Assignments for the course "Fourier Analysis" MATH 425
Lecturer: Hans G. Feichtinger (University Vienna)

Some theoretical background (course notes in English) can be found at the lecturers homepage

`http://www.univie.ac.at/NuHAG/FEICOURS/ss10/ss10.htm`

A lot of literature (details will be given later on) can be retrieved from the workgroup NuHAG homepage at **www.nuhag.eu** , specifically from the BIBTEX collection:

`http://www.univie.ac.at/nuhag-php/bibtex/`

Literature for the course: [**?**, **?**] or Grafakos [**?**].

1 Warm-up

1. How do we multiply natural numbers? What about teaching a young child how to compute $123 \cdot 45$ and why is it (?) clear that it equals $45 \cdot 123$. What is the result? And would that result change if we would work in the hexa-system (with numbers not from $\{0, 1, \dots 9\}$ but only from $\{0, 1, \dots 6\}$?

2. What do you know about the random experiment of throwing a dice: Equal probability (hence 1/6 for each of the possible outcomes: $1, 2, \dots 6$, assuming that the dice is fair).

   What can you say about the possible outcomes of if you ask for the sum of two dices? What is the result and provide possible explanation to it.

3. Hoping that some of you are familiar with (whatever) mathematical software, such as MATLAB, MATHEMATICA, OCTAVE, SCILAB or anything else. Could you run a numerical experiment confirming your expectations (with 2 or more dices summed up); I will certainly demonstrate occasionally things using MATLAB.

4. Think, if you have any idea how to calculate the probability of reaching 100 points in a board game within 30 repetitions of throwing the dice. Is it more or less than 50%?

Material for the course "Fourier Analysis" MATH 425
Lecturer: Hans G. Feichtinger (University Vienna)

2 Background from Linear Algebra

We recall the notation of (finite dimensional) vector spaces which (by definition) have a *basis.* A basis is a *generating set* (every vector can be represented) which is *linear independent*, i.e. every vector has *at most* one representation.

It is well known that a set of column vectors a_k, $1 \leq k \leq d$ in $\mathbb{R}^d$ forms a basis if - put as columns into a $d \times d$ matrix $\mathbf{A}$, this is an invertible matrix. Since

$$\mathbf{A} * x = \sum_{k+1}^{d} x_k a_k,$$

matrix multiplication corresponds just to "taking linear combinations of columns vectors" [1] the uniquely determined coefficients that one needs in order to represent a vector b must be those which allow $\mathbf{A} * x = b$ or obviously $x = \mathbf{A}^{-1} * b$.

A matrix is called unitary (or orthogonal, if it is a real matrix) if the inverse matrix can be obtained "at no cost", i.e. by just taking the transpose conjugate: $\mathbf{U}^{-1} = \mathbf{U}'$ (where MATLAB used the dash to denote the transpose conjugate).

Then every vector is of the form

$$x = \mathbf{U} * \mathbf{U}' * x = \sum_k \langle x, u_k \rangle u_k,$$

recalling that the scalar product of two vectors u and x can be obtained in MATLAB using matrix multiplication $u' * x$.

Geometrically the unitary matrices are so nice because they are characterized by the mutual orthogonality of the vectors plus a normalization condition, which is evidently equivalent to the assumption that $\mathbf{U}' * \mathbf{U} == eye(n)$ (using also the fact that in a group an element cannot have different left and right inverse elements).

Other non-trivial examples of finite-dimensional vector spaces are e.g. the polynomials of a fixed degree, say cubic polynomials. We write $\mathscr{P}_n(\mathbb{C})$ for the polynomials of *degree* n. The are of *order* $n+1$ because the standard monomial basis consisting of $\{1, x, x^2, \ldots, x^n\}$ has $n+1$ elements. Of course the order of this basis is not relevant, and MATLAB chooses them in the reverse order $\{x^n, x^{n-1}, \ldots, x, 1\}$, somehow like 123 is interpreted as $1 \cdot 10^2 + 2 \cdot 10^1 + 3 \cdot 10^0$, i.e. the value of the polynomial $p(t) = 1 \cdot t^2 + 2 \cdot^1 + 3 \cdot t^0$. In MATLAB we could verify by using the command

```
polyval([1,2,3],10);
```

[1] matrix multiplication from the right on row vectors provides a way to do linear combinations of row vectors of a row matrix!

In general the polyval-routine allows to evaluate the polynomal with coefficients a at a sequence of points $t = [t_1, t_2, \ldots, t_k]$. If $k = n + 1$, i.e. if the order of the polynomial equals the number of points we have the same number of points as we have coefficients. Of course these points can be also in the complex plane, i.e. the evaluations can take place anywhere (pairwise distinct points) in the complex plane.

From this it follows (some [linear] algebra, e.g. involving the Gauss elimination method) that *every polynomial* that the mapping a to $(p(t_k))_{k=1}^{n+1}$ is a *linear* and invertible matrix, called the Vandermonde matrix.

For example we have for the sequence $t = [0, 1, 2, 3]$

```
>> vander(0:3)
ans =
     0     0     0     1
     1     1     1     1
     8     4     2     1
    27     9     3     1
```

We will see soon that the DFT (Discrete Fourier transform) is - up to a different choice of the monomial basis for the polynomials of order N - just the Vandermonde matrix corresponding to the set of unit roots of order N.

Recall that a number $z \in \mathbb{C}$ is called a *unit root of order* N if $z^N = 1$. Clearly only complex numbers from the unit cirlce $U = \{z \,|, |z| = 1\}$ need to be discussed. It is easy to see that any power of the *primitive root of order* N, namely $\omega := \omega_N = e^{2\pi i/N}$, e.g. $\{\omega^0, \omega, \ldots \omega^{N-1}\}$, is a possible choice, but in fact these are all the possible points.

```
bas = 2 * pi *  (0: 1/N : 1) ;  roots = exp(i*bas);  plot(roots);
hold on; plot(roots,'r*'); hold off; axis square; figure(gcf);
```

In this particular case (and more or less only in this also geometrically perfect situation) one can find that the Vandermonde (now Fourier) matrix is unitary.

This can be verified directly by observing that the scalar products reduce to the calculation of a sum of the form

$$\sum_{k=0}^{N-1} q^k = \frac{1 - q^N}{1 - q} = 0$$

for any unit root of order N except for $q = 1$.

Another important property of the system of columns of the FFT-matrix is the following: if one takes a pointwise product of two such columns one obtains another column of this matrix. In other words, if we call the columns of F (the matrix is symmetric in the sense of being equal to its transpose, so rows are the same as columns!) pure frequencies, labelled with $\chi_0, \ldots \chi_{N-1}$. Then one has

$$\chi_k \cdot \chi_l = \chi_r \quad \text{with} \quad r = k + l(modr),$$

in particular it is clear that $\chi_0 == 1$ is the multiplicative identity and hence $\chi_{N-k} = conj(\chi_k)$ is the inverse element to χ_k. We will call these functions χ_k the *pure frequencies*. In abstract harmonic analysis they are called the characters of the group $\mathbf{Z}_N$ (the cyclic group of unit roots of order N, which is of course isomorphic to $\mathbb{Z}/N\mathbb{Z}$.

3 Classical Fourier Series, ONBs in Hilbert spaces

The classical theory of Fourier series has been introduced by J.B. Fourier around 1809, in order to study the heat equation (allowing him a separtion of variables). He claimed that *every* periodic function has a Fourier series expansions, without really specifying in which sense (from a modern point of view).

Recall that a complex-valued function f on $\mathbb{R}$ is $\mathbb{Z}$-periodic if we have $f(x+k) = f(k)$ for all $k \in \mathbb{Z}$.

