GRADIENT CONVERSION BETWEEN TIME AND FREQUENCY DOMAINS USING WIRTINGER CALCULUS

Hugo Caracalla

Analysis-Synthesis Team
IRCAM
Paris, France
hugo.caracalla@ircam.fr

Axel Roebel

Analysis-Synthesis Team
IRCAM
Paris, France
axel.roebel@ircam.fr

ABSTRACT

Gradient-based optimizations are commonly found in areas where Fourier transforms are used, such as in audio signal processing. This paper presents a new method of converting any gradient of a cost function with respect to a signal into, or from, a gradient with respect to the spectrum of this signal: thus, it allows the gradient descent to be performed indiscriminately in time or frequency domain. For efficiency purposes, and because the gradient of a real function with respect to a complex signal does not formally exist, this work is performed using Wirtinger calculus. An application to sound texture synthesis then experimentally validates this gradient conversion.

1. INTRODUCTION

Mathematical optimization is a recurring theme in audio signal processing: many sound synthesis algorithms, for instance physics-based synthesis [1] or sound texture synthesis [2, 3], require the solving of non-linear equations, which in turn becomes a function optimization problem. Formally, this means seeking the minimum of a real-valued cost function $\mathcal{C}$ over the ensemble of its inputs.

From here on several solvers exist, including the widely used gradient descent algorithms. As the name implies, this kind of algorithm requires the computation of the gradient of $\mathcal{C}$ to iteratively progress toward a minimum. But a problem arises if the parameters of $\mathcal{C}$ are complex-valued: from the Cauchy-Riemann equations follows that a real-valued non-constant function with complex parameters is not differentiable. This means that if the parameters of our cost function $\mathcal{C}$ are complex-valued, for example when the cost is evaluated from a spectrum, the gradient of $\mathcal{C}$ does not exist.

Recently, this situation was encountered in the context of sound texture synthesis. In this synthesis algorithm the sub-bands envelopes of a white noise signal are adapted so that a set of their statistics reach some desired values, previously extracted from a sound texture. Current implementations of this algorithm either perform the optimization over each sub-band envelope individually while performing additional steps in parallel to compute a corresponding time signal [2], or perform it over the magnitude of the short-term Fourier transform of the white noise signal, combined with a phase retrieval algorithm allowing to recreate the synthesized time signal [3]. But both methods require additional steps to recreate the outputted signal, be it reconstruction from the sub-bands or phase retrieval, which in turn tend to damage the optimization performed over the sub-bands envelopes.

Such a problem could be avoided altogether if the optimization was performed directly over the base time signal, although this would mean that the gradient of the cost function would have to be calculated with respect to said time signal. Since this cost is defined over the envelopes of the sub-band signals statistics, computing its gradient with respect to those envelopes is usually straightforward, but converting this gradient in a gradient with respect to the base time signal means going back up the whole envelope extraction process. Because this procedure involves manipulating spectra and because the sub-band signals are complex-valued, this situation typically falls in the case mentioned above where the formal gradients of the cost function with respect to those complex-valued signals do not exist.

Hence the goal of the work presented in this paper is twofold: establishing a relation between a gradient of a cost function with respect to a signal and with respect to its spectrum, all the while using a formalism that both allows the manipulation of those otherwise non-existent complex gradients and does so in the most efficient way possible.

The formalism chosen to be worked with is Wirtinger calculus, first introduced in [4], which offers an extension of complex differentiability in the form of Wirtinger derivatives and Wirtinger gradients. Even though it can be encountered in articles such as [5] (although without any mention of the name Wirtinger), it has been scarcely used in signal processing since then. Because this formalism is not well known, this paper starts off with an introduction to it, followed by our work on gradient conversion between time and frequency domains, and then the application and experimental validation of this conversion to the case of sound texture synthesis.

2. WIRTINGER CALCULUS AND NOTATIONS

2.1. General mathematical notations

Let us first introduce the mathematical notations that are used throughout this paper.

Arrays are denoted by bold lower case letters as $\boldsymbol{c}$, while their m-th element is denoted c_m.

