A Survival Guide to

Dynamical Systems

Revised and Reissued 2013 Dover Edition

Shlomo Sternberg

Partial scrutiny,

Comments, Suggestions and Errata

José Renato Ramos Barbosa

2016

Departamento de Matemática
Universidade Federal do Paraná
Curitiba - Paraná - Brasil
jrrb@ufpr.br

==
==
1
==
==
Comment, p. 11
The tangent to the curve $G = \{(x, y) \in \mathbb{R}^2 \mid y = P(x)\}$ at $(x, y) = (x_n, P(x_n))$ is given by

$$T_n = \left\{(x, y) \in \mathbb{R}^2 \mid P'(x_n)(x - x_n) + (-1)(y - P(x_n)) = 0\right\}.$$

Therefore, for $(x, y) \in T_n$ with $y = 0$, we have $x = x_{n+1}$ from Newton's method.
==
Comment, p. 12, l. -1
As a matter of fact, the inequality is strict, that is,

$$|x_{new} - z| < \frac{\mu}{2}.$$

Hence, since

$$x_{new} \in \left(z - \frac{\mu}{2}, z + \frac{\mu}{2}\right),$$

if we denote the next iterate by x_{newer}, that is,

$$x_{newer} := x_{new} - \frac{f(x_{new})}{f'(x_{new})},$$

it follows that

$$|x_{newer} - z| < \frac{\mu}{4}.$$

This iteration goes on and each *Newton step* more than halves the distance to z.
==
Erratum, p. 13, 1st sentence of **1.2.3**
Change to "Now let f **be** a function ...".
==
Comment, p. 21, ll. -2 and -1
To start the sequence of Newton steps, **Proposition 1.2.1** considers $x_0 = 0$ as the initial guess (starting point). Then, in case $P(z) = 0$ holds for some z with $x_n \to z$, (1.14) means

$$|P(x_0) - P(z)| \leq K^{-5},$$

that is, $P(0)$ within K^{-5} of $P(z)$.
==
Comments, p. 22

- (1.26)
  $x_0 = 0$ is **the** initial guess after a shifting of coordinates and it is most likely that $P(x_0) \neq 0$ (when considering P as a single function of x). $u = \mathbf{0}$ is supposed to be a zero of $P(u, 0)$.

- 1st sentences right after (1.26)
  The inequality with $x_0 = 0$ guarantees that (1.14) holds. Hence **Proposition 1.2.1** guarantees the existence of a zero: there is a point z which is the limit of the sequence of Newton steps with initial guess $x_0 = 0$:

  $$P(u, z) = 0, \quad x_n \to z \quad \text{and} \quad x_0 = 0.$$

  Therefore **1.2.3** implies that the convergence to z also takes place for the sequence of Newton steps with initial guess $x_0 \neq 0$ provided that $|x_0 - z|$ is sufficiently small. (Furthermore, since $|z| \leq |x_0 - z| + |x_0|$, notice that $|z|$ is also small enough. So both $x_0 \neq 0$ and z must be sufficiently close to $x_0 = 0$!)

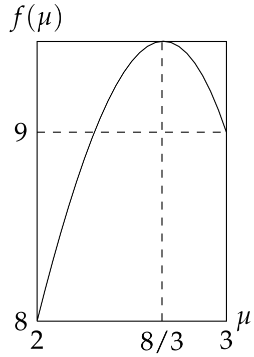
==
Errata, p. 23

- 1st line after (1.29)
  "... $u = 0$, ..." should be "... $u = \mathbf{0}$, ...".

- l. -9
  "..., say $\in \mathbb{R}^n$, ..." should be "..., say $u \in \mathbb{R}^n$, ...".

==

==
==
2
==
==
Erratum, p. 34
2.1.1 does not deal with the $\mu = 1$ case.
==
Comment, p. 38, ll. -6 and -5
The value of $f(\mu) = \mu^2(4-\mu)$, since $f'(\mu) = \mu(8-3\mu)$ and $f''(8/3) < 0$, increases from $f(2) = 8$ to $f(8/3) > 9$ and decreases from $f(8/3)$ to $f(3) = 9$.



==
Erratum, p. 40, Figure 2.5
$L_\mu^{(2)}$ should be $L_\mu^{\circ 2}$.
==
Comments, p. 40
As a matter of fact, (2.3) refers to the discriminant Δ of the quadratic function (which precedes it) with zeros

$$p_{2\pm} = \frac{-\left(\mu^2+\mu\right) \pm \sqrt{\Delta}}{-2\mu^2}.$$

2.1.7 The double root comes from the fact that $\Delta = 0$ for $\mu = 3$.

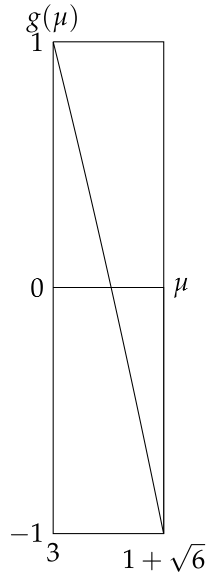
2.1.8 l. -1
For $\mu > 3$, since $\sqrt{(\mu+1)(\mu-3)} = \sqrt{\mu^2-2\mu-3} < \sqrt{\mu^2-2\mu+1} = \mu - 1$,

$$\begin{aligned} 0 < \frac{1}{\mu} &= \frac{1}{2} + \frac{1}{2\mu} - \frac{\mu-1}{2\mu} \\ &< p_{2-} \\ &< p_{2+} \\ &< \frac{1}{2} + \frac{1}{2\mu} + \frac{\mu-1}{2\mu} = 1. \end{aligned}$$

==
Erratum, p. 41
$L_\mu^{(2)}$ should be $L_\mu^{\circ 2}$.
==
Comments, p. 41, last two sentences of **2.1.8**

$$\begin{aligned} \left(L_3^{\circ 2}\right)'(p_{2\pm}) &= \left(L_3^{\circ 2}\right)'\left(\frac{2}{3}\right) \\ &= L_3'(2/3)L_3'(2/3) \\ &= (-1)(-1) \\ &= 1. \end{aligned}$$

Furthermore, consider the graph of $g(\mu) = -\mu^2 + 2\mu + 4$ for $\mu \in \left[3, 1+\sqrt{6}\right]$.



==

Erratum, p. 43, **2.1.10**

Shouldn't p_{2+} be p_{2-}?

==

Erratum, p. 45, Figure 2.9

Change μ to r or vice versa.

==

Comment, p. 49, $\mu''(0)$

If $\gamma(x) = (x, \mu(x))$ and $\Gamma(x, \mu)$ denotes $-\frac{P_x}{P_\mu}$, then $\mu'(x) = \Gamma(\gamma(x))$. Hence

$$\begin{aligned}
\mu''(0) &= \gamma'(0) \cdot \nabla\Gamma(\gamma(0)) \\
&= 1 \cdot \Gamma_x(0, \mu(0)) + \mu'(0) \cdot \Gamma_\mu(0, \mu(0)) \\
&= \Gamma_x(0, 0) + 0 \\
&= -\frac{P_{xx}(0,0) \cdot P_\mu(0,0) - P_x(0,0) \cdot P_{x\mu}(0,0)}{\left(P_\mu(0,0)\right)^2} \\
&= -\frac{P_{xx}(0,0)}{P_\mu(0,0)} + \frac{0 \cdot P_{x\mu}(0,0)}{\left(P_\mu(0,0)\right)^2} \\
&= -\frac{\partial^2 P/\partial x^2}{\partial P/\partial \mu}(0,0).
\end{aligned}$$

==

Comment, p. 52, first paragraph of **2.3.2**

See the paragraph starting with "Before embarking ..." on p. 47.

==

Comment, p. 54, **Step I**

(2.7) holds since $P(x, \mu)$ is well-defined via the first paragraph of **2.3.2**.

==

Comment, p. 56

l. 6

The chain rule implies that

$$\frac{d}{dx}\left(\frac{\partial F^{\circ 2}}{\partial x}(x, \nu(x))\right)\bigg|_{x=0} = \frac{\partial^2 F^{\circ 2}}{\partial x^2}(0,0) \cdot 1 + \frac{\partial^2 F^{\circ 2}}{\partial \mu \partial x}(0,0) \cdot \nu'(0).^{1}$$

[1] Also, see $\phi'(x)$, **Step VI**.

l. 7
By considering l. 6, p. 53, and l. -3, p. 52, it follows that

$$\frac{\partial F^{\circ 2}}{\partial x}(0,0) = \left(\frac{\partial F}{\partial x}(0,0)\right)^2 = 1 = (\lambda(0))^2.$$

==
Comment, pp. 56–7, 1st sentence of **2.4**
Suppose (2.10) holds. Therefore, if $'$ denotes differentiation with respect to x and $p = \frac{1}{2}$, then

$$\left(L_\mu^{\circ 2^{n-1}}\right)'(p) = L_\mu'(p) L_\mu'\left(L_\mu(p)\right)\cdots L_\mu'\left(L_\mu^{\circ(2^{n-1}-1)}(p)\right) = 0$$

since $L_\mu'(p) = 0$, which implies that p is a superattractive periodic point of period 2^{n-1}.[2]
==
Comment, p. 58, ll. 9–12
Consider s_n is a solution of (2.10) for which $\frac{1}{2}$ has period $2^{i-1} < 2^{n-1}$. Therefore $L_{s_n}^{\circ 2^{i-1}}(1/2) = 1/2$ and, since $\left(L_{s_n}^{\circ 2^{i-1}}\right)'(1/2) = 0$, it follows that $s_n = s_i$ with $i < n$!

