Note on “A Remez Exchange Algorithm for Orthonormal Wavelets”

Olivier Rioul and Pierre Duhamel *

April 1994

Abstract

The purpose of this note is to give a rigorous prrof of convergence of the modified Remez exchange algorithm proposed in the paper entitled “A Remez exchange algorithm for orthonormal wavelets,” by O. Rioul and P. Duhamel, *IEEE Trans. Circuits Systems II*, vol. 41, no. 8, Aug. 1994, pp. 550–560.

1 Introduction

Let us recall the two optimization problems described in the paper and introduce useful notations. The pass-band is denoted by $BP = [0, \omega_p]$. For variable $y = \cos^2 \omega$, we have $y \in I \iff \omega \in BP$. For any continuous function $f(\omega)$ depending on variable ω we write $f(y)$ the dependency of f on variable y. The ambiguity should be easily resolved from the context. The following functional norm is used:

$$\|f\| = \max_{\omega \in BP} |f(\omega)| = \max_{y \in I} |f(y)|$$

1.1 Problem # 1

This is the initial optimization problem presented in section II.A of the paper: Given L, $0 < K \leq L/2$, and transition bandwidth BP, find the best trigonometric polynomial

$$P(\omega) = 1 + \sum_{n=1}^{L/2} a_n \cos(2n-1)\omega$$

($L/2$ variables a_n) such that δ (the tolerance in the pass-band) is minimized subject to the constraints of magnitude specification

$$\|2 - \delta - P\| \leq \delta$$

*Olivier Rioul is with France Télécom, Centre National d’Études des Télécommunications, PAB/STC/SGV, 38–40 rue du Général Leclerc, 92131 Issy-Les-Moulineaux, France. Pierre Duhamel is with Ecole Nationale Supérieure des Télécommunications, ENST/SIG, 46 rue Barrault, 75013 Paris, France.

and Kth order flatness (equations (10)-(11) in the paper). The latter constraints leaves

$$N = L/2 - K$$

independent variables a_n.

Since this is a classical linear program whose set of constraints is not empty (maximally flat solution $P_K(\omega)$ satisfies the constraints for some $\delta > 0$) optimal solutions $\bar{P}(\omega)$ exist. Note: $\bar{P}(\omega)$ is *a priori* not unique. Any optimal solution $\bar{P}(\omega)$ has pass-band tolerance equal to

$$\bar{\delta} = \min_{P(\omega)} \delta$$

which is unique by definition.

1.2 Problem # 2

Reformulation of problem # 1 is done in section III of the paper. The main difference with problem # 1 is that δ is fixed to some value. Given this value of δ, problem # 2 is to find the best $P(\omega)$ which minimizes $\|2 - \delta - P\|$ subject to the K flatness constraints.

As explained in the paper, this is equivalent to the following.

$$\min_{R(y)} \|E\|$$

where $R(y)$ is a polynomial of degree $N - 1$ and

$$E(y) = W(y)(D(y) - R(y)).$$

Here

$$D(y) = \frac{D_0(y) - \delta}{W(y)}$$

where $D_0(y)$ and $W(y) \geq 0$ are continuous functions of y and do not depend on δ.

From section IV of the paper, to each value of δ corresponds

- A unique optimal solution $P^*(\omega)$ or $R^*(y)$ characterized by $N + 1$ alternations $y1 < y2 < \cdots < y_{N+1}$ in I:

$$E(y_i) = \pm(-1)^i \|E\|$$

- A number

$$\delta^* = \|E^*\| = \min_{R(y)} \|E\| = \min_{P(w) flat} \|2 - \delta - P\|$$

Note that implicitly there is an application '*' that maps $\delta > 0$ to $\delta^* > 0$.

Keep in mind notations $\bar{P}$, $\bar{\delta}$, P^*, E^*, δ^*. They will be constantly used in this note.

2 Preliminaries

The issue of the paper is to solve problem # 1 using a Remez exchange algorithm for problem # 2. In this section we state the precise connections between these two problems.

