A closer look at the cascade algorithm: convergence, compact support, reproducing polynomials, vanishing moments

Volker Strehl

December 9, 2015

Outline of the main ideas and results

- The cascade iteration for a suitable finite filter $\boldsymbol{h} = (h_\ell, \ldots, h_L)$ converges in the $\mathcal{L}^2$-sense to a function $\phi(t)$, the scaling function
- The convergence is easier to handle in the frequency domain (making use of the convolution theorem!)
- Cascade iteration
  - shows that the scaling function $\phi(t)$ and the wavelet function $\psi(t)$ vanish outside finite intervals of length $L - \ell$ (*compact support*)
  - preserves the ONST property, which guarantees orthogonality for $\phi(t)$ and $\psi(t)$ and its scaled and translated versions
  - shows that necessarily $\widehat{\phi}(0) = 1$ and $\widehat{\psi}(0) = 0$
- Low-pass conds. (*vanishing moments*) like $\widehat{\psi}^{(n)}(0) = 0$ $(0 \leq n < N)$
  - ensure exact reproduction of low-degree polynomials in the approximation spaces and makes them "transparent" in the wavelet spaces
  - provide sharp bounds for the size of wavelet coefficients of (locally) smooth functions

- The filters
  - $\boldsymbol{h} = (h_k)_{k=\ell..L}$ a finite filter satisfying the orthogonality conditions

$$\sum_k h_k h_{2k-m} = \delta_{m,0} \quad \text{and} \quad \sum_k h_k = \sqrt{2}$$

  - Such a filter is called a (finite) *quadrature mirror filter* (QMF)
  - $\boldsymbol{g} = (g_k)_{k=1-L..1-\ell}$ the dual filter to $\boldsymbol{h}$, defined by

$$g_k = (-1)^k h_{1-k},$$

    satisfying automatically the orthogonality conditions

$$\sum_k g_k g_{2k-m} = \delta_{m,0}$$

  - Orthogonality of $\boldsymbol{h}$ and $\boldsymbol{g}$ is a consequence

$$\sum_k h_k g_{2k-m} = 0$$

- The Fourier picture
  - for the filter $\boldsymbol{h}$

$$m_0(s) = \frac{1}{\sqrt{2}} H(-2\pi s) = \frac{1}{\sqrt{2}} \sum_k h_k e^{-2\pi iks} = \frac{1}{\sqrt{2}} \widehat{\sum_k h_k \delta_k}(s)$$

$$m_0(0) = \frac{1}{\sqrt{2}} \sum_k h_k = 1$$

  - for the filter $\boldsymbol{g}$

$$m_1(s) = \frac{1}{\sqrt{2}} G(-2\pi s) = \frac{1}{\sqrt{2}} \sum_k g_k e^{-2\pi iks} = \frac{1}{\sqrt{2}} \widehat{\sum_k g_k \delta_k}(s)$$
$$= e^{-2\pi i(s+1/2)} \cdot \overline{m_0(s+1/2)}$$

  - orthogonality

$$|m_0(s)|^2 + |m_0(s+1/2)|^2 \equiv 1$$
$$|m_1(s)|^2 + |m_1(s+1/2)|^2 \equiv 1$$
$$m_0(s) \cdot \overline{m_0(s+1/2)} + m_1(s) \cdot \overline{m_1(s+1/2)} \equiv 0$$

  - low/highpass

$$m_0(1/2) = m_1(0) = 0$$

- The scaling equation

$$(S) \qquad \phi(t) = \sqrt{2} \sum_k h_k \, \phi(2t - k)$$

- The wavelet equation

$$(W) \qquad \psi(t) = \sqrt{2} \sum_k g_k \, \phi(2t - k)$$

- The scaling equation in the frequency domain

$$(\widehat{S}) \qquad \widehat{\phi}(s) = m_0(s/2) \cdot \widehat{\phi}(s/2)$$

- The wavelet equation in the frequency domain

$$(\widehat{W}) \qquad \widehat{\psi}(s) = m_1(s/2) \cdot \widehat{\phi}(s/2)$$

- The *cascade mapping* in the time domain

$$\mathcal{C} : f(t) \longmapsto (\mathcal{C}f)(t) = \begin{cases} \sqrt{2}\sum_k h_k\, f(2t-k) \\ = D_2\left(\sum_k h_k\, f(t-k)\right) \\ = D_2\left(\sum_k h_k\, \delta_k(t) \star f(t)\right) \end{cases}$$

