Mathematical Background to a Digital World
Fourier based methods in digital signal processing

Hans G. Feichtinger
hans.feichtinger@univie.ac.at
www.nuhag.eu

April 28th, 2011
PAM Seminariet
Uppsala University

Hans G. Feichtinger: personal background

- Study of mathematics and physics, Univ. Vienna
- PhD in abstract harmonic analysis (function spaces)
- initiator of image processing research network;
- various guest professorships (e.g. USA, 1989/90);
- early 90's: founding of NuHAG (= the Numerical Harmonic analysis Group, jointly with K. Gröchenig and T. Strohmer);
- `www.nuhag.eu`: Platform of the group
- last decade: EU-networks, Marie Curie Projects, etc.
- view myself as a **abstract harmonic analyst converted to an application oriented and computational HA.**

A short survey of NuHAG Projects: see www.nuhag.eu

- **ESO** data reduction pipeline (stacking and processing of astronomical images), instrument independent software;
- **UnlocX**: FET project [EU]: new forms of signal repres., medical imaging and sound shaping (electro-cars);
- **SISE**: FWF, nat. focus network, communication engineers;
- **Audio-Miner**: WWTF-project (Monika Dörfler: Mathematics and Music!, see Notices AMS), sound objects;
- **MULAC**: WWTF-project (Peter Balazs, Acoust. Res. Inst. Austrian Academy of Sciences), Gabor multiplier, MP3;
- **SPEVOTAP**: Margit Pap: Indiv. Marie Curie Fellowship on complex analysis and atomic decompositions;
- **EUCETIFA: 2005-2009**: Marie Curie Excellence Grant

The Talk contains Comments on

- The role of **abstraction** and mathematics;
- The relevance of information representation, for efficient storage (compression), handling, transmission;
- The advantages of **digital information representation**;
- Some mathematical ideas in the background;
- A variety of applications (audio, mobile communication, image representation);

Note: By representing information (sound, images) in a digital way we do not only change the representation, but we also allow ourselves to store, transmit, modify this information (think of *video on demand*, GOOGLE, PICASA, etc.), search, sort ...

THREE THESES

1. **Mathematics is part of many activities of our daily life!** Most often it goes un-noticed, and it is not just about *numbers* and *computers!* You will see a number of examples in this talk.
2. **Mathematics is abstract in a good sense**, i.e. we don't have to count apples or pears using a different number system. Abstractness means concentration on the *essential structures* needed, and thus contributes enormously to the *universality* of mathematics.
3. **Mathematics is an exciting subject**, with many new (and complex) structures to be investigated, although 2×2 remains 4 after all. Much to do!

Typology of Mathematicians: Comparison with Climbers

There are many ways to do *mathematics* (but most of what I am saying applies to other sciences as well):

- extreme climbing == top research (the "famous");
- professional tour guides == (academic) teachers;
- recreational mountaineering == mathematics working in different disciplines of science and technology;
- the majority of people who dislike strain and sweating == the majority of people who have learned to dislike mathematics because it has brought pain and fear at school!

However, if on is getting a good training mathematics (just like mountaineering) gives you a good (mental) shape and a lot of joy, because you can do things *very much on your own*.

Comparison with Musicians:

- Composers/singers == top scientist (the "famous ones");
- professional musicians == university professors
- musical school teachers == mathematics teachers
- laymen musicians == Hobby-mathematicians
- the general public, which maybe likes music and likes to listen to good music == IN CONTRAST: the few ones who consider mathematics a useful and interesting subject

This example shows very well that a great majority of people like music, they enjoy it, they have a positive opinion about it, although they my typically not produce their own music, or even care about score, transcription, musicology and so on.

Mathematics in our Daily Life: the Digital Revolution

1. CD-Player, MP3-Player, DVD-Player, Digi-Cam;
2. WLAN, mobile phone, digital radio;
3. satellites for weather forecast; flight scheduling;
4. databases, search engines;
5. tomography, medical imaging;
6. computational Physics (material sciences, car crash simulation);
7. computational Chemistry (new chem. compounds);
8. pharmacy (new remedies, computer simulation);
9. **In almost all these areas mathematical methods and simulations, carried out on computers, play an essential role!**

Numbers and Efficient Computation

The easiness of computations depends on the representation of numbers. Our decimal system (which easily allows to sort numbers according to size, and to carry out addition easily was a big advantage over the Roman number system!)

```
XXXXVIIII + XVII = LXVI
```

is much harder to carry out than the decimal addition

```
49 + 17 = 66
```

And WIKIPEDIA tells us that **Beda Venerabilis** (death anno 735) introduced the symbol 0 only early middle age.

Different types of Numbers and Multiplication

If I am asking you about the product of the following two numbers, it may take you a while:

$$1.714285714285714 \cdot 0.538461538461538 = \; ? \tag{1}$$

On the other hand, the multiplication

$$\frac{12}{7} \cdot \frac{7}{13} = \frac{12}{13} = 1 - 1/13 \tag{2}$$

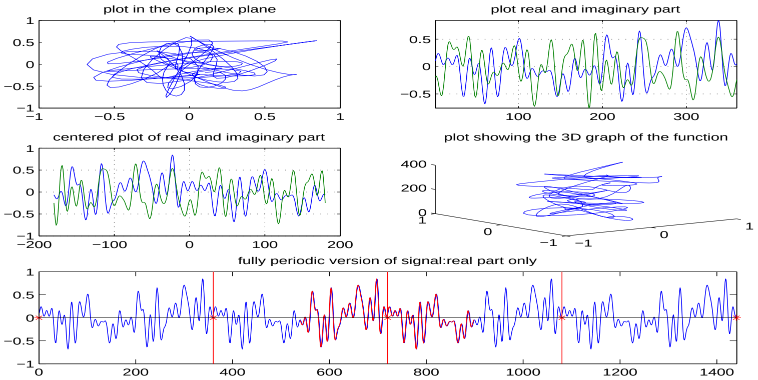
is easily performed. A similar thing would be to ask for