Some MATLAB experiments:

```
>> n = 360; F = fft(eye(360));
>> F = fft(eye(360));
>> plot(1:n,real(F(13,:)),1:n, imag(F(13,:)));

or also:
plot(fft(eye(17)));  axis square; axis off; figure(gcf);
```

This shows how FFT works (applies the FFT to the collection of unit vectors, producing such the matrix F which satisfies

$$y = fft(x) = F * x$$

for every column vector of length N, and the plotting routine which connects all the points within one column. Alternatively, you might plot as follows:

```
>> plot(fft(eye(32)),'b*');
>> axis square; axis off; figure(gcf);
```

Convolution is a very natural and important kind of general multiplication. Given two polynomals $p(x)$ and $q(x)$ is is clear that the product of these two polynomials $r(x) := p(x) \cdot q(x)$ is another polynomial, and the degrees add (degree of r = degree of p+ degree of q).

The coefficients of $r(x)$ are obtaine by the so-called **Cauchy product** formula, which is more or less the way how one multiplies long numbers: Let us call the set of coefficients for p, q, r by a, b, c respectively. Then we have

$$c_k = \sum_{j+l=k} a_j b_l,$$

because this is what comes out after sorting the entries after pairwise multiplication of terms according to which monomial the product belongs.

We will see later that this principle is also relevant for probability.

Let us also recall the concept of a Hilbert space.

Definition 1. A vector space $\mathscr{H}$ is called a Hilbert space, if it is endowed with a scalar product, such that the corresponding norm

$$\|f\|_{\mathscr{H}} := \sqrt{\langle f, f \rangle}$$

turns $\mathscr{H}$ into a Banach space, i.e. a normed space (this is always true) which is complete, in the sense that every Cauchy sequence is convergent.

Our prototype will be (first of all) the space of periodic, square integrable functions on the unit interval (identified with the torus), i.e. $\boldsymbol{L}^2(\mathbb{U})$, with the scalar product

$$\langle f, g \rangle := \int_0^1 f(x)\overline{g(x)}dx.$$

Of course we get the completeness of this space only due to the use of the *Lebesgue integral*. Note that the norm is to be understood in the sense that $\|f\|_2 = 0$ then we have $f(x) = 0$ a.e. (almost everywhere, i.e. up to a set of Lebesgue measure zero)[2]

Since we have no an infinite dimensional vector space (for any N there are more than N linear independent functions in this space, e.g. pairwise disjointly supported continuous functions with small support) we have to find an infinite set which could be an orthormal basis for $\boldsymbol{L}^2(\mathbb{U})$, similar to the discrete case. Let us try to consider the system $\chi_k(t) := e^{2piikt}$, for $k \in \mathbb{Z}, t \in \mathbb{R}$ or just $t \in [0, 1]$.

It is a good exercise to check that it is orthonormal, either by using the real and imaginary parts (cos and sin functions) or better the complex relation, so one has to verify

$$\int_0^1 e^{2piikt}e^{-2piilt}dt = \int_0^1 e^{2pii(k-l)t} = \delta l, k$$

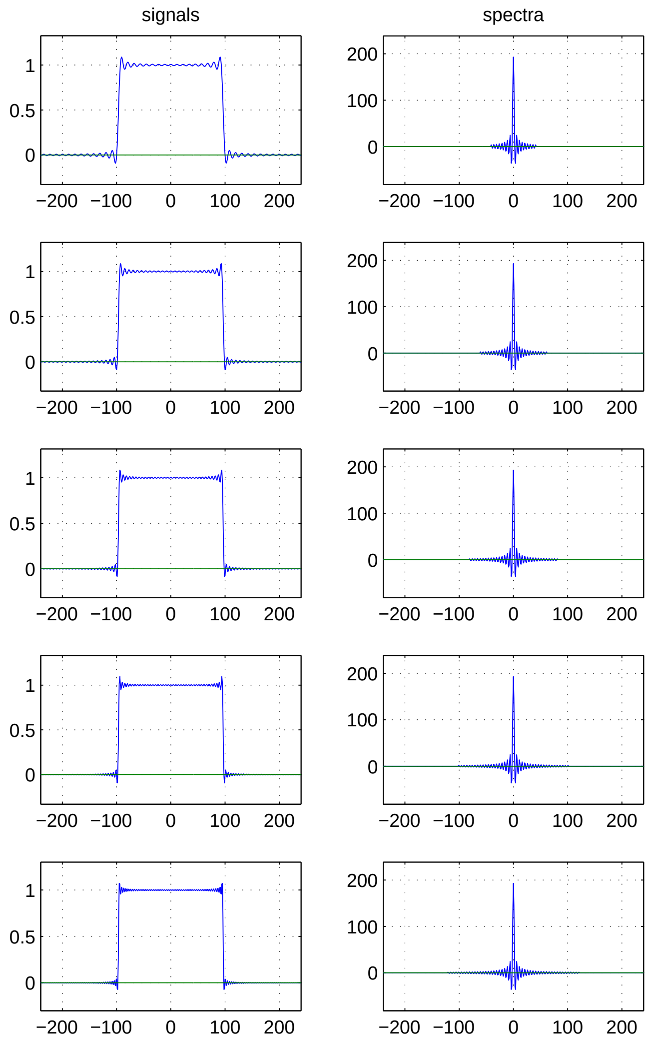
(Kronecker's delta: $\delta l, k = 1 \qquad k = l$ and 0 otherwise.

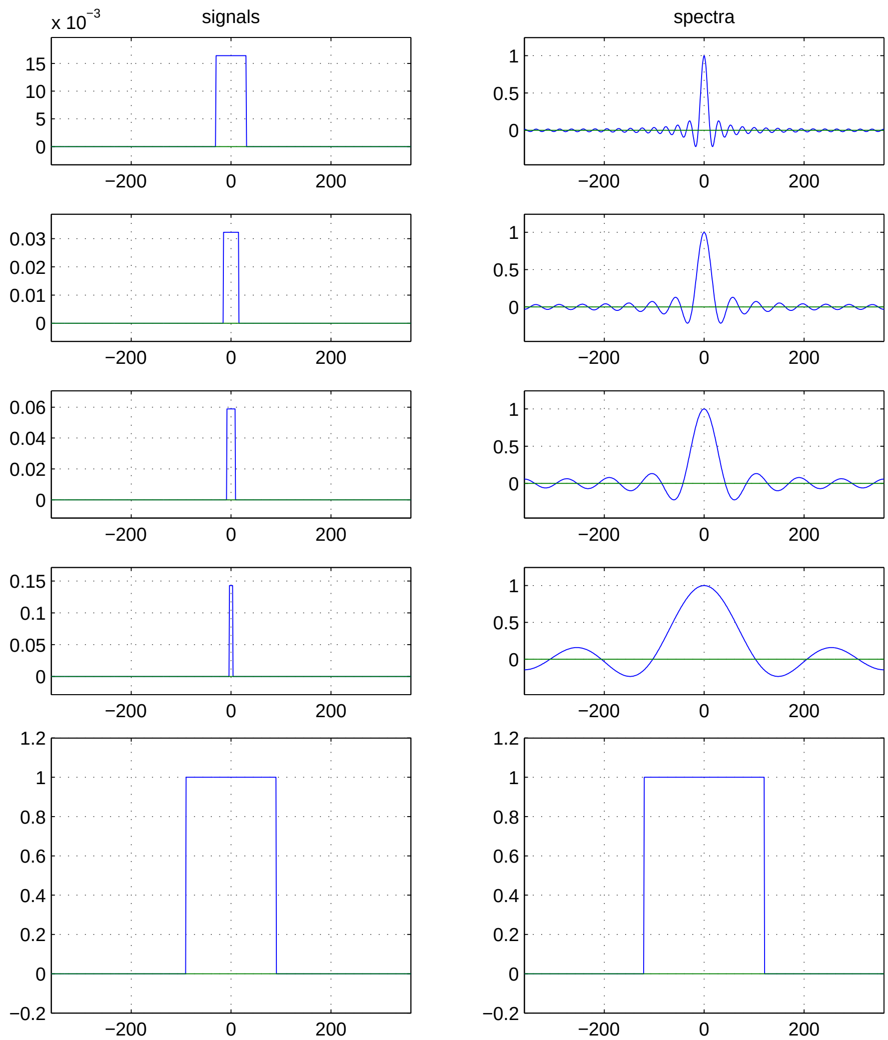
#BLOCK[136]

