## 1.1 Ergodic Transformations

In this lecture we will define the notion of *ergodicity*, or *metric indecomposability.* Ergodic measure-preserving transformations are the building blocks of all measure-preserving transformations (as prime numbers are building blocks of natural numbers). Moreover, ergodicity will play an important role in Birkhoff ergodic theorem.

Let $(X, \mathscr{A}, \mu)$ be a *finite* measure space. In this section (and more in general when we want to talk of ergodic transformations) we will assume that $(X, \mathscr{A}, \mu)$ is a *probability space.* This is not a great restriction, since if $\mu(X) < \infty$, if we consider $\mu/\mu(X)$ (that is, the measure *rescaled* by $\mu(X)$), then $\mu/\mu(X)$ is a measure with total mass 1 and $(X, \mathscr{A}, \mu/\mu(X))$ is a probability space. Let $T : X \to X$ be a measure-preserving transformation.

**Definition 1.1.1.** A set $A \subset X$ is called *invariant* invariant under $T$ (or simply invariant if the transformation is clear from the context) if

$$T^{-1}(A) = A.$$

Remark that in the definition we consider preimages $T^{-1}$. This is important if the transformation is not invertible.

**Exercise 1.1.1.** If $T$ is invertible, show that $A$ is invariant if and only if $T(A) = A$.

**Example 1.1.1.** Assume that $T$ is invertible and that $x \in X$ is a periodic point of period $n$. Then

$$A = \{x, T(x), \ldots, T^{n-1}(x)\} \tag{1.1}$$

is an invariant set.

**Definition 1.1.2.** A measure preserving transformation $T$ on a probability space $(X, \mathscr{A}, \mu)$ is *ergodic* if and only if for any set measurable $A \in \mathscr{A}$ such that $T^{-1}(A) = A$ either $\mu(A) = 0$ or $\mu(A) = 1$, that is *all invariant sets are trivial* from the point of view of the measure.

**Remark 1.1.1.** A transformation which is *not* ergodic is reducible in the following sense. If $A \in \mathscr{A}$ is an *invariant* measurable set of positive measure $\mu(A) > 0$, then we can consider the restriction $\mu_A$ of the measure $\mu$ to $A$, that is the measure defined by

$$\mu_A(B) = \frac{\mu(A \cap B)}{\mu(A)}, \qquad \text{for all } B \in \mathscr{A}.$$

It is easy to check that $\mu_A$ is again a probability measure and that it is invariant under $T$ (Exercise). Remark that we used that $\mu(A) > 0$ to renormalize $\mu_A$. Similarly, also the restriction $\mu_{X\setminus A}$ of the measure $\mu$ to the complement $X\setminus A$, given by

$$\mu_{X\setminus A}(B) = \frac{\mu(X\setminus A \cap B)}{\mu(X\setminus A)}, \qquad \text{for all } B \in \mathscr{A},$$

is an invariant probability measure (Exercise). Remark that here we used that $\mu(X\setminus A) > 0$ since $\mu(A) < 1$. Thus, we have decomposed $\mu$ into two invariant measures $\mu_A$ and $\mu_{X\setminus A}$ and one can study separately the two dynamical systems obtained restricting $T$ to $A$ and to $X\setminus A$. In this sense, non ergodic transformations are *decomposable* while ergodic transformations are *indecomposable.*

As prime numbers, that cannot be written as product of prime numbers, are the basic building block used to decompose any other integer number, similarly ergodic transformations, that are indecomposable in this metric sense, are the basic building block used to study any other measure-preserving transformation.

**Exercise 1.1.2.** Let $(X, \mathscr{A})$ be a measurable space and $T : X \to X$ be a transformation.

(a) Check that if $\mu_1$ and $\mu_2$ are probability measures on $(X, \mathscr{A})$, then any linear combination

$$\mu = \lambda\mu_1 + (1-\lambda)\mu_2, \qquad \text{where} \quad 0 \le \lambda \le 1,$$

is again a measure. Check that it is a probability measure.