$$\frac{1}{x}, \quad \text{given} \quad x = 1.714285714285714. \tag{3}$$

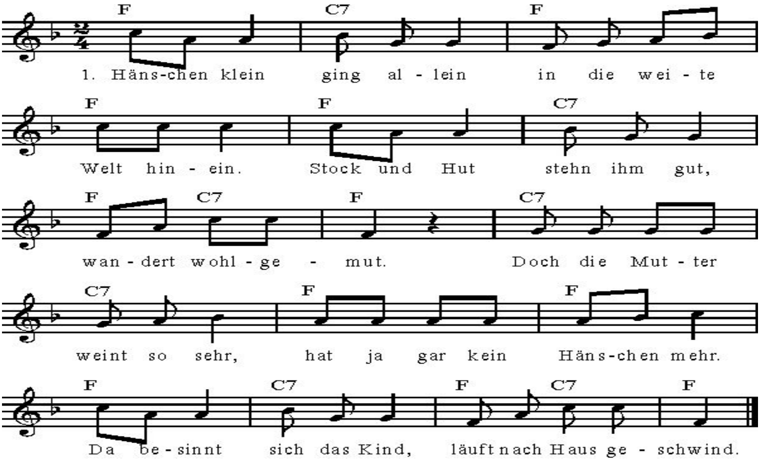
which by the same reasoning is just $7/12$.

Representing complex valued functions

There are real functions: $t \mapsto f(t)$, which assign to each time some real value (say the temperature curve). There is however a much larger system ("field") of numbers, the complex numbers $\mathbb{C}$, and correspondingly complex-valued functions.



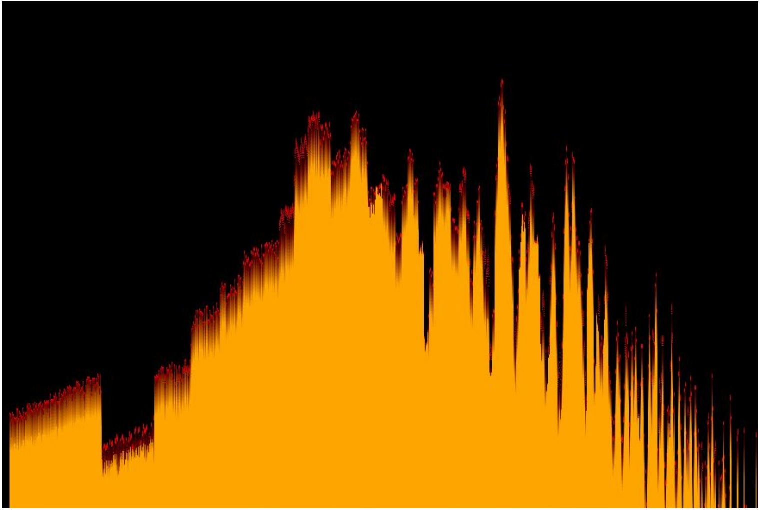
A simple Musical Score



Visualization in the Media-Player

The next slide will show a snap-shot from the Windows Media-Player.
It is one of the visualizations of music, and it provides the change of energy over time and frequency, in other words, if at a given time (a specific part of a piece of music) there is high amplitude at a given frequency - like in the image given below - then at this moment this particular frequency is most prominent.
It is particularly clear if e.g. a flute is playing, or a female singer.
The profile of such an image (i.e. the distributions of harmonics, etc.) indicates the timbre of the piece of music resp. instrument playing.

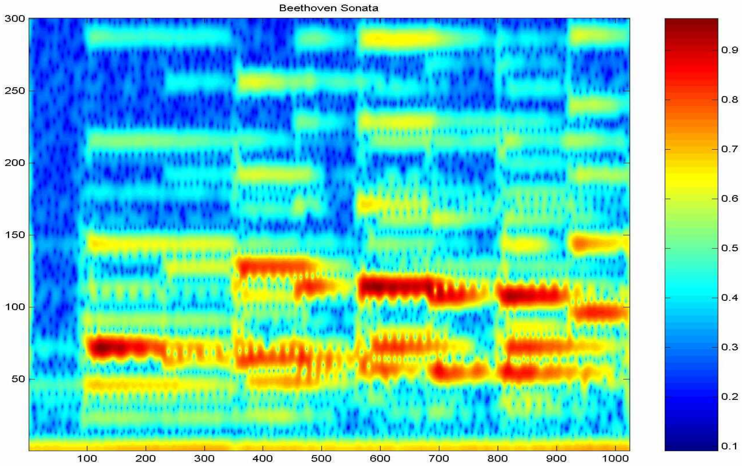
Visualization in the Media-Player



A Typical Musical STFT

For a mathematical interpretation of this “sliding window Fourier transform” (the Fourier = frequency analysis of a piece of music) is performed in time, over segments of the incoming time-signal. However, since the FFT is performing a Fourier decomposition of the periodic signal arising from such a segment, it is not advisable to just take the segment itself, because most often the beginning and the end of the finite signal do not match well. It is therefore much better, if the localization is done by smooth window function, such as a plateau-type function with side-lobes which are like a raised cosine squared (say) (using the formula $\sin^2(t) + \cos^2(t) = 1$). Such considerations are also part of MP3. We display this STFT with the time-axis horizontally, and the frequency axis vertically.