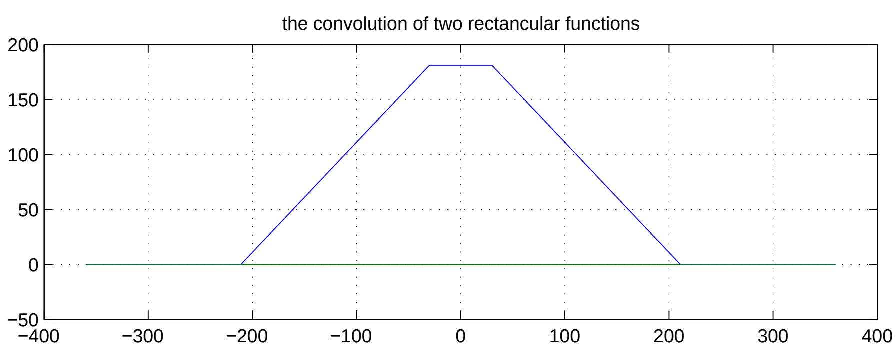
4 Some MATLAB Plots

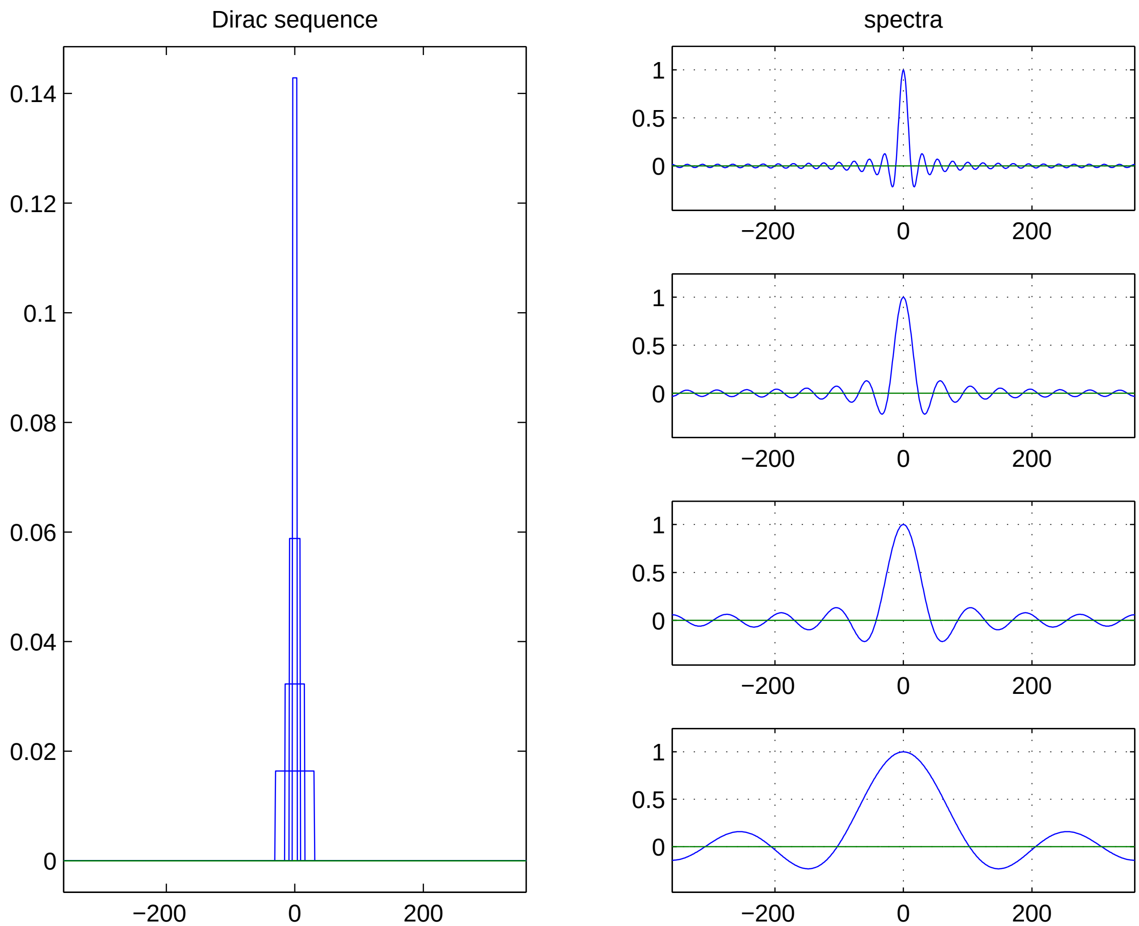
First demonstrating the Fourier partial sums (symmetrically around zero) for the rectangular function (or in fact sequence)

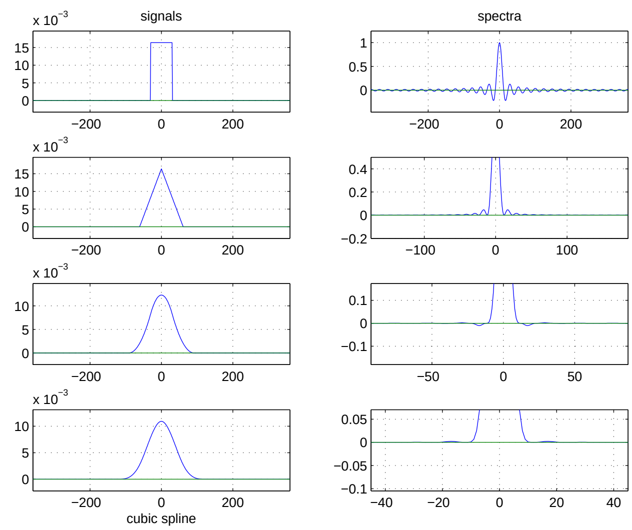
[2]Thus strictly speaking we are only dealing with equivalence classes of measurable functions up to functions which are zero almost everywhere, but for pratical cases this fine difference can be widely ignored, and *in particular* the problem is not in existence if one restricts the attention only to *continuous* periodic function, because if a continuous function has norm zero then it is clear that it has to be zero everywhere!

