# Appendix L

# Introduction to Fourier Methods

## L.1. The Circle Group and its Characters

We will denote by $\mathbf{S}$ the unit circle in the complex plane:

$$\begin{aligned}\mathbf{S} &= \{(x, y) \in \mathbf{R}^2 \mid x^2 + y^2 = 1\} \\ &= \{(\cos\theta, \sin\theta) \in \mathbf{R}^2 \mid 0 \leq \theta < 2\pi\} \\ &= \{z \in \mathbf{C} \mid |z| = 1\} \\ &= \{e^{i\theta} \mid 0 \leq \theta < 2\pi\}.\end{aligned}$$

We note that $\mathbf{S}$ is a group under multiplication, i.e., if $z_1$ and $z_2$ are in $\mathbf{S}$ then so is $z_1 z_2$. and if $z = x + yi \in \mathbf{S}$ then $\bar{z} = x - yi$ also is in $\mathbf{S}$ with $z\bar{z} = \bar{z}z = 1$, so $\bar{z} = z^{-1} = 1/z$. In particular then, given any function $f$ defined on $\mathbf{S}$, and any $z_0$ in $\mathbf{S}$, we can define another function $f_{z_0}$ on $\mathbf{S}$ (called $f$ translated by $z_0$), by $f_{z_0}(z) = f(zz_0)$. The set of piecewise continuous complex-valued functions on $\mathbf{S}$ will be denoted by $H(\mathbf{S})$.

The map $\kappa : t \mapsto e^{it}$ is a continuous group homomorphism of the additive group of real numbers, $\mathbf{R}$, onto $\mathbf{S}$, and so it induces a homeomorphism and group isomorphism $[t] \mapsto e^{it}$ of the compact quotient group $\mathbf{K} := \mathbf{R}/2\pi\mathbf{Z}$ with $\mathbf{S}$. Here $[t] \in \mathbf{K}$ of course denotes the coset of the real number $t$, i.e., the set of all real numbers $s$ that differ from $t$ by an integer multiple of $2\pi$, and we note that we can always choose a unique representative of $[t]$ in $[0, 2\pi)$. $\mathbf{K}$ is an additive model for $\mathbf{S}$, and this makes it more convenient for some purposes. A function on $\mathbf{K}$ is clearly the "same thing" as a $2\pi$-periodic function on $\mathbf{R}$. In particular, the complex vector space $H_{2\pi}(\mathbf{R})$ of piecewise continuous $2\pi$-periodic complex-valued functions on $\mathbf{R}$ can be identified with the space $H(\mathbf{K})$ of piecewise continuous complex-valued functions on $\mathbf{K}$. So, if to an element $f$ of $H(\mathbf{S})$ we associate the element $\tilde{f}$ in $H_{2\pi}(\mathbf{R})$ defined by $\tilde{f}(\theta) = f(e^{i\theta})$, then this clearly establishes a natural linear isomorphism between these two vector spaces of functions, and it is customary to use this isomorphism to "identify" these spaces. Note that if $z_0 = e^{i\theta_0}$ and $g = f_{z_0}$, then $\tilde{g}(\theta) = \tilde{f}(\theta + \theta_0)$. The correspondence $f \leftrightarrow \tilde{f}$ allows us to define $f$ being differentiable at $z = e^{i\theta}$ to mean that $\tilde{f}$ is differentiable at $\theta$, and we define $f'(z) = \tilde{f}'(\theta)$, so that $\widetilde{f'} = \tilde{f}'$. We will write $C^k(\mathbf{S})$ for the linear subspace of $H(\mathbf{S})$ consisting of functions with $k$ continuous derivatives, and as usual we will write $f^{(k)}$ for the $k$-th derivative of an element $f$ of $C^k(\mathbf{S})$. If $f \in H(\mathbf{S})$, we define the integral (or average) of $f$ over $\mathbf{S}$ by the formula:

$$\int_S f(z)\,dz := \frac{1}{2\pi}\int_0^{2\pi} \tilde{f}(\theta)\,d\theta.$$

and we can use this integral over $\mathbf{S}$ to define an inner product on $H(\mathbf{S})$ by:

$$\langle f, g\rangle := \int_S f(z)\overline{g(z)}\,dz.$$

and an associated "$L^2$ norm":

$$\|f\|_2 := \langle f, f\rangle^{\frac{1}{2}} = \left(\int_S |f(z)|^2\,dz\right)^{\frac{1}{2}}.$$

We will also need the so-called "sup" norm on $H(\mathbf{S})$, defined by:

$$\|f\|_\infty := \sup\{|f(z)| \mid z \in \mathbf{S}\},$$

and we note the obvious inequality $\|f\|_2 \leq \|f\|_\infty$.