For a differentiable function $f\colon \mathbb{R} \to \mathbb{C}$, its real derivative at $a \in \mathbb{R}$ is denoted:

$$\frac{\partial f}{\partial x}(a) \tag{1}$$

For a differentiable function $f\colon \mathbb{R}^M \to \mathbb{C}$, its real gradient at $\boldsymbol{a} \in \mathbb{R}^M$ is denoted:

$$\frac{\partial f}{\partial \boldsymbol{x}}(a) = \left(\frac{\partial f}{\partial x_1}(a), \frac{\partial f}{\partial x_2}(a), ..., \frac{\partial f}{\partial x_M}(a)\right) \tag{2}$$

The discrete Fourier transform (DFT) is denoted by $\mathcal{F}\colon \mathbb{C}^M \to \mathbb{C}^N$, with $M \leq N$, while the inverse DFT (iDFT) is denoted by $\mathcal{F}^{-1}\colon \mathbb{C}^N \to \mathbb{C}^M$. For a vector $\boldsymbol{c} \in \mathbb{C}^M$, its spectrum is denoted by a bold upper case letter $\boldsymbol{C} \in \mathbb{C}^N$. As such we have:

$$C_n = \mathcal{F}(c)_n = \sum_{m\in[0,M-1]} c_m e^{-j\frac{2\pi mn}{N}} \tag{3}$$

And:

$$c_m = \mathcal{F}^{-1}(C)_m = \frac{1}{N}\sum_{n\in[0,N-1]} C_n e^{j\frac{2\pi mn}{N}} \tag{4}$$

2.2. Wirtinger Calculus

We now introduce Wirtinger calculus and summarize some of its more useful properties, as can be found in [5, 6, 7].

2.2.1. *Wirtinger derivatives and gradients*

Any complex number $c \in \mathbb{C}$ can be decomposed as $a + jb$ with $(a, b) \in \mathbb{R}^2$ its real and imaginary parts. Similarly, any function $f\colon \mathbb{C} \to \mathbb{C}$ can be considered as a function of $\mathbb{R}^2 \to \mathbb{C}$ with $f(c) = f(a, b)$. If it exists, the derivative of f at c with respect to the real part of its input is denoted by:

$$\frac{\partial f}{\partial x}(c) \tag{5}$$

This concords with our previous notations, since this derivative is also the real derivative of f when its domain is restrained to $\mathbb{R}$. If it exists, the derivative of f at c with respect to the imaginary part of its input is denoted by:

$$\frac{\partial f}{\partial y}(c) \tag{6}$$

If f is differentiable with respect to both the real and imaginary part of its input we call it differentiable in the real sense. This property is weaker than $\mathbb{C}$-differentiability since it lacks the Cauchy-Riemann conditions, but is sufficient when looking to optimize f since it means we could always manipulate it as a function of $\mathbb{R}^2 \to \mathbb{C}$, whose optimization does not require any differentiability over $\mathbb{C}$.

This is where Wirtinger calculus intervenes: it is a way of manipulating both partial derivatives of a function of $\mathbb{C} \to \mathbb{C}$ that is only differentiable in the real sense without going through the trouble of treating them individually. In addition to this, Wirtinger calculus acts as a bridge toward $\mathbb{C}$-differentiability since it overlaps with $\mathbb{C}$-derivation when the complex derivative exists.

For $f\colon \mathbb{C} \to \mathbb{C}$ differentiable in the real sense, its Wirtinger derivative (henceforth W-derivative) at $c \in \mathbb{C}$ is denoted by and defined as:

$$\frac{\partial f}{\partial z}(c) = \frac{1}{2}\left(\frac{\partial f}{\partial x}(c) - j\frac{\partial f}{\partial y}(c)\right) \tag{7}$$

While its conjugate Wirtinger derivative (henceforth W^*-derivative) is denoted by and defined as:

$$\frac{\partial f}{\partial z^*}(c) = \frac{1}{2}\left(\frac{\partial f}{\partial x}(c) + j\frac{\partial f}{\partial y}(c)\right) \tag{8}$$

Manipulating those two derivatives is equivalent to manipulating the two real partial derivatives, and we can with ease switch from one expression to the other if need be: since the partial derivatives are enough to optimize a function, so is the Wirtinger derivative.