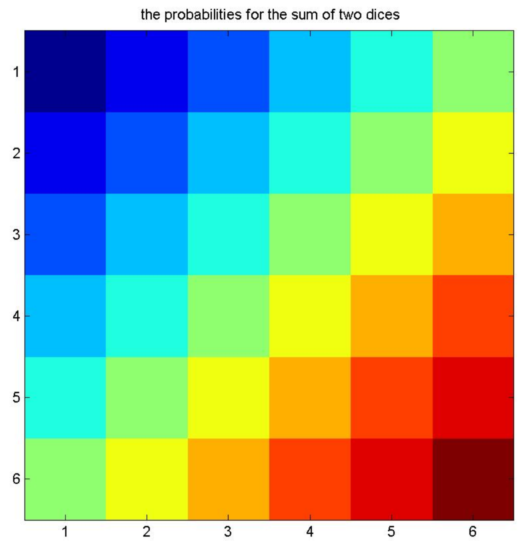

signals
spectra
1
0.5
0
200
100
0
−200 −100 0 100 200


signals
spectra
x 10^−3


the convolution of two rectancular functions

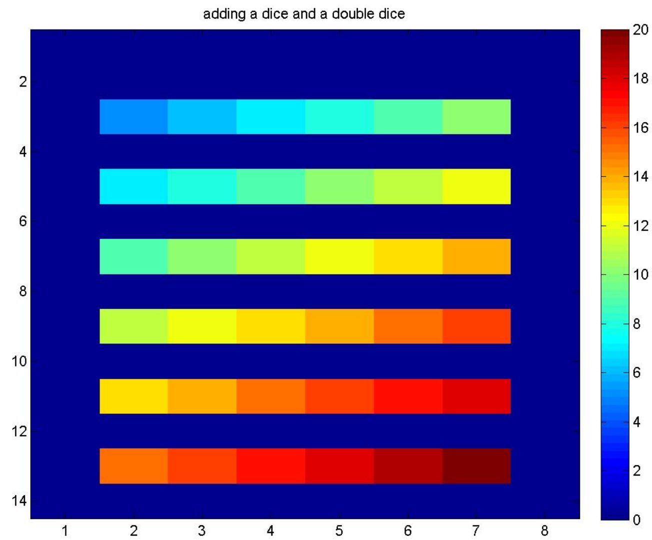

Dirac sequence
spectra
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
−200
0
200
1
0.5

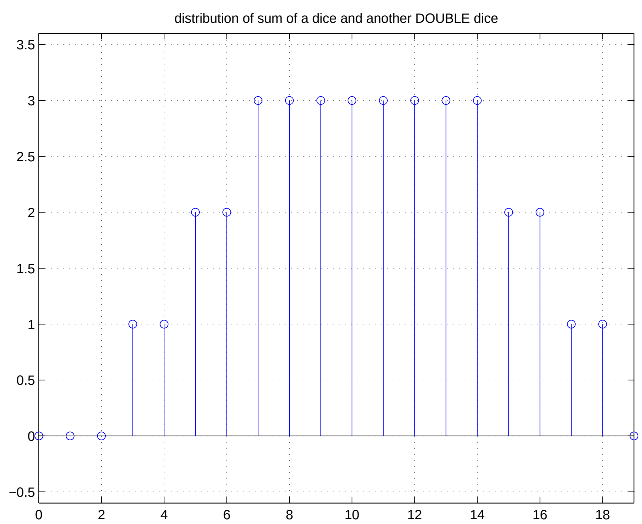

signals
spectra
x 10^−3
cubic spline
−200
0
200
15
10
5
1
0.5
0.4
0.2
−0.2
−100
100
0.1
−0.1
−50
50
0.05
−0.05
−40
−20
20
40



Each color represents one of the values from 2 to 12, and of course the number of tiles which appear in that given color has to be seen as an event of probability 1/36. Since the middle diagonal (sum equals 7) appears 6 times we find that the probability to get 7 points from two independent dices equals 1/6 (like to get 6 from one dice).

A variation of the argument





5 Possible Material for the future:

Recalling basic facts about the complex exponential function

Week two of the course in Canterbury University (July 20/21st, 2010), Christchurch, by HGFei

Assumptions (to be completed in the drafts for the earlier course notes):

We should have *Parseval's Identity* or *Plancherel's Theorem*, stating that the Fourier transform (considered on $\boldsymbol{L}^2(\mathbb{U})$) is a unitary mapping (really in the sense of giving a bijective, isometric one-to-one mapping) from the abstract Hilbert space $\boldsymbol{L}^2(\mathbb{U})$ with its scalar product to the prototypical separable Hilbert space $\boldsymbol{\ell}^2(\mathbb{Z})$.

This is in fact "only" the limiting relationship for the fact, that for every trigonometric polynomial with is standard representation

$$p(t) = \sum_{|k| \leq K} \hat{p}(k) e^{2\pi i k t}$$

one has

$$\|p\|_{\boldsymbol{L}^2(\mathbb{U})} = \|\hat{p}\|_{\boldsymbol{\ell}^2} = \sqrt{\sum_{|k| \leq K} |\hat{p}(k)|^2},$$

which is simple a consequency of the *Pythagorean theorem* (for general orthonormal systems in general Hilbert spaces).

Ex.: $f(t)$ is a p -periodic function if and only if $h(t) = f(t/p)$ is a $\mathbb{Z}$-periodic function.

Periodic functions of a general period p, i.e. for p-periodic functions, satisfying $f(x+p) = f(x)$, hence $f(x+kp) = f(x)$ for all $k \in \mathbb{Z}$. $\hat{f}(n)$ then of course gives the contribution of the function which oscillates n times over the fundamental interval $[0, p]$, or if one tries to normalize things for the sake of comparison, one would say that it has n/p oscillations per unit interval (whenever it makes sense).

6 The Fourier Transform in the non-periodic case

The idea of the Fourier transform came up as a generalization of the theory of periodic functions, and is obtained formally by letting the period p go to infinity. So essentially by "hand-waving" arguments one argues that the forward Fourier transform (or **Fourier analysis**) mapping f to the sequence of Fourier coefficients $\hat{f}(n)$ (but they are in fact related to frequencies located at the lattice $\mathbb{Z}/p$) so they get more and more dense as $p \to \infty$. Taking the Fourier inversion result (also called **Fourier Synthesis**) because it gives a series representation of the function as a superposition of pure frequencies [3] as a model one can argue that the non-periodic case should read as follows:

$$\hat{f}(s) = \int_{\mathbb{R}} f(t) \overline{\chi_s(t)} dt = \int_{\mathbb{R}} f(t) e^{-2\pi i s t} dt$$

and an inversion of the form

$$f(t) = \int_{\mathbb{R}} \hat{f}(s) \chi_s(t) ds = \int_{\mathbb{R}} \hat{f}(s) e^{2\pi i s t} ds.$$

[3]Even if the corresponding concept of convergence is not clear at all from the beginning, except for the case of $\boldsymbol{L}^2(\mathbb{U})$, where convergence is obviously in the norm, i.e. in the quadratic mean. Only around 1972 L. Carleson was able to prove that Lusin's conjecture is true: the convergence is in the pointwise almost everywhere sense, for $f \in \boldsymbol{L}^2(\mathbb{U})$, and soon later Hunt was extending this claim to the case of $\boldsymbol{L}^p(\mathbb{U})$.

This is absolutely OK as a heuristics, and makes us wonder, under which conditions one can trust such formal manipulations, let us therefore discuss them in detail.

In fact it is possible to claim that also in this expanded context one has Plancherel's theorem, i.e. one can claim that the Fourier transform is not only kind of invertible, but it is even isometric with respect to the *energy*-norm $\|f\|_2 := \sqrt{\int_{\mathbb{R}} |f(t)|^2 dt}$, and even preserves scalar products to the end that

$$\langle f, g \rangle = \langle \mathcal{F}f, \mathcal{F}g \rangle,$$

or altogether

Theorem 1 (Theorem of Plancherel)**.** *The Fourier transform is an isometric and bijective mapping from $\boldsymbol{L}^2(\mathbb{R})$ onto itself. It is unitary, i.e. the kernel of the inverse mapping is just the conjugate*[4] *of the forward Fourier transform, or equivalently, $\mathcal{F}\mathcal{F}f(t) = f(-t)$.*

In fact, we are immediately facing the following questions:

1. What are the conditions, under which one can assume that these two formulas make sense? As we shall see, the natural assumption is that the absolute integrand, i.e. $|f(t)|$ resp. $|\hat{f}(s)|$ are Lebesgue integrable (since this is the most general and important integral). We shall write $f \in \boldsymbol{L}^1(\mathbb{R})$ resp. $\hat{f} \in \boldsymbol{L}^1(\mathbb{R})$ to express this assumption in an elegant form.