(b) Let $\mu$ be a measure on $(X, \mathscr{A})$ preserved by $T$. Let $A \in \mathscr{A}$ be a measurable set with positive measure $\mu(A) > 0$. Check that by setting

$$\mu_1(B) = \frac{\mu(A \cap B)}{\mu(A)} \quad \text{for all } B \in \mathscr{A}, \qquad \mu_2(B) = \frac{\mu(A^c \cap B)}{\mu(A^c)} \quad \text{for all } B \in \mathscr{A},$$

(where $A^c = X \backslash A$ denotes the complement of $A$ in $X$) one defines two probability measures $\mu_1$ and $\mu_2$. Show that if $A$ is *invariant* under $T$, then both $\mu_1$ and $\mu_2$ are invariant under $T$.

(c) Show using the two previous points that a probability measure $\mu$ invariant under $T$ is *ergodic* if it cannot be written as strict linear combination of two invariant probability measures for $T$, that is as

$$\mu = \lambda\mu_1 + (1-\lambda)\mu_2, \qquad \text{where} \quad 0 < \lambda < 1, \quad \mu_1 \neq \mu_2. \tag{1.2}$$

[The converse is also true, but harder to prove: a measure $\mu$ is ergodic *if and only* if it cannot be decomposed as in (**??**).]

Part $(a)$ of Exercise **??** shows that the space of all probability $T$-invariant measures is *convex* (recall that a set $C$ is a convex if for any $x, y \in C$ and any $0 \le \lambda \le 1$ the points $\lambda x + (1-\lambda)\lambda y$ all belong to $C$). If $C$ is a convex set, the *extremal points* of $C$ are the points $x \in C$ which cannot be expressed as linear combination of the other points, that is, there is no $0 < \lambda < 1$ and $x_1 \neq x_2$ such that $x = \lambda x_1 + (1-\lambda)x_2$. Thus, Part $(b)$ of Exercise **??** shows that *ergodic probability measures are extremal points* of the set of all probability $T$-invariant measures.

Let us now give an example of a *non-ergodic* transformation and one of an ergodic one.

**Example 1.1.2.** [**Rational rotations are not ergodic**] Let $X = S^1$, $\mathscr{B}$ its Borel subsets and $\lambda$ the Lebesgue measure. Consider a rational rotation $R_\alpha : S^1 \to S^1$ where $\alpha = p/q$ with $p, q$ coprime. Consider for example the following set in $\mathbb{R}/\mathbb{Z}$

$$A = \bigcup_{i=0}^{q-1} \left[\frac{i}{q}, \frac{i}{q} + \frac{1}{2q}\right].$$

The set $A$ in $S^1$ is shown in Figure **??**.

The set $A$ is clearly invariant under $R_{p/q}$, since the clockwise rotation $R_{p/q}$ by $2\pi p/q$ sends each interval into another one. Since $A$ is union of $q$ intervals of equal length $1/2q$, $\lambda(A) = 1/2$, so $0 < \lambda(A) < 1$. Thus, since we constructed an invariant set whose measure is neither 0 nor 1, $R_{p/q}$ is *not* ergodic.

[Remark that any point is periodic of period $q$ and since $R_\alpha$ is invertible, any periodic orbit is an invariant set, but it has measure zero. Thus, to show that $R_{p/q}$ is *not* ergodic, we need to construct an invariant set with *positive* measure and here we constructed one by considering the orbit of an interval.]

In the next lecture we will prove that on the other hand *irrational rotations* are ergodic. Thus, a rotation $R_\alpha$ is ergodic if and only if $\alpha$ is irrational.

Let us show that the doubling map is ergodic directly using the definition of ergodicity.