- The *cascade mapping* in the frequency domain

$$\widehat{\mathcal{C}} \;:\; \widehat{f}(s) \;\longmapsto\; \widehat{(\mathcal{C}f)}(s) \;=\; \begin{cases} D_{1/2}\left(\widehat{\sum_k h_k\,\delta_k}(s) \cdot \widehat{f}(s)\right) \\ = \frac{1}{\sqrt{2}}\sum_k h_k\, e^{-2\pi i k(s/2)} \cdot \widehat{f}(s/2) \\ = m_0(s/2) \cdot \widehat{f}(s/2) \end{cases}$$

- Iterating the cascading operation

$$\mathcal{C}^n : f(t) \longmapsto \mathcal{C}(\mathcal{C}^{n-1} f)(t) = (\underbrace{\mathcal{C} \circ \cdots \circ \mathcal{C}}_{n \text{ times}} f)(t)$$

- Iterating the cascading operation in the frequency domain

$$\widehat{C}^n : \widehat{f}(s) \longmapsto \widehat{(\mathcal{C}^n f)}(s) = \begin{cases} m_0(s/2) \cdot m_0(s/4) \cdots m_0(s/2^n) \cdot \widehat{f}(s/2^n) \\ = m^{[n]}(s) \cdot \widehat{f}(s/2^\ell) \end{cases}$$

where

$$m^{[n]}(s) := m_0(s/2) \cdot m_0(s/4) \cdots m_0(s/2^n)$$

is an exponential polynomial of period 2^n ...

- ... but these *cascade multipliers* do not belong to $\mathcal{L}^2(\mathbb{R})$

- Question: Is there a $\phi(t) \in \mathcal{L}^2(\mathbb{R})$ which satisfies the scaling equation

$$(S) \qquad \phi(t) = \sqrt{2}\,\sum_k h_k\,\phi(2t-k) \quad .$$

i.e., does the cascading operator have a fixed point in $\mathcal{L}^2(\mathbb{R})$?

- What one ideally would like to have:

$$\eta^{[0]}(t) = \chi_{[0,1)}(t) \qquad\qquad \widehat{\eta^{[0]}} = \mathrm{sinc}(s)$$

$$\downarrow \qquad\qquad \downarrow$$

$$\eta^{[n+1]}(t) = \begin{cases} \mathcal{C}\eta^{[n]}(t) \\ \sqrt{2}\,\sum_k h_k\,\eta^{[n]}(2t-k) \end{cases} \qquad \widehat{\eta^{[n+1]}}(s) = \begin{cases} m_0(s/2)\cdot\widehat{\eta^{[n]}}(s) \\ = m^{[n]}(s)\cdot\mathrm{sinc}(s/2^n) \end{cases}$$

$$\downarrow_{n\to\infty} \qquad\qquad \downarrow_{n\to\infty}$$

$$\eta^{[\infty]}(t) = \sqrt{2}\,\sum_k h_k\,\eta^{[\infty]}(2t-k) \qquad \widehat{\eta^{[\infty]}}(s) = \begin{cases} m^{[\infty]}(s)\cdot\mathrm{sinc}(s/2^\infty) \\ = m^{[\infty]}(s) \end{cases}$$

- Do limit functions $\eta^{[\infty]}(t)$ and $m^{[\infty]}(s)$ exist in $\mathcal{L}^2(\mathbb{R})$?
  If so, what properties do they have? What tells this about the properties of the wavelet transform based on $\boldsymbol{h}$?

- It appears that if

$$m^{[\infty]}(s) = \prod_{n\geq 1} m_0(s/2^n) = \lim_{N\to\infty} \prod_{1\leq n\leq N} m_0(s/2^n)$$

makes sense and belongs to $\mathcal{L}^1(\mathbb{R})(\mathbb{R}) \cap \mathcal{L}^2$, then its inverse Fourier transform would satisfy

$$\left(m^{[\infty]}\right)^{\vee}(t) = \eta^{[\infty]}(t)$$

and hence would be the $\phi(t)$ as desired ...