==
Erratum, p. 58, l. -1
As a matter of fact, since the superattractive value s_r is the solution of the equation (for μ) $f_\mu^{\circ 2^{r-1}}(X_m) = X_m$,[3] then

$$f_{s_r}^{\circ 2^{r-1}}(X_m) - X_m = 0.$$

Since d_r denotes the difference between X_m and the next nearest point on the superstable 2^r orbit, which is the orbit $\left\{X_m, f_{s_r}(X_m), \ldots, f_{s_r}^{\circ 2^{r-1}-1}(X_m)\right\}$, then

$$d_r = f_{s_r}^{\circ j}(X_m) - X_m \text{ for } j \in \left\{1, \ldots, 2^{r-1}-1\right\}.$$

==
Erratum, p. 61, l. 12
It seems the insertion of 2^{r+1} is unnecessary! It should be just

$$\mathcal{R}(g_k)(y) = \lim(-\alpha)^{r+1} g_{s_{k+r}}^{2^{r+1}}(y/(-\alpha)^{r+1}) = g_{k-1}(y)$$

since $g_k(y) = \lim(-\alpha)^r g_{s_{k+r}}^{2^r}(y/(-\alpha)^r)$ implies that

$$g_{k-1}(y) = \lim(-\alpha)^r g_{s_{k-1+r}}^{2^r}(y/(-\alpha)^r) \underbrace{=}_{\mathbf{r} = r-1} \lim(-\alpha)^{\mathbf{r}+1} g_{s_{k+\mathbf{r}}}^{2^{\mathbf{r}+1}}(y/(-\alpha)^{\mathbf{r}+1}).$$

[2] See p. 26, **1.4.6**.
[3] See the discussion around (2.10), p. 57.

==
==

3

==
==

Comments, p. 63, *Proof* of **Lemma 3.1.2**

- Negating the existence of the greatest point $c \in J$ with $f(c) = a$ leads us to an increasing sequence $\{c_n\}$ in J, which converges to $\sup \{c_n \mid n \in \mathbb{N}\} = s \in J$ since J is compact, such that $f(c_n) = a$ for all n. Then $f(s) = a$ since the sequence $\{f(c_n)\}$ converges to both $f(s)$ and a. Therefore s is the greatest point $c \in J$ with $f(c) = a$!

- "Then we may take $L = [c, d]$."
  If $f([c, d]) = [a', b']$, then The Intermediate Value Theorem guarantees that $[a', b'] \supseteq [a, b]$. Hence, if $a' < a$, there exists $c' > c$ with $f(c') = a$,[4] which is a contradiction. Therefore, $a' = a$. Similarly, $b' = b$.

==

Comments/Errata, p. 64

- **Notation**
  Where do we use $< a, b >$ from this point on?

- *Proof* of **Theo. 3.1.1**
  "$f(I_0) \supset I_1, \quad f(I_1) \supset I_0 \cup I_1$."
  Apply The Intermediate Value Theorem.[5]
  "Finally, since $f(I_1) \supset I_0$**,** there is a compact interval $A_n \subset I_1$ with $f(A_n) = A_{n-1}$."

  (Note the insertion of the second comma - the one in bold.)

  In fact, $I_0 \supset A_{n-1}$.[6]
  "By Lemma 3.1.1, $f^{\circ n}$ has a fixed point, x, in A_n."

  (Note the use of the adopted notation in place of f^n.)

  In fact

$$\begin{aligned} A_n \subset I_1 &= f^{\circ(n-2)}(A_{n-2}) \\ &= f^{\circ(n-2)}(f(A_{n-1})) \\ &= f^{\circ(n-1)}(A_{n-1}) \\ &= f^{\circ(n-1)}(f(A_n)) \\ &= f^{\circ n}(A_n). \end{aligned}$$

  "But $f(x)$ lies in I_0 and all the higher iterates up to n lie in I_1 so the period can not be smaller than n."
  In fact, on the one hand $f(x) \in I_0$ since $x \in A_n$ and $f(A_n) = A_{n-1} \subset I_0$. On the other hand, if $f^{\circ i}(x) = x \in A_n \subset I_1$ for some index $1 < i < n$, then $f^{\circ(i+1)}(x) = f(x) \in I_0$, which is absurd since $f^{\circ 2}(x), \ldots, f^{\circ n}(x) \in I_1$.

==

Erratum, p. 67, l. -3
The denominator $[f'(g(x))g'(x)$ should be $[f'(g(x))]g'(x)$ or $f'(g(x))g'(x)$.

==

Errata, p. 68

[4] Draw a graph illustrating the situation!
[5] Note that $I_0 \cup I_1 = [a, c]$.
[6] See p. 64, l. -10.

- *Proof* of **Lemma 3.2.1**, last sentence
  Write | The reverse **occurs** at a relative maximum. |.

- *Proof* of **Lemma 3.2.3**, first sentence
  'fixed' should be 'critical'.

==

Comment, p. 69

- *Proof* of **Lemma 3.2.3**
  | "By the mean value theorem, there is a point s with $r < s < t$ such that $g'(s) < 1$." |
  In fact, if $f(x) = g(x) - x$, there is a point s in the open interval (r, t) at which

$$\begin{aligned} g'(s) - 1 &= f'(s) \\ &= \frac{f(t) - f(r)}{t - r} \\ &= \frac{g(t) - t - (g(r) - r)}{t - r} \\ &< 0. \end{aligned}$$

- *Proof* of **Lemma 3.2.4**
  Suppose first that the set $C\left(f^{\circ 2}\right)$ of all critical points of $f^{\circ 2}$ is infinite. On the other hand, the set $C(f)$ consisting of all critical points of f is finite by hypothesis. Therefore there can be only finitely many elements of $C\left(f^{\circ 2}\right)$ in $C(f)$ and

$$\left\{x \in C\left(f^{\circ 2}\right) \mid f(x) \in C(f)\right\}$$

  is an infinite set whose image under f cannot be infinite, which is a contradiction by the second sentence of the proof.

- last sentence, | "So there are three possibilities: ..." |
  On the one hand, neither $g(L)$ nor $g(R)$ is in (L, R). In fact, suppose otherwise. So one or both of them are attracted to p by g. Then one or both of the endpoints of (L, R) are attracted to p by g, which contradicts the maximality of (L, R). On the other hand, $g(\{L, R\}) \subset \{L, R\}$.[7] In fact, L and R are accumulation points of (L, R). Hence there are sequences $L_1, L_2, \ldots$ and $R_1, R_2, \ldots$ in (L, R) with $L_n \to L$ and $R_n \to R$ as $n \to \infty$. Then, since g is continuous, $g(L_n) \to g(L)$ and $g(R_n) \to g(R)$ as $n \to \infty$. Now suppose $|g(L) - L| > 0$ and $|g(L) - R| > 0$. So there is some index n_0 such that $g(L_{n_0})$ is not in (L, R).[8] Therefore, since L_{n_0} is attracted to p by g, it follows that $g(L_{n_0})$ is attracted to p by g, which contradicts the maximality of (L, R). Similarly, the supposition that $g(R) \notin \{L, R\}$ contradicts the maximal choice of (L, R).

==

Erratum, p. 70, l. 2

| "... $z, f(z), \ldots f^{\circ m-1}(z)$ ..." | should be | "... $z, f(z), \ldots, f^{\circ(m-1)}(z)$ ..." |.

[7] Notice that, if that inclusion holds, then the last line of the p. 69 means that $g(L) = g(R)$ equals either L or R - one point is fixed and the other is eventually fixed!

[8] As a matter of fact, almost all the terms $g(L_n)$ are not in (L, R).

==
==

4

==
==

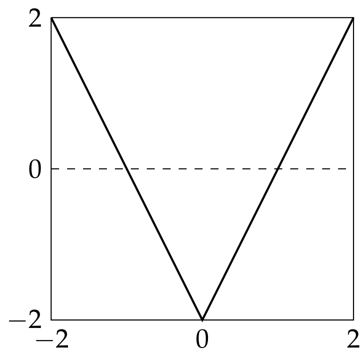
Erratum, p. 79, **Prop. 4.1.1**

"Let $f = ax^2 + bx + d$." should be "Let $f(x) = ax^2 + bx + d$.".

==

Erratum, p. 82, Figure 4.3

The graph of $V(x)$ with the proper scale for the y-axis should be:



==

Comment, p. 84, **The tent map is chaotic**, "..., there is a point $x \in \left[\frac{k}{2^n}, \frac{k+1}{2^n}\right]$ which is mapped into itself."

The diagonal $D = \{(x, y) \in [0,1] \times [0,1] \,|\, y = x\}$ traverses the narrow rectangle $\left[\frac{k}{2^n}, \frac{k+1}{2^n}\right] \times [0,1]$, which contains the graph of the surjective restriction of $T^{\circ n}$ to $\left[\frac{k}{2^n}, \frac{k+1}{2^n}\right]$. Since the image of this restriction is the unit interval $[0,1]$, its graph intersects D.

==

Erratum, p. 85

In the verification of $T \circ S = T \circ T$, seven commas (at the very end of each line) and a period (at the very end of the 8th line) are missing!

==

Erratum, p. 86

Change S^k to $S^{\circ k}$ twice.

==

Comment, p. 88, the sentence right before **Going to the unit circle**

Since S is chaotic on $[0,1]$, the commutative diagram at the bottom of the page 84 gives us an alternative proof that T is chaotic.