Lemma 1 *For problem # 2 we always have*

$$\delta \leq \delta^*$$

Proof: Because of the flatness constraints we always have $P(\omega = 0) = 2$, hence

$$\delta^* = \|2 - \delta - P^*\| \geq |2 - \delta - P^*(0)| = \delta$$

$\square$

Proposition 1 *$\bar{P}(\omega)$, the optimal solution for problem # 1, is unique. It is also the optimal solution for problem # 2 where δ is set to $\bar{\delta}$, and we have $\bar{\delta} = \bar{\delta}^*$.*

Proof: We have $\bar{\delta} \leq \bar{\delta}^*$ from the preceding lemma. But $\bar{\delta}^* = \min_{P(\omega)} \|2 - \bar{\delta} - P\| \leq \|2 - \bar{\delta} - \bar{P}\|$ and the latter quantity is $\leq \bar{\delta}$ because of the constraints of problem # 1 for $\bar{P}(\omega)$. Hence $\bar{\delta}^* = \bar{\delta}$ and $\|2 - \bar{\delta} - \bar{P}\| = \min_{P(\omega)} \|2 - \bar{\delta} - P\|$ for any optimal solution $\bar{P}$ of problem # 1. This means that $\bar{P}(\omega)$ is the optimal solution to problem # 2 for $\delta = \bar{\delta}$, and is therefore unique. $\square$

Proposition 2 *Let $P^*(\omega)$ be the solution to problem # 2 for some δ. Then it satisfies the constraints of problem # 1 if and only if $\delta = \delta^*$. In particular, $\delta \geq \bar{\delta}$.*

Proof: If P^* satisfies the constraints of problem # 1 then $\delta \geq \bar{\delta}$ and

$$\begin{aligned}
\delta^* = \min_{P(\omega)} \|2 - \delta - P\| &\leq \|2 - \delta - \bar{P}\| \\
&\leq \|2 - \bar{\delta} - \bar{P}\| + \delta - \bar{\delta} \\
&\leq \bar{\delta}^* + \delta - \bar{\delta}
\end{aligned}$$

but $\bar{\delta}^* = \bar{\delta}$ from the preceding lemma so we end up with $\delta^* \leq \delta$. But from the first lemma $\delta^* \geq \delta$, so $\delta^* = \delta$.

Conversely, assume $\delta = \delta^*$, i.e., $\delta = \|2 - \delta - P^*\|$. Then because of this equality P^* satisfies the constraints of problem # 1. $\square$

This result shows that there is hope in solving problem # 1 using problem # 2 provided the optimal solution is such that $\delta = \delta^*$.

3 Modified Remez algorithm

Let us summarize the proposed algorithm as it was described in section VI of the paper.

At the nth iteration, we are given $N+1$ critical points y_i^n, $i = 0 ,\ldots, N$. Then equation

$$E_n(y_i^n) = \pm(-1)^i \delta_n$$

where

$$E_n(y) = W(y)(D_n(y) - R(y))$$

and

$$D_n(y)W(y) = D_0(y) - \delta_n$$

suffices to determine δ_n and $R(y) = R_n(y)$ *uniquely.*

Note that this does *not* mean that $R_n(y)$ is the optimal solution to problem # 2 for $\delta = \delta_n$, since $\delta_n < \|E_n(y)\|$ in general.

From here a multiple exchange procedure gives the next critical points y_i^{n+1} in such a way that for all i,

$$E_n(y_i^{n+1}) \geq \delta_n$$

and there exists i_0 such that

$$|E_n(y_0^{n+1})| = \|E_n\|.$$

From here another iteration starts.

The purpose of this note is to show that

1. Convergence holds to $R_\infty(y) = \lim_{n\to\infty} R_n(y)$, corresponding to $P_\infty(\omega)$ whose tolerance in the pass-band is $\delta_\infty = \lim_{n\to\infty} \delta_n$.

2. At convergence, we have $\delta_\infty = \delta_\infty^* = \bar{\delta}$ hence the obtained solution $P_\infty(\omega)$ is indeed the optimal solution $\bar{P}(\omega)$ of initial problem # 1.

4 Analysis of convergence

Lemma 2 *Let $R_n^*(y)$ be the optimal solution of problem # 2 for $\delta = \delta_n$. If $\delta_n = \delta_n^*$ then $R_n(y) = R_n^*(y)$.*

Proof: We have seen that $R_n(y)$ is determined by equation $E_n(y_i^n) = \pm(-1)^i\delta_n$. where $\delta_n = |E_n(y_i^n)|$. Now, this is exactly the alternation theorem for the "discrete" problem

$$\min_{R(y)} \max_{y\in\{y_i^n\}} |E_n(y)|$$

Indeed Chebyshev's alternation theorem still applies for $I = \{y_i^n\}$ where all y_i^n are alternations! Therefore, $R_n(y)$ is the *unique* solution to the discrete problem and we have

$$\begin{aligned}\delta_n &= \min_{R(y)} \max_i |E_n(y_i^n)| \\ &\leq \max_i |E_n^*(y_i^n)| \\ &\leq \|E_n^*\| = \delta_n^*.\end{aligned}$$