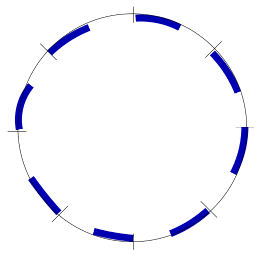

Figure 1.1: An invariant set $A$ with $0 < \lambda < 1$ for a rational rotation $R_{p/q}$ with $q = 8$.

**Example 1.1.3.** [**The doubling map is ergodic**] Let $X = \mathbb{R}/\mathbb{Z}$, $\mathscr{B}$ be Borel sets of $\mathbb{R}$ and $\lambda$ the Lebesgue measure. Let $T : X \to X$ be the doubling map, that is $T(x) = 2x \mod 1$. Let us show that the doubling map is ergodic.

Let $A \in \mathscr{B}$ be an invariant set, so that $T^{-1}(A) = A$. We have to show that $\lambda(A)$ is either 0 or 1. If $\lambda(A) = 1$, we are done. Let us assume that $\lambda(A) < 1$ and show that then $\lambda(A)$ has to be 0. Since we assume that $\lambda(A) < 1$, $\lambda(X\backslash A) > 0$. One can show (see Theorem **??** in the Extra) that a measurable set of positive measure is well approximated by small intervals in the following sense: given $\epsilon > 0$, we can find $n \in \mathbb{N}$ and a dyadic interval $I$ of length $1/2^n$ such that

$$\lambda(I\backslash A) > (1 - \epsilon)\lambda(I), \tag{1.3}$$

that is, the proportion of points in $I$ which are *not* is $A$ is at least $1-\epsilon$.Recall that we showed that if $I$ is a dyadic interval of length $1/2^n$, its images $T^k(I)$ under the doubling map for $0 \le k \le n$ are again dyadic intervals of size $1/2^{n-k}$. In particular, the length $\lambda(T^k(I))$ is $2^k\lambda(I)$ and for $k = n$, $T^n(I) = \mathbb{R}/\mathbb{Z} = X$. Furthermore we can calculate that

$$\lambda(T^n(I\backslash A)) = 2^n\lambda(I\backslash A) = (1 - \epsilon).$$

We now need to show that $T^n(I\backslash A) \subseteq X\backslash A$. To do this suppose $x \in T^n(I\backslash A)$ and $x \in A$. We then have that there exists $y \in I\backslash A$ such that $T^n(y) = x$ and therefore $y \in T^{-n}(A) = A$. This is a contradiction since $y \in I\backslash A$. Therefore there is no such $x$ and $T^n(I\backslash A) \subseteq X\backslash A$. Putting this together means that

$$\lambda(X\backslash A) \ge \lambda(T^n(I\backslash A)) \ge 1 - \epsilon.$$

Since $\lambda(X\backslash A) \ge 1 - \epsilon$ holds for all $\epsilon > 0$, we conclude that $\lambda(X\backslash A) = 1$ and hence $\lambda(A) = 0$. This concludes the proof that the doubling map is ergodic.

To prove directly from the definition that the doubling map is ergodic, we had to use a fact from measure theory that we stated without a proof (the existence of the interval in (**??**), see also Theorem **??** in the Extra). In the following lecture, §3.6, we will see how to prove ergodicity using Fourier series and we will see that is possible to reprove that the doubling map is ergodic by using Fourier series, which gives a simpler and self-contained proof.

**Exercise 1.1.3.** Let $X = [0, 1]$, $\mathscr{B}$ the Borel $\sigma-$algebra, $\lambda$ the $1-$dimensional Lebesgue measure. Let $m > 1$ is an integer and consider the linear expanding map $T_m(x) = mx \mod 1$. Show that $T_m$ is ergodic.
[*Hint*: mimic the previous proof that the doubling map is ergodic.]

