From the Hydrogen Atom to the mp3 files

DOCTORAL CLASS on MODERN MATHEMATICS FOR CONTEMPORARY SOUND MODELING

Pietro Polotti

DIPARTIMENTO DI INFORMATICA DELLA UNIVERSITA' DI VERONA
Vision, Image Processing, and Sound Laboratory

From the Hydrogen Atom to the mp3 files

Quantum Mechanics

Wave Fields

Orthogonal set of functions

Digital (Discrete) World

Perception

Pietro Polotti

G. Essl and S. Zambon

- “PDE we use are quite related to Quantum Mechanics” (Essl)
- **Bessel** functions: Membrane oscillation.
- ANALOGY

  OPTOMECHANICAL OPTICS MECHANICS

- Also... From Astronomy and Acoustics (**Thompson**) to Quantum Mechanics and DSP
- **Legendre** functions to decompose Sphere modes
- **Laguerre** Transform in order to obtain Frequency warping

Adrien-Marie Legendre
(1752-1833, Paris)

- ***Astronomer and mathematician***

- He published about celestial mechanics with papers such as "Recherches sur la figure des planètes" (1784), which contains the Legendre polynomials

- An *elliptic function* is an analytic function from **C** to **C** which is doubly periodic. That is, for two independent values of the complex number w, the functions $f(z)$ and $f(w + z)$ are the same.
  It can also be regarded as the inverse function to certain integrals (called *elliptic integrals*) of the form,

$$\int \frac{dz}{\sqrt{R(z)}}$$

  where R is a polynomial of degree 3 or 4.

Edmond Nicolas Laguerre

(1834, 1886 Bar-le-Duc, France)

•*mathematician*

- **Laguerre polynomials which are solutions of the Laguerre differential equations**
- **Laguerre Transform**

Friedrich Wilhelm Bessel

(1784 - 1846) German mathematician and astronomer

- ***Confined waves and modes of oscillation***

- ***Linear*** $\omega = kc = \frac{n\pi c}{L}$ $kL = n\pi$ L string length

And thus music….

- ***Bidimensional*** $\frac{1}{\lambda^2} = \frac{n^2}{4a^2} + \frac{m^2}{4b^2}$ a, b side length of a rectangle

Friedrich Wilhelm Bessel

- *Capacitor in an oscillating circuit.*

  E denotes the electric field between the two plates as a function of the distance from the centre of the plate r.

  By iterative approximations one obtains:

$$E = E_0 e^{j\omega t}\left[1 - \frac{1}{(1!)^2}\left(\frac{\omega r}{2c}\right)^2 + \frac{1}{(2!)^2}\left(\frac{\omega r}{2c}\right)^4 - \frac{1}{(3!)^2}\left(\frac{\omega r}{2c}\right)^6 + \ldots\right] = E_0 e^{j\omega t} J_0\left(\frac{\omega r}{c}\right)$$

- *At high frequencies → Resonant cavity*

J.J. Thompson

(1856 - 1940)

- Also Thompson: from Astronomy and Acoustics to Quantum Mechanics
- In 1897 he discovered the first subatomic particle, a component of all atoms, the **electron**.

Electrons play

- Acoustic waves
- Stationary wave equation
- Stationary oscillations modes (Bessel functions in the case of circular surfaces)

- Probability waves
- Hamiltonian equation ot the stationary states
- Atomic orbitals

The electron “hits” (interacts with) the atomic “volume” (atomic force field) “generating” probability waves, representable as the combination of “modes” (eigenfunctions) of the Energy-Hamiltonian Equation.

From the Hydrogen Atom to the mp3 files

An introduction to 2^{nd} degree differential equations

Pietro Polotti

The crisis of the Classical Physics

- Radiations present particle-like behavior (Photoelectric effect)

- Particles present wave-like behavior (Diffraction)

- Nature is essentially discontinuous (Energy levels of the atomic orbital)

The crisis of the Classical Physics

Classical Mechanics

- Material point
- Trajectory:
  Least Action Principle
- Potential of the force field
  $V(x,y,z)$
- Energy W
- Point velocity $v(x,y,z,W)$

Geometrical Optics

- Wave packet
- Light ray**:**
  Fermat's principle
- Refraction index:
  $\rho=\rho(x,y,z)$
- Frequency ν
- Group velocity $v_g(x,y,z)$

The crisis of the Classical Physics

Classical Mechanics

Least Action Principle

$$S = \int_{path} L dt = \int_{path} (T - V) dt$$

The **Principle of Least Action** demands that the action *S* be a minimum for the path taken by the particle.

Geometrical Optics

Fermat's principle

The path of a ray of light between two points is the path that: minimizes the travel time

Note: **Hamilton** did research in Optics. He tried to see if the Fermat's principle and other aspects of Optics could be "shifted" to mechanics and the result was the Hamiltonian formalism.

Schrödinger Equation

- Problem: Which kind of relationship between energy/potential from one side and mean-frequency/refraction-law from the other is necessary, in order to make the motion of the material point coinciding with the motion of the wave-group?

$$\nabla^2\psi - \frac{2mV}{\hbar^2}\psi + j\frac{2m}{\hbar}\frac{\partial\psi}{\partial t} = 0$$

- **Material** point of mass *m* in a force field with potential energy *V*.

Schrödinger Equation

- By means of a separation of variables: $\psi(x,y,z,t)=u(x,y,z)\varphi(t)$

$$\begin{cases} \left[\nabla^2-\dfrac{2mV}{\hbar^2}\right]u=Wu & \text{Stationary state equation} \\ \dfrac{\partial\varphi}{\partial t}=-\dfrac{j}{\hbar}W\varphi & \text{Temporal evolution} \end{cases}$$

- **Eigenvalue equation** (W_n)
- **Solution** $\psi(x,y,z,t)=\sum_n c_n u_n(x,y,z)e^{\frac{j}{\hbar}W_n t}(t)$
- ***u_n* eigenvector** (stationary waves).

Schrödinger Equation

- One can show that

$$\int u_n^* u_m \, dxdydz = \delta(m\text{-}n)$$

- Thus from the condition $\int |\psi|^2 \, dxdydz = 1$

  it follows:

$$\sum_n |c_n|^2 = 1$$

- $|c_n|^2$ is interpreted as the **probability** of finding the value $\boldsymbol{W_n}$ when measuring the energy of the system.

Schrödinger Equation: The Operator Mechanics

- Let's consider the Hamiltonian: $H = E \Rightarrow \frac{\sum p_i^2}{2m} + V(x,y,z) = E$

- Let's "postulate" the correspondences

$$E \rightarrow j\hbar \frac{\partial}{\partial t}$$

$$p_i \rightarrow -j\hbar \frac{\partial}{\partial x_i}$$

- We obtain an equality between two differential operators

$$-\frac{\hbar^2}{2m}\nabla^2 + V(x,y,z) = j\hbar \frac{\partial}{\partial t}$$

- Multiplying by $-2m/\hbar^2$ we obtain the Schroedinger Equation

Building the Schroedinger Equation from the classical Hamiltonian

$$H = H_c\left(q_j, p_j, t\right) \quad \rightarrow \quad H = H_Q\left(q_j, -i\hbar\frac{\partial}{\partial q_j}, t\right)$$

$$E \quad \rightarrow \quad i\hbar\frac{\partial}{\partial t}$$

$$H_Q\Psi\left(q_i, t\right) = i\hbar\frac{\partial\Psi\left(q_i, t\right)}{\partial t}$$

- The Quantum Hamiltonian is an equation between **operators** that **act** on L^2 functions of the q_i ‘s and the time t.
- The same procedure can be applied for any classical physical observable

General Structure of the Quantum Mechanics

1. Given a **physical observable** *G* the only possible results of a measure of *G* are the eigenvalues *g* of the equation

$$G\varphi_i = g_i\varphi_i$$

2. System preparation: If one measures *G* at t=0 and finds g_l then, immediately after the measure the wave function of the system is:

$$\psi(\vec{x},0) = \varphi_l(\vec{x})$$

3. The system evolution is given by the solution of the equation

$$H\psi(\vec{x},t) = i\hbar\frac{\partial\psi(\vec{x},t)}{\partial t}$$

General Structure of the Quantum Mechanics

4. Future measure forecast are PROBABILISTIC. When considering the generic physical observable

$$\Omega\xi_i = \omega_i \xi_i$$

It is a property of the ξ_i that for any ψ_i it is possible to write:

$$\psi(\vec{x},t) = \sum_i c_i(t)\xi_i(\vec{x},t)$$

Then the $\left|c_n(t)\right|^2$ are the probabilities that in a measure of Ω at time t one can obtain the result ω_n.

Uncertainty Principle (Werner Heisenberg, 1925)

- Classical Physics admit, at least in principle, the possibility of measuring simultaneously any couple of physical variable.
- As Achilles and the turtle, the more precise the measure instruments the more precise the measure, with non infinitesimal limitation
- **BUT Reality is different**

$$\Delta q_i \Delta p_i \geq h$$

Linear Harmonic Oscillator

- Steady state equation (classical Hamiltonian)

$$\frac{d^2u}{dx^2}+\frac{2m}{\hbar^2}\left(E-2\pi m\upsilon_0^2x^2\right)u=0$$

$x=$ displacement,

$u=$ energy eigenfunctions

$V(x)=2\pi m\upsilon_0^2x^2$

chracteristic frequency $\upsilon_0=\frac{1}{2\pi}\sqrt{\frac{x}{m}}$

- Confluent Hypergeometric.

- Considering the asymptotic behavior of the solution, one finds that the solution include as a factor Hermit polynomial and from this follows that the energy must be:

$$E_n=h\upsilon_0\left(n+\frac{1}{2}\right)$$

Towards the Hydrogen Atom

- Schroedinger equation for steady states

$$-\frac{\hbar^2}{2m}\nabla^2 u - [E - V(r)]u = 0$$

- Central force field

$$V(\vec{r}) = V(r)$$

- Variable Separation: Radial and Angular components

$$u(r,\theta,\varphi) = R(r)\Theta(\theta)\Phi(\varphi)$$

Towards the Hydrogen Atom

- Variable Separation: Radial and Angular component equations:

$$\frac{1}{R}\frac{d}{dr}\left(r^2\frac{dR}{dr}\right)-\frac{2m}{\hbar^2}\left[E-V(r)\right]r^2=C$$

$$\frac{1}{\sin(\theta)}\frac{d}{d\theta}\left(\sin(\theta)\frac{d\Theta}{d\theta}\right)-\frac{2m}{\hbar^2}\left[C-\frac{V(r)}{\sin(\theta)^2}\right]\Theta=0$$

$$\frac{d^2\Phi}{d\varphi^2}+\lambda\Phi=0$$

Angular Momentum

$$\frac{1}{\sin(\theta)}\frac{d}{d\theta}\left(\sin(\theta)\frac{d\Theta}{d\theta}\right)-\frac{2m}{\hbar^2}\left[C-\frac{V(r)}{\sin(\theta)^2}\right]\Theta=0$$

It is an **hypergeometric equation** with fuchsian points: $z = -1, 1, \infty$

The solutions are:

$$\Theta_{m,l}(\theta)=AP_l^m(\theta)+BQ_l^m(\theta)$$

With m and l parameters of the hypergeometric and P_l and Q_l are **Legendre functions** of the 1st and 2nd kind, respectively.