▷ **Exercise L–1.** If $f \in H(\mathbf{S})$ is continuously differentiable show that $\int_S f'(z)\,dz = 0$.

▷ **Exercise L–2.** Use the periodicity of $f$ to show that, for any real number $\theta_0$,

$$\int_S f(z)\,dz = \frac{1}{2\pi}\int_{\theta_0}^{\theta_0+2\pi} \tilde{f}(\theta)\,d\theta.$$

Hints:

1) First prove the easy case, $\theta_0$ is an integral multiple of $2\pi$.
2) Using 1) show that, without loss of generality, we can assume that $0 \leq \theta < 2\pi$

3) Using $\int_0^{2\pi} = \int_0^{\theta_0} + \int_{\theta_0}^{2\pi}$ and $\int_{\theta_0}^{\theta_0+2\pi} = \int_{\theta_0}^{2\pi} + \int_{2\pi}^{\theta_0+2\pi}$, the rest should be easy.

▷ **Exercise L–3.** Use the preceding exercise to deduce that for any $z_0 \in \mathbf{S}$,

$$\int_S f_{z_0}(z)\,dz = \int_S f(z)\,dz.$$

(This property of the integral is called "translation invariance".)

A continuous function $f$ from $\mathbf{S}$ into the non-zero complex numbers is called a *character* (of $\mathbf{S}$) if it "preserves products", i.e., if $f(z_1 z_2) = f(z_1)f(z_2)$. We will denote the set of characters by $\hat{\mathbf{S}}$. One very obvious character (the "identity" character) is $e_0$, defined by $e_0(z) = 1$ for all $z$ in $\mathbf{S}$. It is easy to see that if $f$ is a character then $f^{-1}$, defined by $f^{-1}(z) = 1/f(z)$ is also a character, and that if $f_1$ and $f_2$ are characters then so is $f_1 f_2$, defined by $f_1 f_2(z) = f_1(z)f_2(z)$.

▷ **Exercise L–4.** Check that these definitions of identity, inverse, and product make $\hat{\mathbf{S}}$ into a commutative group (the group of characters of $\mathbf{S}$).

Of course $\int_S e_0(z)\,dz = \int_S 1\,dz = 1$, but we now note a very important fact: *If $f$ is any character other than the identity then* $\int_S f(z)\,dz = 0$. To see this, recall the definition of $f$ translated by $z_o$, $f_{z_0}(z) = f(z_o z)$. If $f$ is a character then $f(z_o z) = f(z_0)f(z)$, and translation in the domain becomes translation in the range, $f_{z_0}(z) = f(z_0)f(z)$. If we combine this with translation invariance of the integral, we find

$$\int_S f(z)\,dz = \int_S f_{z_0}(z)\,dz = \int_S f(z_0)f(z)\,dz = f(z_0)\int_S f(z)\,dz.$$

If $f$ is not the identity character, $e_0$, then there is some $z_o$ for which $f(z_o) \neq 1$, and so, as claimed, $\int_S f(z)\,dz = 0$.

Are there any characters other than the identity? Yes, namely

$$e_n(z) = z^n$$

for any $n$ in $\mathbf{Z}$, because

$$e_n(z_1 z_2) = (z_1 z_2)^n = z_1^n z_2^n = e_n(z_1)e_n(z_2).$$

Note that when $n = 0$ this includes the identity character $e_0$, and that in general $e_n e_m = e_{n+m}$, so that in the terminology of group theory, the set $\{e_n\}_{n\in Z}$ is a subgroup of $\hat{\mathbf{S}}$ isomorphic to the integers under addition. Notice, by the way, that in terms of the identification with $2\pi-$periodic functions on the line,

$$\widetilde{e_n}(\theta) = e^{in\theta}.$$

All the characters we have seen map $\mathbf{S}$ into itself. This is no accident! If $f$ is a character and $|f(z)| \neq 1$ for some $z$ in $\mathbf{S}$, then (since $1 = f(1) = f(zz^{-1}) = f(z)f(z^{-1}) = f(z)f(z)^{-1})$ we can assume that in fact $|f(z)| > 1$. But then the sequence $|f(z^n)| = |f(z)|^n$ would be unbounded, contradicting the fact that a continuous complex-valued function on a compact set is bounded.