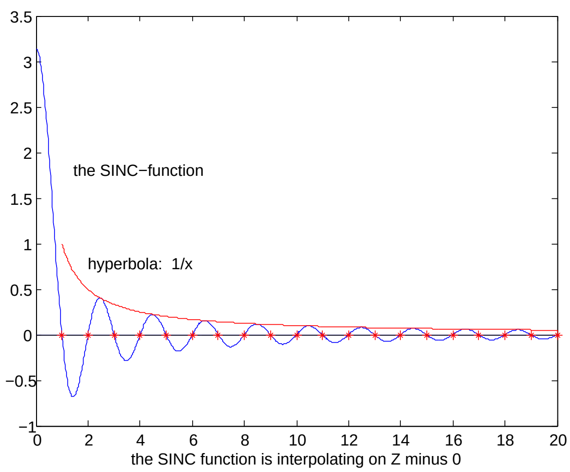
2. But is it always true that $f \in \boldsymbol{L}^1(\mathbb{R})$ implies $\hat{f} \in \boldsymbol{L}^1(\mathbb{R})$? Unfortunately the answer is NO, one may just take the simple box function

   $box(t) = \mathbf{1}_{[-1/2,1/2](t)} = 1$ for $t \in [-1/2, 1/2]$ and zero elsewhere. This function is clearly (even Riemann-) integrable, and its Fourier transform **Exercise!** is the SINC-function

   $$SINC(s) = sin(\pi s)/\pi s,$$

   with the convention $SINC(0) = lim_{s \to 0} SINC(s) = \int_{\mathbb{R}} box(t)dt = 1$.

[4] Since the function $(x, y) \to e^{2\pi i x y}$ is symmetric! with respect to x and y anyway!

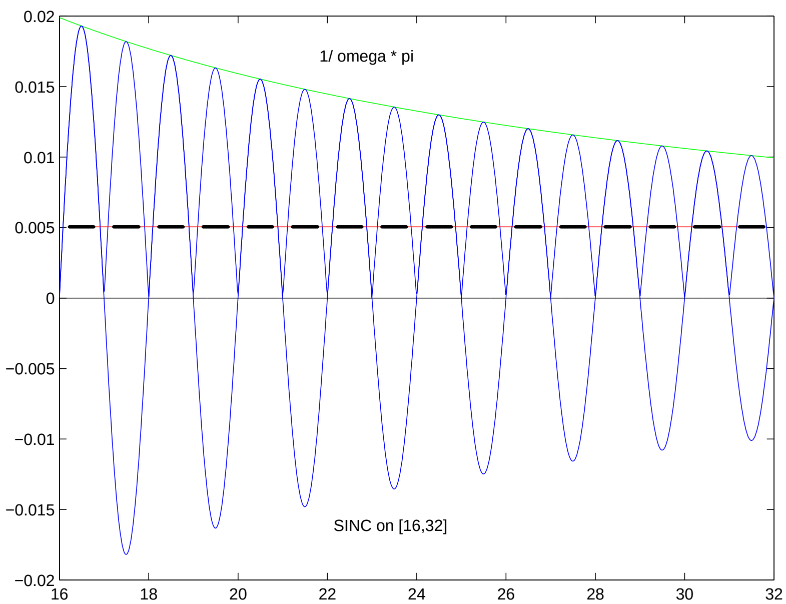


3. In order to prove Plancherel's theorem: Is every $f \in \boldsymbol{L}^2(\mathbb{R})$ also integrable. Over the compact interval such an argument can be made, because one

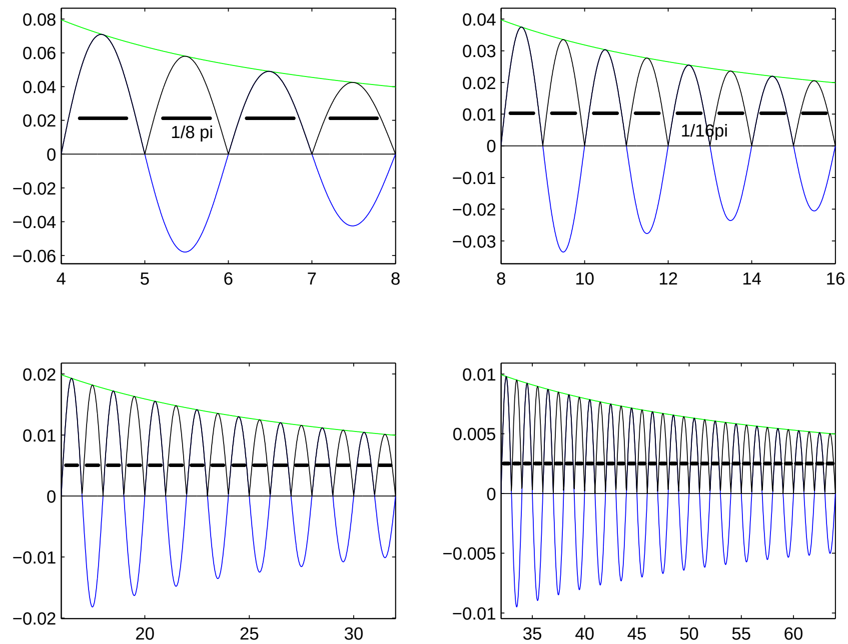
can apply the Cauchy Schwartz inequality, which says (in this context)

$$\int_0^1 |f(t)g(t)|dt \leq \sqrt{\int_0^1 f(t)dt} \cdot \sqrt{\int_0^1 g(t)dt}, \quad f,g \in \boldsymbol{L}^2$$

by choosing $g(t) \equiv 1$, but for the (unbounded) domain $\mathbb{R}$ such an argument is not true anymore, in fact $SINC$ is a good example of a function which is (!obviously: Exercise) square integrable (because $\int_1^\infty \frac{1}{x^2}dx < \infty$), but it is NOT integrable, due the divergence of the integral $\int_1^\infty \frac{1}{x}dx$. Displaying the poor decay behaviour of SINC:



Displaying the poor decay behaviour of SINC over several intervals, each of which picks up a minimum area of $1/\pi 2^k$, for $k = 3, 4, 5, 6$.



So how can we define the (Plancherel-) Fourier transform then?

4. Is there any chance of interpreting the inverse Fourier transform in a norm sense, or in a pointwise sense. In fact, how should the infinite integral be understood. As an *improper* Riemannian integral or as a (absolutely convergent) Lebesgue integral?

Before going into a discussion of the *technical* (positive) side of these issues (i.e. how to live with them and by which techniques the first-sight problems with these issues can be circumvented) let us feature another aspect of the Fourier transform, which makes it so important: it is its compatibility with *convolution*, or in other words to so-called *convolution theorem*, stating that the Fourier transform takes the (costly or complicate) bilinear operation called **convolution** into simple ptw. (pointwise) multiplication.

Again it appears to be natural to assume that the ingredients are taken from $\boldsymbol{L}^1(\mathbb{R})$, because one can show that for $f, g \in \boldsymbol{L}^1(\mathbb{R})$ one has: The convolution integral

$$f * g(x) := \int_{\mathbb{R}} f(x-y)g(y)dy,$$

exists for almost all $x \in \mathbb{R}$, and the function so defined (and spelled as f star g), or f *convolved with* g) satisfies

$$\|f * g\|_1 \leq \|f\|_1 \cdot \|g\|_1 \quad \text{for} \quad f, g \in \boldsymbol{L}^1(\mathbb{R}),$$

where as usual

$$\|f\|_1 := \int_{\mathbb{R}} |f(x)|dx < \infty$$

is the natural norm on $\boldsymbol{L}^1(\mathbb{R})$, turning the normed space $\big(\boldsymbol{L}^1(\mathbb{R}),\, \|\cdot\|_1\big)$ in fact into a Banach space!

