mathbf{y}$ such that

$$\|\mathbf{f}(\mathbf{y}) - \mathbf{f}(\mathbf{x})\| = \mathbf{a} \cdot (\mathbf{f}(\mathbf{y}) - \mathbf{f}(\mathbf{x})) = \mathbf{a} \cdot (\mathbf{f}'(\mathbf{z})(\mathbf{y} - \mathbf{x}))\,.$$

Hence

$$\|\mathbf{f}(\mathbf{y}) - \mathbf{f}(\mathbf{x})\| \quad = \quad |\mathbf{a} \cdot (\mathbf{f}'(\mathbf{z})(\mathbf{y} - \mathbf{x}))| \leq \|\mathbf{f}'(\mathbf{z})(\mathbf{y} - \mathbf{x})\|$$

implying $\|\mathbf{f}(\mathbf{y}) - f(\mathbf{x})\| \leq M \cdot \|\mathbf{y} - \mathbf{x}\|$, where M^2 is the sum of the squares of the entries of $\mathbf{f}'(\mathbf{z})$. (Here M depends on $\mathbf{z}$, and hence on $\mathbf{x}$ and $\mathbf{y}$.) □

Note: If D is convex in Theorem 9.4 and all the partials $D_j f_k$ are bounded on D, there is a constant $A > 0$ such that f satisfies a Lipschitz condition on D, i.e., $\|f(\mathbf{y}) - f(\mathbf{x})\| \leq A \cdot \|\mathbf{y} - \mathbf{x}\|$.

Next, reread the original problem # 8.

Solution: Write the system of equations as the fixed point problem

$$f\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} x \\ y \end{bmatrix}$$

with

$$f\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \frac{1}{3}\begin{bmatrix} y + \sin x \\ x + \cos y \end{bmatrix}$$

The function f maps the complete space $\mathbb{R}^2$ into itself and it has the Jacobian matrix

$$f'(x, y) = \frac{1}{3}\begin{bmatrix} \cos x & 1 \\ 1 & -\sin y \end{bmatrix}$$

and we can see that the constant M for the matrix $f'(x, y)$ satisfies $M \leq \frac{2}{3} < 1$. Thus for any $\mathbf{z}, \mathbf{w} \in \mathbb{R}^2$, since $\mathbb{R}^2$ contains the segment connecting $\mathbf{z}$ and $\mathbf{w}$, we have

$$\|f(\mathbf{z}) - f(\mathbf{w})\| \leq \frac{2}{3}\|\mathbf{z} - \mathbf{w}\|$$

so the function f is a contraction. The statement follows from the Banach contraction theorem.

We give a second application.

Exercise. Prove that the following system of equations has a unique solution.

$$\begin{aligned} 2x &= \sin(y) + \cos(z) \\ 2y &= \cos(x) + \sin(z) \\ 2z &= \sin(x) + \cos(y) \end{aligned}$$

Solution Start by defining

$$F\colon \mathbb{R}^3 \to \mathbb{R}^3\colon \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} \frac{1}{2}\sin(y) + \frac{1}{2}\cos(z) \\ \frac{1}{2}\cos(x) + \frac{1}{2}\sin(z) \\ \frac{1}{2}\sin(x) + \frac{1}{2}\cos(y) \end{pmatrix}.$$

Calculate the Jacobian matrix as a function of x, y, z:

$$(F') = \begin{pmatrix} 0 & \frac{1}{2}\cos(y) & -\frac{1}{2}\sin(z) \\ -\frac{1}{2}\sin(x) & 0 & \frac{1}{2}\cos(z) \\ \frac{1}{2}\cos(x) & -\frac{1}{2}\sin(y) & 0 \end{pmatrix}.$$

No matter what x, y, z we use, the sum of the squares of the entries of F' is always equal to $3/4$. Hence $M = \sqrt{\frac{3}{4}} < 1$, so F is a contraction on the complete metric space $\mathbb{R}^3$. Hence it must have a unique fixed point in $\mathbb{R}^3$, showing that the original system of equations has a unique solution.