### Ergodicity via invariant functions

The following equivalent definition of ergodicity is also very useful to prove that a transformation is ergodic:

**Lemma 1.1.1 (Ergodicity via measurable invariant functions).** *A measure preserving transformation $T : X \to X$ is ergodic if and only if, any measurable function $f : X \to \mathbb{R}$ that is invariant, that is such that*

$$f \circ T = f \qquad \textit{almost everywhere} \quad (\textit{that is,} \;\; f(T(x)) = f(x) \quad \textit{for } \mu - \textit{almost every } x \in X) \tag{1.4}$$

*is $\mu$-almost everywhere constant (that is, there exists $c \in \mathbb{R}$ such that $f(x) = c$ for $\mu$-a.e. $x \in X$).*

*Proof.* Assume first that (**??**) hold. Let $A \in \mathscr{B}$ be an invariant set. Consider its characteristic function $\chi_A$, which is measurable (see Example 3.4.1 in §3.4). Let us check that $\chi_A$ is an invariant *function*, that is $\chi_A \circ T = \chi_A$. Recall that we showed last time that $\chi_A \circ T = \chi_{T^{-1}(A)}$ (see equation (3.14) in §3.4). Thus

$$\chi_A \circ T = \chi_{T^{-1}(A)} = \chi_A, \qquad (\text{since } \; T^{-1}(A) = A).$$

Thus, we can apply (**??**) to $\chi_A$ and conclude that $\chi_A$ is almost everywhere constant. But since an indicatrix function takes only the values 0 and 1, either

$$\chi_A = 0 \; a.e. \qquad \Rightarrow \qquad \mu(A) = \int_A \chi_A \mathrm{d}\mu = 0, \qquad \text{or} \tag{1.5}$$

$$\chi_A = 1 \; a.e. \qquad \Rightarrow \qquad \mu(A) = \int_A \chi_A \mathrm{d}\mu = 1. \tag{1.6}$$

This concludes the proof that $T$ is ergodic.

Let us assume now that $T$ is ergodic and prove (**??**). Let $f : X \to \mathbb{R}$ be a measurable function. Assume that $f \circ T = f$ almost everywhere. One can redefine $f$ on a set of measure zero so that the redefined function, which we will still call $f$, is invariant everywhere, that is $f(T(x) = f(x)$ for all $x \in X$ (see Exercise **??**).

Consider the sets

$$A_t = \{x \in X, \text{ such that } f(x) > t\}, \qquad t \in \mathbb{R}.$$

The set $A_t$ are called *level sets* of the function $f$ and they are measurable since $A_t = f^{-1}((t, +\infty))$ and $f$ is measurable, which by definition means that the preimage of each interval is in $\mathscr{B}$. Let us show that each $A_t$ is *invariant*:

$$\begin{aligned} T^{-1}(A_t) &= \{x \in X, \text{ such that } T(x) \in A_t\} \qquad (\text{by definition of preimage}) \\ &= \{x \in X, \text{ such that } f(T(x)) > t\} \qquad (\text{by definition of } T(x) \in A_t) \\ &= \{x \in X, \text{ such that } f(x) > t\} \qquad (\text{since } f(T(x)) = f(x)) \\ &= A_t. \end{aligned}$$

Thus, since $T$ is ergodic, for each $t \in \mathbb{R}$ either $\mu(A_t) = 0$ or $\mu(A_t) = 1$. If the function $f$ is constant equal to $c$ almost everywhere, then $\mu(A_t) = 1$ for all $t < c$ (since $f(x) = c > t$ for a.e. $x \in X$), while $\mu(A_t) = 0$ for all $t \geq c$ (since $f(x) = c \leq t$ for a.e. $x \in X$). On the other hand, if $f$ is *not* constant almost everywhere, one can find a level set $t_0$ such that $0 < \mu(t_0) < 1$, which is a contradiction with what we just proved. Thus, $f$ has to be constant almost everywhere. This shows that (**??**) holds when $T$ is ergodic. □

**Exercise 1.1.4.** Let $T : X \to X$ be a measure preserving transformation of the measured space $(X, \mathscr{B}, \mu)$. Let $f$ be a measurable function $f : X \to \mathbb{R}$ that is invariant almost everywhere under $T$, that is $f \circ T(x) = f(x)$ for $\mu$-almost every $x \in X$.