- The good news
  - One can show that the infinite product

$$m^{[\infty]}(s) = \prod_{n\geq 1} m_0(s/2^n) = \lim_{N\to\infty} \prod_{1\leq n\leq N} m_0(s/2^n) = \lim_{N\to\infty} m^{[N]}(s)$$

    converges absolutely and uniformly on every finite interval $[-\rho, \rho] \subset \mathbb{R}$. Thus it makes sense to speak of this expression as defining function defined (and continuous) of $\mathbb{R}$
- The bad news
  - The multiplier for N-fold cascading

$$m^{[N]}(s) = m_0(s/2) \cdot m_0(s/4) \cdots m_0(s/2^N)$$

    is a 2^N-periodic function – so the $m^{[N]}(s)$ (as $m_0(s)$) do certainly not belong to $\mathcal{L}^2(\mathbb{R})$ and therefore cannot converge in $\mathcal{L}^2(\mathbb{R})$ towards $m^{[\infty]}(t)$

- Solving the problem
  - A common technique to resolve this kind of problem is to introduce *band-limited* versions of these functions:

$$\mu^{[n]}(s) = m^{[n]}(s) \cdot \chi_{[-2^{n-1},2^{n-1}]}(s)$$

  - Check that

$$(*) \qquad \mu^{[n]}(s) = m_0(s/2) \cdot \mu^{[n-1]}(s/2)$$

  - Band-limiting ensures that $\mu^{[\ell]}(s) \in \mathcal{L}^2(\mathbb{R}) \cap \mathcal{L}^1(\mathbb{R})$, so these functions do have an inverse Fourier transform $(\mu^{[\ell]})^\vee(t)$
  - From $(*)$ one gets cascading in the time domain:

$$(\mu^{[n]})^\vee(t) = \mathcal{C}(\mu^{[n-1]})^\vee(t)$$

  - It now makes sense to consider the sequence $\left(\mu^{[n]}(s)\right)_{\ell\geq 0}$ as a sequence in $\mathcal{L}^2(\mathbb{R})$ and ask for its convergence

- Theorem:
  If $\boldsymbol{h}$ is a QMF as above, for which there exists a constant $c > 0$ such that $|m_0(s)| \geq c$ for all $|s| \geq 1/4$. Then one has $\mathcal{L}^2$-convergence

$$\mu^{[n]}(s) \longrightarrow_{n\to\infty} m^{[\infty]}(s) = \widehat{\phi}(s)$$

- Consequently, by applying the inverse Fourier transform, one gets $\mathcal{L}^2$-convergence

$$(\mu^{[n]})^{\vee}(t) \longrightarrow_{n\to\infty} \phi(t),$$

  where the functions $(\mu^{[\ell]})^{\vee}(t)$ are band-limited approximations of $\phi(t)$
- This can be used to show that, as desired, one has $\mathcal{L}^2$-convergence

$$\widehat{\eta^{[n]}}(s) \to_{n\to\infty} \widehat{\phi}(s) \quad \text{and} \quad \eta^{[n]}(t) \to_{n\to\infty} \phi(t)$$

  as desired

- A function $f : \mathbb{R} \to \mathbb{C}$ has *compact support*, if it vanishes outside an interval $I \subset \mathbb{R}$ of finite length
- From the cascade iteration step for a QMF of length $2N$

$$\eta^{[n+1]}(s) = \mathcal{C}\eta^{[n]}(t) = \sqrt{2}\,\sum_{k=\ell}^{L} h_k\,\eta^{[n]}(2t-k)$$

one gets

  - if $\eta^{[n]}(t)$ vanishes outside the interval $[a_n, b_n]$ then $\eta^{[n+1]}(t)$ vanishes outside the interval $[a_{n+1}, b_{n+1}] = [(a_n+\ell)/2, (b_n+L)/2]$
  - by induction one gets in the limit that $\eta^{[\infty]}(t) = \phi(t)$ vanishes outside the interval $[a_\infty, b_\infty] = [\ell, L]$ of length $L - \ell = 2N - 1$
  - and by (W) it follows that $\psi(t)$ vanishes outside the interval $[-N+1, N]$