Now that we know that for any character $f$, $|f(z)| = 1$, it follows that $f^{-1}(z) = \overline{f(z)}$, so if $g$ is also a character then their inner product is give by:

$$\langle f, g\rangle = \int_S fg^{-1}(z)\, dz.$$

We can immediately deduce from this that:

**L.1.1 Theorem.** *The set of characters is orthonormal.*

For if $f \neq g$, then the product $h = fg^{-1} \neq e_0$ is a character, and we have seen that $\int_S h(z)\, dz = 0$ for a character $h \neq e_0$. Thus $\langle f, g\rangle = \int_S fg^{-1}(z)\, dz = 0$. If $f = g$, then $fg^{-1} = e_0$ and $\int_S e_0(z)\, dz = 1$. In particular, $\{e_n\}_{n\in Z}$ is an orthonormal sequence in $H(\mathbf{S})$. In fact:

**L.1.2 Theorem.** *$\{e_n\}_{n\in Z}$ exhausts the set of characters.*

To see that this is so, first note that if $f$ is a character, then $f$ is determined by its restriction to any neighborhood of 1, That is, if we know $f(e^{it})$ for small $t$, then we also know it for arbitrary $T$, since $f(e^{iT})) = (f(e^{i\frac{T}{n}}))^n$. Secondly, for each real $t$ we can find a real $g(t)$ such that $f(e^{it}) = e^{ig(t)}$, and $g(t)$ is uniquely determined modulo $2\pi$. Since $f(e^{i0}) = e^{i0}$ and $f$ is continuous, there is a unique such function $g(t)$ defined near $t = 0$ with $|g(t)| < \pi$, and clearly this $g$ is continuous. If $t_1$ and $t_2$ are small, then since

$$\begin{aligned} f(e^{i(t_1+t_2)}) &= f(e^{it_1}e^{it_2}) = f(e^{it_1})f(e^{it_2}) = e^{ig(t_1)}e^{ig(t_2)} \\ &= e^{i(g(t_1)+g(t_2))} \end{aligned}$$

it follows that $g(t_1 + t_2) = g(t_1) + g(t_2)$, i.e., $g$ is additive, and since it is continuous it must be of the form $g(t) = ct$ for some real $c$, and clearly, in order for $f$ to be well-defined, $c$ must be an integer.

Since the characters $\{e_n\}_{n \in Z}$ form an orthonormal sequence, it is natural to try to expand functions $f$ on $\mathbf{S}$ into a linear combination of characters. So for $f \in H(\mathbf{S})$ we define a function $\hat{f} : \mathbf{Z} \to \mathbf{C}$ by:

$$\hat{f}(n) = \langle f, e_n \rangle = \int_S f(z)\overline{e_n(z)}\, dz = \frac{1}{2\pi} \int_0^{2\pi} \tilde{f}(\theta) e^{-in\theta}\, d\theta,$$

and call $\hat{f}$ the "Fourier Transform of $f$". We write $f \approx \sum_{n \in Z} \hat{f}(n) e_n$ to suggest that the Fourier Transform $\hat{f}$ of $f$ in some sense decomposes $f$ into an orthogonal linear combination of characters.