Note: By studying the asymptotic behavior one finds the eigenvalues: c= $l(l+1)$

Angular Solution

Legendre polynomials

$$P_l^m(z) = (-)^m \frac{1}{2^l l!}\left(1 - z^2\right)^{m/2} \frac{d^{l+m}}{dz^{l+m}}\left(z^2 - 1\right)$$

Angular Solution

The global angular solutions are:

$$Y_{l,m}(\theta,\varphi) = (-)^{|m|+m/2}\sqrt{\frac{2l+1}{4\pi}\frac{(l-m)!}{(l+m)!}}(\sin\theta)^{|m|}\frac{d^{|m|}P_l(\cos\theta)}{d(\cos\theta)^{|m|}}e^{jm\varphi}$$

These are the **Spherical Functions** of order *l* and grade *m* and they satisfy the orthogonality condition:

.

$$\int_0^{2\pi} d\varphi \int_0^{\pi} \sin\theta d\theta \; Y_{l,m}(\theta,\varphi) Y^*_{l',m'}(\theta,\varphi) = \delta_{l,l'}\delta_{m,m'}$$

Radial Solution

The equation is now:

$$\frac{1}{R}\frac{d}{dr}\left(r^2\frac{dR}{dr}\right)-\frac{2m}{\hbar^2}\left[E-\frac{Ze^2}{r}\right]r^2=l(l+1)$$

where $l(l+1)\hbar^2$ is the angular momentum

Z nucleus charge

It can be reduced to an **hypergeometric equation** and by studying the behavior at the infinity, the solution beomes:

$$R_{n.l}(r')=\left[\left(\frac{2Z}{na_0}\right)^3\frac{(n-l-1)!}{2n\{(n+l)!\}^3}\right]\mathrm{e}^{-r'/2}\,r'^l\,L_{n+l}^{2l(l+1)}(r')$$

$L_{n+l}^{2l(l+1)}(r')$ **Associated Laguerre functions**

Global Solution

$$u_{n,l,m}(r,\theta,\varphi) = Y_l^m(\theta,\varphi)R_{n,l}(r) =$$

$$= A\Phi(\varphi)P_l^{|m|}(\cos(\theta))x^l e^{-\frac{x}{2}} L_{l+n}^{2l+1}(r)$$

Orthonormality:

$$\int u_{n,l,m}(r,\theta,\varphi)u_{n',l',m'}(r,\theta,\varphi)dV = \delta_{n,n'}\delta_{l,l'}\delta_{m,m'}$$

Global Solution

The energy eigenvalues are:

$$E_n = -\frac{m}{2\hbar^2}\frac{Z^2 e^4}{n^2}$$

They depend only on n. This happens only for the spherical symmetry case (1 electron)

l=0	orbitali s
l=1	orbitali p
l=2	orbitali d
l=3	orbitali f
l=4	orbitali g

Bessel function for free particle

The radial component of the solution of the Schroedinger equation for a free particle

$$\nabla^2\psi(r,\theta,\varphi)+\frac{2m}{\hbar^2}E\psi(r,\theta,\varphi)=0$$

is,

$$R_l(r)=\frac{1}{\sqrt{Cr}}\left[AJ_{l+\frac{1}{2}}\left(r\frac{\sqrt{2mE}}{\hbar}\right)+BJ_{-l-\frac{1}{2}}\left(r\frac{\sqrt{2mE}}{\hbar}\right)\right]$$

where the $J_{l+1/2}(\rho)$ are the cylindrical Bessel functions

Bessel functions

- Bessel in the modes of a circular membrane
- Bessel in the FM
- Bessel in the Ambisonics

From the Hydrogen Atom to the mp3 files

An Electrodynamic digression: Resonant cavities

Pietro Polotti

Hilbert space

- **Linear (Vector) space** over a **numeric set.** The sum of two elements of the space and the product by a number of the set belong to the linear space.
- **Operators,** $\phi' = T\phi \in S$
- **Operator algebra** (N.B. in general NON-commutative, thus NON-COMMUTATIVE PHYSICS),
- **Linear operators**
- **Euclideian space** = linear space with a hermitian scalar product

Hilbert space

- Hermitian scalar product $$\left(\psi,\varphi\right)$$

  S x S → **C** a function with some properties:

  a $\left(\psi,a\varphi_1+b\varphi_2\right)=a\left(\psi,\varphi_1\right)+b\left(\psi,\varphi_2\right)$

  b $\left(\psi,\varphi\right)=\left(\varphi,\psi\right)^*$

  c $\left(\varphi,\varphi\right)\geq 0$ "=" iff $\varphi=\omega$

- **Norm** $$\left\|\psi\right\|=\left[\left(\psi,\psi\right)\right]^{1/2}$$

- **Distance** $$d_{\phi,\psi}=\left\|\phi-\psi\right\|$$

Hilbert space

Hilbert Space:= Linear **Euclideian space**, which is also **complete** and **separable**

Definition: *S* is complete if any Cauchy sequence converges to an element of S

Definition: *S* is separable if there exists a numerable set of elements everywhere dense in *S*.
Equivalent to: for any ε>0, there exists a sequence ϕ_n such that $\|\phi - \phi_n\| < \varepsilon$.

Orthonormal set $\{\phi_k\}$: $$(\phi_m, \phi_n) = \delta_{m,n}$$

Complete set $$(\phi, \phi_k) = 0 \quad \text{for every } k \quad \Rightarrow \quad \phi = \omega$$

$\{\phi\}$:= linear variety spanned by $\{\phi_k\}$.

Hilbert Spaces

Two operators T and T^+ are said to be **adjoin,** if they have the same domain and

$$\left(T\phi,\psi\right)=\left(\phi,T^+\psi\right)$$

Definition: if $T = T^+$, T is **Hermitian**.

Hilbert Space

- **Eigenvalue Equation**

$$H\phi_r = h_r\phi_r$$

- The values h_r for which the equation has solution form the **Discrete Spectrum of *H***.

- **Theorem** : The Discrete Spectrum of a **linear Hermitian transformation** is a set of points on the real axis empty, finite or infinite enumerable. **Eigenvectors** corresponding to different eigenvalues **are orthogonal**

(Furthermore the condition of Hermitianity implies real eigenvalues)

Quantum Mechanics Hilbert Spaces

***L*$^{(2)}$(∞)** is the Euclideian space, whose elements are the complex functions of real variables, for which the following holds:

$$\int \ldots \int \left| f\left(q_1, \ldots, q_k\right) \right|^2 < \infty$$

It is possible to make the space a vector (linear) space and an Euclideian space, by defining:

$$f + g = f(q_1, \ldots, q_k) + g(q_1, \ldots, q_k)$$

$$\alpha f = \alpha f(q_1, \ldots, q_k)$$

$$(f, g) = \int \ldots \int f^*(q_1, \ldots, q_k) g(q_1, \ldots, q_k) dq_1 \ldots dq_k$$

It is possible to show that the space is also complete and separable, thus it is a Hilbert space.

DSP Hilbert Spaces

$\mathbf{l^{(2)}(\infty)}$ is the Euclideian space, whose elements are the sequences $a=\{a_k\}$ of complex numbers, and the following holds:

$$\sum_k |a_k|^2 < \infty$$

It is possible to show that the space is also complete and separable, thus it is a Hilbert space.

Theorem: Any ∞–dimensional Hilbert space is isomorphic to $l^{(2)}(\infty)$.

$$(a,b) = \sum_k a_k b_k$$

Completeness

- Completeness conditions

$$\forall \phi \in H \qquad \varphi = \sum_k (\varphi_k, \varphi)\varphi_k + \int d\lambda (\varphi_\lambda, \varphi)\varphi_\lambda$$

$$\forall \phi, \psi \in H \qquad (\psi, \varphi) = \sum_k (\psi, \varphi_k)(\varphi_k, \varphi) + \int d\lambda (\psi, \varphi_\lambda)(\varphi_\lambda, \varphi)$$

- Complete Hermitian operators are called Hypermaximal operators

Dirac Notation

$$|\psi\rangle = a|\alpha\rangle + b|\beta\rangle$$

Antilinear correspondence

$$\langle\psi| = a^*\langle\alpha| + b^*\langle\beta|$$

Scalar product

$$\langle\psi|\phi\rangle$$

Dirac Notation

Eigenvalues equation

$$A|\alpha\rangle = a|\alpha\rangle$$

Orthogonality condition

$$\langle\alpha_r|\alpha_s\rangle = \delta_{r,s}$$

Operator:
Projector onto the span of the α_k

$$P = \sum_k |\alpha_k\rangle\langle\alpha_k|$$

Completeness condition

$$\sum_k |\alpha_k\rangle\langle\alpha_k| = 1$$

Quantum Mechanics General Postulates

- Physical variable → Hypermaximal operator.
- The only possible values of a measure are the eigenvalues.
- Immediately after a measure the system falls into the state corresponding to an eigenvector.
- The system evolves according to the Schroedinger equation.
- Future measures are predictable in terms of the probabilities $|c_k|^2$, where the c_k are the eigenvalues of a certain physical variable.

Quantum Mechanics General Postulates

- We say that the measure projects the system onto one of the possible stationary state of the physical variable.

- Quantum Mechanics is the Physics of the Projection Operators, i.e. of the orthogonal expansion!

General Structure of the Quantum Mechanics addendum

- We say that the measure projects the system onto one of the possible stationary state of the physical variable.

- Quantum Mechanics is the Physics of the **Projection Operators**, i.e. of the **orthogonal expansion**!

Electrons play

Acoustics / Electrodynamics / Quantum Mechanics

Applets

http://www.falstad.com/mathphysics.html

The Point

- DUE TO QUANTUM MECHANICS AND THE CENTRALITY OF THE STEADY STATE EQUATION, THE SOLUTION OF **EIGENVALUE EQUATIONS** AND THE EXPANSION OF THEIR SOLUTIONS ON THE **EIGENVECTOR BASIS** BECOME A CENTRAL PROBLEM.

- THIS FORMALISM FIND EQUIVALENCIES WITH AND CAN BE TRASPOSED TO MANY OTHER FIELDS OF PHYSICS AND OTHER SCIENTIFIC CONTEXTS.

The Point: QM "reminiscences" in DSP

Atomic Orbitals ↔ Subband-Coding

Position/Momentum Uncertainty

↕

Time/Frequency Uncertainty

l^2 : the "DSP Hilbert Space"

- Expansion $$x[n] = \sum_k \langle \varphi_k[n], x[n] \rangle \varphi_k[n] = \sum_k X[k]\, \varphi_k[n]$$
- **Transform** coefficients $$X[k] = \langle \varphi_k[l], x[n] \rangle = \sum_l \varphi_k{}^*[l]\, x[l]$$
- Orthonormality $$\langle \varphi_k[n], \varphi_l[n] \rangle = \delta[k-l]$$
- Energy conservation (Parseval) $$\left\| x[n] \right\|^2 = \left\| X[k] \right\|^2$$

l^2 : the "DSP Hilbert Space"

Theorem: Given an orthonormal set $\{x_1, x_2,\ldots\}$ in a Hilbert space H, the following are equivalent:

- The set of vectors $\{x_1, x_2,\ldots\}$ forms an orthonormal basis for H.
- If $<x_i,y>=0$ for $i=1,2,\ldots$ then $y=0$
- The span of $\{x_1, x_2,\ldots\}$ is dense in H, that is, every vector in H is a limit of a sequence of vectors in the span of $\{x_1, x_2,\ldots\}$
- Parseval's equality: For every y in H: $$\|y\|^2 = \sum_i \left|\langle x_i, y\rangle\right|^2$$

l^2 : the "DSP Hilbert Space"

Given a Hilbert space **H** and a closed subspace V_0, such that:

$$\mathbf{H}= V_0 \oplus W_0,$$

where W_0 is said to be the orthogonal complement of V_0 in **H,** then,if $u \in \mathbf{H}$

$$u=v+w \ ,$$

where $v \in V_0$ and $w \in W_0$

Definition: An operator P is called a **projection operator** onto V_0 if

$$P(u)= P(v+w)=v$$

An operator is a projection operator iff it is
Idempotent: $P^2=P$
Self adjoint: $P^*=P$

l^2 : the "DSP Hilbert Space"

