2

Invariant Measures

We introduce a discrete dynamical system in terms of a measure preserving transformation T defined on a probability space. In other words, we consider $\{T^n : n \in \mathbb{N}\}$ or $\{T^n : n \in \mathbb{Z}\}$ depending on whether T is invertible or not. The approach can handle not only purely mathematical concepts, but also physical phenomena in nature and data compression in information theory. Examples of such measure preserving transformations include an irrational translation modulo 1, the multiplication by 2 modulo 1, the beta transformation, the Gauss transformation, a toral automorphism, the baker's transformation and the shift transformation on the set of infinite binary sequences. A method of classifying them is also introduced.

2.1 Invariant Measures

Definition 2.1. *(i) Let* (X_1, μ_1) *and* (X_2, μ_2) *be measure spaces. We say that a mapping* $T : X_1 \to X_2$ *is* measurable *if* $T^{-1}(E)$ *is measurable for every measurable subset* $E \subset X_2$*. The mapping is* measure preserving *if* $\mu_1(T^{-1}E) = \mu_2(E)$ *for every measurable subset* $E \subset X_2$*. When* $X_1 = X_2$ *and* $\mu_1 = \mu_2$*, we call* T *a* transformation*.*
(ii) If a measurable transformation $T : X \to X$ *preserves a measure* μ*, then we say that* μ *is* T-invariant *(or invariant under* T*). If* T *is invertible and if both* T *and* T^{-1} *are measurable and measure preserving, then we call* T *an* invertible measure preserving transformation*.*

Theorem 2.2. *Let* $(X, \mathcal{A}, \mu)$ *be a measure space. The following statements are equivalent:*
(i) A transformation $T : X \to X$ *preserves* μ*.*
(ii) For any $f \in L^1(X, \mu)$ *we have*

$$\int_X f(x)\,\mathrm{d}\mu = \int_X f(T(x))\,\mathrm{d}\mu\ .$$

(iii) Define a linear operator U_T in $L^p(X,\mu)$ by

$$(U_T f)(x) = f(Tx)\ .$$

Then U_T is norm-preserving, i.e., $||U_T f||_p = ||f||_p$.

Proof. First we show the equivalence of (i) and (ii). To show (ii) $\Rightarrow$ (i), take $f = 1_E$. Then

$$\mu(E) = \int_X f(x)\,\mathrm{d}\mu = \int_X f(T(x))\,\mathrm{d}\mu = \int_X 1_{T^{-1}E}(x)\,\mathrm{d}\mu = \mu(T^{-1}E)\ .$$

To show (i) $\Rightarrow$ (ii), observe first that a complex-valued measurable function f can be written as a sum

$$f = f_1 - f_2 + \mathrm{i}\,(f_3 - f_4)$$

where $\mathrm{i} = \sqrt{-1}$ and each f_j is real, nonnegative and measurable. Thus we may assume that f is real-valued and $f \geq 0$. If $f = 1_E$ for some measurable subset E, then

$$\int_X f(x)\,\mathrm{d}\mu = \mu(E) = \mu(T^{-1}E) = \int_X 1_{T^{-1}E}(x)\,\mathrm{d}\mu = \int_X f(T(x))\,\mathrm{d}\mu\ .$$

By linearity the same relation holds for a simple function f. Now for a general nonnegative function $f \in L^1$ choose an increasing sequence of simple functions $s_n \geq 0$ converging to f pointwise. Then $\{s_n \circ T\}_{n=1}^{\infty}$ is an increasing sequence and it converges to $f \circ T$ pointwise. Hence

$$\int_X f(Tx)\,\mathrm{d}\mu = \lim_{n\to\infty} \int_X s_n(Tx)\,\mathrm{d}\mu = \lim_{n\to\infty} \int_X s_n(x)\,\mathrm{d}\mu = \int_X f(x)\,\mathrm{d}\mu\ .$$

The proof of the equivalence of (i) and (iii) is almost identical. $\square$

Example 2.3. (Irrational translations modulo 1) Let $X = [0,1)$. Consider

$$Tx = x + \theta \pmod 1$$

for an irrational number $0 < \theta < 1$. Since T preserves the one-dimensional length, Lebesgue measure is invariant under T. See the left graph in Fig. 2.1.

Example 2.4. (Multiplication by 2 modulo 1) Let $X = [0,1)$. Consider

$$Tx = 2x \pmod 1\ .$$

Then T preserves Lebesgue measure. Even though the map doubles the length of an interval I, its inverse image has two pieces in general, each of which has half the length of I. When we add them, the sum equals the original length of I. For a generalization, see Ex. 2.11.

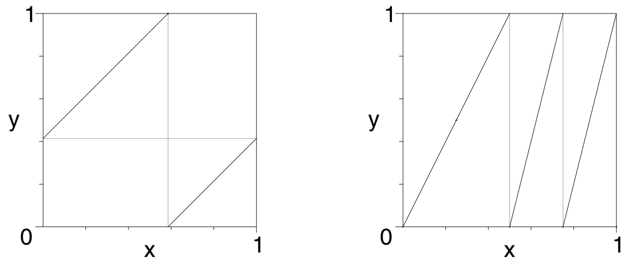

Fig. 2.1. The translation by $\theta = \sqrt{2} - 1$ modulo 1 (left) and a piecewise linear transformation (right)

Example 2.5. (A piecewise linear mapping) Let $X = [0, 1)$. Define

$$Tx = \begin{cases} 2x \pmod 1 \ , & 0 \le x < \frac{1}{2} \ , \\ 4x \pmod 1 \ , & \frac{1}{2} \le x < 1 \ . \end{cases}$$

Then T preserves Lebesgue measure since the inverse image of $[0, a]$ consists of three intervals of lengths $\frac{a}{2}$, $\frac{a}{4}$ and $\frac{a}{4}$. See the right graph in Fig. 2.1.

Example 2.6. (The logistic transformation) Let $X = [0, 1]$. Consider

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

The invariant probability density function of T is given by

$$\rho(x) = \frac{1}{\pi\sqrt{x(1-x)}} \ ,$$

which is unbounded but has its integral equal to 1. See Fig. 2.2. The same density function $\rho(x)$ is invariant under transformations obtained from the Chebyshev polynomials. See Maple Programs 2.6.1, 2.6.2.

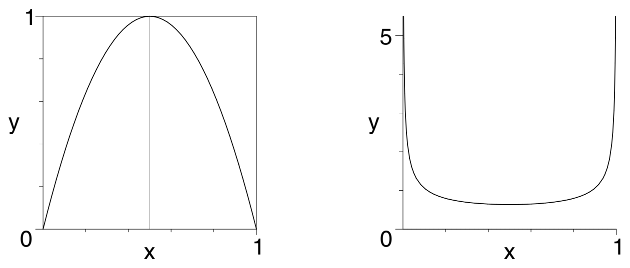

Fig. 2.2. $y = 4x(1 - x)$ (left) and $y = \rho(x)$ (right)

Example 2.7. (The beta transformation) Take $\beta = (\sqrt{5}+1)/2 = 1.618\cdots$, which satisfies $\beta^2-\beta-1=0$. Hence $\beta-1 = 1/\beta = (\sqrt{5}-1)/2$, which is called the *golden ratio*. See Fig. 2.3 where two similar rectangles can be found.

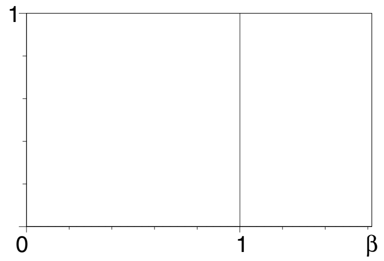

Fig. 2.3. The golden ratio: two rectangles are similar

Let $X = [0,1)$. Consider the so-called β-transformation

$$Tx = \beta x \pmod 1 .$$

Its invariant probability density is given by

$$\rho(x) = \begin{cases} \dfrac{\beta^3}{1+\beta^2}\,, & 0 \le x < \dfrac{1}{\beta}\,, \\ \dfrac{\beta^2}{1+\beta^2}\,, & \dfrac{1}{\beta} \le x < 1\,. \end{cases}$$

For the proof of invariance, it suffices to check the inverse image of $[0,a]$ for every a. See Fig. 2.4 and Maple Program 2.6.3.

There are many other choices for β which lead to a nice formula for $\rho(x)$. For example, try $\beta = 1+\sqrt{2}$. Consult [Pa],[Re].

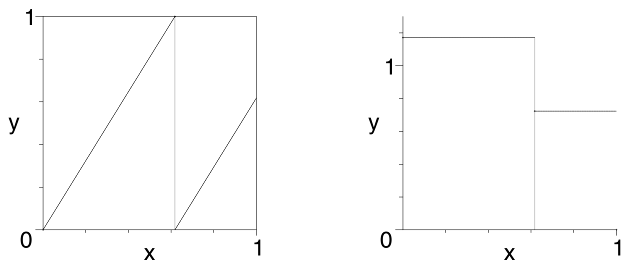

Fig. 2.4. $y = \{\beta x\}$ (left) and $y = \rho(x)$ (right)

Example 2.8. (The Gauss transformation) Let $X = [0, 1)$. To study continued fractions we consider

$$Tx = \frac{1}{x} \pmod 1 , \; 0 < x < 1 , \quad \text{and} \quad T0 = 0 .$$

See Fig. 2.5. Note that $T^n x = 0$ for some n if and only if x is rational. In 1812 C.F. Gauss, in a letter to P.S. Laplace, wrote that the invariant probability density of T is given by

$$\rho(x) = \frac{1}{\log 2} \frac{1}{x+1} ,$$

where log denotes the natural logarithm. For the proof it suffices to show that

$$\int_{T^{-1}(0,a)} \frac{1}{x+1} \, \mathrm{d}x = \int_{(0,a)} \frac{1}{x+1} \, \mathrm{d}x$$

for every $0 < a < 1$. Since

$$T^{-1}(0,a) = \bigcup_{n=1}^{\infty} \left(\frac{1}{n+a}, \frac{1}{n} \right) ,$$

we have

$$\begin{aligned}
\int_{T^{-1}(0,a)} \frac{1}{x+1} \, \mathrm{d}x &= \lim_{N\to\infty} \sum_{n=1}^{N} \left(\log \frac{n+1}{n} - \log \frac{n+1+a}{n+a} \right) \\
&= \lim_{N\to\infty} \left(\log(N+1) - \log(N+1+a) + \log(1+a) \right) \\
&= - \lim_{N\to\infty} \log \left(1 + \frac{a}{N+1} \right) + \log(1+a) \\
&= \log(1+a) = \int_0^a \frac{1}{x+1} \, \mathrm{d}x .
\end{aligned}$$

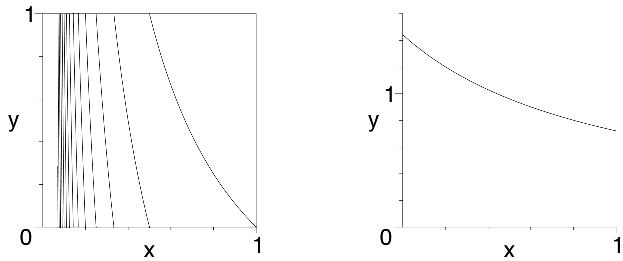

Fig. 2.5. $y = \{1/x\}$ (left) and $y = \rho(x)$ (right)

Example 2.9. Let $X = (-\infty, \infty)$. Consider

$$Tx = x - \frac{1}{x} .$$

Then T preserves Lebesgue measure, i.e.,

$$\int_{-\infty}^{\infty} f(x)\,\mathrm{d}x = \int_{-\infty}^{\infty} f\left(x - \frac{1}{x}\right)\mathrm{d}x$$

for every integrable function f. To prove it, observe that the inverse image of an interval (a, b) under T is the union of two intervals

$$\left(\frac{a-\sqrt{a^2+4}}{2}, \frac{b-\sqrt{b^2+4}}{2}\right) \quad \text{and} \quad \left(\frac{a+\sqrt{a^2+4}}{2}, \frac{b+\sqrt{b^2+4}}{2}\right) ,$$

the sum of whose lengths is equal to $b - a$.