- Orthogonality property of the cascade operator (1)
  - A family of translates $\{\,(T_k f)(t)\,\}_{k\in\mathbb{Z}} = \{\,f(t-k)\,\}_{k\in\mathbb{Z}}$ of an $\mathcal{L}^2$-function $f(t)$ is orthonormal if and only if

$$\sum_{n\in\mathbb{Z}} |\widehat{f}(s+n)|^2 \equiv 1$$

$$\text{Proof :}\quad \langle\, f \mid T_k f\,\rangle = \langle\, \widehat{f} \mid \widehat{T_k f}\,\rangle = \int_{\mathbb{R}} \widehat{f}(s)\,\overline{\widehat{f}(s)}\,e^{2\pi iks} ds$$

$$= \sum_{n\in\mathbb{Z}} \int_n^{n+1} \left|\widehat{f(s)}\right|^2 e^{2\pi iks} ds = \sum_{n\in\mathbb{Z}} \int_0^1 \left|\widehat{f}(s+n)\right|^2 e^{2\pi iks} ds$$

$$= \int_0^1 \sum_{n\in\mathbb{Z}} \left|\widehat{f}(s+n)\right|^2 e^{2\pi iks} ds$$

Hence in terms of Fourier series

$$\sum_{k\in\mathbb{Z}} \langle\, f \mid T_k f\,\rangle\, e^{-2\pi iks} = \sum_{n\in\mathbb{Z}} \left|\widehat{f}(s+n)\right|^2$$

  - Such a family with $\langle\, T_k f \mid T_\ell f\,\rangle = \delta_{k,\ell}\;\; (k,\ell\in\mathbb{Z})$ is called an *orthonormal system of translates* (ONST)

- Orthogonality property of the cascade operator (2)
  - Reminder: orthogonality of the filter $\boldsymbol{h}$ reads as

$$|m_0(s)|^2 + |m_0(s+1/2)|^2 \equiv 1$$

  - If $\{ T_k f \}_{k\in\mathbb{Z}}$ is an ONST, then $\{ T_k (\mathcal{C}f) \}_{k\in\mathbb{Z}}$ is again an ONST

$$\sum_{n\in\mathbb{Z}} \left|\widehat{(\mathcal{C}f)}(s+n)\right|^2 = \sum_{n\in\mathbb{Z}} \left| m_0(\frac{s+n}{2}) \cdot \widehat{f}(\frac{s+n}{2})\right|^2 = \sum_{n\ even} \cdots + \sum_{n\ odd} \cdots$$

$$= \left|m_0(\frac{s}{2})\right|^2 \sum_{n'\in\mathbb{Z}} \left|\widehat{f}(\frac{s}{2}+n')\right|^2 + \left|m_0(\frac{s+1}{2})\right|^2 \sum_{n''\in\mathbb{Z}} \left|\widehat{f}(\frac{s+1}{2}+n'')\right|^2$$

$$= \left|m_0(\frac{s}{2})\right|^2 + \left|m_0(\frac{s+1}{2})\right|^2 \equiv 1$$

because $m_0(s)$ is a 1-periodic function

- Orthogonality property of the cascade operator (3)
  - The family $\left\{ T_k \chi_{[1/2,1/2]}(t) \right\}_{k \in \mathbb{Z}}$ is obviously an ONST
  - Then, by induction, all intermediate families

$$\left\{ T_k \eta^{[n]}(t) \right\}_{k \in \mathbb{Z}} \quad (n = 1, 2, 3, \ldots)$$

    are ONST
  - From $\mathcal{L}^2$-convergence $\eta^{[n]}(t) \to_{n \to \infty} \phi(t)$ one gets:

$$\{ T_k \phi(t) \}_{k \in \mathbb{Z}} \qquad \text{is an ONST}$$

- Integrals
  - If $f(t)$ is integrable, then

$$\int_{\mathbb{R}} (\mathcal{C}f)(t)\,dt = \sqrt{2}\sum_k h_k \int_{\mathbb{R}} f(2t-k)\,dt = \frac{1}{\sqrt{2}}\sum_k h_k \int_{\mathbb{R}} f(t)\,dt = \int_{\mathbb{R}} f(t)\,dt$$

  - hence from the cascade iteration

$$\widehat{\phi}(0) = \int_{\mathbb{R}} \phi(t)\,dt = \ldots = \int_{\mathbb{R}} \chi_{[0,1)}(t)\,dt = 1$$

  - and from the wavelet equation $(\widehat{W})$

$$\int_{\mathbb{R}} \psi(t)\,dt = \widehat{\psi}(0) = m_1(0)\cdot\widehat{\phi}(0) = m_1(0) = 0$$