- Space decomposition $V_0 \perp W_0 \Rightarrow V_0 \oplus W_0 = V_{-1}$
- Design of **orthogonal FIR filter banks**

Sampling Theorem

The set $\mathbf{sinc}(t\text{-}l)$, $l \in \mathbf{Z}$ forms an othonormal basis for the set of functions $f(t)$ bandlimited to $(-\pi,\pi)$, where

$$\mathrm{sinc}(t) = \frac{\sin(\pi t)}{\pi t}$$

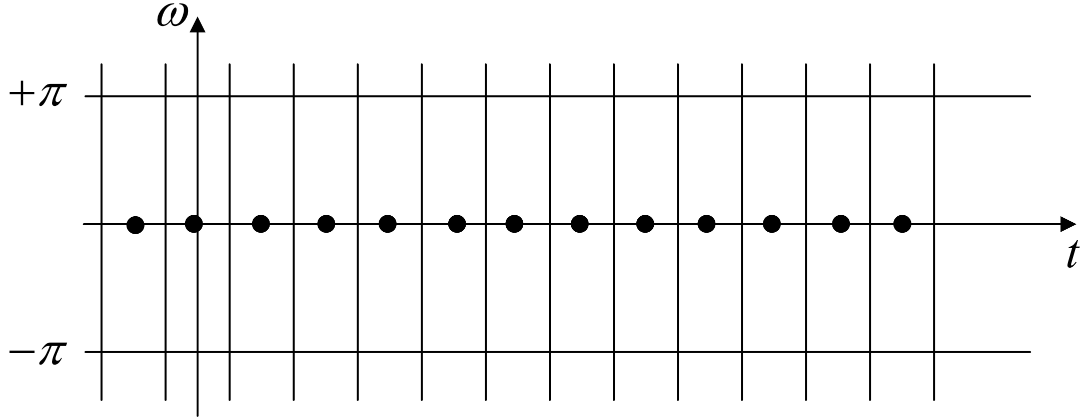
- Even the most simple coding technique (PCM) is done by means of an expansion onto an othonormal set of functions

Sampling Representations

In fact,

$$f(n) = \int_{-\infty}^{+\infty} f(t)\,\text{sinc}(t-n)\,dt = \langle f(\bullet), \text{sinc}(\bullet - n) \rangle$$

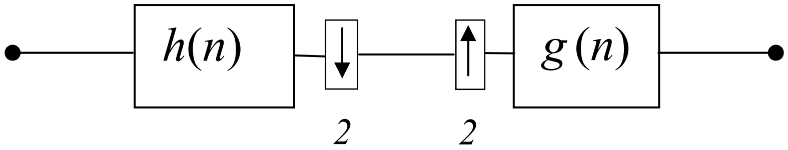
On a time-frequency plane we have



This is a single-band time-frequency representation

Multirate DSP
Orthogonal Filter Banks

- Orthogonal projection onto the subspace **V**



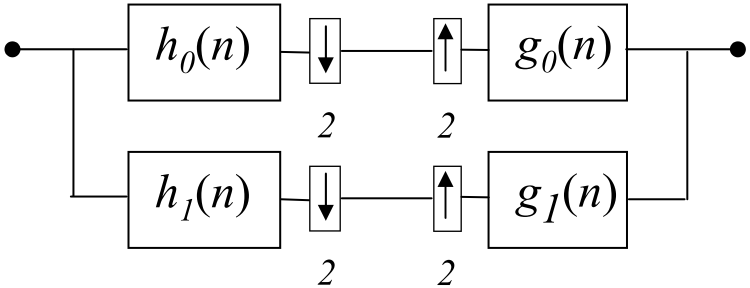
where $\langle g[n] \| g[n-2k] \rangle = \delta[k] \qquad h[n] = g[-n]$

and $\mathrm{V} = \mathrm{span}\{g[n-2k]\}_{k \in Z}$

Multirate DSP
Orthogonal Filter Banks

- Orthogonal projection onto the subspace **V** and **W**

H= V⊕W



H= V⊕W

$$\langle g_i[n], g_l[n] \rangle = \delta[i-l]$$

$$h_i[n] = g_i[-n]$$

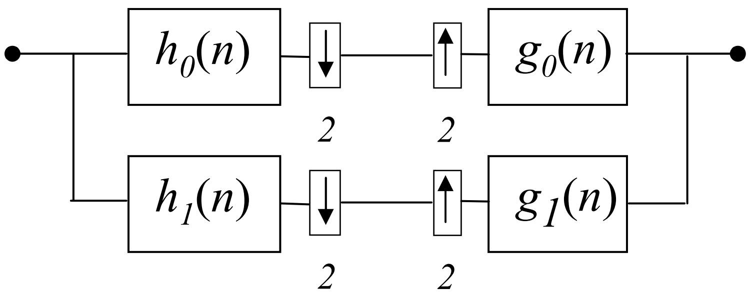
$$\langle g_i[n] \| g_i[n-2k] \rangle = \delta[k]$$

$$\mathrm{V} = \mathrm{span}\{g_0[n-2k]\}_{k\in Z}$$

$$\mathrm{W} = \mathrm{span}\{g_1[n-2k]\}_{k\in Z}$$

Multirate DSP
Orthogonal Filter Banks

In the *z*-domain

H= V⊕W  **H**= V⊕W

Perfect Reconstruction $G_0(\omega)H_0(\omega)+G_1(\omega)H_1(\omega)=I$

Orthogonal system $H_0^*(z)H_0(z)+H_1^*(z)H_1(z)=I$

with $G_0(z)=H_1^*(z)$

Orthogonal Filter Design in l^2

- Aliasing cancelation

$$Y(z) = \frac{1}{2}\left[G_0(z), G_1(z)\right]\begin{bmatrix} H_0(z) & H_0(-z) \\ H_1(z) & H_1(-z) \end{bmatrix}\begin{bmatrix} X(z) \\ X(-z) \end{bmatrix}$$

In order to cancel the contribution of $X(-z)$, for example:

$$\left[G_0(z), G_1(z)\right] = \left[H_1(-z), \quad -H_0(-z)\right]$$

Es. QMF (Esteban Galant 1976)

Orthogonal Filter Design in l^2

In general in order to have Perfect Reconstruction (also biorthogonal):

$$\left[H_0(z)G_0(z)\right]+\left[H_1(z)G_1(z)\right]=2$$

$$\left[G_0(z)H_0(-z)\right]+\left[G_1(z)H_1(-z)\right]=0$$

Orthogonal Filter Design in l^2

- Haar the simplest wavelet:
  - orthogonal
  - linear phase
  - FIR

Haar

- Haar the simplest wavelet:

- Scaling function $$\varphi(t) = \begin{cases} 1 & 0 \le t < 1 \\ 0 & \text{else} \end{cases}$$

- Help the construction of the wavelet since

$$\psi(t) = \varphi(2t) - \varphi(2t-1)$$

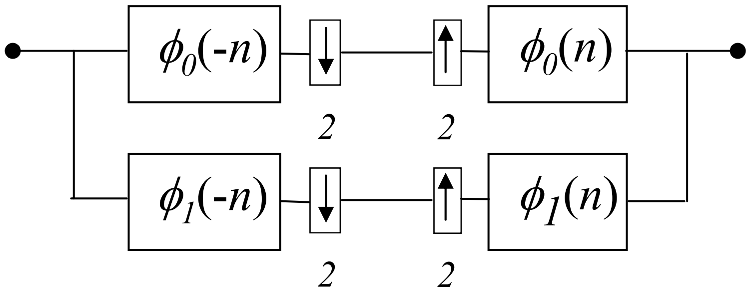
- And satisfies a two-scale equation:

$$\varphi(t) = \varphi(2t) + \varphi(2t-1)$$

Discrete Haar

$$\varphi_{2k}[n] = \begin{cases} \frac{1}{\sqrt{2}} & \text{for } n = 2k, 2k+1 \\ 0 & \text{otherwise} \end{cases}$$

$$\varphi_{2k+1}[n] = \begin{cases} \frac{1}{\sqrt{2}} & \text{for } n = 2k \\ -\frac{1}{\sqrt{2}} & \text{for } n = 2k+1 \\ 0 & \text{otherwise} \end{cases}$$



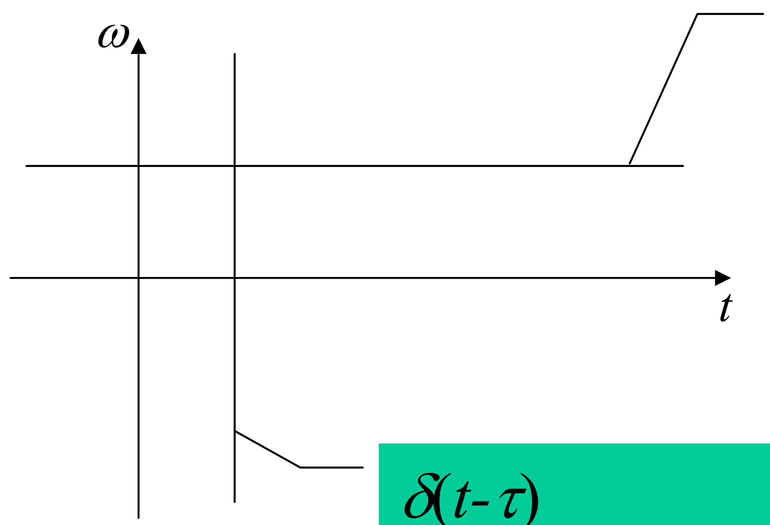
Pietro Polotti

Time vs. Frequency Representations

For the frequency representation the orthogonal set elements are complex sinusoids with

$$e^{j\Omega t} \xrightarrow{FT} 2\pi\delta(\omega - \Omega)$$

On a time-frequency plane we have:



$2\pi\delta(\omega\text{-}\Omega)$
infinitely good in frequency,
infinitely bad in time

$\delta(t\text{-}\tau)$
infinitely good in time,
infinitely bad in frequency

Pietro Polotti

Uncertainty Principle (Heisenberg, 1925)

Time and frequency are not independent variables

One cannot achieve infinite resolution simultaneously in time and frequency:

$$\Delta t \times \Delta \omega \geq \tfrac{1}{2}$$

Where (second central moments):

$$(\Delta t)^2 = \frac{\int_{-\infty}^{+\infty} (t - t_c)^2 \left| f(t) \right|^2 dt}{\int_{-\infty}^{+\infty} \left| f(t) \right|^2 dt} \qquad (\Delta \omega)^2 = \frac{\int_{-\infty}^{+\infty} (\omega - \omega_c)^2 \left| F(\omega) \right|^2 d\omega}{\int_{-\infty}^{+\infty} \left| F(\omega) \right|^2 d\omega}$$

$$t_c = \frac{\int_{-\infty}^{+\infty} t \left| f(t) \right|^2 dt}{\int_{-\infty}^{+\infty} \left| f(t) \right|^2 dt}$$