==

Comment/Errata, p. 90, **Proof of Prop. 4.4.1 with** $\delta = c/4$

- l. 6
  $1 \leq nj - k \leq n$ since $n(j-1) \leq k \leq nj - 1$.
- l. 11
  $\left(f^{-(nj-k)}\left(W_{nj-k}\right)\right.$ should be $\left(f^{-(nj-k)}\left(W_{nj-k}\right)\right)$.
- ll. 13-16 and 20
  Replace f^{nj} by $f^{\circ nj}$ eight times and, at the very beginning, replace f^{nj-k} by $f^{\circ(nj-k)}$ once.
- last sentence, "So ... $m = nj$."
  n should be m in the enunciation of **Prop. 4.4.1**, p. 89.

==

Comments/Erratum, *Proof* of **Prop. 4.5.1**, pp. 91–2

- 2nd sentence, "If (x_0, y_0) ... $y_1 = g(y_0)$."
  (x_1, y_1) satisfies $y = h(x)$ since

$$\begin{aligned} y_1 &= g(y_0) \\ &= g(h(x_0)) \\ &= h(f(x_0)) \\ &= h(x_1). \end{aligned}$$

- 4th sentence, "By hypothesis ... zero."
  By continuity, if $\lim_{n\to+\infty} x_n = p$, then

$$\begin{aligned} f(p) &= \lim_{n\to+\infty} f(x_n) \\ &= \lim_{n\to+\infty} x_{n+1} \\ &= p. \end{aligned}$$

  Hence $p = 0$.

- 9th sentence, "Extend ... $x_2 \leq x \leq x_1$."
  The extension of h to $[x_2, x_1]$ is well-defined since $f^{-1}(x) \in [x_1, x_0]$ for $x \in [x_2, x_1]$.[9]

- 12th sentence, "Continuing ... $h = g^n \circ h \circ f^{-n}$."
  At the very end, $h = g^{\circ n} \circ h \circ f^{-n}$ is more consistent with the adopted notation.

- last sentence
  Let us verify for $[x_3, x_2]$. It follows from:

$$h = g^{\circ 2} \circ h \circ f^{-2} \text{ implies that } h \circ f = g \circ h.$$

  In fact, on the one hand, $f^{-1}(x) \in [x_2, x_1]$ for $x \in [x_3, x_2]$. Furthermore, concerning $[x_2, x_1]$, $g \circ h = h \circ f$ holds. On the other hand, if $x \in [x_3, x_2]$, then

$$\begin{aligned} (h \circ f)(x) &= \left(g^{\circ 2} \circ h \circ f^{-2} \circ f\right)(x) \\ &= \left(g \circ (g \circ h) \circ f^{-1} \circ f^{-1} \circ f\right)(x) \\ &= \left(g \circ (g \circ h) \circ f^{-1}\right)(x) \\ &= g\left((g \circ h)\left(f^{-1}(x)\right)\right) \\ &= g\left((h \circ f)\left(f^{-1}(x)\right)\right) \\ &= \left(g \circ h \circ f \circ f^{-1}\right)(x) \\ &= (g \circ h)(x). \end{aligned}$$

==

Comments/Erratum, p. 94

- 5th sentence,

  "For $-2 \leq c \leq 1/4$, the iterate of any point in $[-p_+, p_+]$ remains in the interval $[-p_+, p_+]$."

  Let $|x| \leq p_+$. On the one hand,

$$x^2 + c \leq p_+^2 + c = \frac{1}{4} + \frac{\sqrt{1-4c}}{2} + \frac{1}{4} - c + c = p_+.$$

[9] $f^{-1}(x_2) = x_1$, $f^{-1}(x_1) = x_0$ and f^{-1} is a continuous strictly increasing function on $[x_2, x_1]$.

On the other hand, since $x^2 + c \geq -p_+$ for $c \geq 0$ (because $-p_+$ is negative) and $x^2 + c \geq c$, it suffices to prove that $c \geq -p_+$ for $-2 \leq c < 0$. In fact, suppose otherwise. Hence

$$c < -\frac{1+\sqrt{1-4c}}{2} \Rightarrow -(2c+1) > \sqrt{1-4c},$$

which does not hold if $2c + 1 \geq 0$, that is, if $c \geq -\frac{1}{2}$. So suppose that $-2 \leq c < -\frac{1}{2}$. Therefore

$$\begin{aligned} -(2c+1) > \sqrt{1-4c} &\Rightarrow 4c^2 + 4c + 1 > 1 - 4c \\ &\Rightarrow 4c(c+2) > 0, \end{aligned}$$

which is a contradiction since $4c < 0$ and $c + 2 \geq 0$ for $-2 \leq c < -\frac{1}{2}$.

- 7th sentence, "To visualize **the** what is going on, ..."
  It should be just "To visualize what is going on, ...".

- 8th sentence, "The bottom of the graph will protrude below the bottom of the square."
  It follows since $c < -p_+$ for $c < -2$. In fact, suppose otherwise.

- last sentence
  A_2 is open since, under a continuous function, the inverse image of an open set (in the codomain) is always an open set (in the domain).

==
Erratum, p. 95, l. -3
In order to be consistent with a previous notation (p. 90, l. 1), $Q_c^{-\circ n}(A_1)$ should be just $Q_c^{-n}(A_1)$.
==
Errata/Comment, p. 100

- l. 1
  Replace **itenerary** by **itinerary**.

- l. 4
  As a matter of fact, by abuse of notation, S may used in place of **Sh** as can be seen on p. 87.

- ll. -1, -3 and -4
  In order to be consistent with a previous notation (p. 90, l. 1), $Q_c^{-\circ n}$ should be just Q_c^{-n}. Similarly, $Q_c^{-\circ(n-1)}$ should be changed.

==
Erratum, p. 101, **Theo. 4.6.2**
Change Q_c^n to $Q_c^{\circ n}$.

==
==
5
==
==
Suggestion, p. 103
The 4th sentence should end as in

$$\boxed{\text{"... in } I_k, k = 1, \ldots, N.\text{"}}$$

since the 3rd sentence considers I_k for $k = 1, \ldots, N-1$ separetely from I_N.
==
Comment, p. 105, **The push forward of a discrete measure**, last sentence
Using (5.2), we have

$$(F_*\mu)(I) = \sum_{y_\ell \in F^{-1}(I)} n\left(y_\ell\right).$$

==
Comment, p. 106

- (5.3) and (5.4)
  - x_k is used, twice, instead of just x, to make clear that the sum is over k;
  - If (5.4) holds, then we can put each of the two densities, ρ and σ, in the other's place by (5.3).
- **Back to the histogram**
  Suppose that $\rho \approx$ constant on I_k. Therefore

$$\begin{aligned} p\left(I_k\right) &= \int_{I_k} \rho(x) dx \\ &\approx \rho(x) \int_{\frac{k-1}{N}}^{\frac{k}{N}} dt = \rho(x) \cdot \frac{1}{N}. \end{aligned}$$

  For N large enough, the previous supposition is a possible one.

==
Comments, p. 108

- **First proof of (i)**
  Note that

$$\begin{aligned} \rho(F(x)) &= \frac{1}{\pi\sqrt{4x(1-x)(1-4x(1-x))}} \\ &= \frac{2}{\pi 4|1-2x|\sqrt{x(1-x)}} \\ &= \frac{1+1}{\pi 4|1-2x|\sqrt{x(1-x)}} \\ &= \frac{\frac{1}{\pi\sqrt{x(1-x)}}}{4|1-2x|} + \frac{\frac{1}{\pi\sqrt{(1-x)x}}}{4|-1+2x|} \\ &= \frac{\rho(x)}{|F'(x)|} + \frac{\rho(1-x)}{|F'(1-x)|}. \end{aligned}$$

- **Second proof of (i)**, 4th sentence, $\boxed{\text{"In other words ... } \nu \text{ has density } \rho(x) \equiv 1.\text{"}}$
  In fact, the measure of I is just the length of I, which means that $\nu(I) = \int_I dx$.

==

Comment, p. 109, **Proof of (ii)**
At the very end, we are dealing with a constant sequence of Riemann sums which converges to both $\int f(t)dt$, whose convergence follows from the continuity of f, and the constant $f(x)$. Then

$$f(x) = \int f(t)dt$$

is just a consequence of the uniqueness of the limit of any sequence.
==
Comment, p. 110, **What about (iii)?**
To fix ideas, let $(X, \mathcal{B}, \mu)$ be a probability space. Therefore:

- X is a set of possible outcomes;
- $\mathcal{B}$ is a set of events, which are sets of outcomes;
- μ assigns probabilities to each event in $\mathcal{B}$.

Furthermore, let $T : X \to X$ be an ergodic transformation with respect to an invariant measure, μ. Then:

- $T_* \mu = \mu$;
- $\forall A \in \mathcal{B}$ such that $T^{-1}(A) = A$, either $\mu(A) = 0$ or $\mu(A) = 1$.

An ergodic theorem describes the limiting behavior of a sequence

$$\frac{1}{n} \sum_{i=0}^{n-1} f \circ T^i$$

(as $n \to \infty$) and depends on the function f (for example, it may be assumed that f is integrable, or square integrable (L^2), or continuous) and the type of convergence used in the theorem (for example, pointwise, L^2 or uniform convergence).
The **Birkhoff Ergodic Theorem** deals with pointwise convergence of the previous sequence of partial sums for integrable functions f if $\mu(X)$ is finite.[10] It says that

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} f\left(T^i(x)\right) = \hat{f}(x) \text{ for almost all } x$$

with

$$\hat{f} = \frac{1}{\mu(X)} \int f\mu.$$

==
Comment, pp. 110–111, "Indeed, applying ... as $n \to \infty$"
Since U is an isometry,

$$\frac{||U^n f||}{n} = \frac{||f||}{n} \to 0$$

as $n \to \infty$.
==
Comments, p. 112

- **Lemma 5.3.2**
  The denseness of $D \cup I$ is a consequence of a well-known result (in Functional Analysis) which says that the orthogonal complement of the orthogonal complement of $S \subset H$ is the closure of S. Therefore

$$\begin{aligned} \overline{D \cup I} &= (D \cup I)^{\perp\perp} \\ &= 0^{\perp} \\ &= H. \end{aligned}$$

[10] For a probability space, $\mu(X) = 1$.