A Typical Musical STFT



Key Players for Time-Frequency Analysis

Time-shifts and Frequency shifts

$$T_x f(t) = f(t-x)$$

and $x, \omega, t \in \mathbb{R}^d$

$$M_\omega f(t) = e^{2\pi i \omega \cdot t} f(t)\,.$$

Behavior under Fourier transform

$$(T_x f)\hat{} = M_{-x}\hat{f} \qquad (M_\omega f)\hat{} = T_\omega \hat{f}$$

The Short-Time Fourier Transform

$$V_g f(\lambda) = V_g f(t,\omega) = \langle f, M_\omega T_t g\rangle = \langle f, \pi(\lambda) g\rangle = \langle f, g_\lambda\rangle, \quad \lambda = (t,\omega);$$

STFT = Short-time Fourier Transform: Basic Properties

The STFT (or windowed FT) provides information about local (smoothness) properties of the signal f by multiplication with some *window function* g and a subsequent Fourier transform. Typically g is Schwartz function (from $\mathcal{S}(\mathbb{R}^d)$) concentrated around the origin, such as the Gaussian, and $f \mapsto V_g f$ is linear:

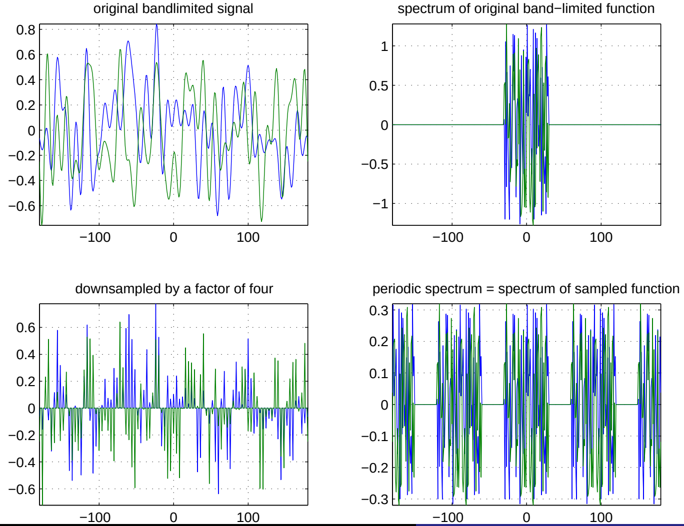
$$V_g f(x,\omega) = \int_{\mathbb{R}^d} f(t)\overline{g(t-x)}e^{-2\pi i t\omega}dt, \quad \text{for} \quad (x,\omega) \in \mathbb{R}^{2d}, \quad (4)$$

In 1927 Weyl pointed out that the translation and modulation operator satisfy the following commutation relation

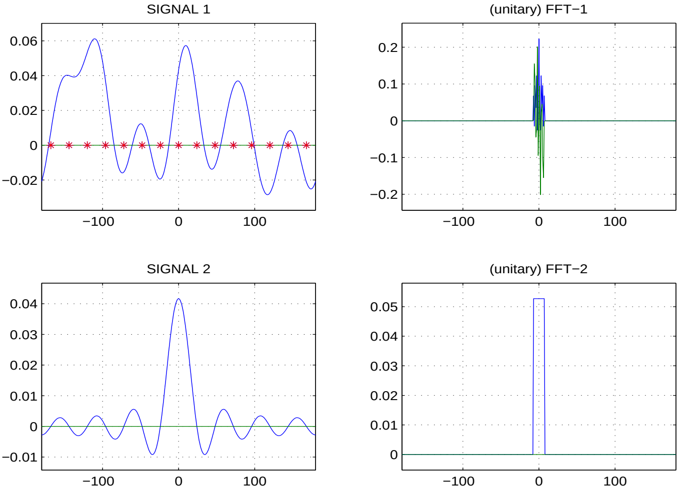
$$T_x M_\omega = e^{-2\pi i x\omega} M_\omega T_x, \quad (x,\omega) \in \mathbb{R}^{2d}. \quad (5)$$

$\{T_x : x \in \mathbb{R}^d\}$ and $\{M_\omega : \omega \in \mathbb{R}^d\}$ are Abelian groups.

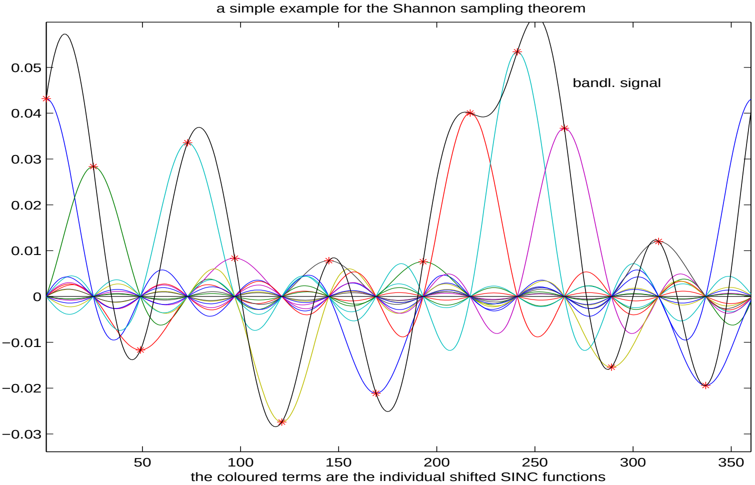
The Shannon Sampling Theorem: basis for CD-players



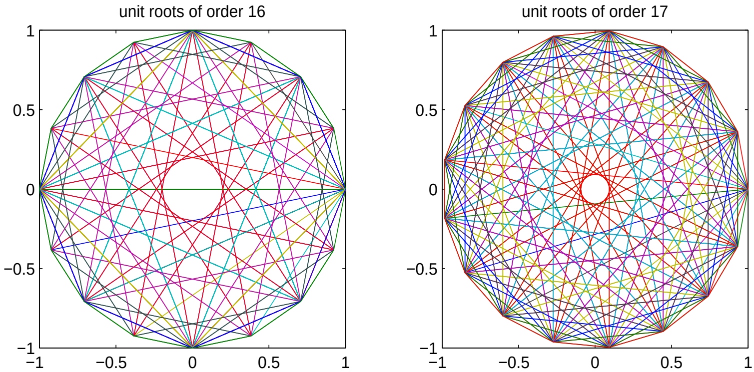
Effect of sampling: periodization of the Spectrum



Reconstruction using shifted SINC-functions



The facts behind the FFT algorithm



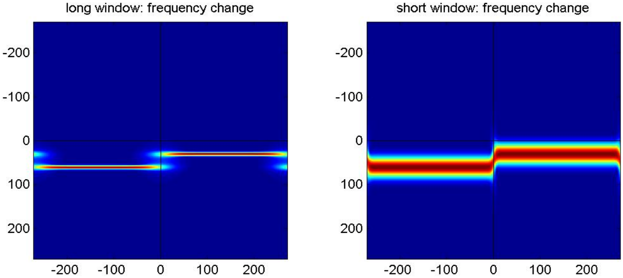
Why do we have the sampling rate 44100 for CDs?