Center time

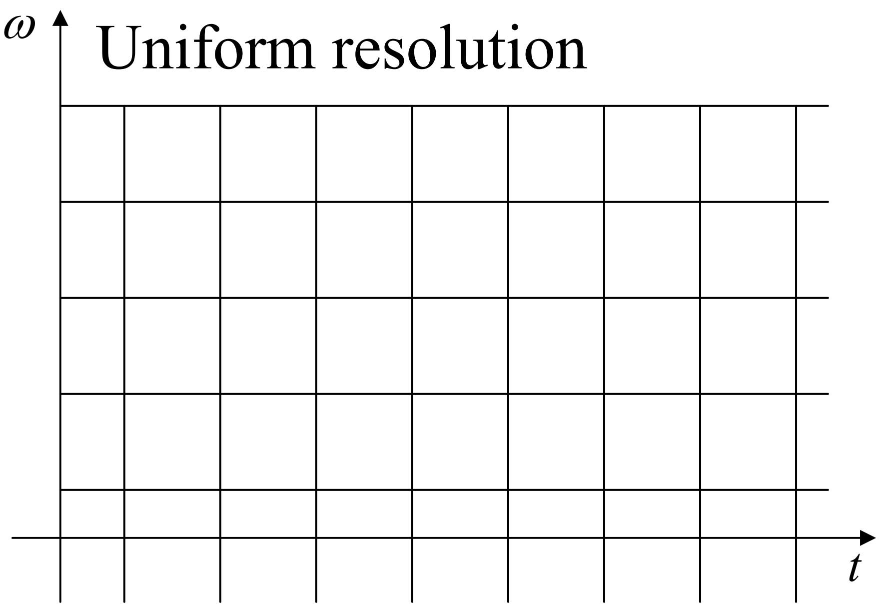
$$\omega_c = \frac{\int_{-\infty}^{+\infty} \omega \left| F(\omega) \right|^2 d\omega}{\int_{-\infty}^{+\infty} \left| F(\omega) \right|^2 d\omega}$$

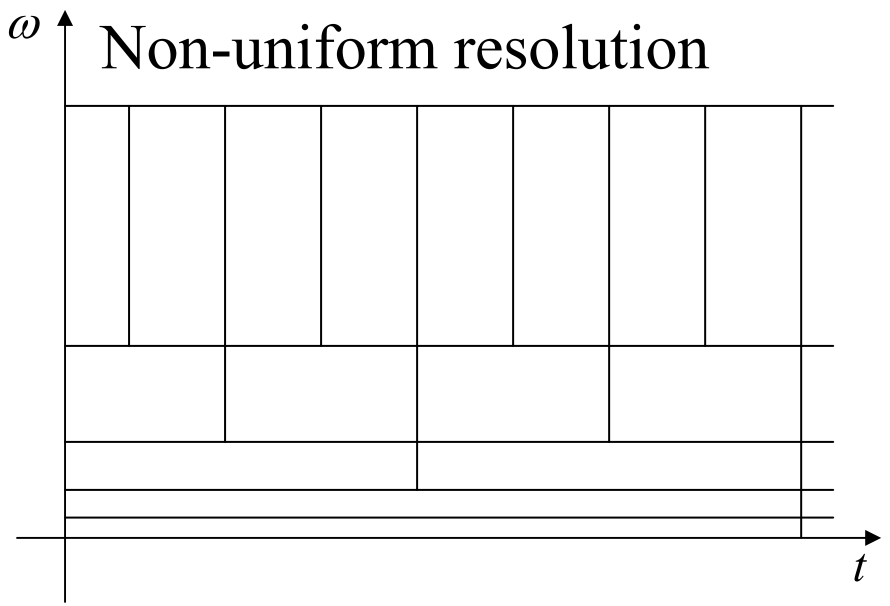
Center frequency

Note: Gaussian shaped signals have minimum uncertainty product

Time-Frequency Representations

By properly scaling, translating and modulating the sinc basis one can construct an arbitrary tiling of the time-frequency plane





However, the sinc function has good frequency localization but bad time localization (uncertainty product is infinite)

Pietro Polotti

Short-Time Fourier Transform

Uniform (Gabor's) expansion is realized by means of the STFT

For discrete-time signals

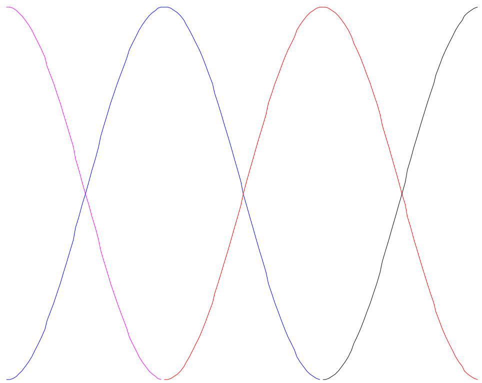
$$F(n,m)=\sum_{k=-\infty}^{+\infty} f(k)w(k-nM)e^{-j\frac{2\pi}{N}km} \qquad \text{(analysis)}$$

$$f(n)=\frac{1}{N}\sum_{m=0}^{N-1} e^{j\frac{2\pi}{N}nm}\sum_{r=-\infty}^{+\infty} F(r,m)\tilde{w}(n-rM) \qquad \text{(synthesis)}$$

with $w(n)$ finite-length N window and $\tilde{w}(n)$ satisfying

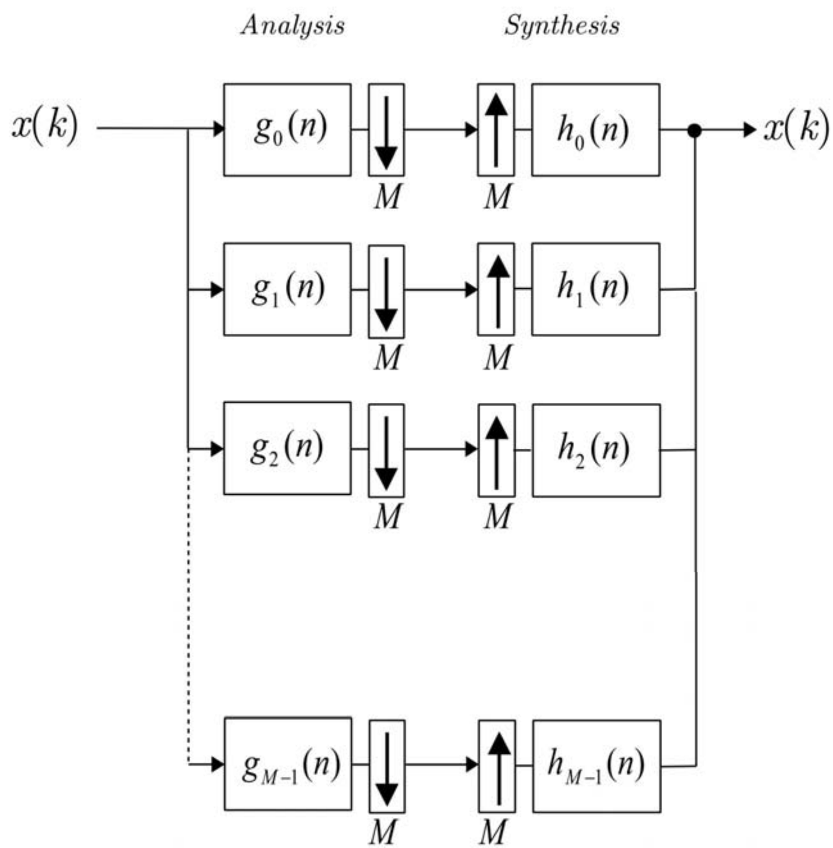
$$\sum_{r=-\infty}^{+\infty} w(n-rM)\tilde{w}(n-rM)=1 \quad \forall n\in\mathbf{Z}$$

This is the overlap-add method for the product of the windows $w(n)\tilde{w}(n)$



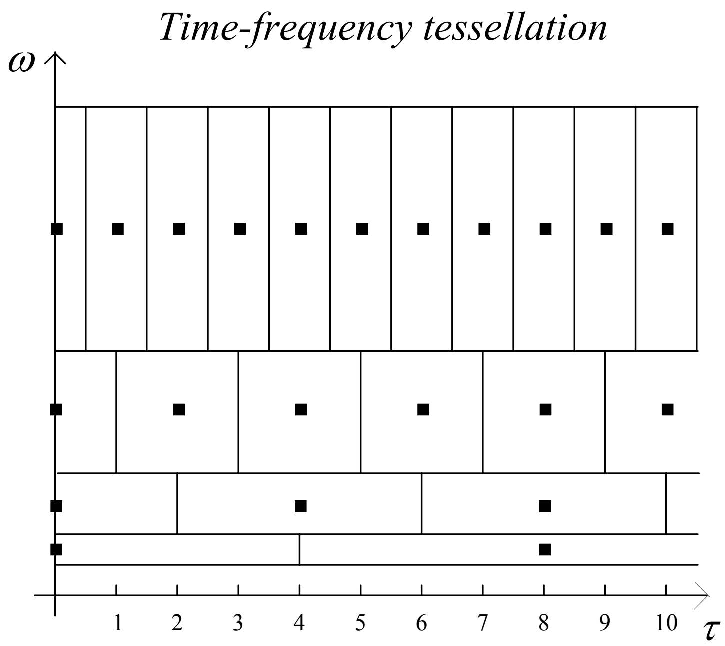
Computation of STFT

Computation may be performed by means of windowed FFTs or by filter bank structure:



The same filter bank structure is shared by other related transforms (e.g., MDCT)

Dyadic Wavelets Time-Frequency Tessellation

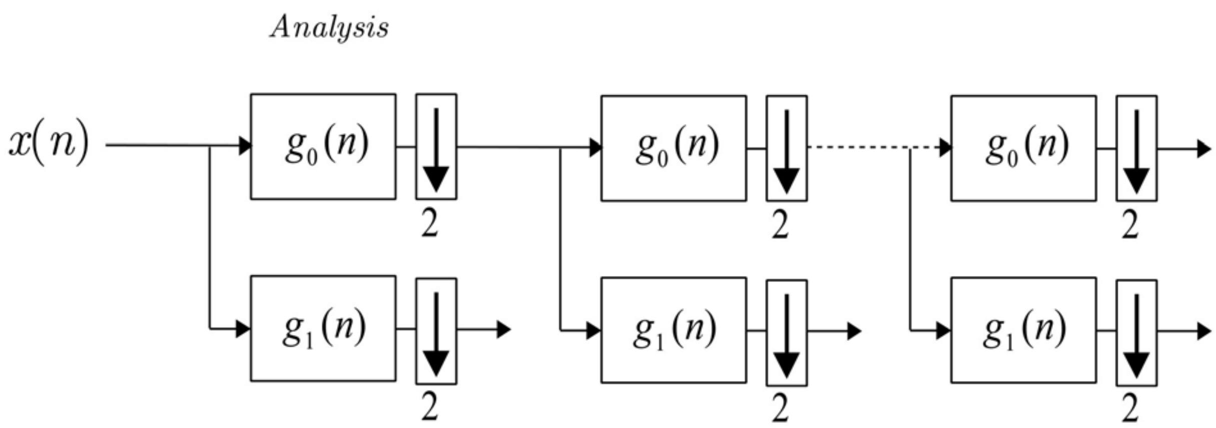


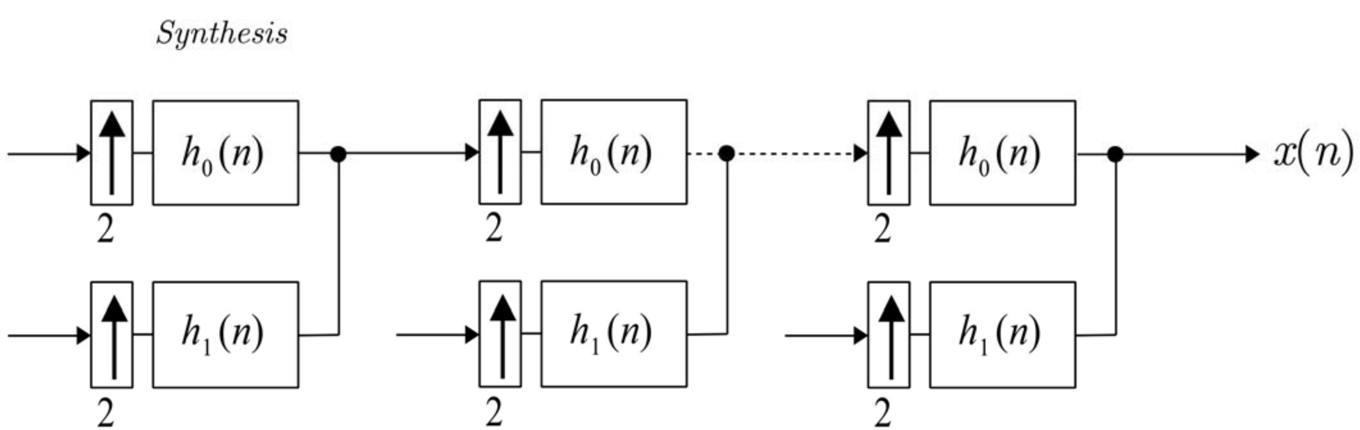
Octave band frequency resolution

Frequency resolution is good at low frequencies and poorer at high frequencies (constant Q)