Suppose $f \in C^1(\mathbf{S})$. If we let $g(z) = f(z)\overline{e_n(z)} \in C^1(\mathbf{S})$ then $g \in C^1(\mathbf{S})$ so $\int_S g'(z)\, dz = 0$ and by the product rule, $g'(z) = f'(z)\overline{e_n(z)} + f(z)(-in\overline{e_n(z)})$. Substituting the latter form in the integral shows us that

$$\widehat{f'}(n) = \int_S f'(z)\overline{e_n(z)}\, dz = in \int_S f(z)\widehat{f'}(n)\, dz = in\hat{f}(n),$$

in other words the Fourier transform takes differentiation to multiplication. By induction, for $f \in C^k(\mathbf{S})$ it now follows that

$$\widehat{f^{(k)}}(n) = (in)^k \hat{f}(n).$$

More generally, suppose $P(X) = a_0 + a_1 X + \ldots + a_k X^k$ is a polynomial of degree $k$ with complex coefficients and we let $D$ denote the first derivative operator. We define a "constant coefficient $k$-th order differential operator", $L = P(D)$, a linear transformation $L : C^k(\mathbf{S}) \to C^0(\mathbf{S})$ by $Lf = P(D)f = a_0 f + a_1 f' + \ldots + a_k f^{(k)}$. Then from the results for $D^k$ above and linearity,

$$\widehat{Lf}(n) = P(in)\hat{f}(n).$$

What's going on here? *Insofar as we can think of* $\{e_n\}$ *as a basis for* $C^\infty(\mathbf{S})$*, all constant coefficient differential operators* $P(D)$ *are simultaneously diagonalized in this basis!* This is part of what makes the Fourier Transform such a powerful tool.

The fact that our characters are eigenvectors of all constant coefficient differential operators is already sufficient reason for our attention. But the class of operators that are diagonalized by characters is even broader. We call a linear subspace $V$ of $H(\mathbf{S})$ *translation invariant*, if $f \in V$ and $z$ in $\mathbf{S}$ implies that $f_z$ is also in $V$. If $V$ is translation invariant, then we say that a linear map $T : V \to V$ *commutes with translations* if for all $f$ in $V$ and $z$ in $\mathbf{S}$, $T(f_z) = (Tf)_z$. As examples, we may check that since $Df_z = (Df)_z$ for any $f \in C^1(\mathbf{S})$ and $z \in \mathbf{S}$, induction and linearity tell us that $C^\infty(\mathbf{S})$ is translation invariant and that any constant coefficient differential operator $P(D)$ commutes with translations.

In fact:

**L.1.3 Theorem.** *if $V$ is any translation invariant linear subspace of $H(\mathbf{S})$ that contains all the characters, and if $T : V \to V$ is any linear map that commutes with translations, then each character $f$ is an eigenvector of $T$ with eigenvalue $(Tf)(1)$.*

**Proof.** The character condition, $f(z)f(\zeta) = f(z\zeta) = f_z(\zeta)$ says that $f_z = f(z)f$ and hence, by linearity of $T$, $(Tf)_z = T(f_z) = f(z)(Tf)$, and evaluating both sides at 1 gives

$$(Tf)(z) = (Tf)_z(1) = f(z)(Tf)(1),$$

or equivalently, $Tf = (Tf)(1)f$. Here is another proof of this important fact: because characters are nowhere equal to zero, we can reformulate the condition that a character $f$ is an eigenfunction of a linear operator $T$, (with eigenvalue $\lambda_f$) as the statement that the quotient function $(Tf/f)$, when evaluated at any $z$, is a constant $\lambda_f$, i.e., is independent of $z$. Also, instead of evaluating the quotent $Tf/f$ at the argument $z$, we can translate it by $z$ and evaluate at the identity, i.e., by definition of translation, $(Tf/f)(z) = (Tf/f)_z(1) = (Tf)_z(1)/f_z(1)$. If we now assume that the operator $T$ is translation invariant, then $(Tf)_z(1)/f_z(1) = (T(f_z))(1)/f_z(1)$. But, because $f$ is a character, translation of $f$ by $z$ is equivalent to scalar multiplication by $f(z)$; i.e., $f_z = f(z)f$, so that by linearity of $T$:

$$\begin{aligned}(Tf/f)(z) &= (T(f_z))(1)/f_z(1)\\ &= f(z)(Tf)(1)/f(z)f(1) = (Tf)(1)/f(1) = (Tf)(1).\end{aligned}$$

proving again that any character $f$ is an eigenfunction of any translation invariant linear operator $T$, corresponding to the eigenvalue $\lambda_f = (Tf)(1)$.