- **Lemma 5.3.3** and **Lemma 5.3.4**
  Let S be the set of elements for which the limit in (5.8) exists. Consider that $f \in H$. So the denseness of S guarantees the existence of a sequence in S which converges to f. Then the closeness of S implies that $f \in S$. Therefore $S = H$.

==

Erratum, p. 113, (5.9)
A dx is missing right after the integrand.

==

Erratum, p. 121
The 2nd **Proof o Prop. 5.4.3** is, in fact, the **Proof o Prop. 5.4.4**.

==
==
6
==
==
Suggestion, p. 129, **6.1**, 1st sentence
Either change $\mathbb{R}$ to $\mathbb{R}^+ \cup \{0\}$, twice, or put the word 'non-negative', which is placed right before the word 'real', right before the word 'function'.
==
Suggestion, p. 130, **Open balls, the topology on a pseudo-metric space** X**.**, 1st sentence
Concerning "..., the **(open) ball of radius** r about x ...", the word 'about' should be written in boldface.
==
Comment, p. 131, the sentence that precedes **An example.**,

"Clearly a Lipschitz map is uniformly continuous."

In fact, for $\epsilon > 0$, consider $\delta \leq \frac{\epsilon}{C+1}$.
==
Comments, p. 132

- **Identifying points at zero distance.**, "It is also open, that is, it maps open sets to open sets."
  If A is open and $\overline{\{x\}} \in A/R$, then $B_r\left(\overline{\{x\}}\right) \subset A/R$ for each $B_r(x) \subset A$.[11] Therefore A/R is open.
- **Prop. 6.1.1**
  $F = p \circ f$ where $p : f(Y) \ni x \mapsto \overline{\{x\}} \in f(Y)/R$.

==
Comment, p. 133, **Complete metric spaces.**, $d\left(\{x_n\}, \{y_n\}\right) := \lim_{n\to\infty} d\left(x_n, y_n\right)$
Since $\mathbb{R}$ is complete, the Cauchy sequence $\{d\left(x_n, y_n\right)\}$ converges.[12] Therefore $d_{X_{seq}}$ is well-defined.
==
Comment, p. 134, last sentence before the definition of a **contraction** map,

"The inequality will continue to hold for this value of C which is known as the Lipschitz constant of f and denoted by $\text{Lip}(f)$."

In fact, on the one hand,

$$d_Y(f(x), f(x)) = 0 = \text{Lip}(f) d_X(x, x) \quad \forall x \in X.$$

On the other hand, as a lower bound of $\{C \,|\, d_Y(f(x_1), f(x_2)) \leq C d_X(x_1, x_2)\, \forall x_1, x_2 \in X\}$,

$$\frac{d_Y(f(x_1), f(x_2))}{d_X(x_1, x_2)} \leq \text{Lip}(f) \quad \forall x_1, x_2 \in X \text{ with } x_1 \neq x_2.$$

==
Erratum/Comment, p. 136

- **Cor. 6.3.1**, *Proof*
  Prop. should be Theo., twice.

[11] In fact, since $d\left(\overline{\{x\}}, \overline{\{y\}}\right) = d(x, y)$, $d\left(\overline{\{x\}}, \overline{\{y\}}\right) < r$ whenever $d(x, y) < r$.

[12] The basic idea to prove that the sequence is Cauchy is based on

$$\begin{aligned} |d\left(x_n, y_n\right) - d\left(x_m, y_m\right)| &= |d\left(x_n, y_n\right) - d\left(y_n, x_m\right) + d\left(y_n, x_m\right) - d\left(x_m, y_m\right)| \\ &\leq |d\left(x_n, y_n\right) - d\left(y_n, x_m\right)| + |d\left(y_n, x_m\right) - d\left(x_m, y_m\right)| \\ &\leq d\left(x_n, x_m\right) + d\left(y_n, y_m\right), \end{aligned}$$

which follows from

$$|d(x,z) - d(z,y)| \leq d(x,y) \iff d(x,z) - d(y,z) \leq d(x,y),\; d(y,z) - d(x,z) \leq d(y,x).$$

for all x, y, z in X.

- **6.4**
Since (S, d_S) and (X, d_X) are metric spaces, then $(S \times X, d_{S\times X})$ is a metric space with

$$d_{S\times X}((s_1, x_1), (s_2, x_2)) = d_S(s_1, s_2) + d_X(x_1, x_2)$$

for all points (s_1, x_1) and (s_2, x_2) in $S \times X$.

==

Errata/Comment, p. 137

- *Proof* of **Theo. 6.4.1**
  - In the middle of the proof, a) is missing right before $\leq$;
  - At the very end of the proof, change Prop. to Cor..

- "..., wish to conclude the existence of an inverse to F, ...", l. -10
Isn't a 1st person (subject) pronoun missing before the word wish?

- "... the mean value theorem ...",[13,14] ll. -5 and -4

If V and W are normed linear spaces, and if F is from $B_r(\alpha) \subset V$ to W with F differentiable and $||dF_\beta|| \leq \epsilon$ for every $\beta \in B_r(\alpha)$, then $||F(\beta + \xi) - F(\beta)|| \leq \epsilon||\xi||$ whenever $\beta, \beta + \xi \in B_r(\alpha)$.

==

Erratum/Comments/Suggestion, **Theo. 6.5.1** and its *Proof*, p. 138

- "..., $\text{Lip}[v] < \lambda < 1$.", 1st sentence of the proof
Either the first $<$ should be the equal sign $=$ or the $=$ of (6.2) on p. 137 should be the inequality sign $<$.

- 3rd sentence of the proof
In fact,

$$\begin{aligned}(\text{id} + v) \circ (\text{id} + w) = \text{id} &\Leftrightarrow \text{id} \circ (\text{id} + w) + v \circ (\text{id} + w) = \text{id}\\ &\Leftrightarrow \text{id} + w + v \circ (\text{id} + w) = \text{id}\\ &\Leftrightarrow w = -v \circ (\text{id} + w).\end{aligned}$$

- "Let X be the space of continuous maps of $\overline{B_s(0)} \to E$ ...", 4th sentence of the proof
Use either

"Let X be the space of continuous maps u from $\overline{B_s(0)}$ to E ..."

or

"Let X be the space of continuous maps $u : \overline{B_s(0)} \to E$ ...".

- "Then X is a complete metric space relative to the sup norm, ...", 5th sentence of the proof
The completeness of X follows from a classical result of Functional Analysis provided that E is complete. In fact, by hypothesis, E is a Banach space.

==

Errata/Suggestions/Comment, p. 139

- **The setup.**, 3rd sentence
"... and y ranges over <u>and</u> open ball ..." should be "... and y ranges over <u>an</u> open ball ...".

[13]See **ADVANCED CALCULUS** (revised edition) by Lynn H. Loomis and Shlomo Sternberg, **Theo. 7.4**, p. 149.
[14]For later use in **Chapter 8**, call it MVT.

- **The method.**, 2nd, 3rd and last sentences
  - "... map from $A \times B \to Y$ ..." should be "... map from $A \times B$ to Y ...";
  - $K(x, y)y = y$ should be $K(x, y) = y$;
  - "The mean value theorem ...", l. -1
    See the previous footnote.

==

Suggestion/Errata, p. 140
Concerning the underlined text:

- "... $||K(x, y_1) - K(x, y_2)|| \leq ||y_1 - y_2||\underline{C}, \quad 0 \leq C < 1, \ \forall \underline{s \in S}$, ..." should be

$$\text{“... } ||K(x, y_1) - K(x, y_2)|| \leq \underline{C}||y_1 - y_2||, \quad 0 \leq C < 1, \ \forall \underline{y_1, y_2 \in B}, \text{ ...”};$$

- "... $y_x \in B$ *such that* $K(x, \underline{y}) = y_x$, ..." should be "... $y_x \in B$ *such that* $K(x, \underline{y_x}) = y_x$, ...".

==

Erratum, p. 141, l. 4
Concerning "..., and the the right ...", delete the extra 'the'.

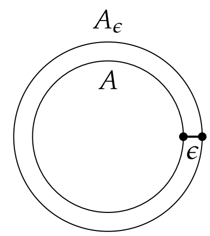
==

Comment/Errata, p. 142

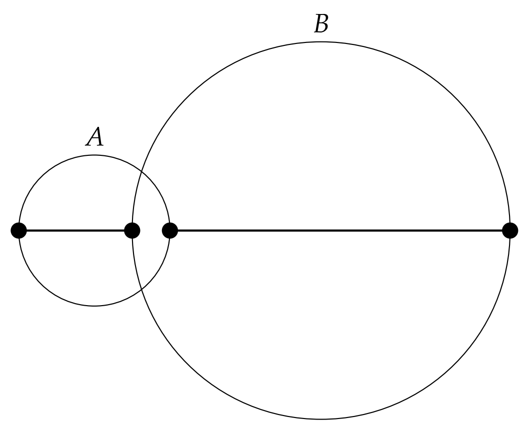
- 1[st] sentence
  - Notice that F is bounded on $\overline{L \times U}$, which contains $L \times U$;
  - Based on the rest of the proof from here, $L = J$.
- At the very end of the proof, $\delta(m + cr) \leq 1$ should be $\delta(m + cr) \leq r$.