- Now consider low-pass properties of the QMF $\boldsymbol{h}$
- as specified by

$$m_0^{(n)}(1/2) = 0 \qquad (0 \leq n < N)$$

- or in terms of the *moment conditions*

$$\sum_k (-1)^k h_k \, k^n = 0 \ (0 \leq n < N)$$

- or else by

$$m_0(s) = \left(\frac{1 + e^{-2\pi i s}}{2}\right)^N \cdot L(s)$$

  with $L(s)$ a trigonometric polynomial (finite Fourier series) of period 1
- If any of these equivalent properties is satisfied,
  then one says that the QMF $\boldsymbol{h}$ has N *vanishing moments*

- By differentiating the scaling equation in the frequency domain form

$$\widehat{\phi}(s) = m_0(s/2) \cdot \widehat{\phi}(s/2)$$

one gets (remember $\widehat{\phi}(0) = 1$)

$$\widehat{\phi}^{(n)}(k) = 0 \quad \text{for } 0 \leq n < N \quad \text{and} \quad k \in \mathbb{Z} \setminus \{0\}$$

- Hence the following Fourier series is a constant

$$(2\pi i)^n \sum_{k \in \mathbb{Z}} \widehat{t^n \phi(t)}(k)\, e^{2\pi i k t} = \sum_{k \in \mathbb{Z}} \widehat{\phi}^{(n)}(k)\, e^{2\pi i k t} = \widehat{\phi}^{(n)}(0)$$

- From Poisson's formula one gets

$$\sum_{k \in \mathbb{Z}} \widehat{t^n \phi(t)}(k)\, e^{2\pi i k t} = \sum_{\ell \in \mathbb{Z}} (t + \ell)^n \phi(t + \ell)$$

- and hence

$$\widehat{\phi}^{(n)}(0) = (2\pi i)^n \sum_{\ell \in \mathbb{Z}} (t + \ell)^n \phi(t + \ell)$$

- Now use the binomial theorem for $0 \leq n < N$

$$
\begin{aligned}
\sum_{\ell \in \mathbb{Z}} \phi(t+\ell) \cdot \ell^n &= \sum_{\ell \in \mathbb{Z}} \phi(t+\ell) \sum_{j=0}^{n} \binom{n}{j} (t+\ell)^j (-t)^{n-j} \\
&= \sum_{j=0}^{n} \binom{n}{j} (-t)^{n-j} \sum_{\ell \in \mathbb{Z}} (t+\ell)^j \phi(t+\ell) \\
&= \frac{1}{(2\pi i)^n} \sum_{j=0}^{n} \binom{n}{j} (-t)^{n-j} \, \widehat{\phi}^{(j)}(0) \\
&= p_n(t)
\end{aligned}
$$

  which is a polynomial of degree n in t

- Note: even though the sum $\sum_{\ell \in \mathbb{Z}}$ is infinite, for any specific $t \in \mathbb{R}$ it contains only a finite number of non-zero summands, since $\phi(t)$ has finite support

- Since the polynomials $p_n(t)$ $(0 \leq n < N)$ are a basis in the vector space of polynomials of degree $< N$, one gets
  - There are polynomials $q_n(t) = \sum_{j=0}^{n} q_{n,j} t^j$ $(0 \leq n < N)$ such that

$$\sum_{j=0}^{n} q_{n,j} \cdot p_j(t) = t^n \quad (0 \leq n < N)$$

  and hence

$$\sum_{\ell \in \mathbb{Z}} \phi(t+\ell) \cdot q_n(\ell) = t^n$$

- Theorem:

  For any polynomial $r(t)$ of degree $< N$ there are constants ρ_ℓ $(\ell \in \mathbb{Z})$ such that

$$\sum_{\ell \in \mathbb{Z}} \phi(t+\ell) \cdot \rho_\ell = r(t)$$

  and by referring to the orthogonality of the translates of $\phi(t)$ one can deduce what these constants really are:

$$\rho_\ell = \langle\, r(t) \,|\, \phi(t+\ell) \,\rangle$$

- Recall, that given a QMF $\boldsymbol{h} = (h_k)$ one has
  - the dual filter $\boldsymbol{g} = (g_k)$ defined by setting $g_k = (-1)^k h_{1-k}$
  - with Fourier series

$$m_1(s) = \frac{1}{\sqrt{2}} \sum_k g_k e^{-2\pi iks} = e^{-2\pi i(s+1/2)} \overline{m_0(s+1/2)}$$