The above image shows, that depending on the signal length (here it is/was 16 ($= 2^4$, i.e. a power of 2) resp. 17 (a prime number)) there is a quite different subgroup structure.
While the unit-roots of order 16 are the union of unit roots of order 8 (2 copies, one rotated by $22,5$ degrees) and so on and so on there is no such basis for a *recursive algorithms* possible.
The sampling rate for HiFi audio should be approx. 10% above the necessary limit for $20kHz$ ($= 40.000$ samples per second) but also a number rich in divisors.
If you don't want to do the prime-factorization during the rest of my lecture, remind me to disclose it to you at the end of the lecture.

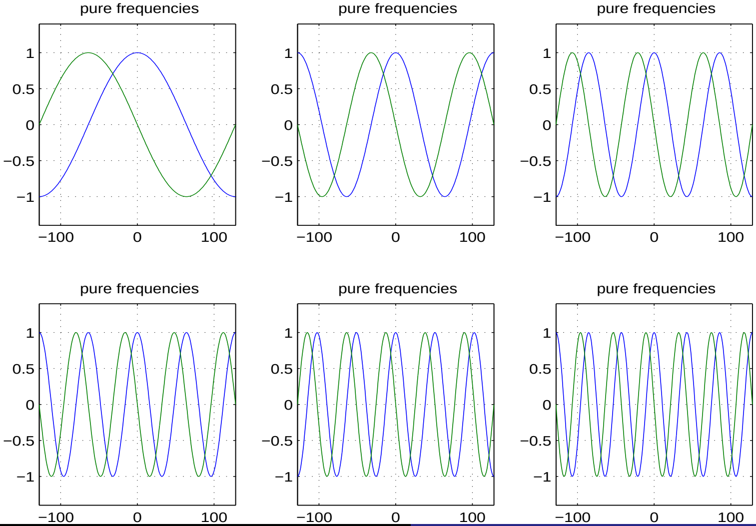
Although the so-called *spectrogram* (official name: the short time Fourier transform, because it describes, quite in the sense of local a musical score, at which times which frequencies contain how much “energy”) contains valuable information about the time-frequency distribution of energy within the signal, one cannot give a meaning to the individual values.
The so-called (Heisenberg) *uncertainty principle* implies that there is no meaning for the energy-content at a *given point in the time-frequency plane*[1]
Depending on the choice of the averaging window one either has good frequency resolution (long windows) or good time resolution (short windows, but poor frequency resolution).

[1] Very much like an atom does not have a color: for the same reason that Audio-Brush project announced a while ago by a young colleague in computer science did not work out well!

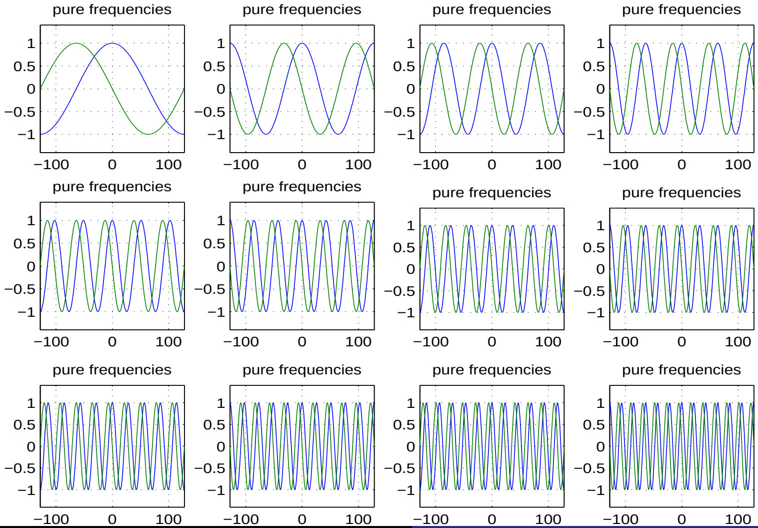
Size of the window and Frequency/Time Resolution