==
==
7
==
==
Comments, pp. 143–4

- Illustration for A_ϵ (where A is a closed ball):



- $d(A,B)$ and $h(A,B)$ are well-defined. In fact, the maxima are achieved since both A and B are bounded.
- Illustration for $d(A,B) \neq d(B,A)$ (where A and B are closed balls):



- **Proof of (7.3).**
  Using set-builder notation and propositional logic, we have
  $$\{\epsilon \mid A \subset C_\epsilon, B \subset D_\epsilon, C \subset A_\epsilon \text{ and } D \subset B_\epsilon\} \subset \{\epsilon \mid A \cup B \subset (C \cup D)_\epsilon \text{ and } C \cup D \subset (A \cup B)_\epsilon\}.$$
  Therefore, the infimum of the latter is less than or equal to the infimum of the former.

==
Comment/Suggestion, p. 145

- **A sketch of the proof of completeness.**
  $A \neq \emptyset$. In fact, since $\{A_n\}$ is Cauchy with respect to h, by selecting a subsequence if necessary, we may assume that, for each index n,
  $$h(A_n, A_{n+1}) < 2^{-n},$$
  that is,
  $$A_n \subset (A_{n+1})_{2^{-n}} \quad \text{and} \quad A_{n+1} \subset (A_n)_{2^{-n}}.$$
  Then, for any natural number N, there is a sequence $\{x_n\}_{n \geq N}$ in (X,d) such that $x_n \in A_n$ and $d(x_n, x_{n+1}) < 2^{-n}$ for each index $n \geq N$. Any such sequence is Cauchy with respect to d and thus converges to some $x \in X$ (due to the completeness of (X,d)).
- **7.1.1**, at the very end of the 2nd line
  'K' should be 'K'.

==
Suggestion/Errata, p. 146, **Theorem 7.2.1**

- Put 'X' right after '*space*';
- '*Lifschitz*' should be '*Lipschitz*';
- '$K_n(a)$' should be '$K_n(A)$';
- At the very end, an '*i*' is missing in '*Lipschtz*'.

==
Errata, p. 148, **7.3.2**

- 1st line after $A_1 = TA_0$
  "... be deleting ..." should be "... by deleting ...";
- 4th line after $A_1 = TA_0$
  'sussive' should be 'successive';
- 7th line after $A_1 = TA_0$
  Put '=' right after 'B_0'.

==
Comment, p. 153, ll. -3 and -2, "..., $M_{t,\epsilon}$ is non-decreasing as $\epsilon \to 0$, ..."
In fact, if $\epsilon \leq \delta$, then a countable cover $\mathcal{B}_\epsilon$ by balls of radius at most ϵ is also a countable cover $\mathcal{B}_\delta$ by balls of radius at most δ. Thus the set of all $m_t(\mathcal{B}_\epsilon)$ is a subset of the set of all $m_t(\mathcal{B}_\delta)$. Therefore

$$M_{t,\epsilon} \geq M_{t,\delta}.$$

==
==
8
==
==

Comment, p. 159, 1st sentence

The theory presented here is built on complete metric spaces, which were defined in chapter 6. So, it is worth recalling that, a Banach space is a normed linear space which is complete with respect to the metric derived from its norm.

==

Erratum/Comment, p. 160

- The equation

$$L(u) := Au - u \circ (A + \phi)$$

is displayed twice. So (8.6) should be the equation number to

$$L(u) = \phi - \psi(\mathrm{id} + u)$$

and the 2nd sentence that starts with

'So we wish ...'

should be deleted.

- l. -1

$$\begin{aligned} \left|\left|L^{-1}\right|\right| \cdot \mathrm{Lip}[\psi] &< \frac{||A^{-1}||}{(1-a)} \cdot \epsilon \\ &< \frac{||A^{-1}||}{1-a} \cdot \frac{1-a}{||A^{-1}||} \\ &< 1 \end{aligned}$$

$$\Downarrow$$

$$2\left|\left|L^{-1}\right|\right| \cdot \mathrm{Lip}[\psi] - \left|\left|L^{-1}\right|\right| \cdot \mathrm{Lip}[\psi] < 1$$

$$\Downarrow$$

$$\begin{aligned} \left|\left|L^{-1}\right|\right| \cdot \mathrm{Lip}[\psi] &< \frac{||L^{-1}|| \cdot \mathrm{Lip}[\psi] + 1}{2} := c \\ &< 1. \end{aligned}$$

==

Errata/Comment, p. 161

- At the very begining, there is a subjacent hypothesis for the conclusion of the existence and uniqueness of the solution to (8.5): X is complete.

- $\left|\left|A^{-1}\right|\right|$, right before the sentence which ends in (8.8), should be $\left|\left|M^{-1}A^{-1}\right|\right|$.

- **The map $A + \phi$ is injective.**

  - The 1st '$\geq$', l. -5, comes from

$$\begin{aligned} ||x|| &= \left|\left|A^{-1}Ax\right|\right| \\ &\leq \left|\left|A^{-1}\right|\right| ||Ax||; \end{aligned}$$

– The 2nd '$\geq$', l. -3, comes from

$$\begin{aligned}||Ax + \phi(x) - Ay - \phi(y)|| + \text{Lip}[\phi]||x-y|| &\geq ||A(x-y) + \phi(x) - \phi(y)|| + ||\phi(x) - \phi(y)|| \\ &\geq ||A(x-y) + \phi(x) - \phi(y) + \phi(y) - \phi(x)|| \\ &\geq ||A(x-y)|| \\ &\geq \frac{1}{||A^{-1}||}||x-y||;\end{aligned}$$

– The 3rd '$\geq$', l. -2, comes from

$$\begin{aligned}\text{Lip}[\phi] < \frac{1-a}{||A^{-1}||} &\Rightarrow \frac{1-a}{||A^{-1}||} - \text{Lip}[\phi] > 0 \\ &\Rightarrow \frac{a}{||A^{-1}||} + \frac{1-a}{||A^{-1}||} - \text{Lip}[\phi] > \frac{a}{||A^{-1}||};\end{aligned}$$

– '(8.4.', l. -1
Put a parentheses right after '4'.

==

Errata/Comments, p. 162

- **The map $A + \phi$ is surjective with continuous inverse.**
  – Remove the period for estimate (8.4).
  – The contraction argument follows from

$$\begin{aligned}\left|\left|A^{-1}(y - \phi(x_1)) - A^{-1}(y - \phi(x_2))\right|\right| &= \left|\left|A^{-1}(\phi(x_2) - \phi(x_1))\right|\right| \\ &\leq \left|\left|A^{-1}\right|\right| \, ||\phi(x_2) - \phi(x_1)|| \\ &\leq \left|\left|A^{-1}\right|\right| \text{Lip}[\phi]||x_2 - x_1|| \\ &\leq \left|\left|A^{-1}\right|\right| \epsilon||x_2 - x_1|| \\ &\leq (1-a)||x_1 - x_2||.\end{aligned}$$

  – 3rd line after estimate (8.4)
  Change "... $A + \phi$ a ..." to "... $A + \phi$ is a ...".

- Concerning the item that completes the proof of the lemma, notice that:
  – $NN^{-1}f = A^{-1}A_s f \circ (A+\phi)^{-1} \circ (A+\phi) = f$ and $N^{-1}Nf = A_s A^{-1} f \circ (A+\phi) \circ (A+\phi)^{-1} = f$;
  – Via the sup norm on Y,

$$\left|\left|N^{-1}f\right|\right| = \sup\left\{\left|\left|N^{-1}fx\right|\right| \,:\, x \in E\right\} \text{ and } ||f|| = \sup\left\{||fy|| \,:\, y \in E\right\}.$$

  Hence, since

$$\begin{aligned}\left|\left|N^{-1}fx\right|\right| &= \left|\left|A_s f \circ (A+\phi)^{-1}x\right|\right| \\ &\leq ||A_s|| \cdot \left|\left|f \circ (A+\phi)^{-1}x\right|\right| \\ &\leq a||f||,\end{aligned}$$

  it follows that $\left|\left|N^{-1}f\right|\right| \leq a||f||$.

==

Comment, pp. 162–3

Concerning the convergence, the geometric series corresponding to $\left(I - N^{-1}\right)^{-1}$ used $\left|\left|N^{-1}\right|\right|| \leq a$, which was obtained by

$$N^{-1}f = A_s f \circ (A+\phi)^{-1},\ ||A_s|| \leq a \Rightarrow \left|\left|N^{-1}f\right|\right| \leq a||f||.$$

Similarly, the geometric series corresponding to $(I-Q)^{-1}$uses $||Q||| \leq a$, which is obtained by

$$Qg = A_u^{-1} g \circ (A+\phi),\ \left|\left|A_u^{-1}\right|\right| \leq a \Rightarrow ||Qg|| \leq a||g||.$$

==

Errata/Comment, p. 163

- Before the completion of the proof of the 1st part of **Theorem 8.1.1**:
  - $||M_u||$ and $||M||$ should be $\left|\left|M_u^{-1}\right|\right|$ and $\left|\left|M^{-1}\right|\right|$, respectively;
  - $M^{-1} = M_s^{-1} \oplus M_u^{-1}$ has the maximum property, that is,

$$\left|\left|M^{-1}\right|\right| = \max\left\{\left|\left|M_s^{-1}\right|\right|, \left|\left|M_u^{-1}\right|\right|\right\}.$$

- Right after the completion of the proof of the 1st part of **Theorem 8.1.1**:
  Since $X \ni u \mapsto \mathcal{F}_0(u) := u(0) \in E$ is a continuous linear functional, if $Ku = u$, then

$$K^n u_0 \to u \Rightarrow \mathcal{F}_0(K^n u_0) \to \mathcal{F}_0(u) = u(0)$$

  for any point $u_0 \in X$.[15] In particular, if $u_0(0) = 0$, then $K^n u_0(0) = 0$ for each index n. So $u(0) = 0$.