Time resolution is good at high frequencies and poorer at lower frequencies

WT Computation





Discrete Cosine Transform (DCT)

PR uniform-band filter banks

DCT I type

$$C[k] = \sum_{n=0}^{N-1} f[n] \cos\left(\frac{\pi}{N} k \left(n + \frac{1}{2}\right)\right)$$

$$f[n] = \frac{C[0]}{N} + \frac{2}{N} \sum_{n=1}^{N-1} C[k] \cos\left(\frac{\pi}{N} k \left(n + \frac{1}{2}\right)\right) \quad \text{for } n = 0,.., N-1$$

DCT IV type

$$C[k] = \sum_{n=0}^{N-1} f[n] \cos\left(\frac{\pi}{N} \left(k + \frac{1}{2}\right)\left(n + \frac{1}{2}\right)\right)$$

$$f[n] = \frac{2}{N} \sum_{n=0}^{N-1} C[k] \cos\left(\frac{\pi}{N} \left(k + \frac{1}{2}\right)\left(n + \frac{1}{2}\right)\right) \quad \text{for } n = 0,.., N-1$$

Modified Discrete Cosine Transform (MDCT)

By means of DCT-I or DCT-IV one can build PR and orthogonal uniform-band cosine modulated filter banks

MDCT from DCT IV type

$$h_k[n] = W[n]\sqrt{\frac{2}{M}}\cos\left(\frac{\pi}{M}\left(k+\frac{1}{2}\right)\left(n+\frac{M+1}{2}\right)\right) \quad \text{length } 2M$$

where $\begin{cases} W^2[n] + W^2[n+M] = 1 \\ W[2M-1-n] = W[n] \end{cases}$ *example* $W[n] = \sin\frac{\pi}{2M}\left(n+\frac{1}{2}\right)$

$$\omega_k = \frac{(2k+1)\pi}{2M} \quad \text{center frequencies of the filters}$$

MR DSP
MDCT completeness and orthogonality conditions

- Completeness

$$\frac{1}{P}\sum_{r=-\infty}^{\infty}\sum_{q=0}^{P-1}W(l-rP)W(l'-rP)\cos\left(\frac{2q+1}{4P}\left(2(l-rP)-P+1\right)\pi\right)\cos\left(\frac{2q+1}{4P}\left(2(l'-rP)-P+1\right)\pi\right)=\delta_{l,l'}$$

- Orthogonality

$$\frac{1}{P}\sum_{l=-\infty}^{\infty}W(l-rP)W(l-r'P)\cos\left(\frac{2q+1}{4P}\left(2(l-rP)-P+1\right)\pi\right)\cos\left(\frac{2q'+1}{4P}\left(2(l-r'P)-P+1\right)\pi\right)=\delta_{q,q'}\delta_{r,r'}$$

Compression Methods

Lossy compression:

1) **MDCT subband decomposition**

2) **STFT** in order to estimate the psychoacoustic model parameters

3) **Dynamic bit allocation** according to the psychoacustic param. (signal-to-mask ratio SMR)

4) **Quantization** and **entropy coding** of **subband signals**

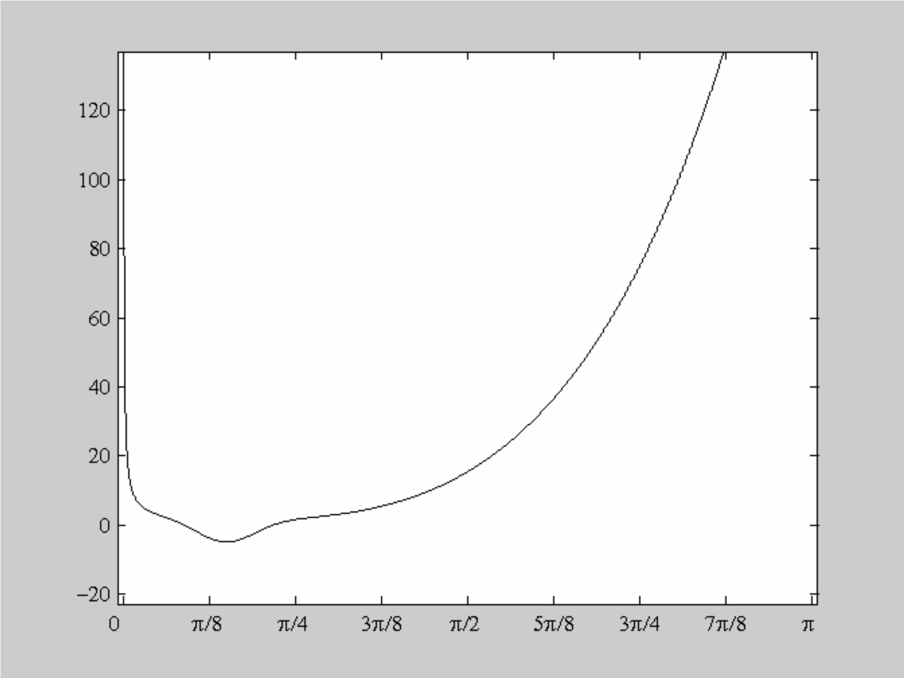
5) **Multiplex** and frame packing

MPEG1-layer 3 (*.mp3)

- MPEG (Motion Picture Experts Group): gruppo di scienziati che studiano codifiche standard per la compressione video e audio
- MPEG1 (layer 1, 2, 3)
- L'orecchio non è in grado di percepire frequenze "deboli" adiacenti a frequenze "forti", che mascherano le prime.
- Le informazioni inerenti le frequenze più deboli vengono eliminate dall'MPEG durante la fase di compressione.

1) Absolute Threshold

Non uniform hearing capabilities along the frequency range



$$AbsTh(f) = 3.64 \cdot (f/1000)^{-0.8} - 6.5 \cdot e^{-0.6(f/1000-3.3)^2} + 10^{-3} \cdot (f/1000)^4 \quad \text{(dB SPL)}$$

2) Basilar Membrane and Critical Bands

- The cochlear “continuos passband filters” are of non-uniform bandwidth

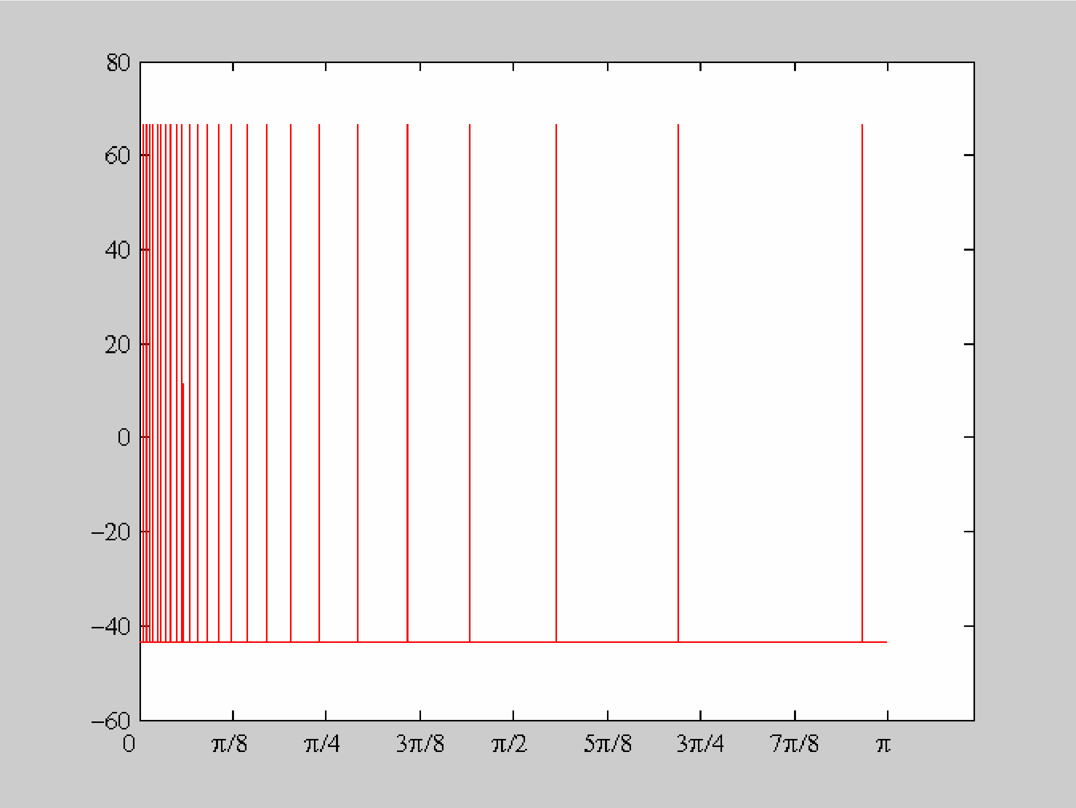
$$BW_c(f) = 25 + 75[1 + 1.4(f/1000)^2]^{0.69} \text{ (Hz)}$$

- Discrete version: Critical Bands → Bark subdivision of FD

$$B(f) = \left[13 \cdot \arctan(0.00076 \cdot f) + 3.5 \cdot \arctan\left[\left(\frac{f}{7500}\right)^2\right] \right] \text{(Bark)}$$