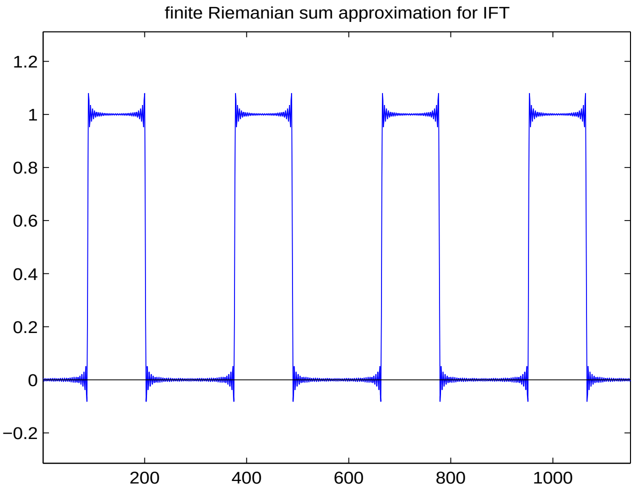
The building blocks: pure frequencies



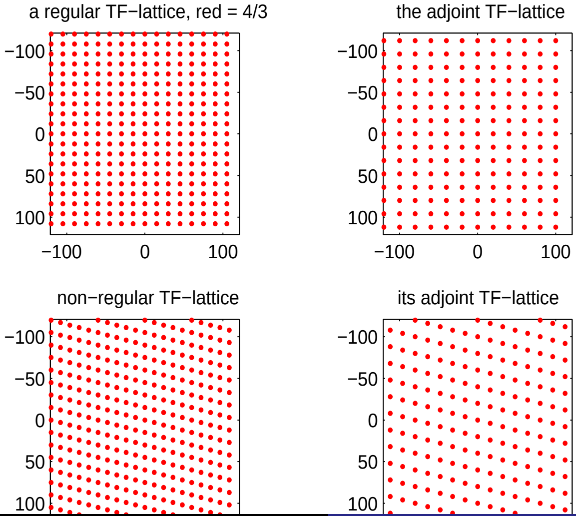
The building blocks: pure frequencies



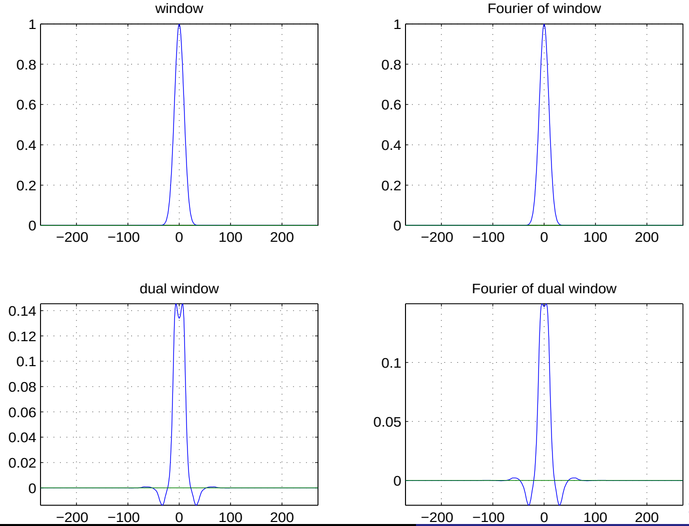
Fourier Series on Step Functions: Gibb's phenomenon



Regularity of Spectrogram: subsampling > Gabor Analysis



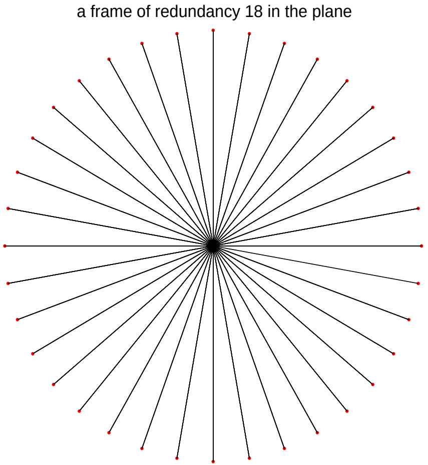
Action of a (frame) multiplier



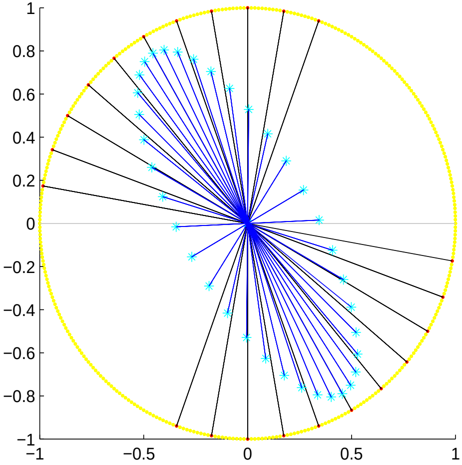
Gabor Transform: Audio I

Gabor Multiplier and Audio Control Desk

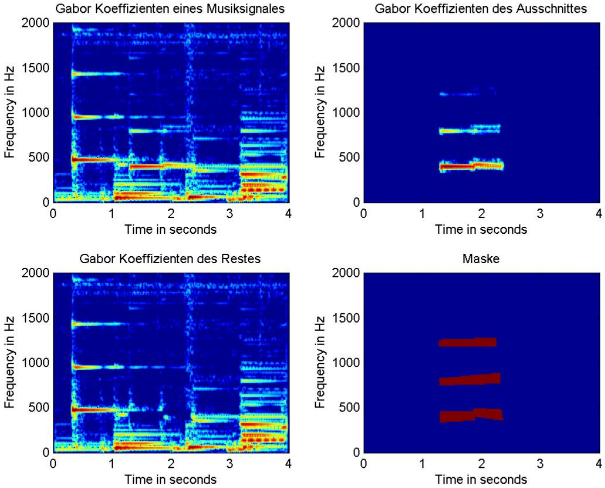
A highly redundant set of generators: a frame