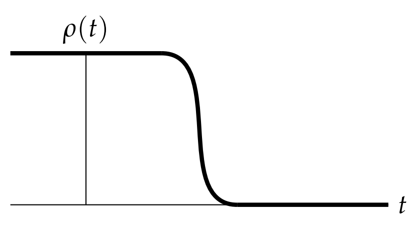
- 1st line of **8.1.2**
  Remove 'differentiable,'.

==

Errata/Suggestions/Comment, p. 164

- **Plan of the proof.**, 1st line after '$K > 2$.'
  Remove the first 'the'.

- **Checking the Lipschitz constant.**
  - The next figure illustrates the graph of a possible $\rho(t)$ and has three parts:
    1. a horizontal semi-line with endpoint $(0.5, 1)$;
    2. a curve from $(0.5, 1)$ to $(1, 0)$ which leaves at an angle of 0° and arrives at an angle of 180°;
    3. a horinzontal semi-line with endpoint $(1, 0)$.



    Notice that $\rho(t) \leq 1$ for all $t \in \mathbb{R}$.
  - l. -2 follows from the MVT (see my footnote 14) since

$$\left|\rho'(t)\right| < K\ \forall t \in \mathbb{R},\ \left|\ ||x_1|| - ||x_2||\ \right| \leq ||x_1 - x_2||,\ ||d\phi_x|| < \frac{\epsilon}{2K}\ \forall x \in B(0,r) \text{ and } \rho\left(\frac{||x_2||}{r}\right) \leq 1.$$

  - l. -1
    Change '$||x_1 - x_2|$' to '$||x_1 - x_2||$'.

[15] See the *Proof* of **Theorem 6.3.1**, p. 135.

==
Erratum, p. 165, last line before (8.10)
Change "... are ..." to "... be ...".
==
Comments/Suggestion/Errata, p. 166

- **More details in the linear case.**
  - As far as $||A^n x|| = ||A^n x_s|| + ||A^n x_u||$ is concerned, if E is a Banach space with a norm $||\cdot||_E$, then

    $$x_s + x_u \mapsto ||x_s + x_u||_\oplus := ||x_s||_E + ||x_u||_E$$

    defines a norm on E.
  - Consider $c = \frac{1}{a}$ since

    $$||x_u|| = \left|\left|A_u^{-1}A_u x_u\right|\right| \leq \left|\left|A_u^{-1}\right|\right| \cdot ||A_u x_u|| \leq a\,||A_u x_u|| \Rightarrow ||Ax_u|| \geq \frac{1}{a}\,||x_u||\,.$$

    Furthermore, observe that $c > 1$.

- **Back to the general case.**
  - 3rd sentence
    "... and homeomorphism ..." should be "... and **a** homeomorphism ...".
  - 4th sentence
    Firstly, change:
    * 'U' to 'B_r';
    * '$h(x) \in S(A)$' to '$h(x) \in S \cap V$'.

    Now, notice that r and V can be taken such that V is bounded.[16] Then, as B_r in the linear case,

    $$A^n y \in V \ \ \forall n \geq 0 \Leftrightarrow y \in S \cap V.$$

    Therefore the 4th sentence can be rewritten as

    $$\begin{aligned} f^n(x) \in B_r \ \forall n \geq 0 &\Leftrightarrow A^n h(x) \in V \ \forall n \geq 0 \\ &\Leftrightarrow h(x) \in S \cap V \\ &\Leftrightarrow A^n h(x) \to 0. \end{aligned}$$

    (Concerning the last $\Leftrightarrow$, the proof of $\Rightarrow$ follows from (8.11), p. 165, since $S = W^s(0, A)$. On the other hand, since $h(x) \in V$ by definition, it remains to show that $h(x) \in S$ in order to prove $\Leftarrow$. In fact, if $h(x) = y_s \oplus y_u$, then $||A^n y_s|| + ||A^n y_u|| \to 0$. Hence $y_u = 0$, which means that $h(x) = y_s \in S$.)
  - l. -2
    In fact, H is the restriction of h^{-1} to $S \cap V$.

==
Errata/Comment, p. 167

- **An important remark.**
  - 2nd line before (8.14)
    "... $p \in f^{-n}[B_r^s(p)]$ ..." should be "... $x \in f^{-n}[B_r^s(p)]$ ...".
  - 1st line after (8.14)
    At the end of p. 166, it was proved that $B_r^s(p)$ is a topological submanifold via h, which is a homeomorphism. Then

    $W^s(p)$ is a topological submanifold

    by (8.14). **Theorem 8.2.1** confirms it with a map which is Lipschitz, a stronger condition than continuity.

[16] Consider $r' < r$, $V' := h\left(\overline{B_{r'}}\right)$ and the appropriate restriction of h.

- 4th line after (8.14)
  At the very end, change 'Her' to 'Here'.
- 6th line after (8.14), "... in[**?**,Shub]"
  * a space between 'in' and '[' is missing;
  * the bibliografical reference related to '**?**' is missing;
  * a period is missing right after ']'.
- l. -1
  At the very end, a 'c' is missing in 'charater'.

• **The Lipschitzian case.**, 5 lines before (8.15)
Change 's' to 'r'.

===

Errata/Suggestions/Comment, p. 168

• **Theo. 8.2.1.**

- l. 1
  Concerning $\epsilon(a)$ and $\delta(a, \epsilon, r)$, change 'a' to 'c' (since the bound a, p. 167, is replaced by c from now on).
- l. 3
  '$g : E_u(r) \to E_s(r)$' should be '$g : U(r) \to S(r)$'.
- l. 6
  Change 'f^{-1}' to 'f'.
- l. 9
  * Change '$U(p)$' to '$U(r)$';
  * It is worth noting that, as a "cograph",

$$\text{graph}(g) \subset S(r) \times U(r),$$

which is identified with

$$S(r) \oplus U(r) \subset S \oplus U.$$

Furthermore, graph(g) has the subspace topology inherited from E and it has a single coordinate chart given by

$$(\text{graph}(g), p_u)$$

where p_u is the projection onto U. On the other hand, $p_u^{-1}(x) = (g(x), x)$ is continuous if and only if g is continuous, which is the case here since g is Lipschitz by *(i)*.
To summarize: graph(g) is a topological submanifold.

• **The idea of the proof.**

- l. 4
  A comma is used to represent 'such that', which is usually represented by a vertical bar (|) or a slash (/) or a colon (:). So, it should be, for example,

$$\text{graph}(v) = \{(v(x), x) \mid x \in U(r)\}.$$

- l. 6
  The shorthand notation represents

$$f[\text{graph}(v)] = \{f(v(x), x) \mid x \in U(r)\} = \{(f_s(v(x), x), f_u(v(x), x)) \mid x \in U(r)\}.$$

- l. -9
  Change "... map of $U(r) \to E$ ..." to "... map $U(r) \to E$ ..." or "... map of $U(r)$ to E ...".
- l. -7
  Change "... map of $U(r) \to U$." to "... map $U(r) \to U$." or "... map of $U(r)$ to U."

- l. -6
  It should be

$$f[\text{graph}(v)] = \left\{ \left(f_s \circ (v, id)([f_u \circ (v, id)]^{-1}(y)), y \right) \,\middle|\, y \in U(r) \right\} = \text{graph} \left[G_f(v) \right].$$

==

Errata/Suggestion, p. 169

- 2 lines before **Lemma 8.2.1**
  "... direc ..." should be "... direct ...".
- **Lemma 8.2.2**
  - 2 lines before (8.18)
    '$f_u \circ (v, id) : E_u(r) \to E_u$' should be '$f_u \circ (v, id) : U(r) \to U$'.
  - 2 lines after (8.18)
    - '$f_u - A_u$' should be '$f_u \circ (v, id) - A_u$';
    - The last '$<$' should be '$\leq$'.
  - 3 lines after (8.18)
    "By the Lipschitz implicit function theorem ..." should be

    "By the Lipschitz inverse function theorem, Theorem 6.5.1, ...".

==
==
9
==
==
Erratum/Comment, p. 176, **9.1.1**

- 1st sentence
  "... matrix square ..." should be "... square matrix ...".
- last two sentences
  Since $T \geq 0$ by definition, $T^l \geq 0$ for each non-negative integer l. Therefore, once K is big enough, the irreducibility hypothesis guarantees that, for any i, j, there is a $k = k(i, j)$ such that

$$\left[(I+T)^K\right]_{ij} = \left[\sum_{\substack{l=0 \\ l \neq k}}^{K} \binom{K}{l} \left(T^l\right)_{ij}\right] + \binom{K}{k} \left(T^k\right)_{ij}$$

  is positive.

==
Comment, p. 176, **Theorem 9.1.1.***5*
As usual, '$S \leq T$' means that $T - S$ is a non-negative matrix.
==
Comments, p. 177

- (9.1)
  Consider $j \in \{1, \ldots, n\}$ and

$$\frac{(Tz)_j}{z_j} = \min\left\{\frac{(Tz)_i}{z_i} \,:\, 1 \leq i \leq n \text{ with } z_i \neq 0\right\}.$$

  Since $z \geq 0$ and, by definition, $T \geq 0$, it follows that $(Tz)_i \geq 0$ for each $i \in \{1, \ldots n\}$. So, whatever i we choose,

$$\frac{(Tz)_j}{z_j} \cdot z_i \leq (Tz)_i$$

  holds, even if $z_i = 0$. Thus

$$\frac{(Tz)_j}{z_j} z \leq Tz.$$

  Then $\{s \,:\, sz \leq Tz\} \neq \emptyset$ and

$$\frac{(Tz)_j}{z_j} = \max\{s \,:\, sz \leq Tz\}.$$

  In fact, suppose otherwise. Hence, there is a number s such that $Tz \geq sz$ and $s > \frac{(Tz)_j}{z_j}$. Therefore

$$\begin{aligned}(Tz)_j &\geq sz_j \\ &> (Tz)_j\,!\end{aligned}$$

- 1st sentence after (9.1)
  Whenever convenient, as far as $L(z)$ is concerned, we may restrict our attention to the case where $z \in C$. As an example, a maximum value of L on C is, in fact, the maximum value of L on all of Q.