Since $\delta_n = \delta_n^*$ it follows that $\min_{R(y)} \max_i |E_n(y_i^n)| = \max_i |E_n^*(y_i^n)|$. This means that $R_n^*(y)$ is also the optimal solution to the discrete problem, hence by uniqueness $R_n(y) = R_n^*(y)$. □

Proposition 3 *As long as we did not converge, $\delta_n < \bar{\delta}$, hence δ_n is bounded for all n. Moreover $\delta_n < \|E_n\|$.*

Proof: Suppose $\delta_n \geq \bar{\delta}$. From proposition II.3, this implies $\delta_n = \delta_n^*$. Then by the preceding lemma we would have $R_n(y) = R_n^*(y)$ and therefore $\delta_n = \delta_n^* = \|E_n^*\| = \|E_n\|$. But then in the multiple exchange procedure we would find $y_i^{n+1} = y_i^n$: the critical points, hence δ_n and $R_n(y)$ are stationary, which means that we have converged (in a finite number of steps).

Since from the preceding discussion $\delta_n < \delta_n^*$, and $\delta_n^* = \|E_n^*\| \leq \|E_n\|$ it follows that $\delta_n < \|E_n\|$. □

Lemma 3 *There exist $N+1$ numbers $\lambda_i \geq 0$ satisfying $\sum_i \lambda_i = 1$ such that*

$$\delta_n = \sum_{i=0}^{N} \lambda_i |E_n(y_i^n)|$$

where $E_n(y) = W(y)(D_n(y) - R(y))$, for any *polynomial $R(y)$ of degree $\leq N-1$.*

Proof: This follows, of course, from the definition of δ_n if $R(y) = R_n(y)$. This lemma states that $R_n(y)$ can in fact be replaced by any polynomial $R(y)$ of degree $\leq N-1$. Since by definition, $E_n(y_i^n)$ has alternating signs, if suffices to choose λ_i such that $\sum_i (-1)^i \lambda_i W(y_i) R(y_i) = 0$ for any $R(y)$.

Using Lagrangian interpolation formula we have $R(y_0^n) = \sum_{i=1}^N L_i(y_0^n) R(y_i^n)$ for any $R(y)$ of degree $\leq N-1$, where $L_i(y) = \prod_{j \neq i} \frac{y - y_i}{y_j - y_i}$. Set $a_i = \prod_{j \neq i} (y_j^n - y_i^n)$ where index j goes from 0 to N. Then $L_i(y_0^n) = -a_0/a_i$ and we have $\sum_{i=0}^N (1/a_i) R(y_i^n) = 0$ for any $R(y)$. A solution is given by $\lambda_i = |\mu_i|/(\sum |\mu_i|)$ where $1/\mu_i = a_i W(y_i^n)$ has the same sign as $\pm(-1)^i$ since a_i's have alternating signs and $W(y) \geq 0$. □

Remark. The equation used for δ_n in the paper follows from this derivation by setting $R(y) \equiv 0$.

Proposition 4 *As long as we did not converge, $\delta_n < \delta_{n+1}$. From the lemma IV. 2, it follows that δ_n, a bounded increasing sequence, converges as $n \to \infty$.*

Proof: Use the preceding lemma for the expression of δ_{n+1}, where we set $R(y) = R_n(y)$. We obtain

$$\delta_{n+1} = \sum_i \pm(-1)^i \lambda_i W(y_i^{n+1})(D_{n+1}(y_i^{n+1}) - R_n(y_i^{n+1}))$$

Since $W(y)D_{n+1}(y) = W(y)D_n(y) + \delta_n - \delta_{n+1}$, we obtain

$$\delta_{n+1} = \sum_i \lambda_i |E_n(y_i^{n+1})| + (\delta_n - \delta_{n+1}) \sum_i \varepsilon_i \lambda_i$$

where $\varepsilon_i = \pm(-1)^i$.

After multiple exchange described in section III, we have

$$\delta_{n+1} \geq \delta_n + \lambda_{i_0}(\|E_n\| - \delta_n) + (\delta_n - \delta_{n+1}) \sum_i \varepsilon_i \lambda_i$$

It follows that

$$\delta_{n+1} - \delta_n \geq \alpha_{n+1}(\|E_n\| - \delta_n)$$

where $\alpha_{n+1} = \lambda_{i_0} / \sum_i (1 - \varepsilon_i)\lambda_i > 0$. By proposition IV.2, $\|E_n\| > \delta_n$, so the proof is complete. □

5 Finding the solution to the initial problem

Let $\delta_\infty = \lim_{n\to\infty} \delta_n$. From propositions IV.2 and 4 this limit exists, is positive and is $\leq \bar{\delta}$. In the sequel we prove that $\delta_\infty = \bar{\delta}$, and the final result will follow.