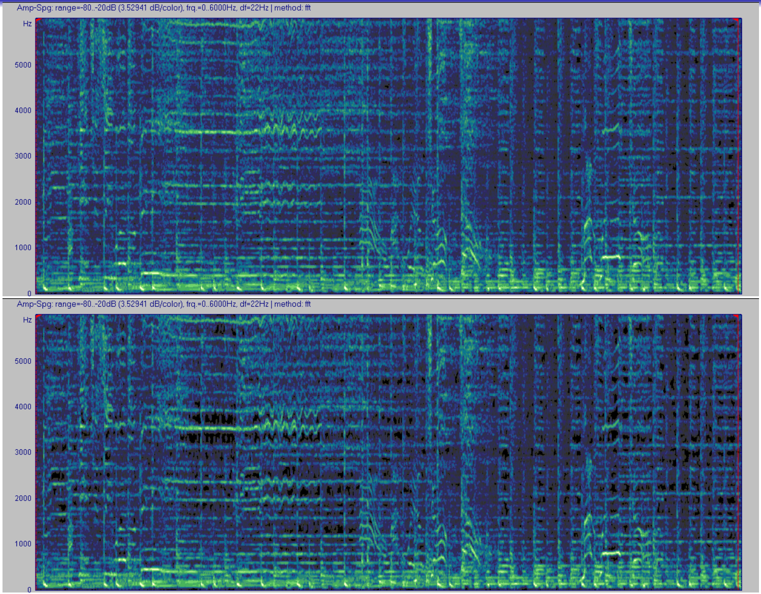
Action of a (frame) multiplier



Gabor multipliers and Masking

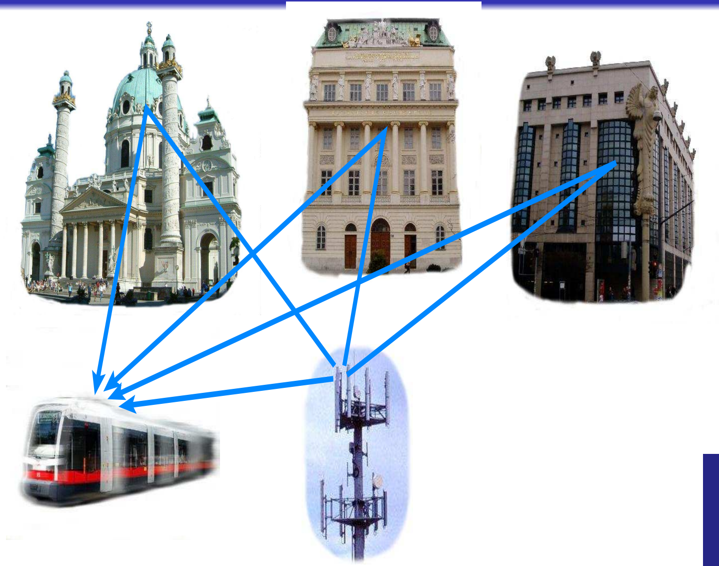


Masking visualized in the time-frequency plane: MP3!



Mobile Communication

Mobile Communication



Mobile Communication: Generalities

Mobile communication in the digital age has a number of **mathematical questions** in the background (which had to be solved, and which are still subject of research, involving communication engineers, mathematicians and computer scientists).

Member of NuHAG are holding two **patents** in this area (a third one is under preparation) and presently a number of companies (including Swedish ones) are investigating the possibility of making use of these inventions in their communication products.

Maybe in contrast to common belief there are many **open questions**, including data models, algorithms for efficient data handling, parallel processing (there are more and more multi-core processors on the market, butnot enough software using it optimally), algorithms for DSPs and GPUs, ...

Communication in Analogue Times

Most of us (say the older ones) remember the times where long distance calls where quite difficult and expensive, and at a low quality. Music was stored on tapes and shellacs (when I was a boy of say 12 years, my grandfather was buying the first stereo record for me, which was a great listening experience!), with a variety of problems:

Each time the record was played contributed to its (physical) deterioration (the most beloved records are the ones which are finally in the worst shape).

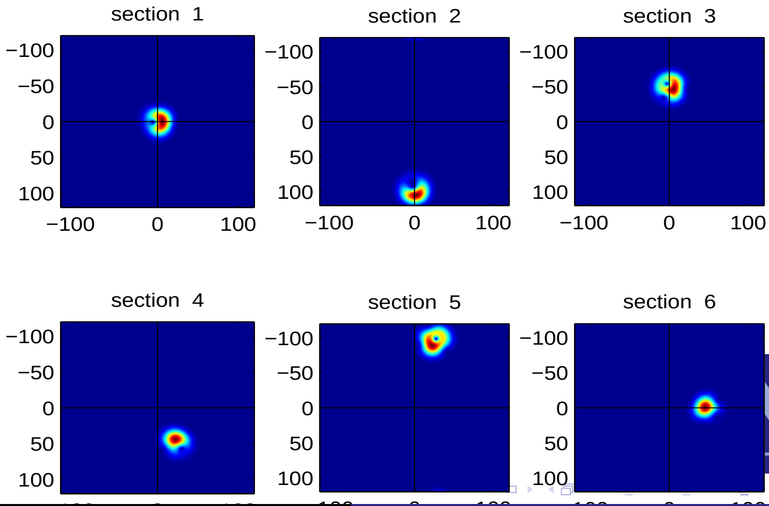
The analogue transmission lines (telephone cables) require analogue amplifiers which are also amplifying the noise (until the cumulative noise buries the underlying signal).

Even earlier (beginning of radio) the only way to have music in the radio was a life performance in front of a micro.

Communication in the Digital World

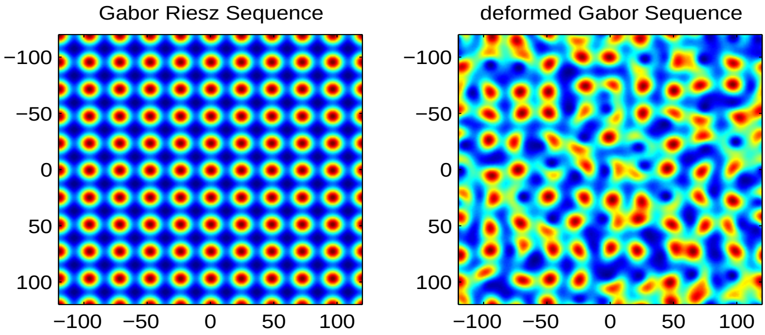
Information is coming in the form of discrete bits. If noise is added, and the receiver obtains a sequence of numbers such as $0.1, 0.2, 0.8, -.1, 0.9, 0.87, 0.12$ he/she can assume that it should be in reality $0, 0, 1, 0, 1, 1, 0$, by proper quantization.
The mathematical way in which bits are taken has to account for the fact that we are not just storing a step function approximation of the original signal.
Amplifying an image (e.g. in order to print a poster) is not the same as printing in large pixel format! (but is much more mathematically involved than most user would think of).
Transmission lines allow for repeated corrections, bringing always back (using also coding theory) the perfect data-set.