==
Erratum/Comment, p. 178

- l. 4
  '$|\lambda|y_i|$' should be '$|\lambda||y_i|$'.

- **Showing that $\lambda_{\max}(T^\dagger) = \lambda_{\max}(T)$.**
  This result is clearly correct since taking transpose does not change the eigenvalues of a matrix. However, in order to show that the result holds, the book's argument is made in such a way that it is also used for **Proving the first two assertions in item 4 of the theorem.**, p. 179.

==
Comments/Errata/Suggestion, p. 179

- **Proving the first two assertions in item 4 of the theorem.**
  - 1st paragraph
    * The subjacent hypothesis is that w is as described on p. 178;
    * At the very end, remove the extra 'then' in "... then then ...".
  - 2nd paragraph
    Change '$n-1$' to 'k'.
  - last paragraph
    * Concerning "If $\mu = \lambda_{\max}$ then $w^\dagger(Ty - \lambda_{\max}y) = 0$ but $Ty - \lambda_{\max}y \le 0$ ...", only '$Ty - \lambda_{\max}y \le 0$' comes from the hypothesis '$\mu = \lambda_{\max}$', whereas '$w^\dagger(Ty - \lambda_{\max}y) = 0$' has to do with the first displayed formula of p. 179;
    * '4)' and '2)' should be '4.'and '2.'.

- In order to maintain the notation compatible with **Theorem 9.1.1.**_6_ and p. 180, in the last two paragraphs,

  'T_i' and 'Λ_i' should be '$T_{(i)}$' and '$\Lambda_{(i)}$'.

==
Comments/Erratum, p. 180

- 1st displayed formula
  See $\lambda I - T$ as a parametric curve with parameter λ, det as a real function of n^2 real variables and $\det(\lambda I - T)$ as the composite function $(\det \circ \text{curve})(\lambda)$. Therefore

$$\begin{aligned}\frac{d}{d\lambda}\det(\lambda I - T) &= \nabla \det(\lambda I - T) \cdot \frac{d}{d\lambda}\text{curve}(\lambda) \\ &= \nabla \det(\lambda I - T) \cdot I,\end{aligned}$$

  where both factors can be seen as vectors in $\mathbb{R}^{n^2}$. Furthermore, notice that

$$(\nabla \det(\lambda I - T))_{ii} = \frac{\partial}{\partial \lambda_i}\det(\lambda I - T)$$

  for each $i = 1, \ldots, n$.

- **Showing that $\lambda_{\max}$ has algebraic (and hence geometric) multiplicity one.**
  - 1st sentence
    Since T is an $n \times n$ matrix, if $\lambda_{i,1}, \ldots, \lambda_{i,n-1}$ are the eigenvalues of $T_{(i)}$ and $j \in \{1, \ldots, n-1\}$, then:
    * $\left|\lambda_{i,j}\right| < \lambda_{\max}$ by **Theorem 9.1.1.**_6_;
    * $\det\left(\lambda I - T_{(i)}\right) = \prod_{j=1}^{n-1}\left(\lambda - \lambda_{i,j}\right)^{n_j}$, where n_j is a positive integer which depends on $\lambda_{i,j}$.

    Therefore, since the roots of a polynomial with real coefficients occur in conjugate pairs,

$$\det\left(\lambda_{\max} I - T_{(i)}\right) > 0.$$

– '2)' should be '2.'.[17]

- **Proof of the last assertion of the theorem.**, 1st sentence
  Clearly, $T^k y = \lambda^k y$ if $Ty = \lambda y$.

==
Errata, p. 181

- l. 2
  A period is missing right after 'x';
- l. -6
  A double apostrophe is missing right after 'v_j'.

==
Comment/Erratum, p. 182

- **Paths and powers.**, 2nd paragraph
  (The result follows by induction on ℓ.) The case $\ell = 1$ is trivial. Now suppose that the result holds for some $\ell > 1$, so that the entries of A^ℓ are as claimed. Consider any path of lenght $\ell + 1$ from v_j to v_i. Hence there is an adjacent vertex, v_k, to v_i on this path. Delete v_i. So the remaining path is a path of lenght ℓ from v_k to v_j. By induction, the number of such paths is given by $\left(A^\ell\right)_{kj}$. On the other hand, each such v_k corresponds to a 1 for A_{ik}. Now consider

$$\begin{aligned}\left(A^{\ell+1}\right)_{ij} &= \left(AA^\ell\right)_{ij} \\ &= \sum_{k=1}^{n} A_{ik} \left(A^\ell\right)_{kj}\end{aligned}$$

  with A of order n.

- **9.2.2**, l. 6
  'postive' should be 'positive'.

==
Comments/Erratum, p. 183

- 2nd paragraph, 1st sentence, "The paths ... have total lengths at most $3(n-1)$."
  In fact, $V = \{v_1, \ldots, v_n\}$.

- **Proof of the lemma.**

  – Concerning the equivalence '$\Longleftrightarrow$', it may seem that the '$\Longrightarrow$' part does not hold for $i \geq b$. However, since j is an arbitrary non-negative integer, we may consider $j = \alpha a + \beta$ with α and β non-negative integers. Therefore, for $i = r$ with $0 \leq r < b$,

$$\begin{aligned}n &= ra + (\alpha a + \beta)b \\ &= (\alpha b + r)a + \beta b.\end{aligned}$$

  – Before the last sentence, in "... mod b, So ...", the comma should be a period.

- **The Frobenius form of an irreducible non-primitive matrix.**

  – $C_i \neq \emptyset \; \forall i$.
  In fact, consider $u_0 = v$. Since irreducibility is equivalent to strong connectedness, there is a path joining u_0 to u_0. Call it u_0-cycle. By the definition of p, the length of u_0-cycle is an integer multiple of p. Then $u_0 \in C_0$. Now consider $u_1 \neq u_0$ on u_0-cycle such that u_1 is adjacent to u_0. Specifically, suppose that u_1 goes to u_0. Clearly $u_1 \in C_1$. Next consider $u_2 \notin \{u_0, u_1\}$ on u_0-cycle such that u_2 is adjacent to u_1. (Notice that u_2 goes to u_1.) Clearly $u_2 \in C_2$. Similar reasoning shows that $u_3 \in C_3, \ldots, u_{p-1} \in C_{p-1}$.

[17] See **Theorem 9.1.1**.2.

- Concerning the first sentence after the definition of C_i, suppose there is a vertex $u \in C_{i_1} \cap C_{i_2}$ with $i_1 \neq i_2$. Hence there are two paths from u to v of lengths n_1 and n_2 with
$$n_j \equiv i_j \mod p, \quad j = 1, 2. \tag{*}$$
On the other hand, since A is irreducible, there is a path of length n joining v to u. Now, if we combine each one of the paths joining u to v with the path joining v to u, by the definition of p, we get two cycles of lengths
$$n_j + n \equiv 0 \mod p, \quad j = 1, 2.$$
Therefore
$$n_1 \equiv n_2 \mod p,$$
which is a contradiction since the remainders on dividing n_1 and n_2 by p are i_1 and i_2, respectively, by (*).
- If $u \in C_0$, there is no path of length 0 from u to v.
In fact, there is no cycle of length 0 by the definition of p!

==

Comment, pp. 183–4, **The Frobenius form of an irreducible non-primitive matrix.**, from "This means ..." on
Let (u, w) be the first edge in a path from u to v. By the convention for edges, p. 181, the edge goes from $u = v_j$ to $w = v_i$ with $A_{ij} \neq 0$. Then u is related to the j-th column of A, whereas w is related to the i-th row of A. Therefore, after relabeling the vertices as required on p. 183, the vertices in C_k are related to consecutive columns of PAP^{-1}, $k = 0, \ldots, p-1$. So let us analyze the $p = 4$ example (since it explains the general case) where C_k is related to A_k, $k = 1, 2, 3$, and
$$C_0 = \left\{v_1, \ldots, v_{\#(C_0)}\right\}$$
is related to A_4.

- First, $\left(PAP^{-1}\right)_{1j} = 0$, $j = 1, \ldots, \#(C_0)$. In fact, suppose otherwise. Hence there exists such an index j for which $\left(v_j, v_1\right)$ is an edge. However, since $v_1 \in C_0$, there is a path of length
$$n \equiv 0 \mod p$$
from v_1 to v. So there is a path of length
$$n + 1 \equiv 1 \mod p$$
from v_j to v. Then $v_j \in C_0 \cap C_1$, which is a contradiction since $C_0 \cap C_1 = \emptyset$.
- Second, as in the previous case, if $i \in \{2, \ldots, \#(C_0)\}$, then $\left(PAP^{-1}\right)_{ij} = 0$, $j = 1, 2, \ldots, \#(C_0)$.

So the first $\#(C_0)$ rows of PAP^{-1} rule out the elements of C_0 as ending vertices of edges whose starting vertices are such elements. Therefore A_4 is the bottom left block of PAP^{-1}.

- Finally, similar arguments show that A_k is where it should be, $k = 1, 2, 3$, once that the $[n - (\#(C_0))] \times [n - (\#(C_0))]$ bottom right block of PAP^{-1} is null. In fact, suppose otherwise. Therefore we get to $C_k \cap C_\ell \neq \emptyset$ for some $\ell \neq k$!