(a) Consider the set

$$E = \bigcup_{n \in \mathbb{N}} T^{-n}(N), \qquad \text{where} \quad N = \{ \ x \text{ such that } f(T(x)) \neq f(x) \ \}.$$

Show that $\mu(E) = 0$ and that $T^{-1}(E) \subset E$.

(b) Define a new function $\tilde{f}$ by setting:

$$\tilde{f}(x) = \begin{cases} f(x) & \text{if } x \notin E \\ 0 & \text{if } x \in E, \end{cases}$$

Show that $f = \tilde{f}$ $\mu$-almost everywhere and that $\tilde{f} \circ T = \tilde{f}$ holds everywhere, that is $\tilde{f}(T(x)) = \tilde{f}(x)$ for *all* $x \in X$.

**Exercise 1.1.5.** Let $R_\alpha : \mathbb{R}/\mathbb{Z} \to \mathbb{R}/\mathbb{Z}$ be a rational rotation, where $\alpha = p/q$ and $p, q$ are coprime. Show that $R_\alpha$ is *not* ergodic by exhibiting a non-constant invariant function.

One can show that in Lemma **??** instead then considering all functions which are measurable, it is enough to check that (**??**) holds for all integrable functions $f \in L^1(\mu)$ or for all square-integrable functions $f \in L^2(\mu)$ (the definition of these spaces was given in §3.4 We have the following two variants of Lemma **??**:

**Lemma 1.1.2 (Ergodicity via invariant integrable functions).** *Let $(X, \mathscr{B}, \mu)$ be a probability space and $T : X \to X$ a measure preserving transformation. Then $T : X \to X$ is ergodic if and only if*

$$\textit{for all } f \in L^1(X, \mathscr{B}, \mu), \quad f \circ T = f \ \mu - a.e. \quad \Rightarrow \quad f \ \mu - a.e. \ \textit{constant.}$$

**Lemma 1.1.3 (Ergodicity via invariant square integrable functions).** *Let $(X, \mathscr{B}, \mu)$ be a probability space and $T : X \to X$ a measure preserving transformation. Then $T : X \to X$ is ergodic if and only if*

$$\textit{for all } f \in L^2(X, \mathscr{B}, \mu), \quad f \circ T = f \ \mu - a.e. \quad \Rightarrow \quad f \ \mu - a.e. \ \textit{constant.}$$

**Exercise 1.1.6.** Let $(X, \mathscr{B}, \mu)$ be a measured space and $T : X \to X$ be a *measure-preserving transformation*. Consider the space $L^2(X, \mathscr{B}, \mu)$ of square integrable functions.

Let $U_T : L^2(X, \mathscr{B}, \mu) \to L^2(X, \mathscr{B}, \mu)$ be given by

$$U_T(f) = f \circ T.$$

(a) Check that $L^2(X, \mathscr{B}, \mu)$ is a vector space and $U_T$ is *linear*, that is for all $f, g \in L^2(\mu)$ and $a, b \in \mathbb{R}$,

$$af + bg \in L^2(X, \mathscr{B}, \mu) \quad \text{and} \quad U_T(af + bg) = aU_T(f) + bU_T(g).$$

(b) Show that $U_T$ preserves the $L^2(\mu)-$norm, that is for any $f \in L^2(\mu)$ we have $\|U_T(f)\|_2 = \|f\|_2$. Deduce that if $d : L^2(\mu) \times L^2(\mu) \to \mathbb{R}^+$ is the distance given by

$$d(f, g) = \|f - g\|_2, \quad \text{for all} \quad f, g \in L^2(\mu),$$

$U_T$ is an *isometry*, that is

$$d(U_T(f), U_T(g)) = d(f, g) \qquad \text{for all} \quad f, g \in L^2(\mu);$$

(c) Verify that if a function is constant almost everywhere, then it is an eigenvector of $U_T$ with eigenvalue 1. Assume in addition that $(X, \mathscr{B}, \mu)$ is a *probability space*. Show that $T$ is ergodic if and only if the *only* eigenfunctions $f \in L^2(\mu)$ of $U_T$ corresponding to the eigenvalue 1 are constant functions.