Communication using Gaborian Riesz Bases

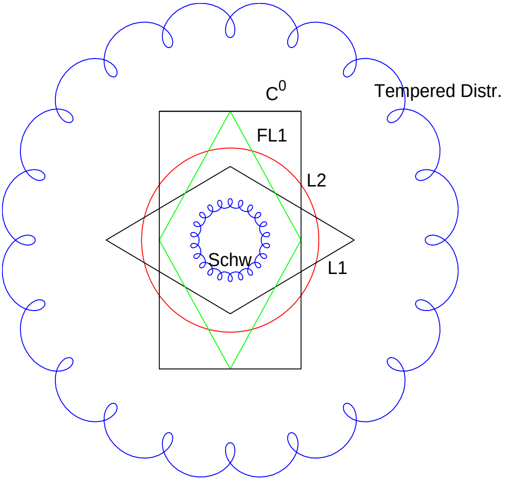


The received Gabor Riesz basis

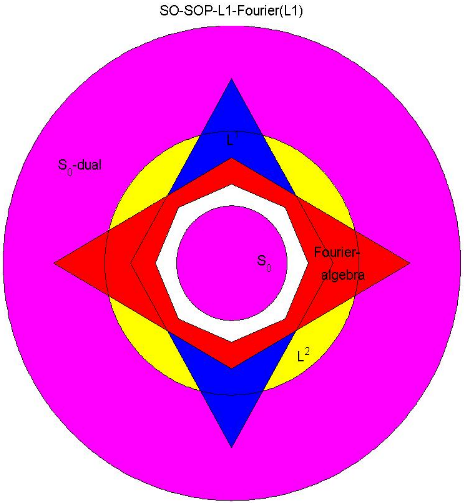
The received atoms carry the information



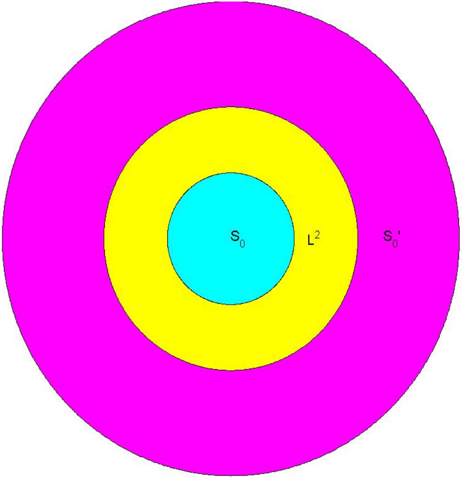
The classical view on the Fourier Transform: $\mathcal{S}'(\mathbb{R}^d)$



The TF-approach: Feichtinger's algebra $\mathbf{S}_0(\mathbb{R}^d) = \mathbf{M}^{1,1}$



The Banach Gelfand Triple $(\mathbf{S}_0, \mathbf{L}^2, \mathbf{S}_0')$



Much more to be done: theory and applications

Although it looks as if there is already a lot of deep mathematics around which just has to be *applied in the real world* we should rather think of the situation that we have in music:
Even thought there are tons of old melodies and scores in the archives which could/should be re-recorded for digital archives, it is the variety of new possibilities, of producing new sounds, of modifying things, or developing new *sound-scapes*, of combining rhythmic patterns with melodic constructions.
Just think (aside from our valued classical music, created by Bach, Haydn, Mozart and Beethoven) of *world music*, the idea of poly-rhythmic patterns (as opposed to polyphonic music), if the restriction of sound-production to regular temporal patterns and a fairly small set of standard instruments.

Image Processing, Digital Cameras and DVDs

Image processing applications are nowadays part of our daily life. Kids take pictures using their mobile phones, they download pictures from the internet, they mail digital photos to each other, the view digital movies on their notebooks and they are proud of their new digital cameras having millions of pixels.

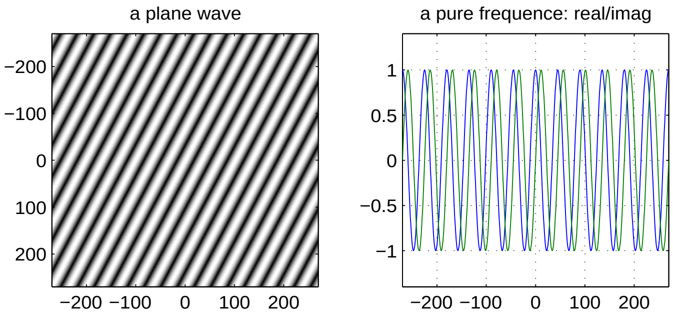
But where is there any mathematics? Isn't it just technology, and computer science??

In fact there is a lot built into all these devices, and most of the technical solutions have interesting mathematical background.

The answer to this is FFT2

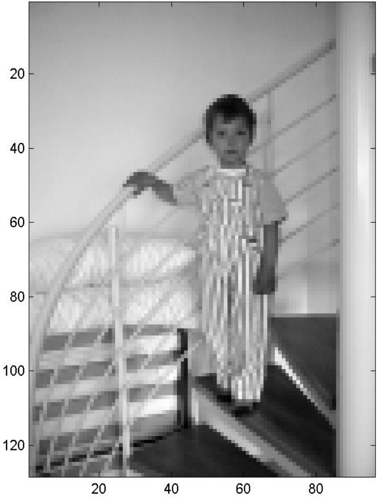
Decomposition of images into plane waves!
It can be obtained by an iterated one-dimensional FFT, by applying it to all the rows and then to all the columns of the resulting matrix. So for an $n \times n$-image one needs $2n$ FFTs of length n. The *order* in which these operations are performed does not matter. In fact, one can show that the operation can be written as matrix multiplication from the left (column operation) combined with matrix mutliplication from the right (row operations). Since matrix multiplication is known to be *associative* the order of these operation does not matter.