Example 2.10. (D. Lind) Let $X = (-\infty, \infty)$. Consider

$$Tx = \frac{1}{2}\left(x - \frac{1}{x}\right) .$$

Then T has a finite invariant density function

$$\rho(x) = \frac{1}{1+x^2} .$$

To prove it, first observe that the inverse image of an interval (a, b) is the union of two intervals

$$(a - \sqrt{a^2+1}, b - \sqrt{b^2+1}) \quad \text{and} \quad (a + \sqrt{a^2+1}, b + \sqrt{b^2+1}) .$$

Next, use the identity

$$\arctan(a + \sqrt{a^2+1}) + \arctan(a - \sqrt{a^2+1}) = \arctan a .$$

The transformation T comes from the Newton's method of finding the roots of $f(x) = x^2 + 1$. The tangent line to the graph $y = f(x)$ at $x = c$ is of the form $y = f'(c)(x - c) + f(c)$. It intersects the x-axis at

$$x = c - \frac{f(c)}{f'(c)} .$$

The iteration algorithm

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

is interpreted as

$$Tx = x - \frac{x^2+1}{2x} = \frac{1}{2}\left(x - \frac{1}{x}\right) .$$

If there existed a real root of $f(x) = 0$, then x_n would converge to a limit. Since $f(x) \neq 0$ for any real x, the iteration algorithm does not converge to a limit and the corresponding behavior of T is random. See Fig. 2.6.

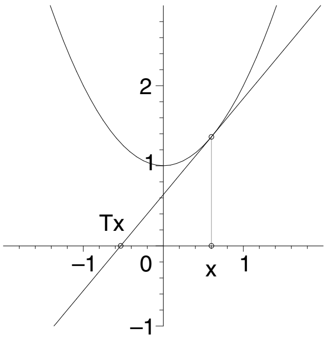

Fig. 2.6. The Newton method does not work for $f(x) = x^2 + 1$

Example 2.11. (Endomorphisms of compact groups) Let G be a compact group with the left Haar measure μ such that $\mu(G) = 1$. A continuous onto endomorphism $\phi : G \to G$ preserves μ. This is a generalization of Ex. 2.4. For the proof, define

$$\nu(E) = \mu(\phi^{-1}(E))$$

for $E \subset G$. Then ν is also a probability measure. Let $xA = \{xa : a \in A\}$ for $x \in G$, $A \subset G$. To show the rotation invariance, take an arbitrary element of G. Since ϕ is onto, an element is of the form $\phi(x)$ for some x. Since

$$\begin{aligned} y \in \phi^{-1}(\phi(x)E) &\Leftrightarrow \phi(y) \in \phi(x)E \\ &\Leftrightarrow \phi(x)^{-1}\phi(y) = \phi(x^{-1}y) \in E \\ &\Leftrightarrow x^{-1}y \in \phi^{-1}(E) \\ &\Leftrightarrow y \in x\phi^{-1}(E) \ , \end{aligned}$$

where the symbol $\Leftrightarrow$ denotes logical equivalence, we have $\phi^{-1}(\phi(x)E) = x\phi^{-1}(E)$. Hence

$$\nu(\phi(x)E) = \mu(\phi^{-1}(\phi(x)E)) = \mu(x\phi^{-1}(E)) = \mu(\phi^{-1}(E)) = \nu(E) \ .$$

Thus ν is also a left Haar measure. By the uniqueness we have $\mu = \nu$ and $\mu(E) = \mu(\phi^{-1}(E))$. The proof for the right Haar measure is almost identical.

Example 2.12. (The baker's transformation) Imagine a baker mixing flour and water when trying to make dough. The essence of the process can be explained as follows: Let $X = [0,1] \times [0,1]$ and define $T : X \to X$ by

$$T(x,y) = \begin{cases} \left(2x, \frac{1}{2}y\right) \ , & 0 \le x < \frac{1}{2} \ , \\ \left(2x-1, \frac{1}{2}y + \frac{1}{2}\right) \ , & \frac{1}{2} \le x \le 1 \ . \end{cases}$$

First, press down the unit square, cut in the middle, and move the right half to the top of the left half. The transformation T preserves the two-dimensional Lebesgue measure on the unit square. See Fig. 2.7.

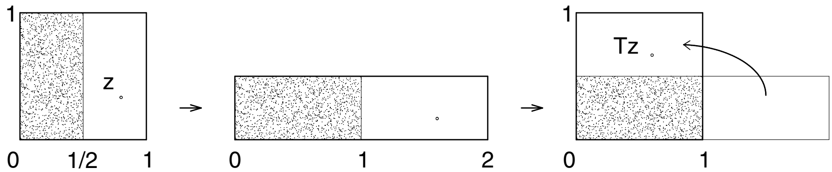

Fig. 2.7. The baker's transformation

We introduce an approach in plotting a graph or a diagram, which may be called the *mathematical pointillism*: A geometric object is described by plotting sufficiently many points on it. For example, to sketch a graph of a function $f : [a, b] \to \mathbb{R}$ we choose sufficiently many, say $S = 1000$, values x_j and plot $(x_j, f(x_j)) \in \mathbb{R}^2$, $1 \le j \le S$. If f is continuous and the points x_j are more or less densely distributed in $[a, b]$, then the points $(x_j, f(x_j))$ as a collection look like a connected graph to the naked eye. Various versions of this idea are used throughout the book.

Let $C = \{(x, y) : 0 \le x \le \frac{1}{2}, 0 \le y \le 1\}$. To sketch the images of C under iterations of T we employ the mathematical pointillism. Take $S = 4000$ points in C. The successive images of those S points under T, T^2 and T^3 are plotted from left to right in Fig. 2.8. See Maple Program 2.6.4.

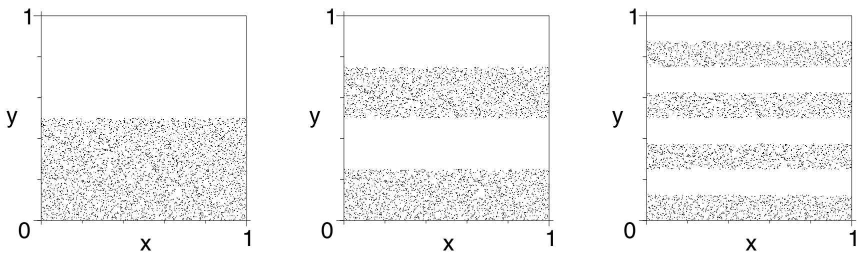

Fig. 2.8. Successive images of a rectangle C under the baker's transformation

Example 2.13. (A toral automorphism) Let $\mathbb{T}^2$ denote the two-dimensional torus with Lebesgue measure, which is regarded as the unit square whose boundaries are identified in the standard way, i.e., left with right and top with bottom. Take

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$$

and consider a transformation $T : \mathbb{T}^2 \to \mathbb{T}^2$ given by

$$T(x, y) = (2x + y, x + y) \pmod 1 .$$

Since $\det A = 1$, we see that T is invertible and preserves Lebesgue measure. Thus T is a toral automorphism. It is called the Arnold cat mapping because V.I. Arnold often draws in his books a face of a cat on the unit square and describes its deformation as T is iterated to indicate how T acts geometrically in the directions of eigenvectors. See [ArA].

Let $D = \{(x, y) : 0 \le x \le \frac{1}{2}, 0 \le y \le \frac{1}{2}\}$ and take $S = 4000$ points in D. We plot the images of those points under T, T^2 and T^3 from left to right in Fig. 2.9 employing the mathematical pointillism. See Maple Program 2.6.5.

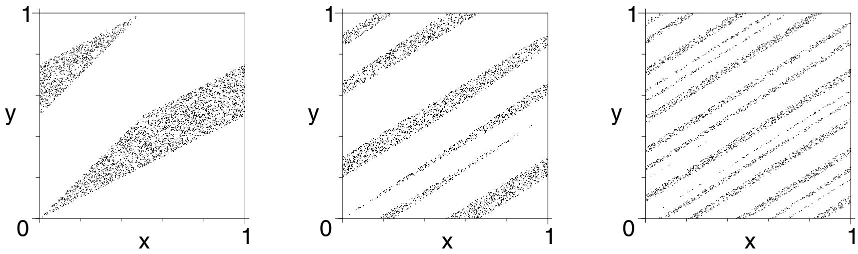

Fig. 2.9. Successive images of a square D under the Arnold cat mapping

Example 2.14. (Lambda transformations) Let $X = [0, 1]$. For $0 < c < 1$ define a *Λ-transformation* τ_c by

$$\tau_c(x) = \begin{cases} \dfrac{1}{c}x\,, & 0 \le x \le c\,, \\ -\dfrac{1}{1-c}x + \dfrac{1}{1-c}\,, & c < x \le 1\,. \end{cases}$$

Then τ_c preserves Lebesgue measure since the inverse image of an interval E is a union of two intervals and the sum of their lengths is the length of E. See Fig. 2.10. The graph of τ_c looks like the uppercase Greek letter Λ.

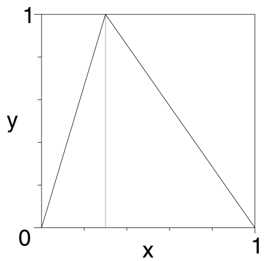

Fig. 2.10. The Λ-transformation τ_c for $c = 3/10$

Example 2.15. (Truncated Λ-transformations) Let $X = [0,1]$. For $\frac{1}{2} < a < 1$, let $b = (2a-1)/a$ and define

$$T_a(x) = \begin{cases} \dfrac{1-a}{b}\, x + a\ , & 0 \le x \le b\ , \\ \dfrac{a}{1-a}(-x+1)\ , & b < x \le 1\ . \end{cases}$$

See the left graph in Fig. 2.11 for $a = \frac{2}{3}$, $b = \frac{1}{2}$. Observe that $T_a([0,a]) = [a,1]$ and $T_a([a,1]) = [0,a]$. Thus, if μ is a T_a-invariant probability measure, then

$$\mu([0,a]) = \frac{1}{2} = \mu([a,1])\ .$$

Note that if μ is invariant under T_a then it is also invariant under $(T_a)^2$. The restrictions of $(T_a)^2$ to $[0,a]$ and $[a,1]$, with corresponding invariant measures $\frac{1}{a}\,\mathrm{d}x$ and $\frac{1}{1-a}\,\mathrm{d}x$, are essentially Λ-transformations, and they preserve the normalized Lebesgue measures on $[0,a]$ and $[a,1]$, respectively. (These two intervals are ergodic components of $(T_a)^2$. See the right graph in Fig. 2.11. For the definition of ergodicity, see Chap 3.) In summary, for $\frac{1}{2} < a < 1$ we have

$$\rho(x) = \begin{cases} \dfrac{1}{2a}\ , & 0 < x < a\ , \\ \dfrac{1}{2(1-a)}\ , & a < x < 1\ . \end{cases}$$

The T_a-invariance of $\mathrm{d}\mu = \rho(x)\,\mathrm{d}x$ may be checked directly: it suffices to prove $\mu(T^{-1}E) = \mu(E)$ for the intervals E such that $E \subset [0,a]$ or $E \subset [a,1]$.