The case can then be extended to functions of several variables, such as arrays: if f denotes a function of $\mathbb{C}^M \to \mathbb{C}$ and is differentiable in the real sense, meaning here differentiable with respect to the real and imaginary parts of all of its inputs, we define the W and W^*-gradients of f similarly to their real counterpart:

$$\frac{\partial f}{\partial \boldsymbol{z}}(c) = \left(\frac{\partial f}{\partial z_1}(c), \frac{\partial f}{\partial z_2}(c), ..., \frac{\partial f}{\partial z_M}(c)\right) \tag{9}$$

$$\frac{\partial f}{\partial \boldsymbol{z}^*}(c) = \left(\frac{\partial f}{\partial z_1^*}(c), \frac{\partial f}{\partial z_2^*}(c), ..., \frac{\partial f}{\partial z_M^*}(c)\right) \tag{10}$$

Once more, and as shown in [6], in the case of a real-valued function (such as a cost function) knowing either the W or W^*-gradient of the function is sufficient to minimize it.

2.2.2. *Link with $\mathbb{C}$-differentiability*

If f denotes a function of $\mathbb{C}^{\mathbb{M}} \to \mathbb{C}$ the following property holds:

$$f \text{ is } \mathbb{C}\text{-differentiable} \iff \begin{cases} f \text{ is differentiable in the real sense} \\ \frac{\partial f}{\partial \boldsymbol{z}^*} = \boldsymbol{0} \end{cases} \tag{11}$$

In the case where f is $\mathbb{C}$-differentiable, both its complex and W-gradients are equal: this is what makes of Wirtinger calculus a proper extension of $\mathbb{C}$-differentiability.

2.2.3. *Linearity*

The W and W^*-derivations are both linear, meaning for f and g two functions of $\mathbb{C}^M \to \mathbb{C}$ differentiable in the real sense and for $(\alpha, \beta) \in \mathbb{C}^2$:

$$\frac{\partial(\alpha f + \beta g)}{\partial \boldsymbol{z}} = \alpha\frac{\partial f}{\partial \boldsymbol{z}} + \beta\frac{\partial g}{\partial \boldsymbol{z}} \tag{12}$$

$$\frac{\partial(\alpha f + \beta g)}{\partial \boldsymbol{z}^*} = \alpha\frac{\partial f}{\partial \boldsymbol{z}^*} + \beta\frac{\partial g}{\partial \boldsymbol{z}^*} \tag{13}$$

2.2.4. *Function composition*

For f and g two functions of $\mathbb{C} \to \mathbb{C}$ differentiable in the real sense, the Wirtinger chain rule gives:

$$\frac{\partial f\circ g}{\partial z} = \left(\frac{\partial f}{\partial z}\circ g\right) \times \frac{\partial g}{\partial z} + \left(\frac{\partial f}{\partial z^*}\circ g\right) \times \frac{\partial g^*}{\partial z} \tag{14}$$

$$\frac{\partial f\circ g}{\partial z^*} = \left(\frac{\partial f}{\partial z}\circ g\right) \times \frac{\partial g}{\partial z^*} + \left(\frac{\partial f}{\partial z^*}\circ g\right) \times \frac{\partial g^*}{\partial z^*} \tag{15}$$

We now extend this in the case of functions of several variables: if this time f denotes a function of $\mathbb{C}^M \to \mathbb{C}$ and g denotes a function of $\mathbb{C}^N \to \mathbb{C}^M$, both being differentiable in the real sense, for $n \in [1, N]$ the chain rule gives:

$$\frac{\partial f\circ g}{\partial z_n} = \sum_{m\in[1,M]} \left(\frac{\partial f}{\partial z_m}\circ g\right) \times \frac{\partial g_m}{\partial z_n} + \left(\frac{\partial f}{\partial z_m^*}\circ g\right) \times \frac{\partial g_m^*}{\partial z_n} \tag{16}$$

$$\frac{\partial f\circ g}{\partial z_n^*} = \sum_{m\in[1,M]} \left(\frac{\partial f}{\partial z_m}\circ g\right) \times \frac{\partial g_m}{\partial z_n^*} + \left(\frac{\partial f}{\partial z_m^*}\circ g\right) \times \frac{\partial g_m^*}{\partial z_n^*} \tag{17}$$