2D-Gabor Transform: Plane Waves



favorite test-image

A rough version of this image

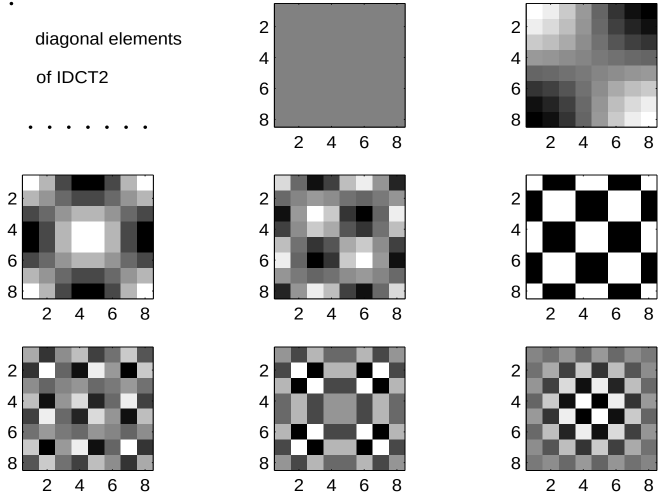


A few building blocks for JPEG-compression

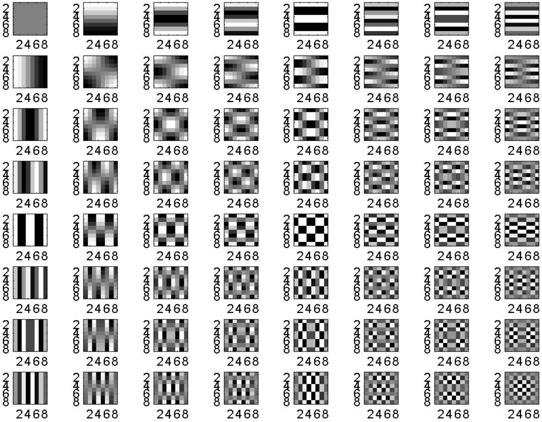
diagonal elements

of IDCT2

. . . .



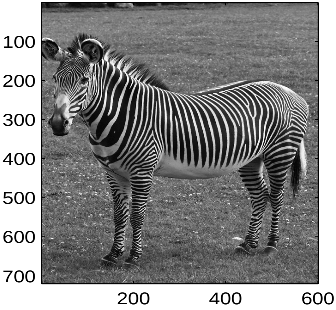
A few building blocks for JPEG-compression



Gabor Analysis for Image Processing

There is a contribution to the "Handbook of Image Processing" (which just appeared), based on the Master Thesis of Stephan Paukner, now working in/for industry.
According to $2D$-Gabor Analysis one decomposes the image into local plane wave patches.

Image Compression using Gabor: Testimage



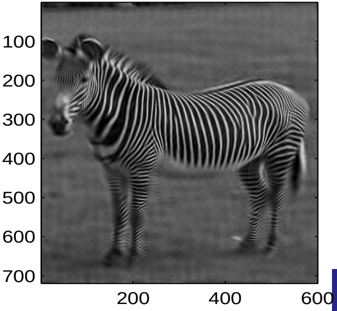
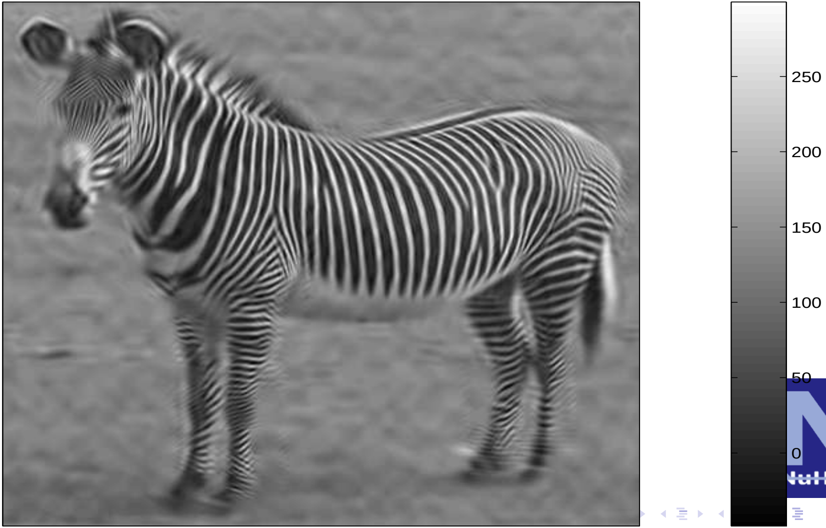
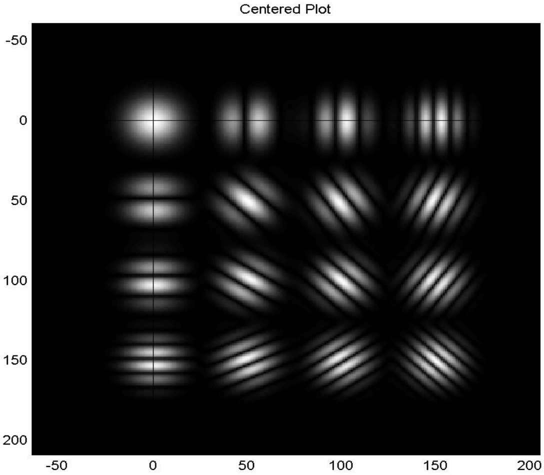


Image Compression: data reduction by thresholding



Building blocks of 2D Gabor decomposition



Removal of Fringe Patterns (ESO)