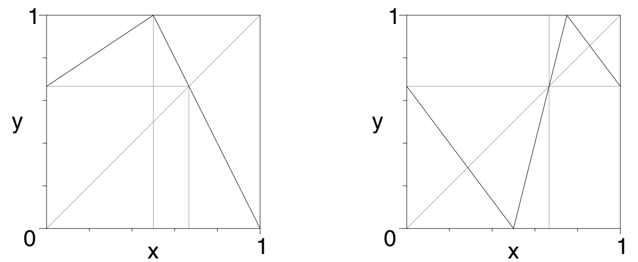

Fig. 2.11. The truncated Λ-transformation T_a (left) and $(T_a)^2$ (right) for $a = 2/3$

2.2 Other Types of Continued Fractions

There are continued fractions arising from transformations other than the Gauss transformation.

Example 2.16. Let $X = \left(-\frac{1}{2}, \frac{1}{2}\right)$. Define

$$Tx = \left|\frac{1}{x}\right| - \left[\left|\frac{1}{x}\right| + \frac{1}{2}\right], \quad x \neq 0 .$$

This is the continued fraction to the nearest integer. Rieger ([Rie1],[Rie2]) showed that the invariant probability density is given by

$$\rho(x) = \begin{cases} \dfrac{1}{\ln\beta}\dfrac{1}{x+\beta+1}, & -\dfrac{1}{2} < x < 0 , \\ \dfrac{1}{\ln\beta}\dfrac{1}{x+\beta}, & 0 < x < \dfrac{1}{2} , \end{cases}$$

where $\beta = \frac{\sqrt{5}+1}{2}$. See Fig. 2.12 where $y = Tx$ is plotted away from $x = 0$.

Nakada [Na] considered a generalization of the form

$$Tx = \left|\frac{1}{x}\right| - \left[\left|\frac{1}{x}\right| + 1 - \alpha\right], \quad x \neq 0 ,$$

on the interval $(\alpha - 1, \alpha)$ for $\frac{1}{2} \le \alpha \le 1$.

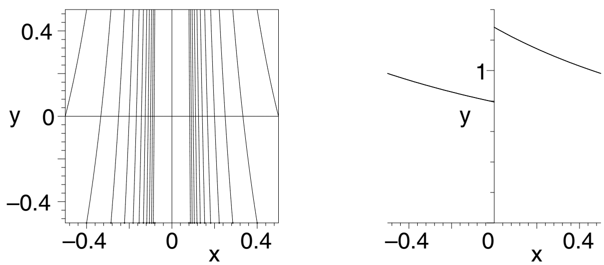

Fig. 2.12. $y = Tx$ (left) and $y = \rho(x)$ (right) for $Tx = |1/x| - [\,|1/x| + 1/2\,]$

Example 2.17. Let $X = \left(-\frac{1}{2}, \frac{1}{2}\right)$. Define

$$Tx = \frac{1}{x} - \left[\frac{1}{x} + \frac{1}{2}\right], \quad x \neq 0 .$$

Put $\alpha = \frac{\sqrt{5}-1}{2} = 0.6180\cdots$ and $\beta = \frac{3-\sqrt{5}}{2} = 0.3819\cdots$. Nakada, Ito and Tanaka [NIT] showed that the invariant probability density is given by

$$\rho(x) = C\frac{1}{(1+\alpha|x|)(1-\beta|x|)}$$

where $C = -2\ln(1+\sqrt{5}) + 2\ln(3+\sqrt{5})$. The transformation T is a real version of a complex continued fraction introduced by Hurwitz [Hur]. See Fig. 2.13.

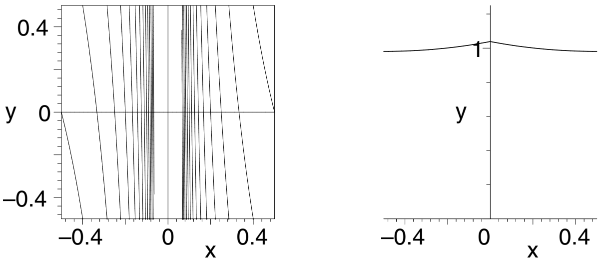

Fig. 2.13. $y = Tx$ (left) and $y = \rho(x)$ (right) for $Tx = 1/x - [1/x + 1/2]$

Example 2.18. Put $\alpha = \frac{\sqrt{5}-1}{2}$ and $\beta = \frac{3-\sqrt{5}}{2}$. Let $X = (-\alpha, \alpha)$. Define

$$Tx = \frac{1}{x} - \mathrm{sgn}(x)\left[\frac{1}{|x|} + \beta\right], \quad x \neq 0,$$

where $\mathrm{sgn}(x) = \dfrac{x}{|x|} \in \{\pm 1\}$, $x \neq 0$. In [NIT] it is shown that the invariant probability density is given by

$$\rho(x) = \begin{cases} C\dfrac{1}{2-x}\,, & -\alpha < x < -\beta\,, \\[2ex] C\dfrac{4}{(2-x)(2+x)}\,, & -\beta < x < \beta\,, \\[2ex] C\dfrac{1}{2+x}\,, & \beta < x < \alpha\,, \end{cases}$$

where C is the normalizing constant given in Ex. 2.17. See Fig. 2.14 where $y = Tx$ is plotted away from $x = 0$. Consult Maple Program 2.6.6.

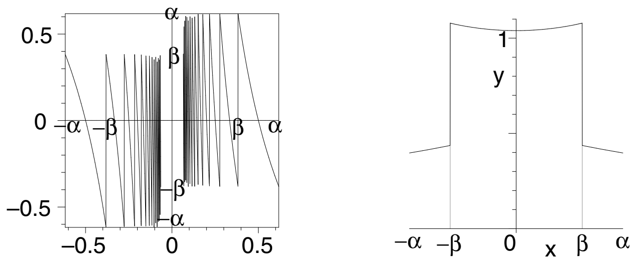

Fig. 2.14. $y = Tx$ (left) and $y = \rho(x)$ (right) for $Tx = 1/x - \mathrm{sgn}(x)\,[1/|x| + \beta]$

For more examples of various types of continued fractions consult [DK].

2.3 Shift Transformations

In this section we study shift transformations, which can be used in digitizing a transformation defined on a continuous space such as an interval. They are the main object of investigation in relation to data compression in Chap. 14.

Consider infinite strings made up of k symbols. For notational convenience we choose k symbols from $\{0, 1, \ldots, k-1\}$ or $\{1, 2, \ldots, k\}$. This choice is convenient for computer simulations, too. Put $X = \prod_1^\infty \{1, \ldots, k\}$. An element x of X is denoted by $(x_1, x_2, x_3, \ldots)$ or simply $(x_1 x_2 x_3 \ldots)$. If $k = 2$ then x is said to be a *binary* sequence. Let $p_1, \ldots, p_k$ be nonnegative numbers such that $p_1 + \cdots + p_k = 1$. For $t \geq 1$ let define a cylinder set (or a block) of length n by

$$[a_1, \ldots, a_n]_{t,\ldots,t+n-1} = \{x \in X : x_{t+1} = a_1, \ldots, x_{t+n} = a_n\} .$$

Define μ on cylinder sets by

$$\mu([a_1, \ldots, a_n]_{t,\ldots,t+n-1}) = p_{a_1} \cdots p_{a_n} .$$

A probability measure on X, again denoted by μ for notational simplicity, is uniquely defined on the σ-algebra generated by cylinder sets. We call μ the $(p_1, \ldots, p_k)$-*Bernoulli measure* and X is the *Bernoulli shift space*. The *one-sided Bernoulli shift transformation* on X, defined by

$$(x_1 x_2 x_3 \ldots) \mapsto (x_2 x_3 x_4 \ldots)$$

preserves the measure μ. Similarly, we define the *two-sided Bernoulli shift transformation* by

$$(\ldots \overset{*}{x}_0 x_1 x_2 \ldots) \mapsto (\ldots \overset{*}{x}_1 x_2 x_3 \ldots)$$

on $\prod_{-\infty}^\infty \{1, \ldots, k\}$ where $*$ denotes the 0th coordinate in a sequence.

Let $X = \prod_1^\infty \{1, \ldots, k\}$. Let $P = (P_{ij})$ be a $k \times k$ stochastic matrix. Suppose that $\boldsymbol{\pi} = (\pi_i) > 0$ satisfies $\sum_i \pi_i = 1$ and $\boldsymbol{\pi} P = \boldsymbol{\pi}$. Define μ by

$$\mu([a_1, \ldots, a_n]_{t,\ldots,t+n-1}) = \pi_{a_1} P_{a_1 a_2} \cdots P_{a_{n-1} a_n} .$$

Then there exists a unique shift invariant probability measure, again denoted by μ, on the σ-algebra generated by the cylinder sets. We call μ the *Markov measure* and X the *Markov shift space*.

Let $\Pr(B|A)$ denote the conditional probability of an event B given that an event A has occurred. (Informally we write $\Pr(A \to B)$.) Then P defines the transition probability

$$\Pr(x_{n+1} = j \mid x_n = i) = P_{ij} .$$

Markov shifts are Bernoulli shifts if the rows of P are identical. When there is no danger of confusion, a shift means either T or (X, T). In general, a

Markov measure of order $m \geq 1$ is defined by the conditional probability $\Pr(b|a_1, \ldots, a_m)$ of seeing b given that we have just seen $a_1, \ldots, a_m$. Note that the sequence of appearance is $a_1, \ldots, a_m, b$. The numbers $\Pr(b|a_1, \ldots, a_m)$ satisfy

$$\sum_{b=1}^{k} \Pr(b|a_1, \ldots, a_m) = 1 .$$

Only the case $m = 1$ is considered in this book. For more information consult [KemS], and for a related result see [Ham].

We identify a Bernoulli measure or a Markov measure with a measure on $[0, 1]$ through the binary expansion. In other words, a binary sequence $(b_1 b_2 b_3 \ldots)$ is identified with $\sum_{n=1}^{\infty} b_n 2^{-n}$. (See Maple Programs 2.6.7, 2.6.8.) If $p \notin \{0, \frac{1}{2}, 1\}$, then the $(p, 1-p)$-Bernoulli measure represented on $[0, 1]$ is singular continuous. In Fig. 2.15 for $p = \frac{1}{4}$, cylinder sets of length $n = 9$ in the Bernoulli shift correspond to intervals of the form $[2^{-n}(i-1), 2^{-n}i)$. The probability of each cylinder set is equal to the height of the vertical bar times 2^{-n} so that the area under the histogram is equal to 1. For more examples see Sect. 5.5.