> Right from the beginning of the previous analysis, the use of PAP^{-1} is an abuse of notation since the block decomposition, which can be written as PAP^{-1}, is obtained as the conclusion of the preceding arguments.

Now let s be the size of a matrix A. If $\wp$ is a permutation of $S = \{1, \ldots, s\}$, there is a permutation matrix P associated to $\wp$. In fact,
$$P = \sum_{i=1}^{s} e_{\wp(i)i}$$
where the i-th term of the sum is the unit matrix which has a 1 in the $\wp(i), i$ position as its only nonzero entry, $i = 1, \ldots, s$. Left multiplication by P permutes the entries of a column vector X using $\wp$. For example, if $s = 3$,

i	1	2	3
$\wp(i)$	2	3	1

and

$$X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix},$$

then

$$\begin{aligned} P &= e_{\wp(1)1} + e_{\wp(2)2} + e_{\wp(3)3} \\ &= e_{21} + e_{32} + e_{13} \\ &= \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \end{aligned}$$

and

$$\begin{aligned} PX &= \begin{bmatrix} x_3 \\ x_1 \\ x_2 \end{bmatrix} \\ &= x_3 e_1 + x_1 e_2 + x_2 e_3 \\ &= x_3 e_{\wp(3)} + x_1 e_{\wp(1)} + x_2 e_{\wp(2)}. \end{aligned}$$

As a matter of fact, whatever s and $\wp$ we consider,

$$\begin{aligned} PX &= \left(\sum_{i=1}^{s} e_{\wp(i)i} \right) \left(\sum_{j=1}^{s} x_j e_j \right) \\ &= \sum_{i,j=1}^{s} x_j e_{\wp(i)i} e_j \\ &= \sum_{i=1}^{s} x_i e_{\wp(i)i} e_i \\ &= \sum_{i=1}^{s} x_i e_{\wp(i)} \end{aligned}$$

since, whenever possible,

$$e_{ij} e_k = \begin{cases} e_i & \text{if } j = k; \\ 0 & \text{if } j \neq k. \end{cases}$$

On the other hand, it is a well-known fact that $P^{-1} = P^t$. Hence, concerning PAP^{-1}, not only P acts on each column of A in the same way as $\wp$ acts on S, but also P^t acts on each row of PA as if it were a permutation of indices. In fact, right multiplication by P^t permutes the entries of an arbitrary $1 \times s$ row vector Y since

$$\begin{aligned} YP^t &= X^t P^t \\ &= (PX)^t \end{aligned}$$

if $X = Y^t$. The relabeling related to

$$\bigcup_{k=0}^{p-1} C_k$$

turns out to be a permutation $\wp$ with associated permutation matrix P such that, by construction, PAP^{-1} is in the block form as previously described.

===

Comment, p. 184, last paragraph before **Proposition 9.2.1**, 1st sentence
Consider $D = D(i)^k$, $E = D(i+1)^k$ and $T = (RS)^{k-1}R$. Hence

$$D = ST \text{ and } E = TS.$$

Furthermore, $d_{ij} = \sum_k s_{ik} t_{kj}$ is the (i, j)-entry of D, where s_{ik} and t_{kj} are entries of S and T, respectively, and a similar representation holds for an arbitrary entry of E. Therefore,

$$\begin{aligned}
\operatorname{trace}(D) &= \sum_i d_{ii} \\
&= \sum_i \sum_k s_{ik} t_{ki} \\
&= \sum_i \sum_k t_{ki} s_{ik} \\
&= \sum_k \sum_i t_{ki} s_{ik} \\
&= \sum_k e_{kk} \\
&= \operatorname{trace}(E).
\end{aligned}$$

==

Comments/Errata, pp. 185–6, **9.3**

- 1st paragraph
  - For the existence of y, see **Showing that** $\lambda_{\max}\left(T^{\dagger}\right) = \lambda_{\max}(T)$**.**, p. 178. Hence, if $A = T$, then $r = \eta$. Now consider $y = w^{\dagger}$.
  - A ')' is missing right after the word 'number'.
  - $y \cdot x = \alpha > 0$. If $\alpha \neq 1$, then $\left(\alpha^{-1} y\right) \cdot x = 1$.
- 2nd paragraph
  - 1st sentence
    - For A $n \times n$, R is the column space of

$$\begin{aligned}
x \otimes y^{\dagger} &:= \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \begin{bmatrix} y_1 & \cdots & y_n \end{bmatrix} \\
&= \begin{bmatrix} x_1 y_1 & \cdots & x_1 y_n \\ \vdots & \cdots & \vdots \\ x_n y_1 & \cdots & x_n y_n \end{bmatrix} \\
&= \begin{bmatrix} y_1 x & \cdots & y_n x \end{bmatrix}.
\end{aligned}$$

    - $H^2 = H$ follows immediately from the definition of outer product, along with the assumption that $y \cdot x = 1$.[18]
  - 2nd sentence
    In fact, $H(I - H) = 0$.
  - 3rd sentence
    Change y to $y^{\dagger}$.
  - 4th and 5th sentences
    - They are, in fact, the same sentence. Delete one of them!

---

[18] In fact,

$$\begin{aligned}
\left(x \otimes y^{\dagger}\right)\left(x \otimes y^{\dagger}\right) &= xyxy \\
&= xy \\
&= x \otimes y^{\dagger}.
\end{aligned}$$

* R and N are invariant under A. In fact, on the one hand, let $u \in R$. Hence $u = sx$ where s is a scalar. Then

$$\begin{aligned} Au &= sAx \\ &= (sr)x \in R. \end{aligned}$$

On the other hand, let $v \in N$. Thus $Hv = 0$. So

$$\begin{aligned} HAv &= AHv \\ &= 0. \end{aligned}$$

Therefore $Av \in N$.

- 3rd paragraph, 1st sentence, 1st clause
  In fact, consider the Perron-Frobenius theorem.[19]

- **Theorem 9.3.1.**
  It means that, eventually, the k-iterate of a typical starting vector concerning the iteration matrix P, say $P^k v_0$ for k large enough, will lie in the direction of the positive eigenvector x associated with $\lambda_{\max} = r$.

==

Comment, p. 188, **An imprimitive Leslie matrix**
L is not primitive since its natural powers have one of the following forms:

$$\begin{pmatrix} 0 & 0 & * \\ * & 0 & 0 \\ 0 & * & 0 \end{pmatrix}, \begin{pmatrix} 0 & * & 0 \\ 0 & 0 & * \\ * & 0 & 0 \end{pmatrix} \text{ and } \begin{pmatrix} * & 0 & 0 \\ 0 & * & 0 \\ 0 & 0 & * \end{pmatrix}.$$

==

Comment, p. 189, **9.5**

- 1st paragraph
  Let T, $\lambda_{\max}$ and x be as in **Theorem 9.1.1**. Therefore, if $M = T$, since $\mathbf{1}$ is a multiple of x,[20] it follows that

$$\begin{aligned} Mx = \lambda_{\max} x &\Longrightarrow M(\alpha \mathbf{1}) = \lambda_{\max}(\alpha \mathbf{1}) \\ &\Longrightarrow M\mathbf{1} = \lambda_{\max} \mathbf{1} \\ &\Longrightarrow \lambda_{\max} = 1. \end{aligned}$$

- 2nd paragraph
  Consider

$$\mathbf{p}^\dagger = \begin{bmatrix} \pi_1 \\ \vdots \\ \pi_n \end{bmatrix}$$

  such that

$$\mathbf{p}M = \mathbf{p}.$$

  Without loss of generality, suppose

$$\sum_{i=1}^{n} \pi_i = 1.$$

  (If $\sum_1^n \pi_i = \pi$, replace $\mathbf{p}$ by $\mathbf{q} = \frac{1}{\pi}\mathbf{p}$.) Therefore, by **Theorem 9.3.1**,

$$\begin{aligned} \lim_{k \to \infty} \left(\frac{1}{1} M^k \right) &= \mathbf{1} \otimes \mathbf{p} \\ &= \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \begin{bmatrix} \pi_1 & \cdots & \pi_n \end{bmatrix}. \end{aligned}$$

[19] Specifically, **Theorem 9.1.1**.*7*.
[20] See **Theorem 9.1.1**.*3*.

==

Errata, p. 190, **9.6.2**

- 2nd paragraph, 2nd sentence
  "... is dangling row ..." should be "... is a dangling row ...";
- Last sentence
  "... dangling mode ..." should be "... dangling node ...".

==

==
==
10
==
==
Erratum, p. 200, **Non-real eigenvalues**
Concerning the diagonal of A, 0 should be a.
==
Comment, p. 202, **variation of constants formula**

$$\begin{aligned}\frac{d}{dt}\left(e^{-tH}x(t)\right) &= -He^{-tH}x(t) + e^{-tH}\frac{d}{dt}x(t) \\ &= -He^{-tH}x(t) + e^{-tH}\big(Hx(t) + f(t)\big) \\ &= e^{-tH}f(t)\end{aligned}$$

since $He^{-tH} = e^{-tH}H$. Therefore

$$e^{-tH}x(t) - x_0 = \int_0^t e^{-sH}f(s)ds.$$

Now multiply both sides of this equation by e^{tH}.
==
Comment, p. 211, (10.6)

$$\begin{aligned}\dot{y} &= \ddot{x} + \frac{d}{dt}F(x) \\ &= -f(x)\dot{x} - x + \dot{x}\frac{d}{dx}F(x) \\ &= -x\end{aligned}$$

since $\frac{d}{dx}F(x) = f(x)$ by property a. of f.
==