The proof of the *submultiplicativity* of the $\boldsymbol{L}^1$-norm. It is a nice exercise in the change of variables (based on Fubini's theorem, justying the change in the order of integration).

Proof.

$$\|f * g\|_1 = \int_{\mathbb{R}^d} |f * g(x)| dx \leq \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} |g(x-y)||f(y)| dy dx$$

which upon a change of the order of integration becomes

$$\|f * g\|_1 \leq \int_{\mathbb{R}^d} |f(y)| [\int_{\mathbb{R}^d} |g(x-y)| dx] dy,$$

which in turn by changing the integration from x to $z = x - y$ (with $dz = dx!$, resp. using the translation invariance of the integral over the whole $\mathbb{R}^d$) gives us

$$\|f * g\|_1 \leq \int_{\mathbb{R}^d} |f(y)| \|g\|_1 dy = \|f\|_1 \cdot \|g\|_1,$$

as was claimed.

□

We can summarize this fact by claiming that $\big(\boldsymbol{L}^1(\mathbb{R}^d),\, \|\cdot\|_1\big)$, endowed with convolution as "multiplication" , is a commutative **Banach algebra**

Theorem 2. *The Banach space* $\big(\boldsymbol{L}^1(\mathbb{R}^d),\, \|\cdot\|_1\big)$*, endowed with* $*$ *(convolution) as multiplication, becomes a Banach algebra (i.e. this multiplication is continuous).*

EXERCISE: Show that $\|f_n - f\|_1 \to 0$ and $\|g_n - g\|_1 \to 0$ for $n \to \infty$ implies $\|f_n * g_n - f * g\|_1 \to 0$.

There are various rules concerning the compatibility of operators (translation, modulation and dilation) with convolution. Especially the compatibility of dilation in an area preserving way gives an *automorphism* if Banach algebra, since

$$St_\rho(f * g) = St_\rho f * St_\rho g. \tag{1}$$

and as a special case ($\rho = -1$) one has

$$(f * g)^{\checkmark} = f^{\checkmark} * g^{\checkmark}. \tag{2}$$

Next we are going to verify the so-called *convolution theorem*, which in short says that the Fourier transform turns convolution into pointwise (ptw) multiplication, or in formulas

$$\widehat{f * g} = \hat{f} \cdot \hat{g}, \quad \forall f, g \in \boldsymbol{L}^1(\mathbb{R}^d). \tag{3}$$

By approximation one can get a similar result for $g \in \boldsymbol{L}^1(\mathbb{R}^d)$ and $f \in LtRd$, with the Fourier transform then being take in the Fourier-Plancherel sense, using the fact that then $g * f \in \boldsymbol{L}^2(\mathbb{R}^d)$ and

$$\|g * f\|_2 \leq \|g\|_1 \cdot \|f\|_2. \tag{4}$$

Proof. We start with the definition of the Fourier transform and work our way through, making use of the exponential function laws

$$e^{z_1+z_2} = e^{z_1} \cdot e^{z_2} \quad \forall z_1, z_2 \in \mathbb{C}. \tag{5}$$

$$\widehat{f * g}(s) = \int_{\mathbb{R}^d} f * g(t) e^{2\pi i s \cdot t} dt$$

[5] or written out

$$\widehat{f * g}(s) = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} g(t-y) f(y) dy \cdot e^{-2\pi i s \cdot t} dt,$$

which upon (again!) a change in the order of integration becomes

$$\widehat{f * g}(s) = \int_{\mathbb{R}^d} f(y) \int_{\mathbb{R}^d} g(t-y) \cdot e^{-2\pi i s \cdot t} dt\, dy.$$

Replacing now $e^{-2\pi i s \cdot t}$ by $e^{-2\pi i s \cdot (t-y)} \cdot e^{-2\pi i s \cdot y}$ (based on the exponential law recalled above) we arrive at the expression

$$\int_{\mathbb{R}^d} f(y) \cdot e^{-2\pi i s \cdot y} \int_{\mathbb{R}^d} g(t-y) \cdot e^{-2\pi i s \cdot (t-y)} dt\, dy, \int_{\mathbb{R}^d} f(y) \cdot e^{-2\pi i s \cdot y} \hat{g}(s) dy = \hat{g}(s) \cdot \hat{f}(s),$$

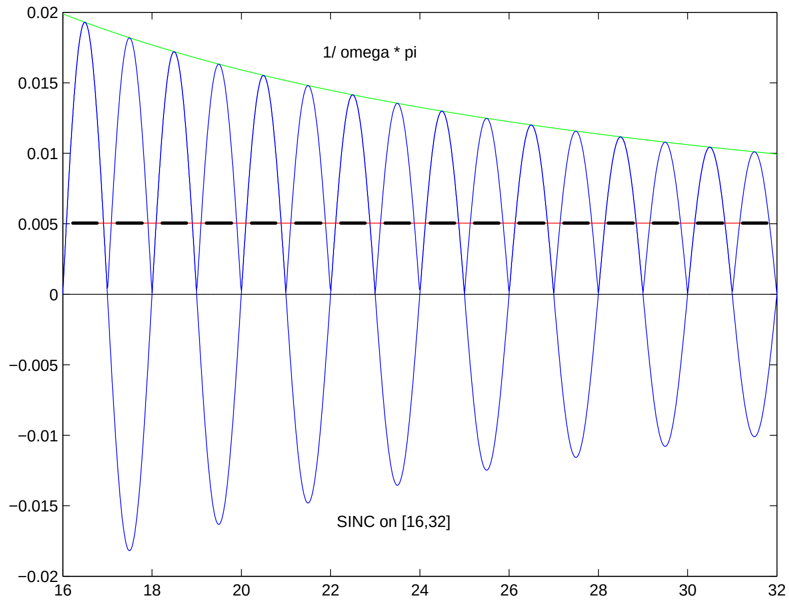
where again the decisive step is justified by substituting $z = x - y$ in the inner integral, which is possible for each fixed (!) $y \in \mathbb{R}^d$. $\square$

7 hints to the literature

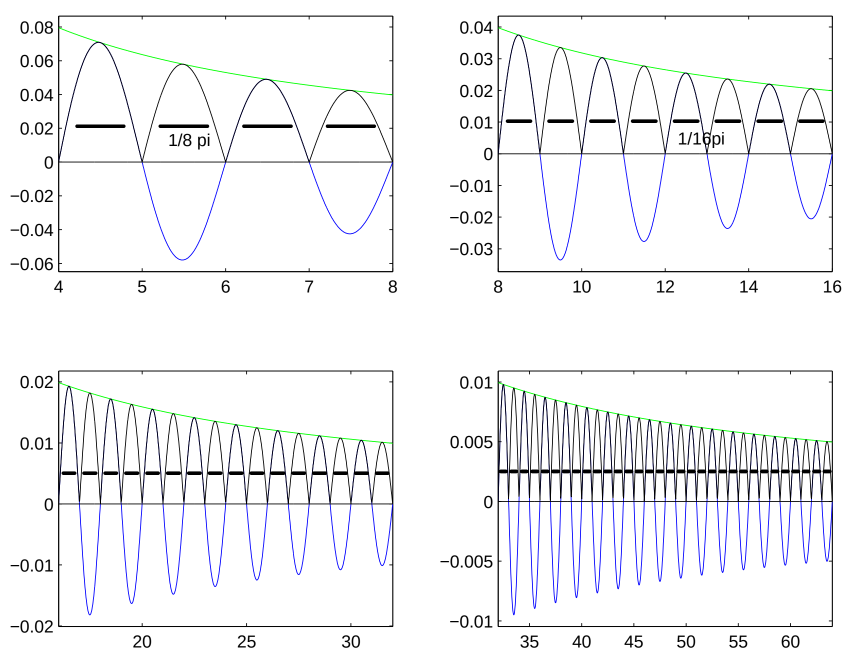
Details are found in most books in Fourier analysis, often making a point in stressing the relevance of Lebesgue integration theory for the business.