Of course $\{e_n\}$ is *not* a basis in the algebraic sense and for a general element $f$ of $H(\mathbf{S})$, the infinite sum $\sum_{n\in Z} \hat{f}(n)e_n$ is not well-defined. However for any positive integer $N$, $f_{_N} = \sum_{|n|\leq N} \hat{f}(n)e_n$ is a well-defined, infinitely differentiable function in $H(\mathbf{S})$ and we can investigate in what sense and to what extent $f_{_N}$ approximates $f$ as $N$ tends to infinity. There are numerous frameworks in which such approximation results hold—and a full discussion of this subject could and does fill respectable volumes. Below we shall only take up a few of the most basic results in this direction. First though let's review some elementary facts and terminology concerning an arbitrary orthonormal sequence $\{v_n\}$ in a vector space $V$ with an inner product $\langle\ ,\ \rangle$ and associated norm $\|v\| = \langle v, v\rangle^{\frac{1}{2}}$.

Let $V_n$ denote the subspace of $V$ spanned by $v_1, \ldots, v_n$, and let $V_\infty$ denote the union of the $V_n$ (i.e., the linear subspace consisting of all finite linear combinations of the $v_i$). If an element $f$ in $V$ belongs to $V_n$, then of course $f = \sum_{i=1}^n \langle f, v_i\rangle v_i$. But what does the sum

$$g = \sum_{i=1}^n \langle f, v_i\rangle v_i$$

represent for an arbitrary element $f$ of $V$, not necessarily in $V_n$? It is easy to see that it is the "orthogonal projection of $f$ on $V_n$", i.e., the unique element of $V_n$ such that $f - g$ is orthogonal to everything in $V_n$. Indeed, $f = g + (f - g)$ is an identity, and for any $v_i$,

$$\langle (f - g), v_i\rangle = \langle f, v_i\rangle - \langle g, v_i\rangle = \langle f, v_i\rangle - \langle f, v_i\rangle = 0,$$

so by linearity, $f - g$ is orthogonal to everything in the span of the $v_i$, i.e., $V_n$. If $g'$ were another such element of $V_n$, then $g - g' = (f - g') - (f - g)$ is the difference of vectors that are both orthogonal to everything in $V_n$. Therefore $g - g'$ itself is orthogonal to everything in $V_n$, and since $g - g'$ is in $V_n$, $g - g' = 0$.

If $h$ is any element of $V_n$ then $f - h = (f - g) + (g - h)$, and since $f - g$ and $g - h$ are orthogonal, it follows from the Pythagorean identity that

$$\|f - h\|^2 = \|f - g\|^2 + \|g - h\|^2 \geq \|f - g\|^2,$$

with equality only if $h = g$. It is now immediate that $g$ can be characterized as *the unique element of* $V_n$ *that is closest to* $f$. Taking $h = 0$ in the above and using $\|g\|^2 = \sum_{i=1}^n |\langle f, v_i\rangle|^2$, we get

$\sum_{i=1}^{n} |\langle f, v_i\rangle|^2 \leq \|f\|^2$. Since this holds for all $n$, we have the so-called "Bessel Inequality"

$$\sum_{i=1}^{\infty} |\langle f, v_i\rangle|^2 \leq \|f\|^2 < +\infty$$

(so in particular $\lim_{n\to\infty} |\langle f, v_n\rangle| = 0$). If Bessel's inequality is an equality, it is referred to as the "Parseval Identity".

**L.1.4 Theorem (and Definition).** *The following three conditions are equivalent, and if any one and hence all of them hold then we call the orthonormal sequence $\{v_n\}$ a complete orthonormal sequence for $V$.*