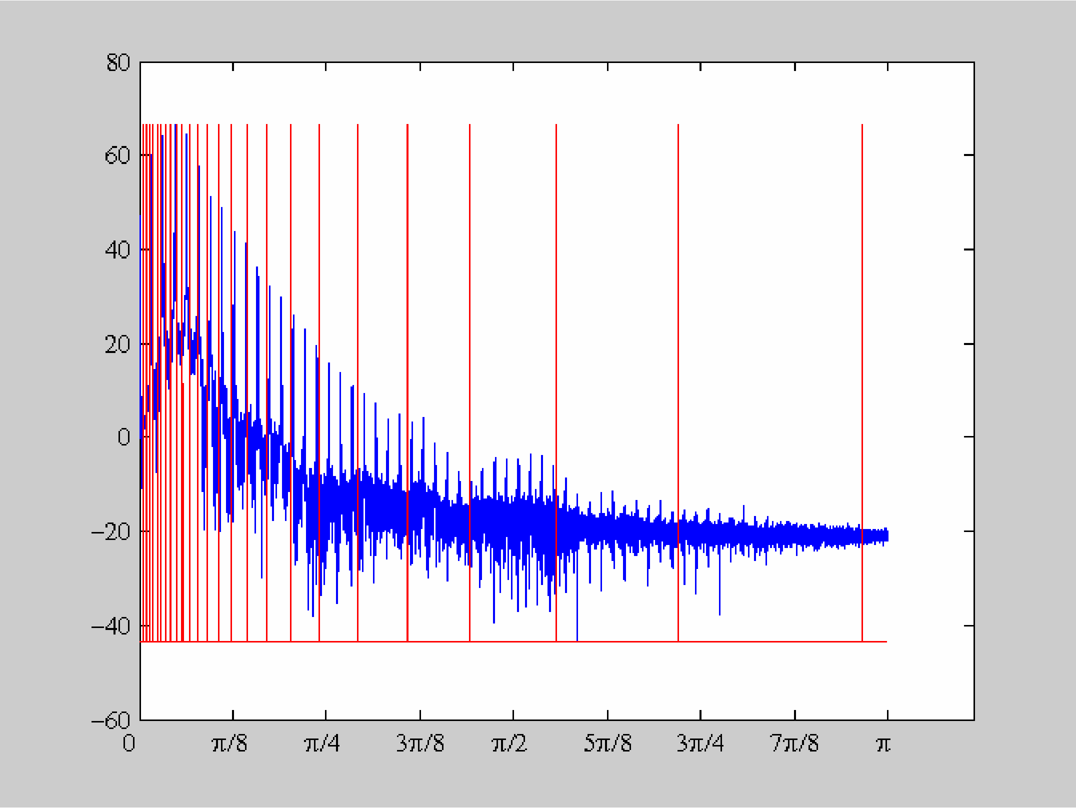
Critical Band Subdivision

Representing the ear as a passband filter bank with non-uniform bandwidth



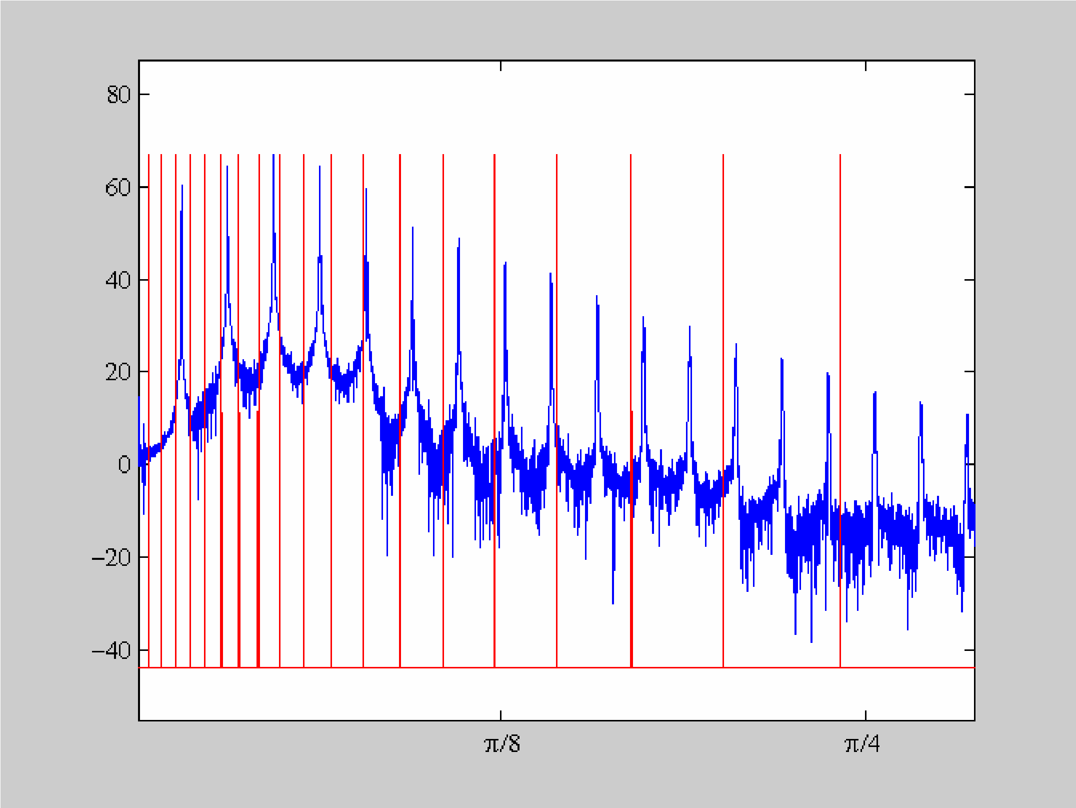
Critical Band Partition of a Trumpet Spectrum

Non uniform distribution of the partials in the Bark subdivision



Pietro Polotti

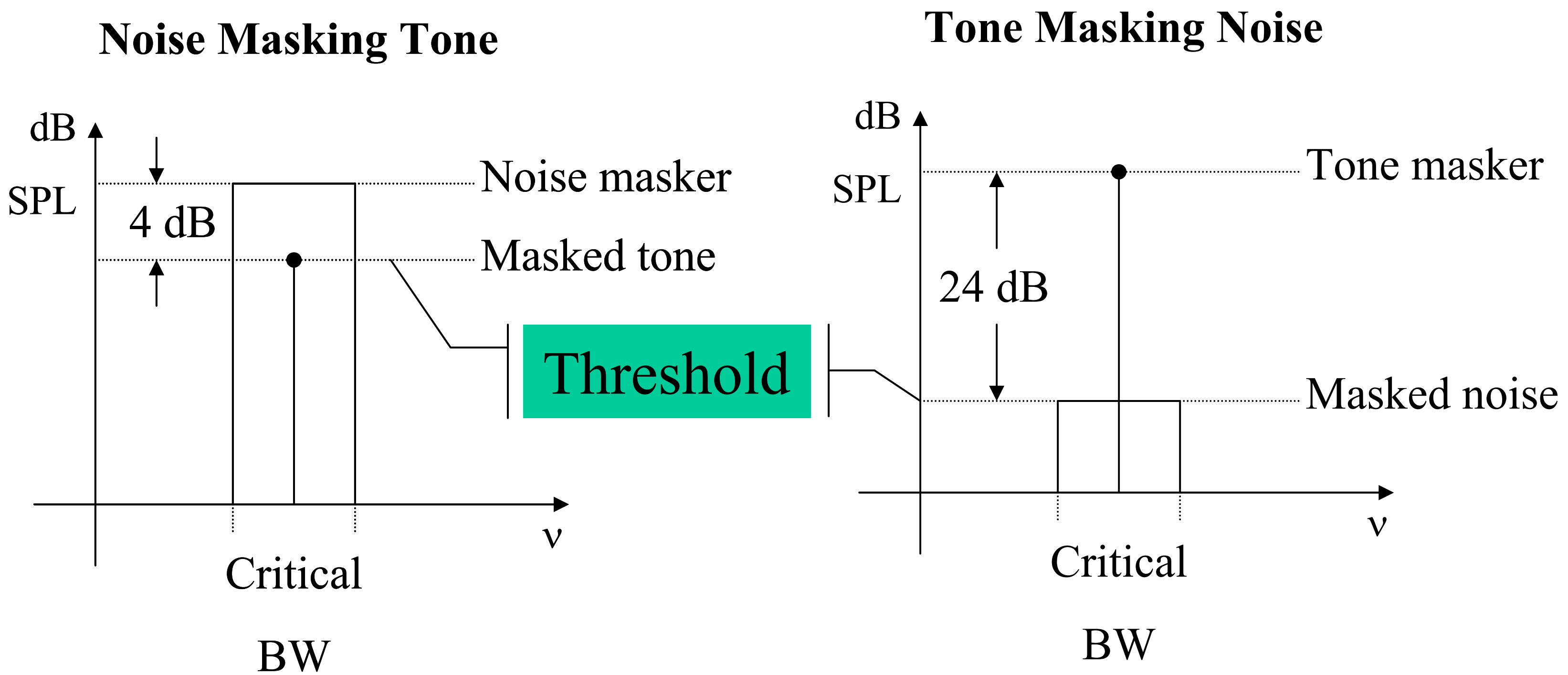
Critical Band Partition of a Trumpet Spectrum: a Detail



Masking

One sound is made inaudible due to the “simultaneous” presence of another sound

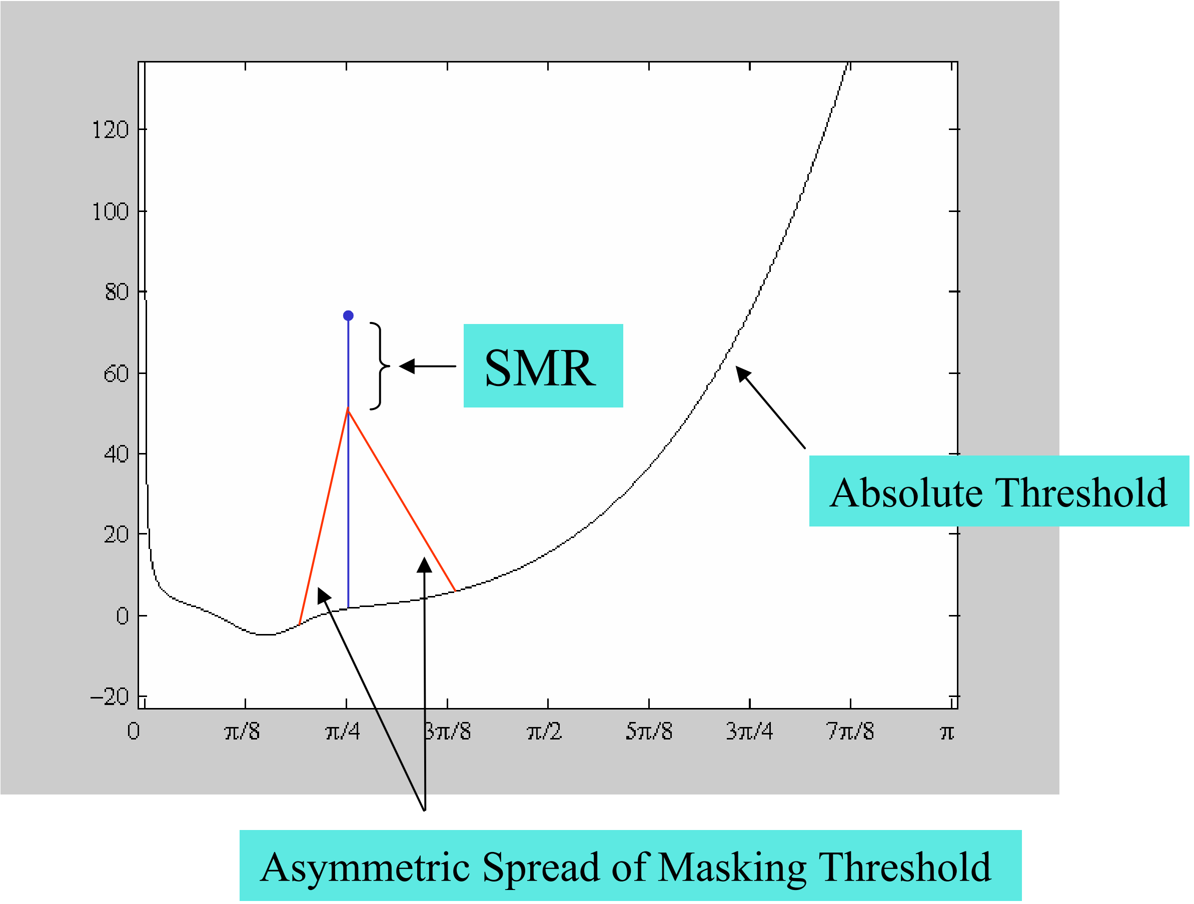
The presence of a strong noise or tone masker creates a sufficient excitation of the basilar membrane at the critical band location to block detection of a weaker signal