[*Hint:* Both Part $(b)$ and $(c)$ of the exercise consist only of recalling and re-interpreting definitions.]

The operator $U_T$ is known as the *Koopman operator* associated to $T$. Many ergodic properties can be equivalently rephrased in terms of properties of the operator $U_T$, as ergodicity in Part $(c)$. The study of the properties of $U_T$ and its spectrum (for example, its eigenvalues) is known as *spectral theory of dynamical systems.*

**Extra: Lebesgue density points**

In the proof that the doubling map is ergodic, we used the following Theorem from measure theory, known as Lebesgue density Theorem.

Let $X = \mathbb{R}^n$ and $\lambda$ be $n-$dimensional Lebesgue measure. Let $A \subset \mathbb{R}^n$ be a Borel measurable set. Let $B(x, \epsilon)$ denote the ball of radius $\epsilon$ at $x$. The *density* of $A$ at $x$, denoted by $d_x(A)$, is by definition

$$d_x(A) = \lim_{\epsilon \to 0} \frac{\lambda(A \cap B(x,\epsilon))}{\lambda(B(x,\epsilon))}.$$

A point $x \in A$ is called a *Lebesgue density point* for $A$ if the density $d_x(A) = 1$. Thus, if $x$ is a density point, small intervals containing $x$ intersect $A$ on a large proportion of their measure, tending to 1 as the size of the interval tends to zero.

**Theorem 1.1.1** (Lebesgue density)**.** *Let $X = \mathbb{R}^n$ and $\lambda$ be $n-$dimensional Lebesgue measure. If $A \subset \mathbb{R}^n$ is a Borel measurable set with positive measure $\lambda(A) > 0$, almost every point $x \in A$ is a Lebesgue density point for $A$.*

This theorem implies that measurable sets can be well approximated by small intervals: on a small scale, measurable sets fill densely the space, so if $I$ is sufficiently small and intersects the set $A$, most of the points in $I$ are contained in $A$ (only a set of points whose measure is a proportion $\epsilon$ of the total measure is left out). Similarly, other small intervals will be missed almost completely by the set $A$, so that the set can be approximated well by a union of small intervals.

**Exercise 1.1.7.** Deduce from the Lebesgue density Theorem the fact that we used in the proof that the doubling map is ergodic, that is: if $\mu(X \backslash A) > 0$, given any $\epsilon > 0$, we can find $n \in \mathbb{N}$ and a dyadic interval $I$ of length $1/2^n$ such that

$$\lambda(I \backslash A) > (1-\epsilon)\lambda(I).$$

## 1.2 Ergodicity using Fourier Series

In the previous lecture §3.5 we defined *ergodicity* and showed from the definition that the doubling map is ergodic. In this lecture we will show how to use Fourier Series to show that certain measure preserving transformations defined on $\mathbb{R}/\mathbb{Z}$ or on the torus $\mathbb{T}^n = \mathbb{R}^n/\mathbb{Z}^n$ are ergodic. Proving ergodicity using Fourier Series turns out to be very simple and elegant. We first give a brief overview of the basics of Fourier Series.