2.2.5. Complex conjugate

If f denotes a function differentiable in the real sense, the following property holds:

$$\left(\frac{\partial f}{\partial \boldsymbol{z}}\right)^* = \frac{\partial f^*}{\partial \boldsymbol{z}^*} \tag{18}$$

This straightforwardly implies that if f is real-valued we have:

$$\left(\frac{\partial f}{\partial \boldsymbol{z}}\right)^* = \frac{\partial f}{\partial \boldsymbol{z}^*} \tag{19}$$

Meaning that in the case of functions with real output, such as cost functions, it is strictly equivalent to manipulate the W and the W^*-gradients.

3. CONVERSION BETWEEN TIME AND FREQUENCY DOMAINS

Those properties can now be put to use in order to convert the W-gradients of a real-valued cost function between time and frequency domains.

3.1. From time to frequency

Let us suppose a cost function $\mathcal{E}$, differentiable in the real sense and defined as:

$$\begin{aligned} \mathcal{E}\colon \mathbb{C}^M &\to \mathbb{R} \\ z &\mapsto \mathcal{E}(z) \end{aligned} \tag{20}$$

The value of the W-gradient of $\mathcal{E}$ is supposed known at a given point $\boldsymbol{c} \in \mathbb{C}^M$. Our goal is now to evaluate this gradient at $\boldsymbol{C} \in \mathbb{C}^N$, the DFT of $\boldsymbol{c}$. More rigorously, this amounts to say that we wish to evaluate at $\boldsymbol{C}$ the W-gradient of $\tilde{\mathcal{E}}$, defined as:

$$\begin{aligned} \tilde{\mathcal{E}}\colon \mathbb{C}^M &\to \mathbb{R} \\ z &\mapsto \mathcal{E}(\mathcal{F}^{-1}(z)) \end{aligned} \tag{21}$$

According to the chain rule of Wirtinger calculus stated in 2.2.4 we have for $n \in [0, N-1]$:

$$\begin{aligned} \frac{\partial \tilde{\mathcal{E}}}{\partial z_n}(\boldsymbol{C}) = \sum_{m\in[1,M]} &\left(\frac{\partial \mathcal{E}}{\partial z_m}(\mathcal{F}^{-1}(\boldsymbol{C})\right) \times \frac{\partial \mathcal{F}_m^{-1}}{\partial z_n}(\boldsymbol{C}) \\ &+ \left(\frac{\partial \mathcal{E}}{\partial z_m^*}(\mathcal{F}^{-1}(\boldsymbol{C}))\right) \times \frac{\partial \left(\mathcal{F}_m^{-1}\right)^*}{\partial z_n}(\boldsymbol{C}) \end{aligned} \tag{22}$$

But $\mathcal{F}^{-1}$ is $\mathbb{C}$-differentiable, meaning that according the property of Wirtinger calculus stated in 2.2.2 its W^*-gradient is null. From (18) then follows that the W-gradient of $(\mathcal{F}^{-1})^*$ is also null, which leaves us with:

$$\frac{\partial \tilde{\mathcal{E}}}{\partial z_n}(\boldsymbol{C}) = \sum_{m\in[1,M]} \frac{\partial \mathcal{E}}{\partial z_m}(\boldsymbol{c}) \times \frac{\partial \mathcal{F}_m^{-1}}{\partial z_n}(\boldsymbol{C}) \tag{23}$$

From (4) we easily derive:

$$\begin{aligned} \frac{\partial \mathcal{F}_m^{-1}}{\partial z_n}(\boldsymbol{C}) &= \frac{1}{N}\frac{\partial}{\partial z_n}\left(\sum_k z_k e^{j\frac{2\pi mk}{N}}\right)\Bigg|_{\boldsymbol{z}=\boldsymbol{C}} \\ &= \frac{1}{N} e^{j\frac{2\pi mn}{N}} \end{aligned} \tag{24}$$