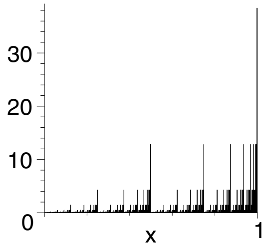

Fig. 2.15. Approximation of the $(1/4, 3/4)$-Bernoulli measure by a histogram

2.4 Isomorphic Transformations

How can we compare and classify measure preserving transformations defined on probability spaces? First we discuss the general concept of equivalence relation among objects under study and then present a method of classifying measure preserving transformations.

Among a collection of objects under investigation, different objects may share the same mathematical structures. They are regarded as being identical. An *equivalence relation* on a set X is a rule that specifies whether two elements $x, y \in X$ are the identical from a given mathematical viewpoint. In this case, we say that x and y are *equivalent*, and write $x \sim y$. Rigorously, we require the following three conditions:

(i) (reflexivity) $x \sim x$ for $x \in X$,
(ii) (symmetry) if $x \sim y$ then $y \sim x$ for $x, y \in X$, and
(iii) (transitivity) if $x \sim y$ and $y \sim z$ then $x \sim z$ for $x, y, z \in X$.

An *equivalence class* containing x is the set of all y satisfying $x \sim y$, and is denoted by $[x]$. If $[x] \cap [y] \neq \varnothing$ then $[x] = [y]$, and so X is the disjoint union of equivalence classes. Any element from an equivalence class is called a *representative* of that equivalence class. If we want to study a mathematical property shared by all elements in the same equivalence class, then it is enough to study it with only one representative from each class.

For example, we have equivalence relations in linear algebra: two vector spaces V and W are isomorphic if there exists a bijective mapping $\phi : V \to W$ such that $\phi(v_1 + v_2) = \phi(v_1) + \phi(v_2)$ and $\phi(cv) = c\phi(v)$ for $v_1, v_2, v \in V$ and a scalar c. Two finite-dimensional vector spaces are isomorphic if and only if they have the same dimension. Once the equivalence of vector spaces is understood, we want to compare two linear transformations $T : V \to V$ and $S : W \to W$ represented by matrices A and B, respectively, where V and W are isomorphic vector spaces. This is usually done by using similarity: $A \sim B$ if and only there exists an invertible matrix P such that $A = P^{-1}BP$.

Now we study how to compare measure spaces and measure preserving transformations defined on them. These correspond to comparing vector spaces and matrices, respectively.

Definition 2.19. *(i) Let (X_1, μ_1) and (X_2, μ_2) be measure spaces. A mapping $\phi : X_1 \to X_2$ is said to be* almost everywhere bijective *if there exist $E_1 \subset X_1$ and $E_2 \subset X_2$ such that $\mu_1(E_1) = 0 = \mu_2(E_2)$ and $\phi : X_1 \setminus E_1 \to X_2 \setminus E_2$ is one-to-one and onto.*
(ii) Furthermore, if there exists an almost everywhere bijective mapping $\phi : (X_1, \mu_1) \to (X_2, \mu_2)$ such that ϕ and ϕ^{-1} are measurable and measure preserving, then (X_1, μ_1) and (X_2, μ_2) are said to be isomorphic *and ϕ is called an* isomorphism *of (X_1, μ_1) and (X_2, μ_2). (By the inverse of ϕ, we mean the inverse of $\phi : X_1 \setminus E_1 \to X_2 \setminus E_2$.)*

It is known that there is essentially just one type of a probability measure space. More precisely, if X is a complete metric space with a countable dense subset and if X has a continuous Borel probability measure μ, then (X, μ) is isomorphic to $[0, 1]$ with Lebesgue measure. (Most examples satisfy this condition.) Thus the classification problem for probability spaces is trivial but the construction of an isomorphism ϕ is not always easy.

Example 2.20. (Isomorphic spaces) Let $X = \prod_1^\infty \{0, 1\}$ be the $(\frac{1}{2}, \frac{1}{2})$-Bernoulli shift space, i.e., each symbol has probability $\frac{1}{2}$. Then X is isomorphic to $[0, 1]$ with Lebesgue measure. To see why, take $x = (b_1, b_2, \ldots) \in X$ and define

$$\phi(x) = \sum_{n=1}^{\infty} b_n 2^{-n} ,$$

which is almost everywhere bijective and measure preserving.

Example 2.21. (Isomorphic spaces) The unit interval and the unit square are isomorphic, where the measures under consideration are Lebesgue measures. To see why, take $x \in [0,1]$ and express it in the binary expansion

$$x = \sum_{n=1}^{\infty} b_n 2^{-n} , \quad b_n = 0,1 .$$

For some x the binary expansion is not unique, but the set of such points has measure zero, and we ignore them. Define $\phi : [0,1] \to [0,1]^2$ by

$$\phi(x) = \left(\sum_{n=1}^{\infty} b_{2n-1} 2^{-n}, \sum_{n=1}^{\infty} b_{2n} 2^{-n} \right) .$$

Then ϕ is defined almost everywhere and measure preserving. Similarly, the unit interval is isomorphic to the k-dimensional cube $[0,1]^k$ for any k.

Now we compare two measure preserving transformations.

Definition 2.22. *(i) Let $T_1 : (X_1, \mu_1) \to (X_1, \mu_1)$ and $T_2 : (X_2, \mu_2) \to (X_2, \mu_2)$ be measure preserving. They are said to be* isomorphic *if there exists an isomorphism $\phi : (X_1, \mu_1) \to (X_2, \mu_2)$ such that $\phi \circ T_1 = T_2 \circ \phi$, i.e., the following diagram commutes:*

$$\begin{array}{ccc} (X_1, \mu_1) & \xrightarrow{T_1} & (X_1, \mu_1) \\ {\scriptstyle\phi}\downarrow & & {\scriptstyle\phi}\downarrow \\ (X_2, \mu_2) & \xrightarrow{T_2} & (X_2, \mu_2) \end{array}$$

(It is assumed that $T_1(X_1 \setminus E_1) \subset X_1 \setminus E_1$ and $T_2(X_2 \setminus E_2) \subset X_2 \setminus E_2$ where E_1 and E_2 are given in Definition 2.19.)
(ii) Suppose that ϕ is measure preserving and $\phi \circ T_1 = T_2 \circ \phi$. If ϕ is not necessarily almost everywhere bijective, then T_2 is said to be a factor *of T_1 and ϕ is called a* factor map. *In this case T_1 is called an* extension *of T_2.*
(iii) Sometimes the objects under discussion belong to the topological category: X_1, X_2 are topological spaces, ϕ is a homeomorphism, the measures are Borel measures, and T_1, T_2 are continuous mappings. In this case we often call ϕ a (topological) conjugacy. *If ϕ is not one-to-one, it is called a* (topological) semi-conjugacy.

Example 2.23. (Isomorphic transformations) The $(\frac{1}{2}, \frac{1}{2})$-Bernoulli shift space $X_1 = \prod_1^{\infty} \{0,1\}$ and $X_2 = [0,1]$ are isomorphic via the isomorphism ϕ in Ex. 2.20. Define $T_1 : X_1 \to X_1$ by

$$T_1((b_1, b_2, b_3, \ldots)) = (b_2, b_3, b_4, \ldots) ,$$

and define $T_2 : X_2 \to X_2$ by

$$T_2(x) = 2x \pmod 1 .$$

Then T_1 and T_2 are isomorphic since $\phi \circ T_1 = T_2 \circ \phi$.

Example 2.24. (Isomorphic transformations) Let $X = [0,1]$ and consider two transformations $Tx = 4x(1-x)$ and

$$\Lambda x = \begin{cases} 2x\,, & 0 \le x \le \frac{1}{2}\,, \\ 2-2x\,, & \frac{1}{2} < x \le 1\,, \end{cases}$$

Note that Λ preserves Lebesgue measure $\mathrm{d}x$ and that T preserves $\mathrm{d}\mu = \rho(x)\,\mathrm{d}x$, $\rho(x) = \frac{1}{\pi\sqrt{x(1-x)}}$. Define ϕ by $\phi(x) = \sin^2\left(\frac{\pi}{2}x\right) = \frac{1}{2}(1-\cos(\pi x))$, $0 \le x \le 1$. (See the left graph in Fig. 2.16.) Then $\phi : (X, \mathrm{d}x) \to (X, \mu)$ is measure preserving. To see this, it suffices to observe that

$$\text{length of } \phi^{-1}([0,\phi(a)]) = a = \int_0^{\sin^2(\pi a/2)} \rho(x)\,\mathrm{d}x = \mu([0,\phi(a)])$$

where the second equality is obtained by taking derivatives with respect to a. Since

$$T\phi(x) = 4\sin^2\left(\frac{\pi}{2}x\right)\left(1-\sin^2\left(\frac{\pi}{2}x\right)\right) = \sin^2(\pi x) = \phi(\Lambda x)\,,$$

the following diagram commutes, and hence Λ and T are isomorphic.

$$\begin{array}{ccc} (X,\,\mathrm{d}x) & \xrightarrow{\Lambda} & (X,\,\mathrm{d}x) \\ \phi\big\downarrow & & \phi\big\downarrow \\ (X,\mu) & \xrightarrow{T} & (X,\mu) \end{array}$$

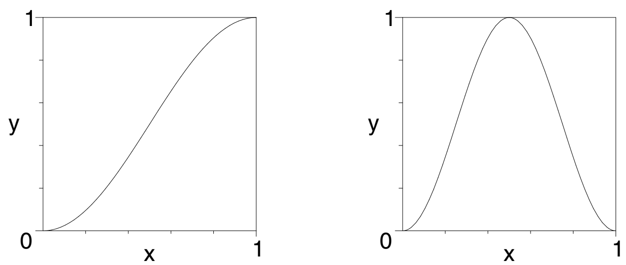

Fig. 2.16. $y = \sin^2(\pi x/2)$ (left) and $y = \sin^2(\pi x)$ (right)

Example 2.25. (Isomorphic transformations) Let $X = [0,1]$ and consider

$$Sx = 2x \pmod 1 \quad \text{and} \quad Tx = 4x(1-x)\,.$$

Note that S preserves Lebesgue measure $\mathrm{d}x$ and that T preserves $\mathrm{d}\mu = \rho(x)\,\mathrm{d}x$, $\rho(x) = \frac{1}{\pi\sqrt{x(1-x)}}$. Define $\psi : (X, \mathrm{d}x) \to (X, \mu)$ by $\psi(x) = \sin^2(\pi x)$.

(See the right graph in Fig. 2.16. Since ψ is almost everywhere two-to-one, it cannot be used to determine whether T and S are isomorphic. But it is a semi-conjugacy as can be seen in the following.) Note that ψ is measure preserving. To see this, it suffices to check that, for $0 \le a \le \frac{1}{2}$,

$$\text{length of } \psi^{-1}([0,\psi(a)]) = 2a = \int_0^{\sin^2(\pi a)} \rho(x)\,\mathrm{d}x = \mu([0,\psi(a)])$$

and, for $\frac{1}{2} \le a \le 1$,

$$\text{length of } \psi^{-1}([0,\psi(a)]) = 2(1-a) = \int_0^{\sin^2(\pi a)} \rho(x)\,\mathrm{d}x = \mu([0,\psi(a)])$$

by taking derivatives with respect to a. Since

$$T\psi(x) = 4\sin^2(\pi x)(1-\sin^2(\pi x)) = \sin^2(2\pi x) = \psi(Sx) ,$$

the following diagram commutes:

$$\begin{array}{ccc} (X,\,\mathrm{d}x) & \xrightarrow{S} & (X,\,\mathrm{d}x) \\ \psi\big\downarrow & & \psi\big\downarrow \\ (X,\mu) & \xrightarrow{T} & (X,\mu) \end{array}$$

Combining Exs. 2.23, 2.24, 2.30, we see that the four transformations Λ, T, S and the $(\frac{1}{2},\frac{1}{2})$-Bernoulli shift transformation are isomorphic. See Maple Program 2.6.9.