*1) The Parseval Identity:*

$$\sum_{i=1}^{\infty} |\langle f, v_i\rangle|^2 = \|f\|^2$$

*holds for all $f$ in $V$.*

*2) For each $f$ in $V$, the sequence of partial sums, $f_n = \sum_{i=1}^{n} \langle f, v_i\rangle v_i$ converges to $f$, i.e.,*

$$\lim_{n\to\infty} \|f - f_n\| = 0.$$

*3) $V_\infty$ is dense in $V$, i.e., given $f$ in $V$ and $\epsilon > 0$, there exist scalars $\alpha_1, \dots, \alpha_n$ such that*

$$\left\| f - \sum_{i=1}^{n} \alpha_i v_i \right\| < \epsilon.$$

**Proof.** The equivalence of 1) and 2) follows from the Pythagorean identity

$$\left\| f - \sum_{i=1}^{n} \langle f, v_i\rangle v_i \right\|^2 = \|f\|^2 - \sum_{i=1}^{n} |\langle f, v_i\rangle|^2,$$

and the equivalence of 2) and 3) follows from the fact that $g = \sum_{i=1}^{n} \langle f, v_i\rangle v_i$ is the linear combination of $v_1, \dots, v_n$ closest to $f$, so that

$$\left\| f - \sum_{i=1}^{n} \langle f, v_i\rangle v_i \right\| \leq \left\| f - \sum_{i=1}^{n} \alpha_i v_i \right\|. \quad \blacksquare$$

Let $H_N(\mathbf{S})$ denote the $2n+1$ dimensional subspace of $H(\mathbf{S})$ spanned by the characters $\{e_n \mid -N \leq n \leq N\}$, and let $H_\infty(\mathbf{S})$ denote the union of all the $H_N(\mathbf{S})$, i.e., the space spanned by all the characters. Elements of $H_\infty(\mathbf{S})$ are sometimes called finite Fourier series, or trigonometric polynomials. Note that for $f \in H(\mathbf{S})$, $f_N$ is the orthogonal projection of $f$ on $H_N(\mathbf{S})$ and so is the unique best approximation to $f$ (in the $L^2$ sense described above) among all finite linear combinations of $e_{-N}, \ldots, e_N$. And in order to prove that the characters $e_n$ are a complete orthonormal sequence for $H(\mathbf{S})$, so that these approximations to $f$ in fact always converge to $f$ in the $L^2$ sense, it will suffice to prove that $H_\infty(\mathbf{S})$ is $L^2$-dense in $H(\mathbf{S})$. This fact is important historically, and remains so in both theory and applications. We now sketch a proof as a series of exercises.

$\triangleright$ **Exercise L–5.** A complex-valued function $f$ on $\mathbf{C}$ is said to be a "polynomial in $z$ and $\bar{z}$" if there is complex polynomial in two variables $P(X, Y)$ such that $f(z) = P(z, \bar{z})$ for all $z \in \mathbf{C}$. Show that $H_\infty(\mathbf{S})$ is just the space of functions on $\mathbf{S}$ that are restrictions to $\mathbf{S}$ of such functions $f$.

**L.1.5 The Stone-Weierstrass Aproximation Theorem (Special Case).** *Any continuous complex-valued function on a closed, bounded subset of $\mathbf{C}$ is the uniform limit of a sequence of polynomials in $z$ and $\bar{z}$.*

**L.1.6 Corollary.** *$H_\infty$ is dense in $C^0(\mathbf{S})$ with respect to the sup norm and (since $\| \; \|_2 \leq \| \; \|_\infty$) also with respect to the $L^2$ norm, i.e., given any $f \in C^0(\mathbf{S})$ and $\epsilon > 0$, there is a $g$ in $H_\infty(\mathbf{S})$ such that $\|f - g\|_2 \leq \|f - g\|_\infty < \epsilon$.*

**L.1.7 Theorem.** *The set $\{e_n\}_{n \in Z}$ is a complete orthonormal sequence for $H(\mathbf{S})$, so that for every $f \in H(\mathbf{S})$, $f_N$ converges to $f$ in the $L^2$ sense, i.e., $\lim_{N \to \infty} \|f - f_N\|_2 = 0$.*

**L.1.8 Corollary (Parseval Theorem).** *The map $f \mapsto \hat{f}$ extends uniquely to an isometry of $L^2(\mathbf{S})$ with $L^2(\mathbf{Z})$.*

$\triangleright$ **Exercise L–6.** Prove the preceding theorem by showing that $H_\infty(\mathbf{S})$ is dense in $H(\mathbf{S})$ with respect to the $L^2$ norm. Since we already know that $H_\infty(\mathbf{S})$ is $L^2$-dense in $C^0(\mathbf{S})$, it will suffice to show that $C^0(\mathbf{S})$

is $L^2$-dense in $H(\mathbf{S})$, i.e., that if $f \in H(\mathbf{S})$ then one can find a continuous $g$ with $\|f - g\|_2$ arbitrarily small. Now recall that elements of $H(\mathbf{S})$ are by definition piecewise continuous, and first handle the case of an $f$ that is continuous on a single subinterval of $[0, 2\pi)$ and zero on the rest.