Displaying the poor decay behaviour of SINC:

[5] On $\mathbb{R}^d$ we use the scalar product $s \cdot t := \sum_{k=1}^{d} s_k t_k$ to define pure frequencies!



Displaying the poor decay behaviour of SINC over several intervals, each of which picks up a minimum area of $1/\pi 2^k$, for $k = 3, 4, 5, 6$.



One can say that in some sense the Shannon Sampling theorem is a consequence of Plancherel's Theorem (telling us that the Fourier transform is a *uni-*

tary linear operator on $\left(\boldsymbol{L}^2(\mathbb{R}^d), \|\cdot\|_2\right)$ and the fact that the *pure frequencies* define an orthonormal system for the $\boldsymbol{L}^2$-space over the fundamental domain, e.g. the unit cube $[-1/2, 1/2]^d$ within $\mathbb{R}^d$, viewed as the subspace of $\boldsymbol{L}^2(\mathbb{R}^d)$ which is obtained by multiplying (on the FT side) all the $\boldsymbol{L}^2$-functions with the indicator function (box function) of the fundamental domain.

In order to formulate it in a way that is compatible with the usual conventions let us call Λ to be a lattice, i.e. a discrete subgroup of $\mathbb{R}^d$, or more concretely $\Lambda = A(\mathbb{Z}^d)$ for some non-singular $d \times d$-matrix. Furthermore $\Lambda^\perp$ is by definition the so-called *orthogonal lattice*, which is the set of all pure frequencies χ_s such that $\chi_s(\lambda) \equiv 1$ for all $\lambda \in \Lambda$.

Theorem 3. *Assume that $\Omega \subset \mathbb{R}^d$ is a bounded subset with the property Ω is disjoint from all its (non-trivial) $\Lambda^\perp$-translates, i.e. $\lambda^\perp + \Omega \cap \Omega = \emptyset$ for all $\lambda^\perp \neq 0, \lambda^\perp \in \Lambda^\perp$.*

Then every Ω-band-limited function $f \in \boldsymbol{L}^2(\mathbb{R}^d)$, i.e. every f which can be written as an inverse [Plancherel] Fourier transform of the form

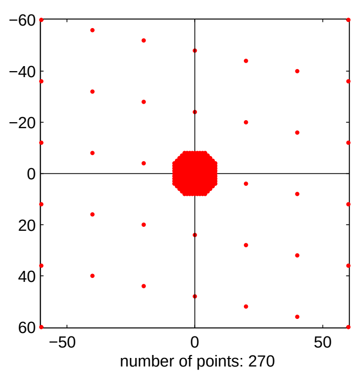
$$f(t) = \int_\Omega h(s) e^{2\pi i s \cdot t} ds \tag{6}$$

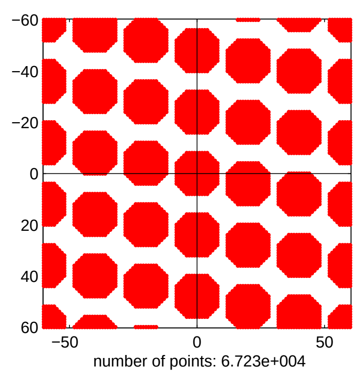
[6] *can be completely be recovered from the samples $(f(\lambda))_{\lambda \in \Lambda}$ by the following series expansion (called the Shannon Sampling expansion): For a suitable chosen constant $C_\Lambda > 0$, one has:*

$$f(t) = C_\Lambda \sum_{\lambda \in \Lambda} f(\lambda) T_\lambda SINC_\Omega(t) = C_\Lambda \sum_{\lambda \in \Lambda} f(\lambda) SINC_\Omega(t - \lambda), \tag{7}$$

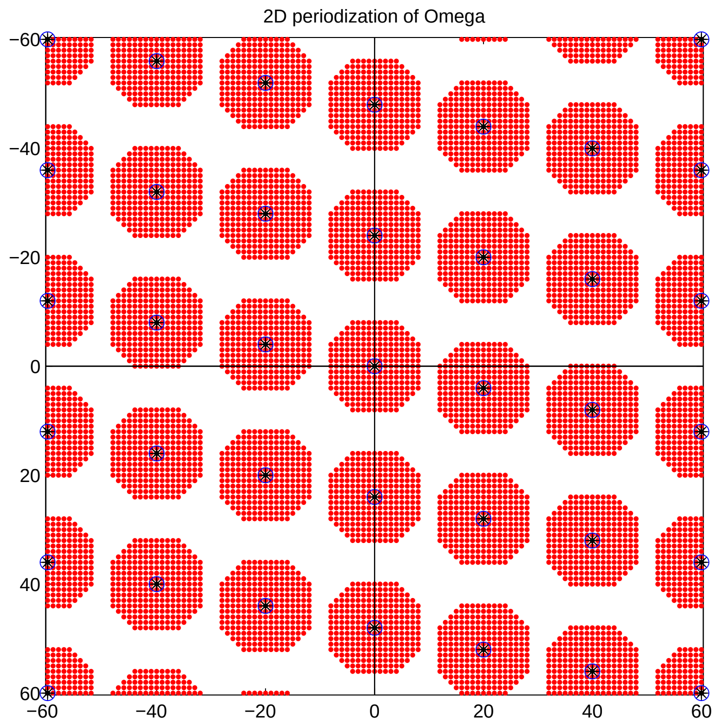
where $SINC_\Omega = IFFT(\mathbf{1}_\Omega)$. The convergence takes place in the $\left(\boldsymbol{L}^2(\mathbb{R}^d), \|\cdot\|_2\right)$-sense, but also uniformly.

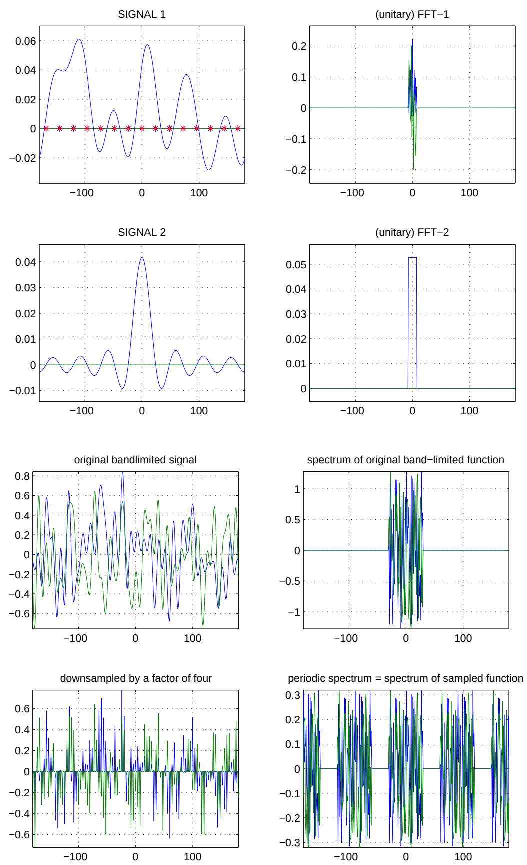
Remark 1. For the standard choice, i.e. $\Omega = [-1/2, 1/2]^d$ we simply have $SINC_\Omega(t_1, \cdots, t_d) = SINC(t_1) \cdots SINC(t_d)$, with the classical SINC-function, $SINC = IFFT(\mathbf{1}_{[-1/2,1/2]})$. Moreover we have $\Lambda^\perp = \mathbb{Z}^d$ for $\Lambda = \mathbb{Z}^d$, and $\Lambda^\perp = \frac{1}{a}\mathbb{Z}^d$ for $\Lambda = a \cdot \mathbb{Z}^d$.