Example 2.26. (Isomorphic transformations) Consider $Sx = 2x \pmod 1$ on the unit interval $[0,1)$, which is identified with the unit circle by $x \mapsto \mathrm{i}\,\mathrm{e}^{2\pi\mathrm{i}x}$. Note that x is the angle between the tangent line to $|z| = 1$ at $z = \mathrm{i}$ and the line connecting i and $\mathrm{i}\,\mathrm{e}^{2\pi\mathrm{i}x}$. By the standard stereographic projection ϕ of the unit circle onto $\mathbb{R}$, as x increases from 0 to 1, a point on $|z| = 1$ starts from $z = \mathrm{i}$ and moves counterclockwise while the corresponding point on $\mathbb{R}$ moves from $-\infty$ to $+\infty$. The straight line connecting i and $\mathrm{i}\,\mathrm{e}^{2\pi\mathrm{i}x}$ is given by

$$\mathrm{i} + t(\mathrm{i}\,\mathrm{e}^{2\pi\mathrm{i}x} - \mathrm{i}) = -t\sin 2\pi x + \mathrm{i}\,(1 + t(\cos 2\pi x - 1)) , \quad -\infty < t < \infty .$$

It intersects the real axis when $t = -1/(\cos 2\pi x - 1)$, and hence

$$\phi(x) = \sin 2\pi x/(\cos 2\pi x - 1) = -\cot \pi x .$$

See Fig. 2.17.

A measure μ on $\mathbb{R}$ is induced by ϕ, i.e.,

$$\mu((-\infty, -\cot \pi a)) = a$$

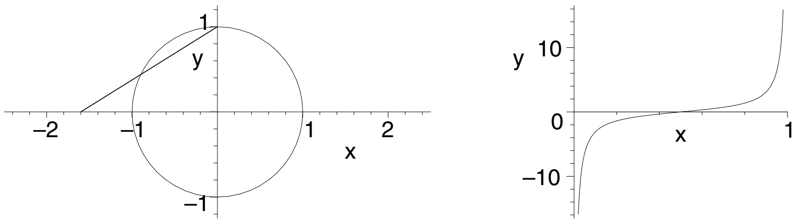

Fig. 2.17. The stereographic projection of the unit circle onto the real line (left) and $y = -\cot \pi x$, $0 < x < 1$ (right)

for $0 \le a \le 1$, so that $\phi : ([0, 1), \mathrm{d}x) \to (\mathbb{R}, \mu)$ is measure preserving where $\mathrm{d}x$ is the normalized Lebesgue measure. Let us find $\rho(x)$ such that

$$\mathrm{d}\mu = \rho \, \mathrm{d}x \, .$$

Since

$$\int_{-\infty}^{-\cot \pi a} \rho(x) \, \mathrm{d}x = a \, ,$$

by taking derivatives with respect to a, we obtain

$$\rho(-\cot \pi a) \left(-\frac{-\sin^2 \pi a - \cos^2 \pi a}{\sin^2 \pi a} \right) \pi = 1 \, .$$

Hence

$$\rho(x) = \frac{1}{\pi} \frac{1}{1 + x^2} \, .$$

Define

$$T = \phi \circ S \circ \phi^{-1} \, .$$

Then T preserves μ on $\mathbb{R}$ and the following diagram commutes:

$$\begin{array}{ccc} (X, \mathrm{d}x) & \xrightarrow{S} & (X, \mathrm{d}x) \\ \phi \big\downarrow & & \phi \big\downarrow \\ (\mathbb{R}, \mu) & \xrightarrow{T} & (\mathbb{R}, \mu) \end{array}$$

Hence S and T are isomorphic. It remains to find the explicit formula for T. Since

$$T(-\cot \pi x) = -\cot 2\pi x = \frac{1}{2}(\tan \pi x - \cot \pi x) \, ,$$

we have

$$Tx = \frac{1}{2} \left(x - \frac{1}{x} \right) \, .$$

We may start with T and μ and find S later. See Ex. 2.10.

Example 2.27. (Isomorphic transformations) Let $X = [0,1]^2$ and T be the baker's transformation. Let $Y = \prod_{-\infty}^{\infty}\{0,1\}$ be the two-sided $(\frac{1}{2},\frac{1}{2})$-Bernoulli shift space and let Σ be the left shift transformation on Y. For $x = (a,b) \in X$ we have the binary expansions

$$a = \sum_{j=0}^{\infty} a_j 2^{-j-1}\ , \quad b = \sum_{j=1}^{\infty} b_j 2^{-j}\ .$$

Then define $\phi : X \to Y$ by

$$\phi(x) = \phi((a,b)) = (\ldots, b_{-2}, b_{-1}, \overset{*}{a_0}, a_1, a_2, \ldots)$$

the symbol $*$ indicates the 0th position. Note that $\phi \circ T = \Sigma \circ \phi$. Thus T is isomorphic to Σ. See Ex. 2.33 for more information.

2.5 Coding Map

If we look for a not necessarily continuous conjugacy mapping, then the following idea can be used in some cases.

Definition 2.28. *Let T be measure preserving on a probability space (X,μ). A partition $\mathcal{P} = \{E_0, E_1\}$ of X is* generating *with respect to T if subsets of the form*

$$E_{i_1} \cap T^{-1}E_{i_2} \cap \cdots \cap T^{-(n-1)}E_{i_n}$$

generate all the measurable subsets in X. (We may consider a partition $\mathcal{P}$ consisting of any number of subsets. For example, if there are 3 subsets in $\mathcal{P}$, then we use the ternary expansion in the following discussion.)

Let T be measure preserving on a probability space (X,μ). Suppose that we have a generating partition $\mathcal{P} = \{E_0, E_1\}$ of X. Assume that

$$\bigcap_{n=1}^{\infty} T^{-(n-1)}E_{i_n}$$

contains at most a single element with probability one for any choice of a sequence $i_1, i_2, i_3, \ldots$, which would imply that the coding map ϕ defined in the following is one-to-one almost everywhere. (This assumption is satisfied in the following examples. See p.274 in [Pet].)

Let

$$Y = \prod_{1}^{\infty}\{0,1\}$$

and let $\Sigma : Y \to Y$ be the left shift transformation

$$\Sigma : (i_1, i_2, i_3, \ldots) \mapsto (i_2, i_3, i_4, \ldots)\ .$$

For $x \in X$ define the *coding map* $\phi : X \to Y$ by

$$\phi(x) = (i_1, i_2, i_3, \ldots) \in Y$$

where

$$T^{n-1}x \in E_{i_n}, \quad n \geq 1 .$$

Often $(i_1, \ldots, i_n)$ and $(i_1, i_2, i_3, \ldots)$ are called the $(\mathcal{P}, n)$-name and $\mathcal{P}$-name of x, respectively.

Let $[i_1, \ldots, i_n]$ denote a cylinder set in Y, i.e.,

$$[i_1, \ldots, i_n] = \{(y_1, y_2, y_3, \ldots) \in Y : \, y_k = i_k, \, 1 \leq k \leq n\} .$$

Then

$$\phi(E_{i_1} \cap T^{-1}E_{i_2} \cap \cdots \cap T^{-(n-1)}E_{i_n}) = [i_1, \ldots, i_n] .$$

Define a probability measure ν on Y by

$$\nu([i_1, \ldots, i_n]) = \mu(E_{i_1} \cap T^{-1}E_{i_2} \cap \cdots \cap T^{-(n-1)}E_{i_n}) .$$

Then $\phi : (X, \mu) \to (Y, \nu)$ is measure preserving and

$$\phi \circ T = \Sigma \circ \phi .$$

Thus T and Σ are isomorphic. This method may be used to convert an orbit of a transformation T defined on a continuous space into a digitized sequence, which is easier to study in many applications.

Now consider the case when $X = [0, 1]$. Here is how to visualize ϕ based on the mathematical pointillism. Define $\gamma : Y \to [0, 1]$ by

$$\gamma((i_1, i_2, i_3, \ldots)) = \sum_{n=1}^{\infty} i_n 2^{-n} .$$

Define a set function ν_0 on the images of cylinder sets under γ by

$$\nu_0(\gamma([i_1, \ldots, i_n])) = \nu([i_1, \ldots, i_n]) .$$

Then there is a unique probability measure ν_0 on $[0, 1]$ such that $\nu_0(\gamma(E)) = \nu(E)$ for $E \subset Y$. Even when μ is absolutely continuous, ν_0 can be singular. Put $\psi = \gamma \circ \phi$. Then $\psi \circ T = S \circ \psi$ where $Sx = 2x \pmod 1$, and the following diagram commutes:

$$\begin{tikzcd}
(X,\, \mu) \arrow[r, "T"] \arrow[d, "\phi"'] & (X,\, \mu) \arrow[d, "\phi"'] \\
(Y,\, \nu) \arrow[r, "\Sigma"] \arrow[d, "\gamma"'] & (Y,\, \nu) \arrow[d, "\gamma"'] \\
([0,1],\, \nu_0) \arrow[r, "S"] & ([0,1],\, \nu_0)
\end{tikzcd}$$

In later chapters, to simplify the notation, we sometimes write $[i_1, \ldots, i_n]$ to denote the inverse image

$$\phi^{-1}([i_1, \ldots, i_n]) = E_{i_1} \cap T^{-1}E_{i_2} \cap \cdots \cap T^{-(n-1)}E_{i_n}$$

if there is no danger of confusion.

Example 2.29. Let $X = [0,1]$ and $Tx = 2x \pmod 1$. Take the partition $\mathcal{P} = \{E_0, E_1\}$ where $E_0 = [0, \frac{1}{2})$, $E_1 = [\frac{1}{2}, 1)$. Then we have $\phi(x) = (b_1, b_2, b_3, \ldots)$ where $x = \sum_{n=1}^{\infty} b_n 2^{-n}$, $b_n \in \{0,1\}$. Since $\psi = \gamma \circ \phi = id$, it is clear that ν_0 is Lebesgue measure. See Fig. 2.18 for the graph of $\psi : [0,1) \to [0,1)$, which is nothing but $y = x$.

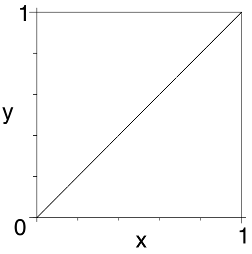

Fig. 2.18. $y = \psi(x)$ for $Tx = 2x \pmod 1$

Example 2.30. Let T be the Λ-transformation given in Ex. 2.24. Recall that T preserves Lebesgue measure. Take the partition $\mathcal{P} = \{E_0, E_1\}$ where $E_0 = [0, \frac{1}{2})$, $E_1 = [\frac{1}{2}, 1]$. Using the coding map ϕ defined by $\mathcal{P}$, we observe that the Λ-transformation is isomorphic to the $(\frac{1}{2}, \frac{1}{2})$-Bernoulli shift transformation. Note that ν_0 is Lebesgue measure. For the graph of $\psi : [0,1] \to [0,1]$ see Fig. 2.19.