Which re-injected in (23) gives:

$$\frac{\partial \tilde{\mathcal{E}}}{\partial z_n}(\boldsymbol{C}) = \frac{1}{N} \sum_{m\in[1,M]} \frac{\partial \mathcal{E}}{\partial z_m}(\boldsymbol{c}) e^{j\frac{2\pi mn}{N}} \tag{25}$$

$$= \frac{1}{N}\left[\sum_{m\in[1,M]} \left(\frac{\partial \mathcal{E}}{\partial z_m}(\boldsymbol{c})\right)^* e^{-j\frac{2\pi mn}{N}}\right]^* \tag{26}$$

Here we recognize a DFT, leading to the expression of the whole W-gradient of $\tilde{\mathcal{E}}$ as:

$$\frac{\partial \tilde{\mathcal{E}}}{\partial \boldsymbol{z}}(\boldsymbol{C}) = \frac{1}{N}\mathcal{F}\left(\left[\frac{\partial \mathcal{E}}{\partial \boldsymbol{z}}(\boldsymbol{c})\right]^*\right)^* \tag{27}$$

Please note that since $M \leq N$, the sum over m in (25) could not be interpreted as an inverse DFT, resulting in the identification of the conjugate of the DFT in (26).
But since $\mathcal{E}$ (and $\tilde{\mathcal{E}}$) are real-valued functions, according to (19) the previous expression can be formulated in a cleaner way:

$$\frac{\partial \tilde{\mathcal{E}}}{\partial \boldsymbol{z}^*}(\boldsymbol{C}) = \frac{1}{N}\mathcal{F}\left(\frac{\partial \mathcal{E}}{\partial \boldsymbol{z}^*}(\boldsymbol{c})\right) \tag{28}$$

In other words, knowing the W^* (or equivalently the W)-gradient of a cost function at a given point $\boldsymbol{c}$ in time domain, it is possible to convert it over frequency domain to obtain the gradient at the spectrum $\boldsymbol{C}$.

3.2. From frequency to time

Alternatively, the expression (28) can also be reversed to give:

$$\frac{\partial \mathcal{E}}{\partial \boldsymbol{z}^*}(\boldsymbol{c}) = N\mathcal{F}^{-1}\left(\frac{\partial \tilde{\mathcal{E}}}{\partial \boldsymbol{z}^*}(\boldsymbol{C})\right) \tag{29}$$

Which this time allows us to transition from the gradient of the error function at $\boldsymbol{C}$ to the gradient at $\boldsymbol{c}$.

Thus, it is now possible to straightforwardly convert the gradient of a cost function between time and frequency domains: this also means that any gradient descent can be chosen to be performed on a signal or its spectrum, independently of which signal was used in cost computations. Additionally, and since this conversion is performed using only DFTs or iDFTs, we can also make full use of efficient Fourier transform algorithms such as the Fast Fourier Transform (FFT).

We now apply and experimentally validate the present results in a sound synthesis algorithm.

4. APPLICATION

As mentioned in the introduction, the conversion of a gradient between time and frequency domains can be put to use in the case of sound texture synthesis through statistics imposition.

During the synthesis algorithm a base signal such as a white noise is converted to frequency domain where it is filtered using a Fourier multiplier, then taken back to time domain to result in an analytic sub-band of the base signal. The goal of the algorithm is then to impose a given set of values to some selected statistics of those sub-bands magnitudes. Since imposing the statistics via gradient descent directly over the sub-bands and then proceeding on recreating the corresponding time signal would damage the imposition, we seek to perform the gradient descent directly over the base time signal. This requires a conversion of the gradient of the cost function, easily defined with respect to the sub-bands, to a gradient with respect to the base signal.

The case can be formulated this way: we call $\boldsymbol{s}$ a real-valued array of length M representing the base sound signal we wish to alter using gradient descent and $\boldsymbol{S}$ its DFT of length N. $\boldsymbol{S}$ is then windowed in frequency domain by a function $\mathcal{G}$ defined as:

$$\begin{aligned} \mathcal{G} \colon \mathbb{C}^N &\to \mathbb{C}^N \\ \boldsymbol{S} &\mapsto \boldsymbol{W}.\boldsymbol{S} \end{aligned} \tag{30}$$