  - This is indeed an orthogonal filter

$$|m_1(s)|^2 + |m_1(s+1/2)|^2 \equiv 1$$

    and

$$m_1(0) = \frac{1}{\sqrt{2}} \sum_k g_k = \frac{1}{\sqrt{2}} \sum_k (-1)^{k-1} h_k = -\overline{m_0(1/2)}$$

  - Thus if $\boldsymbol{h}$ is low-pass($m_0(1/2) = 0$) then $\boldsymbol{g}$ is high-pass ($m_1(0) = 0$)
  - Orthogonality between $\boldsymbol{h}$ and $\boldsymbol{g}$ is expressed by

$$m_0(s) \cdot \overline{m_0(s+1/2)} + m_1(s) \cdot \overline{m_1(s+1/2)} \equiv 0$$

- Using the *scaling function* $\phi(t)$ belonging to the QMF $\boldsymbol{h}$ one defines the *wavelet function* $\psi(t)$ by the *wavelet equation*

$$(W) \qquad \psi(t) = \sum_k g_k\,\phi_{1,k}(t) = \sqrt{2}\sum_k g_k\,\phi(2t-k)$$

- This is a finite sum, so $\psi(t)$ is also a function with compact support (as mentioned before)
- The frequency picture is

$$\widehat{\psi}(s) = m_1(s/2)\cdot\widehat{\phi}(s/2)$$

- The scaling and wavelet functions $\phi(t)$ and $\psi(t)$ can be translated and dilated as usual

$$\phi_{j,k}(t) = D_{2^j}T_k\phi(t) = 2^{j/2}\phi(2^j t - k)$$
$$\psi_{j,k}(t) = D_{2^j}T_k\psi(t) = 2^{j/2}\psi(2^j t - k)$$

- As shown before, the family $\{ T_k\phi(t) \}_{k\in\mathbb{Z}}$ is an ONST, hence for any fixed $j \in \mathbb{Z}$ the family $\{ \phi_{j,k}(t) \}_{k\in\mathbb{Z}}$ is orthonormal
- From the orthogonality property of $\boldsymbol{g}$ it follows that for any fixed $j \in \mathbb{Z}$ the family $\{ \psi_{j,k}(t) \}_{k\in\mathbb{Z}}$ is orthonormal
- From the orthogonality between $\boldsymbol{h}$ and $\boldsymbol{g}$ it follows that for any fixed $j \in \mathbb{Z}$ the families $\{ \phi_{j,k}(t) \}_{k\in\mathbb{Z}}$ and $\{ \psi_{j,\ell}(t) \}_{\ell\in\mathbb{Z}}$ are orthonormal
- It is then easy to show that the family of all wavelet functions $\{ \psi_{j,k}(t) \}_{j,k\in\mathbb{Z}}$ is an orthonormal family in $\mathcal{L}^2(\mathbb{R})$
- $\mathcal{L}^2$-completeness is not guaranteed!

- Remember that in the QMF cascading situation

$$\widehat{\psi}(0) = \int_{\mathbb{R}} \psi(t)\, dt = 0$$

- Theorem:

  If $\{\, \psi_{j,k}(t) \,\}_{j,k \in \mathbb{Z}}$ is any orthonormal family in $\mathcal{L}^2(\mathbb{R})$ with $\psi(t)$ and $\widehat{\psi}(s)$ integrable, i.e., $\in \mathcal{L}^1(\mathbb{R})$.
  Then

$$\int_{\mathbb{R}} \psi(t)\, dt = 0$$

  Comment: Integrability of $\psi(t)$ assures that the integral exists. Integrability of $\widehat{\psi}(t)$ plus the Riemann-Lebesgue Lemma says that $\psi(t)$ is uniformly continuous on $\mathbb{R}$ and vanishes at infinity

- For a proof see the Lecture Notes

The previous theorem can be extended.