Let $(X, \mathscr{B}, \mu)$ be a measure space. In §3.4 we defined integrals with respect to a measure. Recall that we also introduced the following notation for the spaces of *integrable* and *square-integrable functions*

$$L^1(X, \mathscr{B}, \mu) = L^1(\mu) = \{f : X \to \mathbb{R},\ f \text{ measurable}, \int |f| \mathrm{d}\mu < +\infty\}/\sim,$$
$$L^2(X, \mathscr{B}, \mu) = L^2(\mu) = \{f : X \to \mathbb{R},\ f \text{ measurable}, \int |f|^2 \mathrm{d}\mu < +\infty\}/\sim,$$

where $f \sim g$ if $f = g$ $\mu$-almost everywhere and the norms are respectively given by $\|f\|_1 := \int |f| \mathrm{d}\mu$ and $\|f\|_2 := (\int |f|^2 \mathrm{d}\mu)^{1/2}$.

**Fourier Series**

Let $X = \mathbb{R}/\mathbb{Z}$ with the Lebesgue measure $\lambda$. Consider a function $f : \mathbb{R}/\mathbb{Z} \to \mathbb{R}$ (more in general, one can consider a function $f : \mathbb{R} \to \mathbb{R}$ which is 1−periodic, that is $f(x+1) = f(x)$ for all $x \in \mathbb{R}$). We would like to represent $f$ as superposition of harmonics, decomposing it via the basic oscillating functions

$$\sin(2\pi n x), \quad \cos(2\pi n x), \qquad n = 0, 1, 2, \dots,$$

More precisely, we would like to represent $f$ as a linear combination

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos(2\pi n x) + \sum_{n=1}^{\infty} b_n \sin(2\pi n x).$$

Instead than using this notation, we prefer to use the complex notation, which is more compact. Recall the identity

$$e^{2\pi i x} = \cos(2\pi n x) + i \sin(2\pi n x).$$

Using this identity, one can can show that we can equivalently write

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos(2\pi n x) + \sum_{n=1}^{\infty} b_n \sin(2\pi n x) = \sum_{n=-\infty}^{+\infty} c_n e^{2\pi i n x}, \tag{1.7}$$

$$\text{where } c_n = \begin{cases} \frac{1}{2}(a_n - i b_n) & \text{if } n > 0, \\ a_0/2 & \text{if } n = 0, \\ \frac{1}{2}(a_{-n} + i b_{-n}) & \text{if } n < 0. \end{cases} \tag{1.8}$$

Thus, we look for a representation of $f$ of the form

$$\sum_{n=-\infty}^{+\infty} c_n e^{2\pi i n x}.$$

[Using this complex form, more in general, one can try to represent more in general functions $f : \mathbb{R}/\mathbb{Z} \to \mathbb{C}$. See also Exercise **??**, Part (b)].

**Exercise 1.2.1.** (a) Verify that if $a_n, b_n$ and $c_n$ are related by (**??**), then (**??**) holds.

Assume that

$$f = \sum_{n=-\infty}^{+\infty} c_n e^{2\pi i n x}.$$

(b) Show that $f$ is *real* if and only if $c_{-n} = \overline{c_n}$ for all $n \in \mathbb{Z}$ (where $\overline{z}$ denotes the complex conjugate of $z$);

(c) Show that $f$ is *even* (that is $f(-x) = f(x)$ for all $x$) if and only if $c_n = c_{-n}$ for all $n \in \mathbb{Z}$;

show that $f$ is *odd* (that is $f(-x) = -f(x)$ for all $x$) if and only if $c_n = -c_{-n}$ for all $n \in \mathbb{Z}$.

**Definition 1.2.1.** If $f \in L^1(\mathbb{R}/\mathbb{Z}, \mathscr{B}, \mu)$ we say that the *Fourier series* of $f$ is the expression

$$\sum_{n=-\infty}^{+\infty} c_n e^{2\pi i n x}, \qquad \text{where} \quad c_n = \int f(x) e^{-2\pi i n x} \mathrm{d}\mu,\ n \in \mathbb{Z}.$$

The $c_n$, $n \in \mathbb{Z}$, are called *Fourier coefficients* of $f$. Remark that $c_0 = \int f \mathrm{d}\mu$.