▷ **Exercise L–7.** Weak Riemann-Lebesgue Lemma: Show that,

$$\lim_{n\to\infty} \hat{f}(n) = 0$$

holds for any $f$ in $H(\mathbf{S})$,. (The real Riemann-Lebesgue Lemma says that this is true for $f$ in $L^1(\mathbf{S})$). Deduce that if $f \in C^k(\mathbf{S})$ then $\hat{f}(n) = o(n^{-k})$, i.e., $\hat{f}(n)n^k$ tends to zero. Thus, the more differentiable $f$, the more rapidly its Fourier coefficients tend to zero.

Note that the fact that $f_N$ converges to $f$ in the $L^2$ sense doesn't say anything about convergence in the uniform (i.e., $\| \ \|_\infty$) sense, or even about pointwise convergence. Recalling that the uniform limit of continuous functions is continuous, we see that the Fourier series for $f$ cannot converge uniformly to $f$ unless $f$ is at least in $C^0(\mathbf{S})$. But this is not sufficient, and we complete this discussion by quoting two results indicating that pointwise and uniform convergence of the Fourier series for $f$ is related to the differentiability of $f$.

**L.1.9 Theorem.** *If $f \in H(\mathbf{S})$, then $f_N(z)$ converges to $f(z)$ at all points $z$ of $\mathbf{S}$ where $f$ is differentiable. If $f \in C^1(\mathbf{S})$ then in fact $f_N$ converges uniformly to $f$.*

## L.2. Finite Models for the Circle Group

While the Fourier transform is a powerful analytical tool, it seems at first glance to be a computationally intractable one; there are infinitely many Fourier coefficients $\hat{f}(n)$ needed to specify $f$, and evaluating each of them involves calculating an integral. On the other hand, we know that, for smooth functions, $\hat{f}(n)$ decays very rapidly with $n$, so we can expect that for reasonable applications only a small error will result if we "filter out all the high-frequency modes" from $f$, i.e., if we replace $f$ by $f_N$ where $N$ is sufficiently large. While this is in fact so, it still leaves open the question of finding fast algorithms for computing the finite number of Fourier coefficients necessary for a particular applications. The area of "signal processing" is a good place to get some feel for the problems and the way they are handled. In general, a time varying signal is represented by one or more real or complex valued functions of time. For example a sound or "audio" signal is often represented by a single function (the amplitude of the deviation of pressure from ambient pressure) while a video signal may be represented by three time-dependent intensities (representing the red, green, and blue values of the pixel currently being displayed). Let's consider the simpler case of sound. Suppose we want to store an audio signal with fidelity "as good as the ear can hear". Since the human ear is insensitive to frequencies above roughly 20 KHz, it follows from an important mathematical result (Nyquist's Theorem) that if we sample the amplitude 40 thousand times per second, we will have enough information to reconstruct the signal with "audibly perfect" fidelity. The signal is broken up into "time slices", and we choose units of time so that each time slice has length $2\pi$. Each of these is sampled at a certain sampling frequency (corresponding to at least 40 thousand samples per second) and let's suppose that there are $N$ samples per time slice. Now, we could choose to simply store the N amplitudes samples per time slice but, to process and reconstruct the signal electronically, it is preferable to have available the Fourier transform of the signal during each time slice.

What we shall consider next is how to define a "discrete Fourier transform" or DFT for the sampled signal. By wrapping the time slice $[t_0, t_0 + 2\pi)$ around the circle $\mathbf{S}$, with the map $t \mapsto e^{it}$, the set of sampling instants becomes an $N$-element subset $\mathbf{S}_N$ of $\mathbf{S}$, and the sampled signal becomes a function $f : \mathbf{S}_N \to \mathbf{C}$. (Let's denote the complete signal during the time slice by $F : \mathbf{S} \to \mathbf{C}$, so $f = F|\mathbf{S}_N$.)