Pietro Polotti

3) Tone–Masking-Noise (TMN)

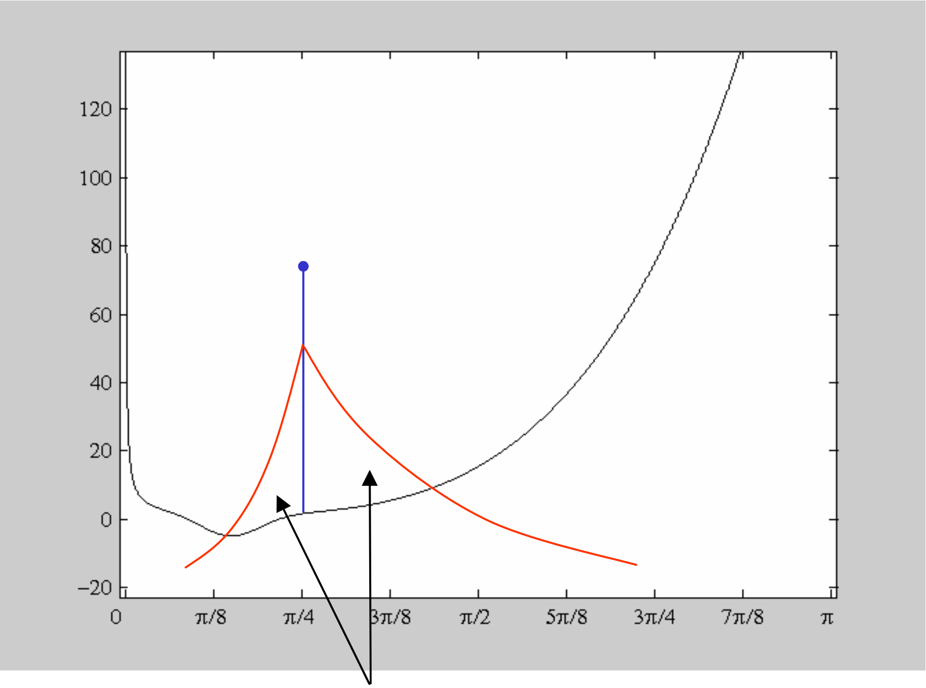
Signal-to-Mask Ratio (SMR) = 24 dB for TMN case



Spread of Masking Threshold through CBs

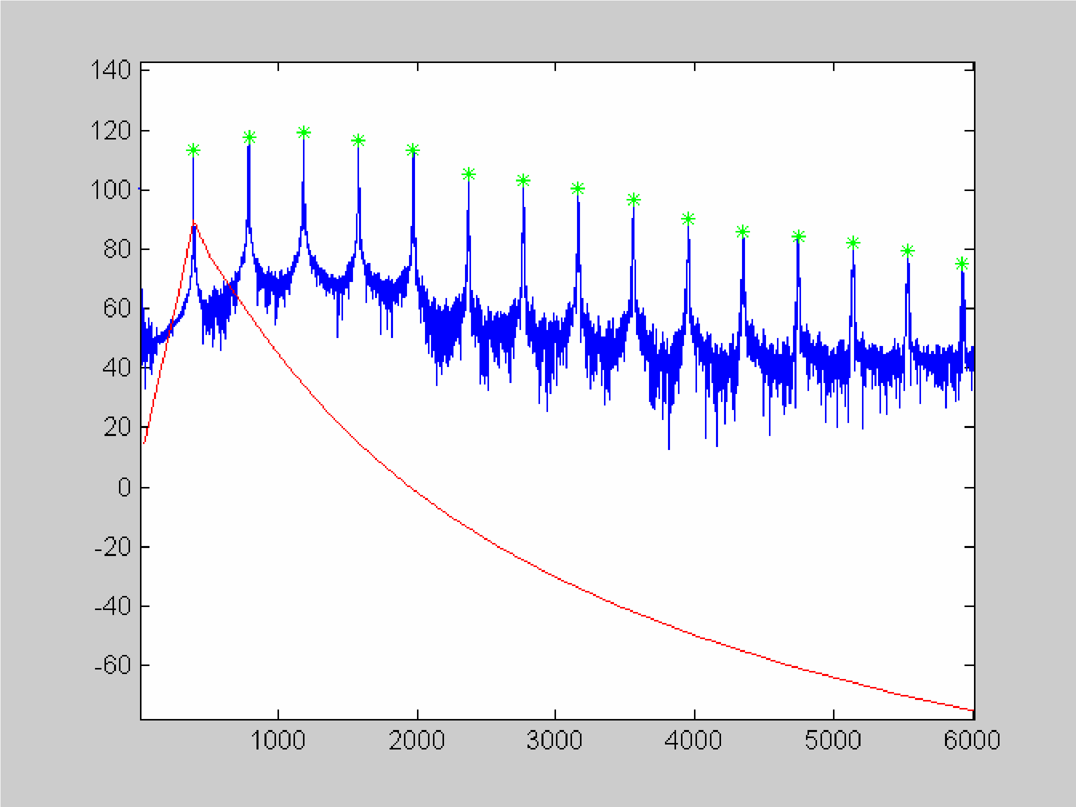
Approximately triangular function

$$M(b) = 15.81 + 7.5 \cdot (b + 0.474) - 17.5 \cdot \sqrt{1 + (b + 0.474)^2} \quad \text{(dB)}$$