We denote by $S_N f$ the $N^{th}$ partial sum of the Fourier series of $f$, given by

$$S_N f(x) = \sum_{n=-N}^{+N} c_n e^{2\pi i n x}.$$

One needs the assumption $f \in L^1(\mu)$ to guarantee that the Fourier coefficients, and hence the Fourier series, is well-defined. The following property of the Fourier coefficients can be easily proved and is very helpful to use to prove ergodicity in certain examples (see Exercise **??**):

**Lemma 1.2.1 (Riemann-Lebesgue Lemma).** *If $f \in L^1(\mu)$, the Fourier coefficients $c_n$ in Definition* **??** *tend to zero in modulus, that is $|c_n| \to 0$ as $|n| \to \infty$.*

Unfortunately, as you should know well from the study of series, the fact that the coefficients tend to zero is not enough to guarantee that the Fourier series converges. It is natural to ask *when* the Fourier series *converges* for all points and when does it actually *represents* the function $f$ from which we started and in which sense. We list below some answers to these questions.

Once we have a representation of $f$ as a Fourier series (in one of the senses here below), one can use Fourier series as a tool which turns out to be very useful in applications. We will use them to show ergodicity, but more in general Fourier series can be used to solve differential equations and have a huge number of applications in applied mathematics.

The following can be proved:

(F1) If $f : \mathbb{R}/\mathbb{Z} \to \mathbb{R}$ is *differentiable* and the derivative is *continuous*, than its Fourier series converges at every point:

$$f(x) = \sum_{n=-\infty}^{+\infty} c_n e^{2\pi i n x}, \qquad \text{for all} \ \ x \in \mathbb{R}/\mathbb{Z}.$$

We say in this case that $S_N f$ *converges pointwise* to $f$.

[Remark that if $f$ is *only continuous* (but not necessarily differentiable), it is *not* necessarily *true* that the Fourier series of $f$ converges to $f$ pointwise. The proof of pointwise convergence can be found in many books in Real Analysis or Harmonic Analysis.]

(F2) If $f \in L^2(\mu)$,

$$\|S_N f - f\|_2 \to 0 \text{ as } N \to +\infty, \tag{1.9}$$

so that $S_N f$ approximate $f$ better and better in the $L^2-$norm. We say in this case that the Fourier series *converges to $f$ in $L^2$*.

[The proof of this statement is not hard and relies entirely on linear algebra. One can prove that $L^2(\mu)$ is a vector space and that the exponentials $e^{2\pi i n}$ form an orthogonal linear bases.]

(F3) If $f \in L^2(\lambda)$, one can actually show a much stronger statement (*Carlson's Theorem*):

$$f(x) = \sum_{n=-\infty}^{+\infty} c_n e^{2\pi i n x} \qquad \text{for a.e.} x \in \mathbb{R}/\mathbb{Z}.$$

[This result if very hard to prove, it was a hard open question and object of research for decades. The proof given by Carleson is very hard and many people have tried to understand it and simplify it.]

We will only use Fourier series for $L^2-$functions. A crucial property of Fourier series that we will use is *uniqueness*:

(F4) If $f \in L^1(\mu)$ and $c_n = 0$ for all $n \in \mathbb{Z}$, then $f = 0$. Recall that if $\mu$ is finite, $L^2(\mu) \subset L^1(\mu)$. As a consequence, if $\mu$ is a probability measure and $f \in L^2(\mu)$

$$\sum c_n e^{2\pi i n x} = \sum c'_n e^{2\pi i n x}$$

where the equality holds in the $L^2$ sense explained in $(F2)$, then $c_n = c'_n$ for all $n \in \mathbb{Z}$. Thus, the coefficients of the Fourier series of a function $f \in L^2(\mu)$ are *unique*.