- Theorem:

  If $\{\, \psi_{j,k}(t) \,\}_{j,k\in\mathbb{Z}}$ is an orthonormal family,
  with $t^N \cdot \psi(t)$ and $s^{N+1} \cdot \widehat{\psi}(s)$ integrable, i.e., $\in \mathcal{L}^1(\mathbb{R})$.
  Then $\psi(t)$ has N *vanishing moments*:

$$\int_{\mathbb{R}} t^n \, \psi(t) \, dt = 0 \quad (0 \le n < N)$$

- In the context of a QMF $\boldsymbol{h}$ and its wavelet function $\psi(t)$ this *vanishing moments* property is indeed equivalent to the earlier statements like

$$m_0^{(n)}(1/2) = 0 \quad (0 \le n < N)$$

  or

$$\sum_k (-1)^k h_k \, k^n = 0 \quad (0 \le n < N)$$

Comments:

- The proof goes by induction over N
- The condition $t^N\,\psi(t) \in \mathcal{L}^1(\mathbb{R})$ guarantees existence of the integrals
- The condition $s^{N+1} \cdot \widehat{\psi}(s) \in \mathcal{L}^1(\mathbb{R})$ says that $\psi(t)$ is *smooth*:
  it has $N+1$ continuous derivatives, remember

$$(2\pi is)^{N+1} \cdot \widehat{\psi}(s) = \widehat{\psi^{(N+1)}}(s)$$

  So $\psi^{(N+1)}(t)$ is uniformly continuous and vanishes as $t \to \pm\infty$ (R-L)
- A practical consequence:
  If $f(t) \in \mathcal{L}^2(\mathbb{R})$ behaves like a polynomial function of degree $< N$ on the support of $\psi_{j,k}$ (a finite interval), then $\langle\, f \mid \psi_{j,k} \,\rangle = 0$,
  i.e., f becomes "invisible" or "transparent" for the detail parts of the wavelet transformation
- The smoothness of a wavelet function $\psi(t)$ corresponds to the number of vanishing moments

- The previous Theorem can be extended even further by making a statement about the size of wavelet coefficients for smooth functions
- Assume:
  - $\psi(t)$ is a (real) wavelet function (coming from an QMF $\boldsymbol{h}$, with compact support $[0, a]$), so that
  - $\{\psi_{j,k}\}_{j,k \in \mathbb{Z}}$ is an orthonormal family
  - Intervals $I_{j,k}$ of length $2^{-j}a$ are obtained from $I_{0,0} = [0, a]$ by dilation and translation
  - $t_{j,k} = 2^{-(j+1)}a + 2^{-j}k$ is the midpoint of $I_{j,k}$
  - $\psi(t)$ has N vanishing moments

- Theorem:

  For any N-times differentiable function $f(t)$ on $\mathbb{R}$ with $f^{(N)}(t)$ bounded, the size of its wavelet coefficients can be bounded:

  i. There exists a constant $C_{N,f} > 0$ such that for all $j, k \in \mathbb{Z}$

  $$|\langle f \mid \psi_{j,k} \rangle| \leq C_{N,f} \cdot 2^{-jN} 2^{-j/2}$$

  ii. More precisely: for large j

  $$|\langle f \mid \psi_{j,k} \rangle| \approx 2^{-jN} 2^{-j/2} \left(\frac{1}{N!} f^{(N)}(t_{j,k}) \int_{-a/2}^{a/2} t^n \psi(t + a/2)\, dt \right)$$

- About the proof
  - Use Taylor's expansion with remainder term for $f(t)$ around $t_{j,k}$, noting that $\int_{I_{j,k}} (t - t_{j,k})^n \psi_{j,k}(t)\, dt = 0$ for $n < N$ to get

$$\langle\, f \mid \psi_{j,k} \rangle = \int_{I_{j,k}} R_N(t)\, \psi_{j,k}(t)\, dt$$

where

$$R_N(t) = \frac{1}{N!}(t - t_{j,k})^N f^{(N)}(\xi)$$

for some ξ between t and $t_{j,k}$

For $t \in I_{j,k}$

$$|R_N(t)| \leq \frac{1}{N!} 2^{-N(j+1)} a^N \max_{t' \in I_{j,k}} |f^{(N)}(t')|$$

With this estimate use Cauchy-Schwarz inequality ...

- What the Theorem says is:
  - Wavelet coefficients for smooth functions decay rapidly as the resolution parameter j increases!
  - This is (as ii. shows) a strictly local phenomenon:
    If a function $f(t)$ is N-times continuously differentiable at some point $t_0 \in \mathbb{R}$, hence $f^{(N)}(t)$ is continuous in some interval J containing t_0, the estimate ii. holds for any j, k for which $I_{j,k} \subseteq J$
  - This means: wavelet coefficients belonging to smooth parts of a signal are usually much smaller than wavelet coefficients for non-smooth parts
  - This has important practical consequences, e.g., for image compression