Lemma 4 *Critical points y_i^n always stay within a certain distance to each other as $n \to \infty$. That is,* $\inf |y_{i+1}^n - y_i^n| > 0$.

Proof: Otherwise there would be a converging subsequence of y_i^n, which we denote by y_i^m, whose limit is $\hat{y}_i$, such that $\hat{y}_i = \hat{y}_{i+1}$. Hence for $i = 0, \ldots, N$, there are at most N distinct values in the set $\{\hat{y}_i\}$. Therefore, there exists $\hat{R}(y)$, a polynomial of degree $\leq N - 1$, such that $\hat{R}(\hat{y}_i) = D_\infty(\hat{y}_i)$, i.e., $\hat{E}_\infty(\hat{y}_i) = 0$.

Then given arbitrarily small ε and for m large enough:

- $E_m(y_i^m) = \pm(-1)^i \delta_m$ where $\delta_m > 2\varepsilon$ (since $\delta_\infty > 0$).
- $|\hat{E}_\infty(y_i^m)| < \varepsilon$ since $\hat{E}_\infty(y)$ is continuous and $\hat{E}_\infty(\hat{y}_i) = 0$.
- $|\delta_\infty - \delta_m| < \varepsilon$ since $\delta_m \to \delta_\infty$.

Therefore,

$$R_m(y_i^m) - \hat{R}(y_i^m) = \frac{1}{W(y_i^m)}(\hat{E}_\infty(y_i^m) - E_m(y_i^m) + \delta_\infty - \delta_m)$$

has same sign as $\pm(-1)^i$. The polynomial $R_m(y) - \hat{R}(y)$ oscillates at $N+1$ distinct points y_i^m, hence it has N distinct zeroes. Since it is of degree $N-1$, we must have $\hat{R}(y) = R_m(y)$. But this is impossible since it implies that $\delta_m = |E_m(y_i^m)| = |\hat{E}_m(y_i^m)| < \varepsilon$, hence $\delta_m \to \delta_\infty = 0$ whereas δ_m is strictly increasing. □

Theorem 1 *We have $\delta_\infty = \bar{\delta}$ and $P_n(\omega)$ in the modified Remez exchange algorithm converges to $\bar{P}(\omega)$, the optimal solution to initial problem # 1.*

Proof: From the proof of proposition IV.4 we have $(\|E_n\| - \delta_n) \leq \frac{1}{\alpha_{n+1}}(\delta_{n+1} - \delta_n)$ where from the expression giving α_{n+1} and from the preceding lemma, there exists $\alpha > 0$ such that $\alpha_{n+1} \geq \alpha > 0$. This shows that $(\|E_n\| - \delta_n)$ tends to zero, hence $\|E_n\| = \|2 - \delta_n - P_n(\omega)\| \to \delta_\infty$.

Now let $P_m(\omega)$ be a converging subsequence of $P_n(\omega)$, whose limit is denoted by $P_\infty(\omega)$. From the preceding discussion we have $\|2 - \delta_\infty - P_\infty(\omega)\| = \delta_\infty$, hence $\delta_\infty^* = \min_{P(\omega)} \|2 - \delta_\infty - P(\omega)\| \leq \delta_\infty$. Therefore, by lemma II.1, $\delta_\infty = \delta_\infty^*$. This implies $\delta_\infty \geq \bar{\delta}$ by proposition II.3. But since $\delta_\infty \leq \bar{\delta}$, we obtain $\delta_\infty = \bar{\delta}$.

Moreover $\delta_\infty = \delta_\infty^*$ can be rewritten as $\|2 - \bar{\delta} - P_\infty\| = \min_{P(\omega)} \|2 - \bar{\delta} - P\|$ which shows that $P_\infty(\omega)$ is the optimal solution of problem # 2 for $\delta = \bar{\delta}$, hence by proposition II.2, $P_\infty(\omega) = \bar{P}(\omega)$. Thus, we have shown that any converging subsequence of $P_n(\omega)$ converges to $\bar{P}(\omega)$.

Now if $P_n(\omega)$ did not converge, there would exist a subsequence $P_m(\omega)$ such that $\|\bar{P}(\omega) - P_m(\omega)\| \geq \varepsilon > 0$ for any m large enough. But from $P_m(\omega)$ we could extract a converging subsequence whose limit would be $\bar{P}(\omega)$, and we would have a contradiction: $\|\bar{P}(w) - \bar{P}(w)\| = 0 \geq \varepsilon > 0$. Therefore the whole sequence $P_n(\omega)$ converges to $\bar{P}(\omega)$. □