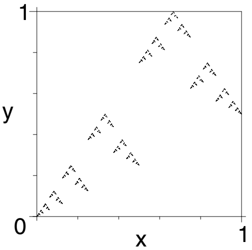

Fig. 2.19. $y = \psi(x)$ for the Λ-transformation

Example 2.31. (Bernoulli transformation) Let $0 < p < 1$. Take the *Bernoulli transformation* T_p on $X = [0,1]$ given by

$$T_p(x) = \begin{cases} x/p\,, & 0 \le x < p\,, \\ (x-p)/(1-p)\,, & p \le x \le 1\,. \end{cases}$$

(Recall that the Bernoulli *shift* transformation is defined on the set of infinite sequences while the Bernoulli transformation is defined on an interval.) Then T_p preserves Lebesgue measure. Choose the partition $\mathcal{P} = \{E_0, E_1\}$ where $E_0 = [0,p)$, $E_1 = [p,1]$. Then ν is the $(p, 1-p)$-Bernoulli measure on $Y = \prod_1^\infty \{0,1\}$. See Fig. 2.20. For $p = \frac{1}{2}$ we have Ex. 2.29.

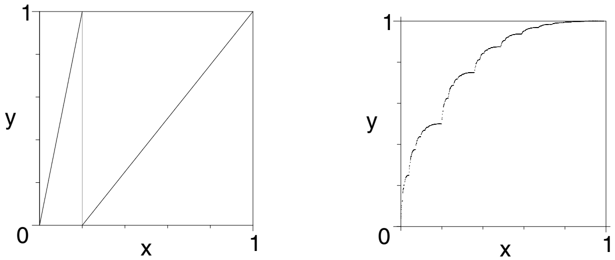

Fig. 2.20. The Bernoulli transformation (left) and $y = \psi(x)$ (right) for $p = 0.2$

Example 2.32. Let $\beta = \frac{\sqrt{5}+1}{2}$ and $Tx = \beta x \pmod 1$. We show that T is isomorphic to a Markov shift transformation. Take the partition $\mathcal{P} = \{E_0, E_1\}$ where $E_0 = [0, \frac{1}{\beta})$, $E_1 = [\frac{1}{\beta}, 1)$. Note that

$$T(E_0) = E_0 \cup E_1\,, \quad T(E_1) = E_0\,.$$

(This can be proved by checking the graph of the β-transformation. See Fig. 2.4.) Hence ν is the Markov measure on $Y = \prod_1^\infty \{0,1\}$ defined by the matrix

$$P = \begin{pmatrix} 1/\beta & 1/\beta^2 \\ 1 & 0 \end{pmatrix}.$$

Note that P is a stochastic matrix since $1/\beta + 1/\beta^2 = 1$. Observe that the binary string '11' does not appear in any sequence $\phi(x)$. (If '1' appears first, then '0' should appear in the next position either by the property of P or by the property of the β-transformation.) Thus the range of ψ is included in the interval $[0, \frac{2}{3}]$ since

$$\psi(x) \le 1 \times 2^{-1} + 0 \times 2^{-2} + 1 \times 2^{-3} + 0 \times 2^{-4} + \cdots = \frac{2}{3}\,.$$

Hence $\nu_0([0, \frac{2}{3}]) = 1$ and $\nu_0((\frac{2}{3}, 1]) = 0$. See Fig. 2.21.

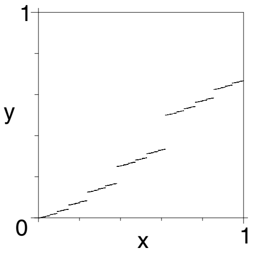

Fig. 2.21. $y = \psi(x)$ for the β-transformation

Example 2.33. Let $X = [0,1] \times [0,1]$ and T be the baker's transformation in Ex. 2.27. Take the partition $\mathcal{P} = \{E_0, E_1\}$ where $E_0 = [0, \frac{1}{2}) \times [0,1]$, $E_1 = [\frac{1}{2}, 1] \times [0,1]$. Let $Y = \prod_{-\infty}^{\infty}\{0,1\}$ be the two-sided $(\frac{1}{2}, \frac{1}{2})$-Bernoulli shift space and let Σ be the left shift transformation on Y. Define $\phi : X \to Y$ by

$$\phi(x) = (\ldots, i_{-1}, \overset{*}{i_0}, i_1, i_2, \ldots)$$

where $T^n x \in E_{i_n}$, $n \in \mathbb{Z}$, and the symbol $*$ indicates the 0th position.

2.6 Maple Programs

In this section we present Maple programs for the study of basic properties of ergodic transformations. Simulations for shift transformations with singular continuous invariant measures are also included. Many of the components of Maple programs in this chapter will be reused later. Thus the explanations in this chapter are more detailed than the ones in later chapters. As the chapter number increases, the level of programs will increase slightly and gradually.

Here is a correspondence of concepts between ergodic theory and Maple programming. See Table 2.1. For a flowchart see Fig. 2.22.

Table 2.1. Comparison of concepts

Ergodic Theory	Maple Programming
a transformation T	a function `T:=x->...`
a starting point x_0	`seed[0]`
iterations of T	a `do` loop
$x_i = T(x_{i-1})$	`seed[i]:=T(seed[i-1])`
an orbit of length n	`for i from 1 to n`
a probability space	a set of inputs

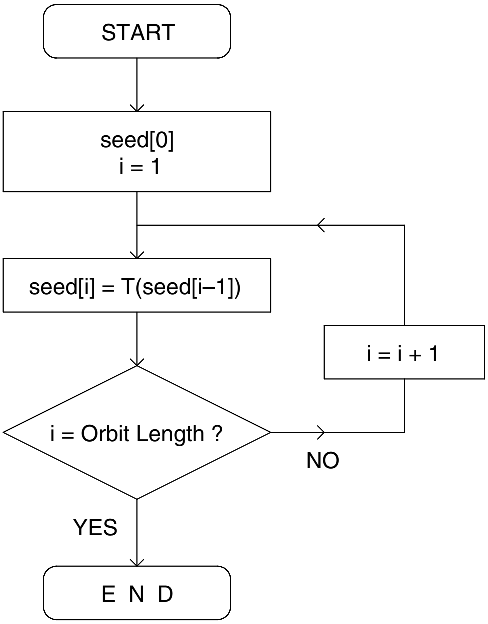

Fig. 2.22. A flowchart for iterations of a transformation T

For the convenience of computer simulations we consider transformations defined on low dimensional Euclidean spaces or the shift transformations on the one-sided shift spaces.

2.6.1 The logistic transformation

Find the invariant density of $Tx = 4x(1-x)$.

```
> with(plots):
```

Define a transformation T.

```
> T:=x-> 4*x*(1-x);
```

Draw the graph $y = Tx$, $0 \le x \le 1$. See Fig. 2.2

```
> g1:=plot(T(x),x=0..1,y=0..1,axes=boxed):
> g2:=listplot([[1/2,0],[1/2,1]],color=green):
> display(g1,g2);
```

Define a probability density function $\rho(x)$.

```
> rho:= x->1/(Pi*sqrt(x*(1-x)));
```

$$\rho := x \to \frac{1}{\pi \sqrt{x\,(1-x)}}$$

Since $\rho(x) \to +\infty$ as $x \to 0+$ or $x \to 1-$, we draw the graph $y = \rho(x)$ only on the interval $[\varepsilon, 1-\varepsilon]$ for some small $\varepsilon > 0$. See Fig. 2.2.

```
> plot(rho(x),x=0.01..0.99);
```

Check whether $\int_0^1 \rho(x)\,\mathrm{d}x = 1$.

```
> int(rho(x),x=0..1);
```

$$1$$

Find the inverse images of b, where b is a point on the positive y-axis.

```
> a:=solve(T(x)=b,x);
```

$$a := \frac{1}{2} + \frac{1}{2}\sqrt{1-b},\ \frac{1}{2} - \frac{1}{2}\sqrt{1-b}$$

In the following definition note that $a_1 \le a_2$.

```
> a1:=1/2-1/2*sqrt(1-b);
```

$$a1 := \frac{1}{2} - \frac{\sqrt{1-b}}{2}$$

```
> a2:=1/2+1/2*sqrt(1-b);
```

$$a2 := \frac{1}{2} + \frac{\sqrt{1-b}}{2}$$

Find inverse images of $b = \frac{2}{3}$ under T.

```
> b:=2/3:
> a1;
```

$$\frac{1}{2} - \frac{\sqrt{3}}{6}$$

```
> a2;
```

$$\frac{1}{2} + \frac{\sqrt{3}}{6}$$

```
> g5:=listplot([[a1,0],[a1,b]],color=green):
> g6:=listplot([[a2,0],[a2,b]],color=green):
> g7:=plot(b,x=0..1,color=green,tickmarks=[2,2]):
> display(g1,g5,g6,g7);
```

See Fig. 2.23.

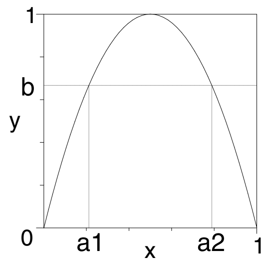

Fig. 2.23. Two inverse images of $b = 1/3$ under the logistic transformation

From now on we treat b again as a symbol.

```
> b:='b';
```

$$b := b$$

```
> a1;
```

$$\frac{1}{2} - \frac{\sqrt{1-b}}{2}$$

```
> a2;
```

$$\frac{1}{2} + \frac{\sqrt{1-b}}{2}$$

Find the measure of $T^{-1}([b,1])$.

```
> measure1:=int(rho(x),x=a1..a2);
```

$$\mathit{measure1} := \frac{2\arcsin(\sqrt{1-b})}{\pi}$$

Find the measure of $[b,1]$.

```
> measure2:=int(rho(x),x=b..1);
```

$$\mathit{measure2} := -\frac{1}{2}\,\frac{-\pi + 2\arcsin(2\,b-1)}{\pi}$$

Compare the two values.

```
> d:=measure1-measure2;
```

$$d := \frac{2\arcsin(\sqrt{1-b})}{\pi} + \frac{1}{2}\,\frac{-\pi + 2\arcsin(2\,b-1)}{\pi}$$

Assume $0 < b < 1$. When we make assumptions about a variable, it is printed with an appended tilde.

```
> assume(b < 1, b > 0):
```

To show $d = 0$ we show $\sin(\pi d) = 0$.

```
> dd:=simplify(Pi*d);
```

$$dd := 2\arcsin(\sqrt{1-b\tilde{}}) - \frac{\pi}{2} + \arcsin(2\,b\tilde{} - 1)$$

```
> expand(sin(dd)):
> simplify(%);
```

$$0$$

This shows that d is an integer. Since $|d| < 1$, we conclude $d = 0$.