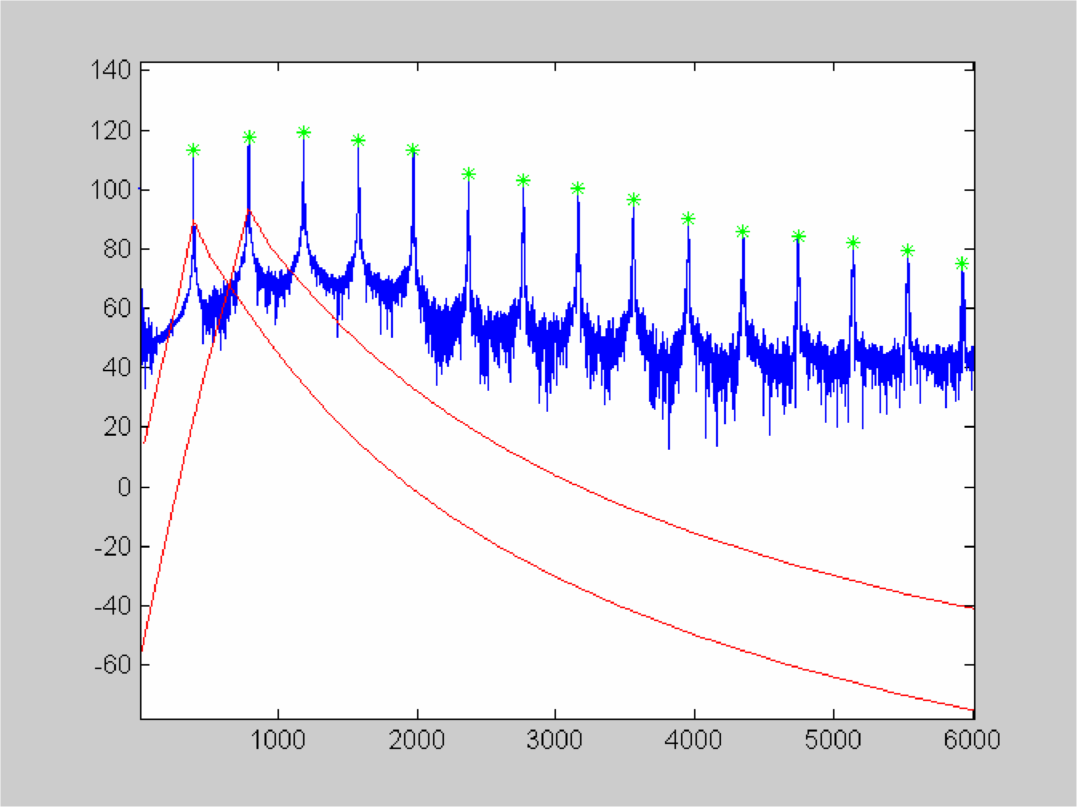


Asymmetric and NL Spread of Masking Threshold

Real case: 1 TMN

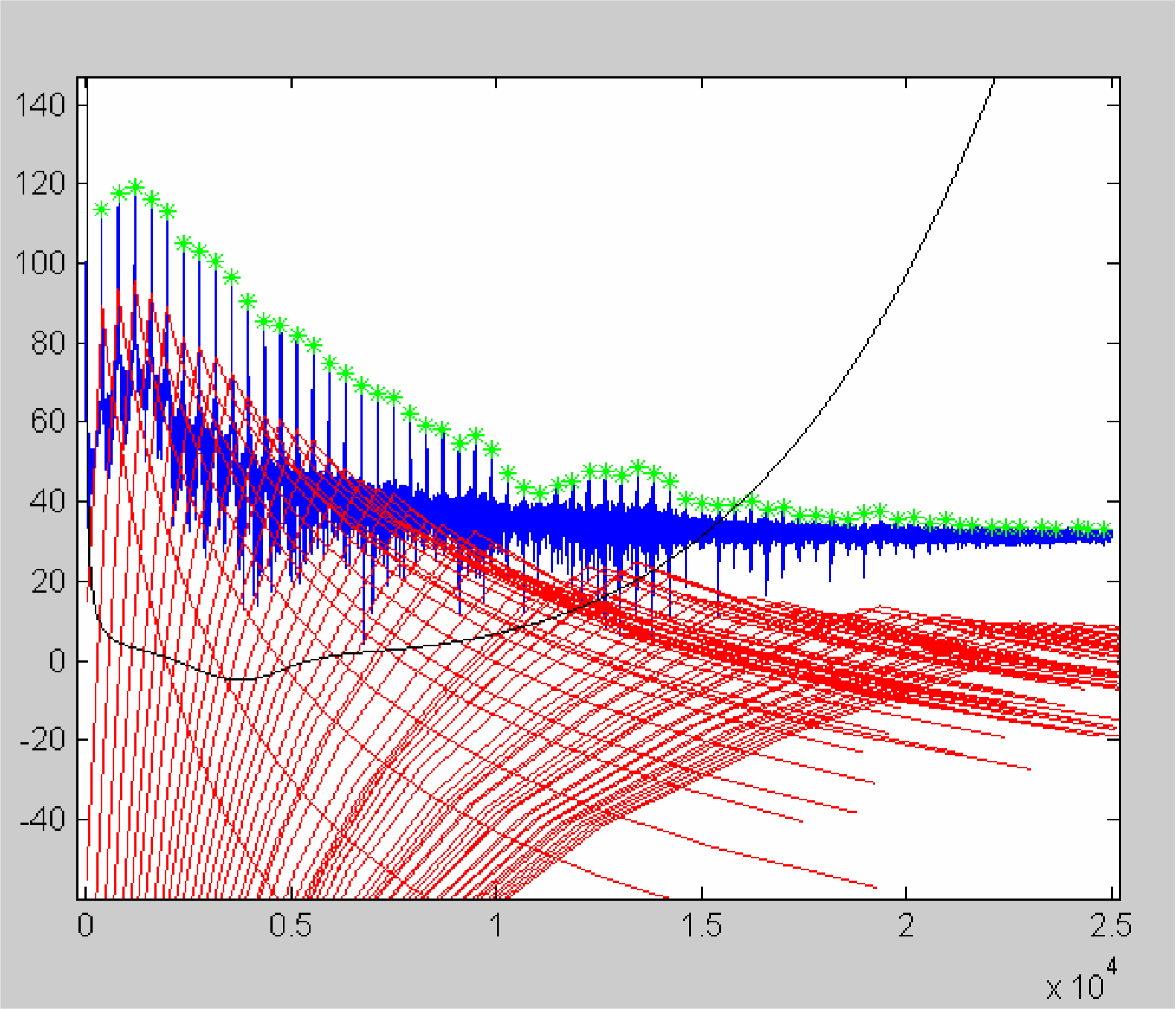


Real case: 2 TMN

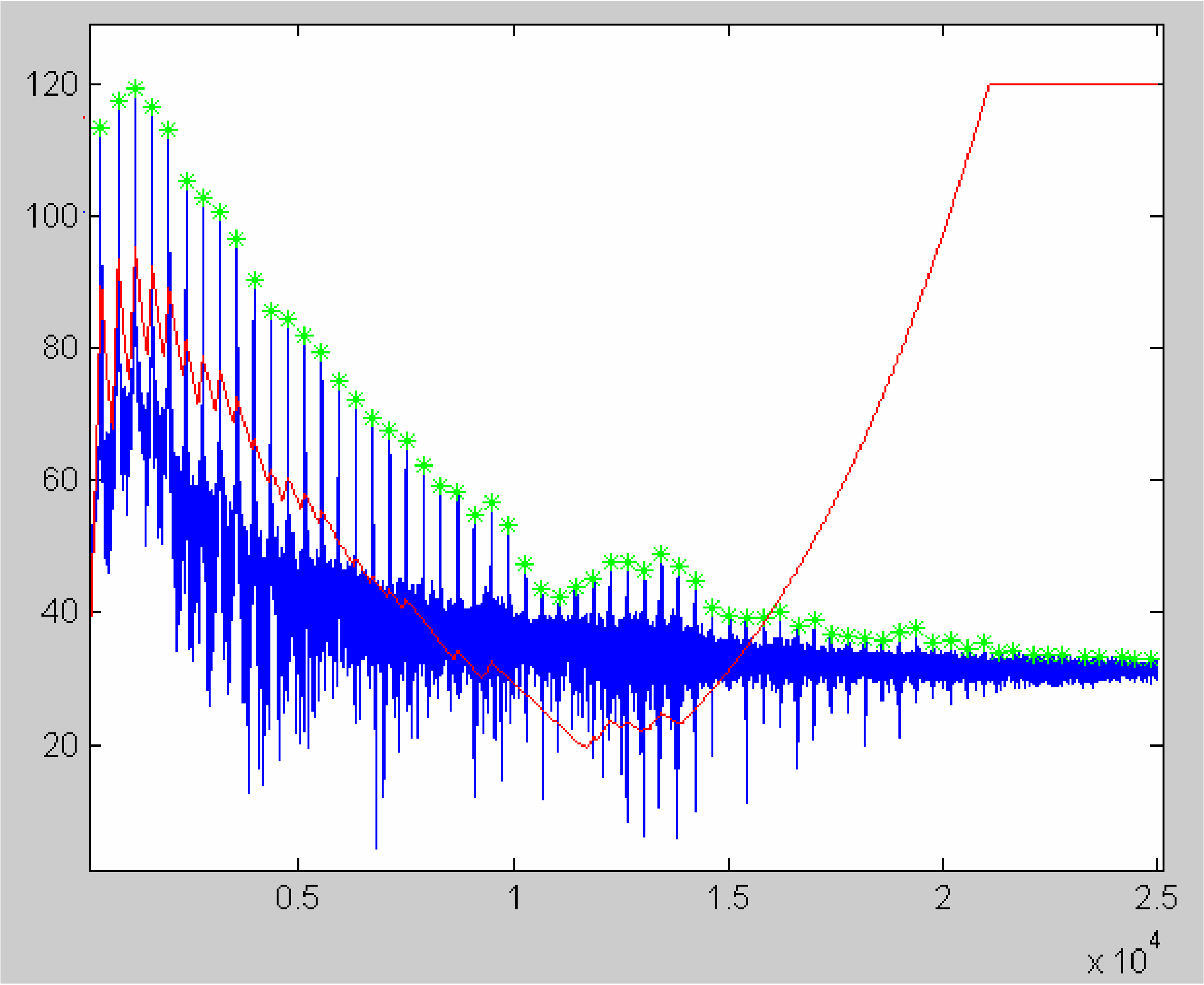


Pietro Polotti

All Maskers

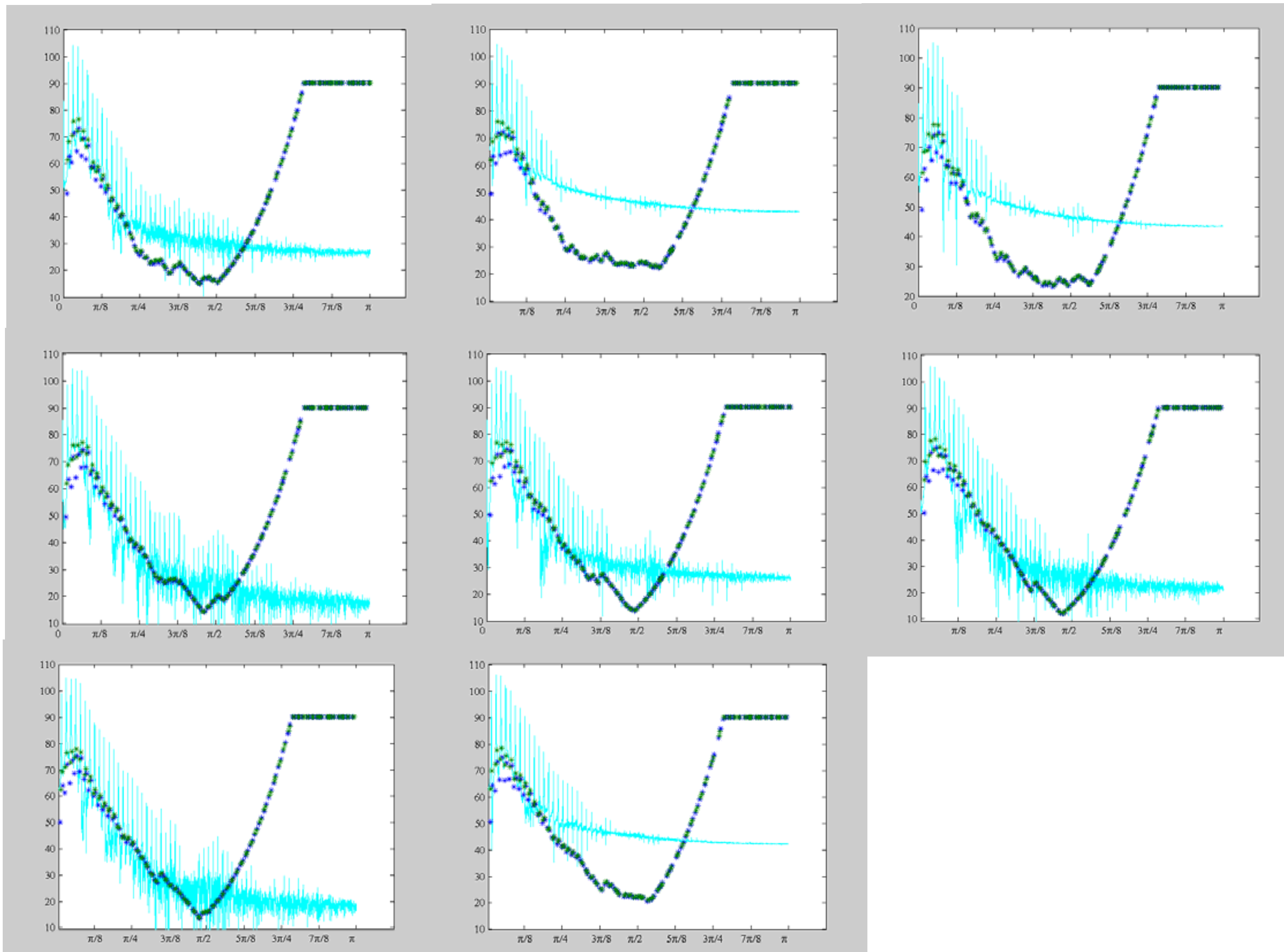


Global Mask



Global masking threshold = Just Noticeable Distortion (JND)

Short-Time Thresholds Evaluation



MPEG1 – layer 1

1) Subdivision of the frequency range by means of a **32-channel MDCT (Subband coding)**

2) An **STFT** is performed in parallel, providing a higher frequency resolution for PSD estimation. Window length = 512. Hop size =384 (12*32, i.e. every 12 samples bit allocation is updated).

3) Signal-to-Mask Ratio computation: For each subband, one considers the maximum of the PSD coefficient **max{*X*(*k*)}** corresponding to that subband. SMR's are set according to this max and mask spreading throughout the critical bands is considered.

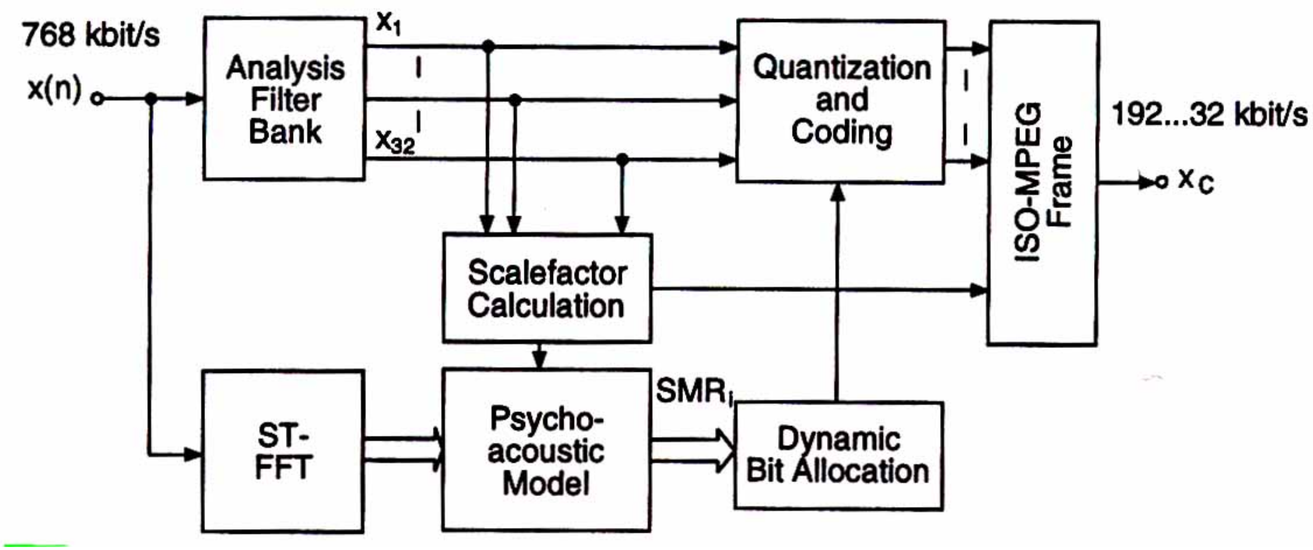
4) Mask spreading: Only masking components that lie in the range –8 / +3 Bark are considered

5) Global masking threshold: sum of all contributing masking components.

6) Bit allocation is carried out in each of the 32 subbands using the SMR.

- Determine the number of bits for each individual subband so that transparent perception is possible.
- This number (simplifying things) corresponds to the difference between the $\max\{X(k)\}$ and the Total Mask Level.

MPEG1 – layer 1



Compression algorithm Scheme