How should we choose the subset $\mathbf{S}_N$? At first glance this seems to be a fairly arbitrary choice. Of course we should take the points fairly evenly spaced, but otherwise it wouldn't seem to matter much what the precise sample points were. But, recall that we used heavily the fact that the circle $\mathbf{S}$ was a group in defining the Fourier transform, and so if we are going to define a Fourier transform for a function on $\mathbf{S}_N$ it is natural to try to choose $\mathbf{S}_N$ to be a subgroup of the circle group. In fact this requirement uniquely specifies $\mathbf{S}_N$ to be the group $\mathbf{S}_N = \{1 = \omega^0, \omega, \dots, \omega^{N-1}\}$ of $N$th roots of unity, where $\omega = \omega_N$ denotes the primitive $N$th root of unity, $\omega = e^{2\pi i/N}$.

▷ **Exercise L–8.** Verify that a finite subgroup of $\mathbf{S}$ must be one of the $\mathbf{S}_N$.

We denote by $H(\mathbf{S}_N)$ the $N$-dimensional complex vector space of complex valued functions on $\mathbf{S}_N$. If $f \in H(\mathbf{S}_N)$, then for any $z_0$ in $\mathbf{S}_N$ we can as before define the "translate", $f_{z_0} \in H(\mathbf{S}_N)$ by $f_{z_0}(z) = f(zz_0)$, and we define the "integral" or average value of $f$ over $\mathbf{S}_N$ by the formula:

$$\int_{S_N} f(z)\,dz := \frac{1}{N}\sum_{z\in S_N} f(z) = \frac{1}{N}\sum_{i=0}^{N-1} f(\omega^i).$$

And again as before, we use this to define an inner product and norm for $H(\mathbf{S}_N)$ by $\langle f, g\rangle := \int_{S_N} f(z)\overline{g(z)}\,dz$ and $\|f\| = \langle f, f\rangle^{1/2}$.

▷ **Exercise L–9.** Prove the translation invariance property of the integral:

$$\int_{S_N} f_{z_0}(z)\,dz = \int_{S_N} f(z)\,dz.$$

Of course we define a character for $\mathbf{S}_N$ to be a product preserving map of $\mathbf{S}_N$ into the non-zero complex numbers, and as with $\mathbf{S}$, we see that the set $\hat{\mathbf{S}}_N$ of characters of $\mathbf{S}_N$ is a group under multiplication, the character group of $\mathbf{S}_N$.

▷ **Exercise L–10.** Use the same argument as before to deduce that the elements of $\hat{\mathbf{S}}_N$ are orthonormal.

Now any character for the circle of course restricts to a character for $\mathbf{S}_N$, and in particular the characters $e_k(z) = z^k$ for $\mathbf{S}$ define characters on $\mathbf{S}_N$, which we still denote by $e_k$.

▷ **Exercise L–11.** Show that $e_j = e_k$ if and only if $j \equiv k \bmod N$, and deduce that $\hat{\mathbf{S}}_N = \{e_0, e_1, \ldots, e_{N-1}\}$.

Thus, $\{e_0, e_1, \ldots, e_{N-1}\}$ is an orthonormal basis for $\mathbf{S}_N$. Now, since $e_j e_k = e_{j+k}$, we can regard the indices of the $\{e_k\}$ as belonging to the group $\mathbf{Z}_N$ of integers modulo $N$, giving a natural isomorphism of $\hat{\mathbf{S}}_N$ with $\mathbf{Z}_N$.

**Definition of the DFT.** We define the *discrete Fourier transform*, or DFT, of an element $f$ of $H(\mathbf{S}_N)$ to be the function $\hat{f} : \mathbf{Z}_N \to \mathbf{C}$ defined by

$$\hat{f}(k) = \langle f, e_k \rangle = \frac{1}{N} \sum_{j=0}^{N-1} f_j e^{-2\pi i k j/N},$$

so that $f = \sum_{j=0}^{N-1} \hat{f}(k) e_k$.