2.6.2 Chebyshev polynomials

Consider the transformations $f : [0,1] \to [0,1]$ obtained from Chebyshev polynomials $T : [-1,1] \to [-1,1]$ with $\deg T \geq 2$ after the normalization of the domains and ranges. When $\deg T = 2$, we obtain the logistic transformation. The transformations share the same invariant density function.

```
> with(plots):
```

We need a package for orthogonal polynomials.

```
> with(orthopoly);
```

$$[G,\ H,\ L,\ P,\ T,\ U]$$

The letter 'T' is from the name of a Russian mathematician P.L. Chebyshev. Many years ago his name was transliterated as Tchebyshev, Tchebycheff or Tschebycheff. That is why 'T' stands for Chebyshev polynomials.

Let $w(x) = (1 - x^2)^{-1/2}$ and let $H = L^2([-1, 1], w\,\mathrm{d}x)$ be the Hilbert space with the inner product given by $(u, v) = \int_{-1}^{1} u(x)\overline{v(x)}w(x)\,\mathrm{d}x$ for $u, v \in H$.

Chebyshev polynomials are orthogonal in H. Let us check it! Choose integers m and n. Take small integers for speedy calculation.

```
> m:=5:
> n:=12:
> int(T(m,x)*T(n,x)/sqrt(1-x^2),x=-1..1);
```

$$0$$

Chebyshev polynomials are defined on $[-1, 1]$ and therefore we normalize the domain so that the induced transformations are defined on $[0, 1]$.

Choose the degree of a Chebyshev polynomial.

```
> Deg:=3;
```

Define a Chebyshev polynomial.

```
> Chev:=x->T(Deg,x):
> plot(Chev(x),x=-1..1);
```

The graph of T is omitted to save space. Normalize the domain and the range of T. In this subsection f denotes the transformation since 'T' is reserved for the Chebyshev polynomial.

```
> f:=x->expand((Chev(2*x-1)+1)/2):
```

Now f is a transformation defined on $[0, 1]$.

```
> f(x);
```

$$16\,x^3 - 24\,x^2 + 9\,x$$

Draw the graph $y = f(x)$.

```
> plot(f(x),x=0..1);
```

See Fig. 2.24.

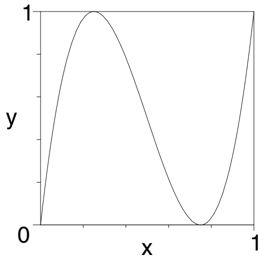

Fig. 2.24. A transformation f defined by a Chebyshev polynomial

Now we prove that the transformations defined by Chebyshev polynomials all share the same invariant density. It is known that $T(n, \cos\theta) = \cos(n\theta)$.

```
> T(Deg,cos(theta))-cos(Deg*theta);
```

$$4\cos(\theta)^3 - 3\cos(\theta) - \cos(3\,\theta)$$

```
> simplify(%);
```

$$0$$

Define a topological conjugacy ϕ based on the formula in [AdMc]. It will be used as an isomorphism for two measure preserving transformations f and g. The formula in Ex. 2.24 may be used, too.

```
> phi:=x->(1+cos(Pi*x))/2;
```

$$\phi := x \to \frac{1}{2} + \frac{1}{2}\cos(\pi\,x)$$

Draw the graph $y = \phi(x)$.

```
> plot(phi(x),x=0..1,y=0..1,axes=boxed);
```

See the left graph in Fig. 2.25.

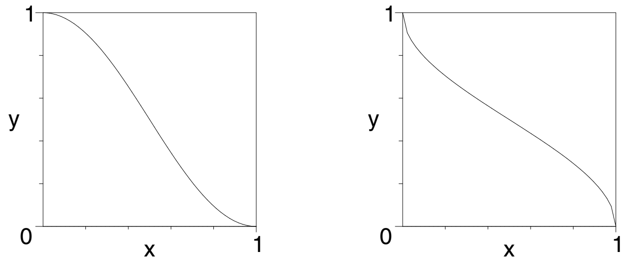

Fig. 2.25. The topological conjugacy $\phi(x)$ (left) and its inverse (right)

Find the inverse of ϕ.

```
> psi:=x->arccos(2*x-1)/Pi;
```

$$\psi := x \to \frac{\arccos(2\,x-1)}{\pi}$$

Draw the graph $y = \psi(x) = \phi^{-1}(x)$.

```
> plot(psi(x),x=0..1,y=0..1,axes=boxed);
```

See the right graph in Fig. 2.25.
Check $\phi(\psi(x)) = x$.

```
> phi(psi(x));
```

$$x$$

Define a transformation $g(x) = \psi(f(\phi(x)))$.

```
> g:=x->psi(f(phi(x)));
```

$$g := x \to \psi(\mathrm{f}(\phi(x)))$$

Draw the graph $y = g(x)$.

```
> plot(g(x),x=0..1,y=0..1,axes=boxed):
```

See Fig. 2.26.

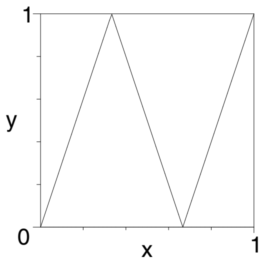

Fig. 2.26. A transformation g conjugate to f through ϕ

It is obvious that g preserves Lebesgue measure $\mathrm{d}x$ on $X = [0, 1]$. Find the invariant probability measure μ for f.

$$\begin{array}{ccc} (X, \mathrm{d}x) & \xrightarrow{g} & (X, \mathrm{d}x) \\ \phi\big\downarrow & & \phi\big\downarrow \\ (X, \mu) & \xrightarrow{f} & (X, \mu) \end{array}$$

Note that the inverse image of $[\phi(x), 1]$ under ϕ is $[0, x]$, which has Lebesgue measure equal to x. For μ to be an f-invariant measure, it must satisfy

$$\mu([\phi(x), 1]) = x$$

for $0 \le x \le 1$. Thus $\mu([y, 1]) = \psi(y)$ and $\mu([0, y]) = 1 - \psi(y)$ for $0 \le y \le 1$.

```
> -diff(psi(y),y);
```

$$\frac{1}{\sqrt{-y^2 + y}\,\pi}$$

Hence

$$\rho(y) = \frac{\mathrm{d}}{\mathrm{d}y}(1 - \psi(y)) = \frac{1}{\sqrt{y(1-y)}\,\pi} \ .$$

2.6.3 The beta transformation

Find the invariant measure of the β-transformation.

```
> with(plots):
> beta:=(1+sqrt(5))/2:
```

Define the transformation T.

```
> T:=x-> frac(beta*x):
```
Define the invariant probability density function $\rho(x)$.
```
> rho:=x->piecewise( 0<=x and x< 1/beta, beta^3/(1+beta^2),
1/beta <= x and x < 1, beta^2/(1+beta^2) );
```
$$\rho := x \to \text{piecewise}(0 \le x \textbf{ and } x < \frac{1}{\beta}, \frac{\beta^3}{1+\beta^2}, \frac{1}{\beta} \le x \textbf{ and } x < 1, \frac{\beta^2}{1+\beta^2})$$
```
> plot(T(x),x=0..1,y=0..1,axes=boxed);
```
See the left graph in Fig. 2.4.
```
> b0:=T(1);
```
$$b0 := -\frac{1}{2} + \frac{\sqrt{5}}{2}$$
```
> plot(rho(x),x=0..1);
```
See the right graph in Fig. 2.4.
```
> int(rho(x),x=0..1);
```
$$\frac{15 + 7\sqrt{5}}{(5+\sqrt{5})(2+\sqrt{5})}$$
```
> simplify(%);
```
$$1$$
When we make assumptions about a variable, a tilde is appended.
```
> assume(b >= 0, b <= 1);
> b;
```
$$b\tilde{}$$
```
> m1:=beta^3/(1+beta^2):
> m2:=beta^2/(1+beta^2):
```
Find the cumulative density function (cdf) for the invariant measure. In the following, `cdf`(b) is the measure of the interval $[0, b]$.
```
> cdf:=piecewise( b<=1/beta, m1*b, b>=1/beta, m1*(1/beta)+
m2*(b-1/beta) ):
> plot(cdf(b),b=0..1);
```
See Fig. 2.27.
Find the inverse image (or inverse images) of a point b on the y-axis.
```
> a1:=solve(beta*x=b,x);
```
$$a1 := \frac{2\, b\tilde{}}{1+\sqrt{5}}$$
```
> a2:=solve(beta*x-1=b,x);
```
$$a2 := \frac{2\,(1+b\tilde{})}{1+\sqrt{5}}$$
In the following, `cdf_inverse`(b) is the measure of the inverse image of $[0, b]$.
```
> cdf_inverse:=piecewise( b <= b0, m1*a1 + m2*(a2-1/beta),
b >= b0, m1*a1 + m2*(1-1/beta) ):
```

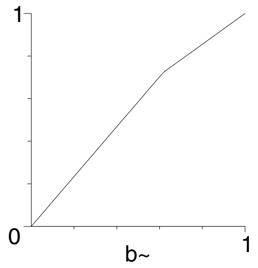

Fig. 2.27. The cumulative density function for the β-transformation

Check the graph, which is not shown here.

```
> plot(cdf_inverse(b),b=0..1);
> simplify(cdf-cdf_inverse);
```

$$0$$

This shows that $\rho(x)$ is invariant under T.

2.6.4 The baker's transformation

Using the mathematical pointillism we draw images of a rectangle under the iterates of the baker's transformation. See Fig. 2.8.

```
> with(plots):
```

Define the baker's transformation.

```
> T:=(x,y)->(frac(2*x), 0.5*y+0.5*trunc(2.0*x));
```

$$T := (x,\, y) \to (\mathrm{frac}(2\,x),\, 0.5\,y + 0.5\,\mathrm{trunc}(2.0\,x))$$

```
> S:=4000:
```

Choose a starting point of an orbit of length S.

```
> seed[0]:=(sqrt(3.0)-1,evalf(Pi-3)):
```

Generate S points evenly scattered in the unit square.

```
> for i from 1 to S do seed[i]:=T(seed[i-1]): od:
> pointplot({seq([seed[i]],i=1..S)},symbolsize=1);
```

See the left plot in Fig. 2.28.

Generate S points evenly scattered in $C = \{(x,y) : 0 \le x \le \frac{1}{2}, 0 \le y \le 1\}$.

```
> for i from 1 to S do
> image0[i]:=(seed[i][1]/2,seed[i][2]): od:
> pointplot({seq([image0[i]],i=1..S)},symbolsize=1);
```

See the right plot in Fig. 2.28.

Find $T(C)$.

```
> for i from 1 to S do image1[i]:=T(image0[i]): od:
> pointplot({seq([image1[i]],i=1..S)},symbolsize=1);
```

See the first graph in Fig. 2.8.