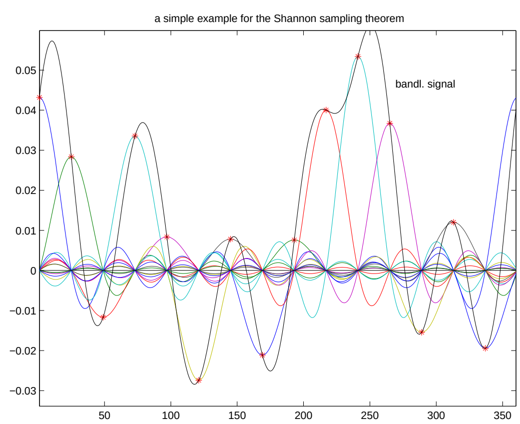




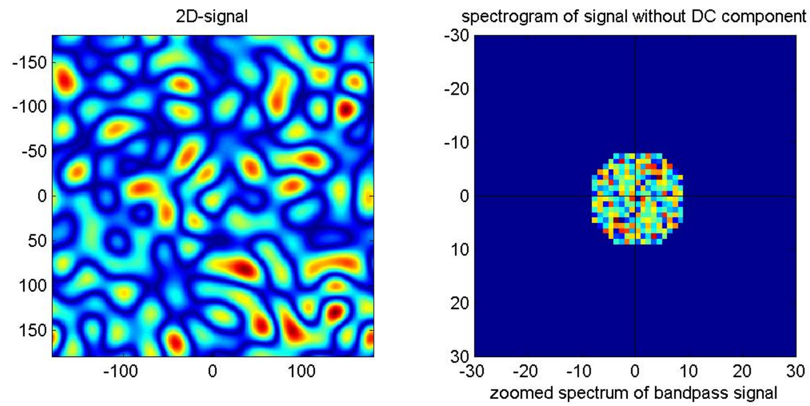
[6] Or in a more practical description: if $h = \hat{f}$ is supported inside Ω, resp. is zero outside of Ω, except possibly for a set of measure zero.


2D periodization of Omega
−60
−40
−20
0
20
40
60
−60
−40
−20
0
20
40
60





A typical 2-D band-limited signal looks as follows:



Such images also give an idea about data compression by sampling. The *true dimension* of the image can be expressed by the number of Fourier coefficients. The number of copies of the spectrum indicate the reduction factor (e.g. in terms of storage).

Not that both images only show a segmet of the full spectral domain!

Material for the course "Fourier Analysis" MATH 425

Lecturer: Hans G. Feichtinger (University Vienna), Assignment

1. Give the formula for the Fourier transform matrix of size $n = 4$. It can be obtained via the MATLAB command `fft(eye(4))` and is essentially the Vandermonde matrix of the unit roots of order 4 (so comparable with `vander([1,i,-1,-i]))` .

2. Verify directly that up to the normalization factor (equals $2 = \sqrt{4}$ here) its rows and columns represent an orthonormal basis for $\mathbb{C}^4$.

3. Make a connection to the so-called *polarization identity* which hold true for arbitrary complex Hilbert spaces $\mathscr{H}$. For any $x, y \in \mathscr{H}$ the scalar product $\langle x, y\rangle_{\mathscr{H}}$ can be expressed as a sum of norms:

$$\langle x, y\rangle_{\mathscr{H}} = \frac{1}{4}\sum_{k=0}^{3} i^k \langle (x + i^k y), (x + i^k y)\rangle_{\mathscr{H}} = \frac{1}{4}\sum_{k=0}^{3} i^k \|x + i^k y)\|^2, \quad (8)$$

where i denotes to complex unit in $\mathbb{C}$.

4. Convolution is a continuous form of the so-called Cauchy product for ordinary polynomial functions. The Cauchy-product of two sequences can be obtained as follows. Turn two given sequences $a, b \in \mathbb{C}^n$ into polynomials, typically by writing

$$p_a(z) = \sum_{k=1}^{na} a_k x^{k-1}, \quad q_b(z) = \sum_{j=1}^{nb} b_j x^{j-1}$$

   and then forming the product polynomial $r(z) := p(z) \cdot q(z) = q(z) \cdot p(z)$. Then

$$c := a * b \quad (= b * a)$$

   is by definition the seqence of coefficients of $r(z)$. Since $p_a(z)$ is of degree $na-1$ and $q_b(z)$ is of degree $nb-1$ it is clear that $r(z)$ has degree $na+nb-1$ (usually one calls the number of coefficients in the polynomial its order, e.g. $order(p(z)) = na$; and one needs at least na values of such a polynomial in order to recover all of its coefficients!).

   This for example the binomial coefficients (describing $(1+x)^n$ as a polynomial) can be obtained by *iterated* Cauchy-products.

   Recalling that the FFT is just producing the values of a polynomial over the unit roots of order N (in clockwise order), try to find out how one can use the FFT to obtain (even high order) binomial coefficients without resorting to long number digit calculates (involving factorials of huge value and their division).

5. By considering the number 123 as value of the polynomial $p(z) = 3+2z+1z^2$ at $z = 10$ one easily can verify that multiplication of integers as we learn it at school is pretty much the same multiplication (in the above sense) of polynomials. Determine 123^3 in this way!

6. There is also a nice analogy to (elementary) probability theory. Let us *code* [7] the possible results of a fair dice, i.e. the equal probability of obtaining the numbers from $\{1, 2, \ldots, 6\}$ by throwing the dice, in the form of the polynomial $p(z) = (z + z^2 + \ldots z^6)/6$ (note that the coefficient to z^0 is assumed to be zero!). Then find out the cofficients of $p^2(z)$ on the other hand and compare with your knowledge about the sum of two fair and independent dices! Which out of the possible numbers in $\{2, 3, \ldots 11, 12\}$ occurs most likely?

8 Continuous Fourier Transform

1. Verify the commutativity of convolution, i.e. $f * g(x) = g * f(x)$;

[7] no other justification then that of a formal analogy which is quite surprising and helpful in order to understand things!

2. Check the details that dilation goes onto compression by the Fourier transform and vice versa (we have the formula $\widehat{St_\rho f} = D_\rho \hat{f})$.

3. Show that $f \mapsto St_\rho f$ defines an automorphism of the Banach algebra $\big(\boldsymbol{L}^1(\mathbb{R}), \|\cdot\|_1\big)$ into itself, i.e. show that St_ρ is a linear mapping for each $\rho > 0$, which preserves the $\boldsymbol{L}^1$-norm ($\|St_\rho f\|_1 = \|f\|_1$), and is also compatible with the multiplication in $\boldsymbol{L}^1(\mathbb{R})$, meaning that

$$St_\rho(f * g) = St_\rho f * St_\rho g.$$

4. Replacement for the duplicate original : In order to derive the Plancherel Formula in the form

$$\langle f, h\rangle = \langle \hat{f}, \hat{h}\rangle \quad \text{for} \quad f, h \in \boldsymbol{L}^2(\mathbb{R}),$$

it is convenient to verify that *applying twice* the mapping $g \mapsto conj(\hat{g})$ is giving the identity operator. Try to argue that this is true by showing that it is valid for all the functions of the form

$$St_\rho g_0, M_t g_0, T_x g_0, T_x D_\rho g_0,$$

where g_0 is the Gauss function (using $\widehat{g_0} = g_0$ and the rules for $\mathcal{F}$.

5. Show that similar claims can be made with respect to modulation operators.

$$M_s f(z) = e^{2\pi i s z} f(z) = \chi_s(z) \cdot f(z)$$

i.e. ptw. multiplication with a *pure frequency*. Again, M_s is a linear operator for each s, we have $M_s \circ M_t = M_{s+t} = M_t \circ M_s$, and finally

$$M_s(f * g) = M_s f * M_s g.$$