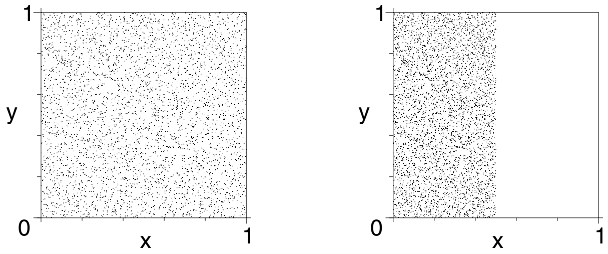

Fig. 2.28. Evenly scattered points in the unit square (left) and a rectangle C (right)

Find $T^2(C)$.

```
> for i from 1 to S do image2[i]:=T(T(image0[i])): od:
> pointplot({seq([image2[i]],i=1..S)},symbolsize=1);
```

See the second graph in Fig. 2.8.
Find $T^3(C)$.

```
> for i from 1 to S do image3[i]:=T(T(T(image0[i]))): od:
> pointplot({seq([image3[i]],i=1..S)},symbolsize=1);
```

See the third graph in Fig. 2.8.
In the preceding three `do` loops, if we want to save computing time and memory, we may repeatedly use `seed[i]:=T(seed[i]):` and plot `seed[i]`, $1 \le i \le S$. See Maple Program 2.6.5.

2.6.5 A toral automorphism

Find the successive images of $D = \{(x, y) : 0 \le x \le \frac{1}{2}, 0 \le y \le \frac{1}{2}\}$ under the Arnold cat mapping.

```
> with(plots):
```

Define a toral automorphism.

```
> T:=(x,y)->(frac(2*x+y),frac(x+y));
> seed[0]:=(evalf(Pi-3),sqrt(2.0)-1):
> S:=4000:
> for i from 1 to S do seed[i]:=T(seed[i-1]): od:
```

Find S points in D.

```
> for i from 1 to S do
> seed[i]:=(seed[i][1]/2,seed[i][2]/2): od:
> pointplot({seq([seed[i]],i=1..S)},symbolsize=1);
```

See Fig. 2.29.
Find $T(D)$.

```
> for i from 1 to S do seed[i]:=T(seed[i]): od:
> pointplot({seq([seed[i]],i=1..S)},symbolsize=1);
```

See the first plot in Fig. 2.9.

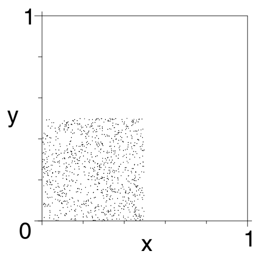

Fig. 2.29. Evenly scattered points in the square D

Find $T^2(D)$.

```
> for i from 1 to S do seed[i]:=T(seed[i]): od:
> pointplot({seq([seed[i]],i=1..S)},symbolsize=1);
```

See the second plot in Fig. 2.9.
Find $T^3(D)$.

```
> for i from 1 to S do seed[i]:=T(seed[i]): od:
> pointplot({seq([seed[i]],i=1..S)},symbolsize=1);
```

See the third plot in Fig. 2.9.

2.6.6 Modified Hurwitz transformation

```
> with(plots):
```

Define two constants α and β.

```
> alpha:=(sqrt(5)-1)/2: beta:=(3-sqrt(5))/2:
```

Define a new command.

```
> bracket:=x->piecewise( 0 <= x, floor(x + beta), 0 > x,
-floor(-x + beta) );
```

Define the transformation T.

```
> T:=x-> 1/x - bracket(1/x);
```

Define a probability density where C is an undetermined normalizing constant.

```
> rho:=x->piecewise(-alpha<=x and x<-beta,C/(2-x),-beta<=x
and x<beta, C/(1-(1/4)*x^2), beta<=x and x<alpha, C/(2+x)):
```

Compute the integral of $\rho(x)$ over the interval $-\alpha < x < \alpha$.

```
> int(rho(x),x=-alpha..alpha);
```

$$2\,C\ln(3+\sqrt{5}) - 2\,C\ln(\sqrt{5}+1)$$

Choose C so that ρ is a probability density function.

```
> C:=1/(-2*ln(sqrt(5)+1)+2*ln(3+sqrt(5)));
```

$$C := \frac{1}{-2\ln(\sqrt{5}+1) + 2\ln(3+\sqrt{5})}$$

Plot the invariant density function $\rho(x)$.

```
> plot(rho(x),x=-alpha..alpha);
```

See Fig. 2.14.

2.6.7 A typical point of the Bernoulli measure

Find a *typical* binary sequence x_0 with respect to the $(\frac{1}{4}, \frac{3}{4})$-Bernoulli measure. Consult Sect. 2.3 and see Fig. 2.30.

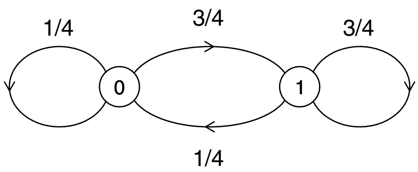

Fig. 2.30. The graph corresponding to the $(\frac{1}{4}, \frac{3}{4})$-Bernoulli shift

Generate four numbers 0,1,2,3 with probability $\frac{1}{4}$ each.

```
> ran:=rand(0..3):
```

Choose the number of significant decimal digits for x_0.

```
> N_decimal:=10000:
> N_decimal*log[2.](10);
```

$$33219.28095$$

```
> N_binary:=33300:
```

This is the number of binary significant digits needed to produce a decimal number with `N_decimal` significant decimal digits.

```
> evalf(2^(-N_binary),10);
```

$$0.5025096306\,10^{-10024}$$

```
> for j from 1 to N_binary do d[j]:=ceil(ran()/3): od:
```

Find the number of the bit '1' in the binary sequence $\{d_j\}$.

```
> num_1:=add(d[j],j=1..N_binary);
```

$$num\_1 := 24915$$

The following number should be close to $\frac{3}{4}$ by the Birkhoff Ergodic Theorem. See Chap. 3 for more information.

```
> evalf(num_1 / N_binary);
```

$$0.7481981982$$

Convert the binary number $0.d_1d_2d_3\ldots$ into a decimal number.

```
> Digits:=N_decimal:
> M:=N_binary/100;
```

$$M := 333$$

To compute $\sum_{s=1}^{33300} d_s 2^{-s}$, calculate 100 partial sums first then add them all. In the following `partial_sum[k]` is stored as a quotient of two integers.

```
> for k from 1 to 100 do
> partial_sum[k]:=add(d[s+M*(k-1)]/2^s,s=1..M): od:
> x0:=evalf(add(partial_sum[k]/2^(M*(k-1)),k=1..100));
```

$$x0 := 0.96477508532181821 5215\ldots$$

This is a *typical* point for the Bernoulli measure represented on $[0, 1]$.

2.6.8 A typical point of the Markov measure

Find a *typical* binary Markov sequence x_0 defined by a stochastic matrix P. Consult Sect. 2.3 and see Fig. 2.31.

```
> with(linalg):
> P:=matrix([[1/3,2/3],[1/2,1/2]]);
```

$$P := \begin{bmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

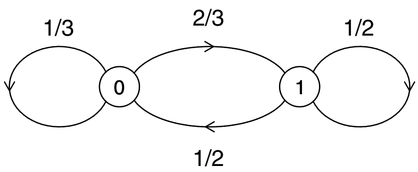

Fig. 2.31. The graph corresponding to P

Choose the positive vector in the following:

```
> eigenvectors(transpose(P));
```

$$[1,\ 1,\ \{\left[1,\ \frac{4}{3}\right]\}],\ [\frac{-1}{6},\ 1,\ \{[-1,\ 1]\}]$$

The Perron-Frobenius eigenvector is given by the following:

```
> v:=[3/7,4/7]:
> evalm(v&*P);
```

$$\left[\frac{3}{7},\ \frac{4}{7}\right]$$

Observe that the rows of P^n converge to $\mathbf{v}$. See Theorem 5.21.

```
> evalf(evalm(P&^20),10);
```

$$\begin{bmatrix} 0.4285714286 & 0.5714285714 \\ 0.4285714286 & 0.5714285714 \end{bmatrix}$$

```
> 3/7.;
```

$$0.4285714286$$

```
> 4/7.;
```

$$0.5714285714$$

Construct a typical binary Markov sequence of length `N_binary`.

```
> ran0:=rand(0..2): ran1:=rand(0..1):
> N_decimal:=10000:
> N_decimal*log[2.](10);
```

$$33219.28095$$

```
> N_binary:=33300;
```

$$N\_binary := 33300$$

```
> evalf(2.0^(-N_binary),5);
```

$$0.50251\,10^{-10024}$$

As in the previous case this shows that the necessary number of binary digits is equal to 33300 when we use 10000 significant decimal digits.

```
> d[0]:=1:
> for j from 1 to N_binary do
> if d[j-1]=0 then d[j]:=ceil(ran0()/2):
> else d[j]:=ran1(): fi
> od:
```

Count the number of the binary symbol 1.

```
> num_1:=add(d[j],j=1..N_binary);
```

$$num\_1 := 19111$$

The following is an approximate measure of the cylinder set $[1]_1$, which should be close to $\frac{4}{7} = 0.5714\ldots$ by the Birkhoff Ergodic Theorem in Chap. 3.

```
> evalf(num_1 / N_binary,10);
```

$$0.5739039039$$

Convert the binary number $0.d_1d_2d_3\ldots$ into a decimal number.

```
> Digits:=N_decimal;
```

$$Digits := 10000$$

In computing `add(d[s]/2^s,s=1..33300)`, we first divide it into 10 partial sums, and calculate each partial sum separately, then finally add them all.

```
> M:=N_binary/10;
```

$$M := 3330$$

In the following `partial_sum[k]` is stored as a quotient of two integers.

```
> for k from 1 to 10 do
> partial_sum[k]:=add(d[s+M*(k-1)]/2^s,s=1..M): od:
> x0:=evalf(add(partial_sum[k]/2^(M*(k-1)),k=1..10));
```

$$x0 := 0.5897685203978253457104972\,1\ldots$$

This is a *typical* point for the Markov measure represented on $[0,1]$.

2.6.9 Coding map for the logistic transformation

Let $E_0 = [0, \frac{1}{2})$, $E_1 = [\frac{1}{2}, 1]$. For $x \in [0,1]$ define a binary sequence b_n by $T^{n-1}x \in E_{b_n}$. We identify $(b_1, b_2, \ldots)$ with $\psi(x) = \sum b_n 2^{-n}$ and sketch the graph of ψ. Consult Sect. 2.5 for the details.

```
> with(plots):
> Digits:=100:
```

Choose the logistic transformation.

```
> T:=x-> 4*x*(1-x):
```

Choose the number of points on the graph.

```
> N:=2000:
> M:=12:
```

Choose a starting point of an orbit.

```
> seed[0]:=evalf(Pi-3):
```

Find $\phi(x)$.

```
> for n from 1 to N+M-1 do
> seed[n]:=T(seed[n-1]):
> if seed[n] < 1/2 then b[n]:=0: else b[n]:=1: fi:
> od:
```

Find $\psi(x)$.

```
> for n from 1 to N do
> psi[n]:=add(b[n+i-1]/2^i,i=1..M): od:
```

We don't need many `Digits` now.

```
> Digits:=10:
> pointplot([seq([seed[n],psi[n]],n=1..N)],symbolsize=1);
```

See Fig. 2.32.

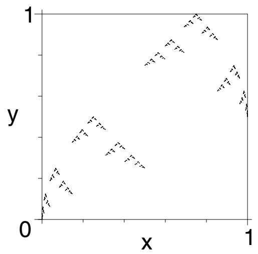

Fig. 2.32. $y = \psi(x)$ for the logistic transformation

http://www.springer.com/978-3-540-23121-9

Computational Ergodic Theory
Choe, G.H.
2005, XX, 453 p. 250 illus., Hardcover
ISBN: 978-3-540-23121-9