With $\boldsymbol{W} \in \mathbb{C}^N$ the spectral weighting array used for band-filtering $\boldsymbol{S}$ and . the element-wise product. We denote the filtered spectrum $\boldsymbol{B} = \mathcal{G}(\boldsymbol{S})$, and $\boldsymbol{b} = \mathcal{F}^{-1}(\boldsymbol{B})$ its inverse DFT of length M: $\boldsymbol{b}$ is thus the complex sub-band of $\boldsymbol{s}$ centered around a frequency chosen via $\boldsymbol{W}$. In this example we want to impose a given value γ to the third order moment of the envelope of $\boldsymbol{b}$, so the cost function $\mathcal{C}$ is chosen as the squared distance between the third order moment of the envelope of $|\boldsymbol{b}|$ the sub-band and γ:

$$\begin{aligned} \mathcal{C} \colon \mathbb{C}^M &\to \mathbb{R} \\ \boldsymbol{b} &\mapsto \left[\left(\sum_{k \in [0, M-1]} |b_k|^3 \right) - \gamma \right]^2 \end{aligned} \tag{31}$$

For simplicity's sake we will consider the raw moment instead of the usual standardized moment that can be found in [2]. Using the rules of Wirtinger's calculus detailed in Section 2.2.4 we straightforwardly obtain the W^*-gradient of $\mathcal{C}$ with respect to the sub-band at $\boldsymbol{b}$:

$$\frac{\partial \mathcal{C}}{\partial \boldsymbol{z}}(\boldsymbol{b}) = 3 \left[\left(\sum_{k \in [0, M-1]} |b_k|^3 \right) - \gamma \right] \boldsymbol{b}.\,|\boldsymbol{b}| \tag{32}$$

But since the gradient descent is made over the base signal $\boldsymbol{s}$ we need to convert the gradient at $\boldsymbol{b}$ over a gradient at $\boldsymbol{s}$. Mathematically speaking, we wish to know the W^*-gradient of the function $\tilde{\mathcal{C}}$ defined as:

$$\begin{aligned} \tilde{\mathcal{C}} \colon \mathbb{C}^M &\to \mathbb{R} \\ \boldsymbol{s} &\mapsto \mathcal{C}\left(\mathcal{F}^{-1}(\mathcal{G}(\mathcal{F}(\boldsymbol{s}))) \right) \end{aligned} \tag{33}$$

Since we know the gradient of $\mathcal{C}$, all we need to do is convert it to the frequency domain, compose it with $\mathcal{G}$ and bring it back to time domain using the rules of Wirtinger calculus and time-frequency gradient conversion.

Using the time to frequency domain conversion expression in (28) we obtain the value of the W^*-gradient of $\mathcal{C} \circ \mathcal{F}^{-1}$ at $\boldsymbol{B}$ as:

$$\frac{\partial \left(\mathcal{C} \circ \mathcal{F}^{-1} \right)}{\partial \boldsymbol{z}^*}(\boldsymbol{B}) = \frac{1}{N} \mathcal{F}\left(\frac{\partial \mathcal{C}}{\partial \boldsymbol{z}^*}(\boldsymbol{b}) \right) \tag{34}$$

From here we need to obtain the gradient of $\mathcal{C} \circ \mathcal{F}^{-1} \circ \mathcal{G}$. Since $\mathcal{G}$ is a simple element-wise product, and since as stated in section 2.2.3 Wirtinger derivation is linear, we directly obtain:

$$\frac{\partial \left(\mathcal{C} \circ \mathcal{F}^{-1} \circ \mathcal{G} \right)}{\partial \boldsymbol{z}^*}(\boldsymbol{S}) = \boldsymbol{W}.\frac{\partial \left(\mathcal{C} \circ \mathcal{F}^{-1} \right)}{\partial \boldsymbol{z}^*}(\boldsymbol{B}) \tag{35}$$

All that is left now is the conversion from frequency to time domain to obtain the W^*-gradient at $\boldsymbol{s}$. Using the result in (29) gives:

$$\frac{\partial \tilde{\mathcal{C}}}{\partial \boldsymbol{z}^*}(\boldsymbol{s}) = N \mathcal{F}^{-1} \left(\frac{\partial \left(\mathcal{C} \circ \mathcal{F}^{-1} \circ \mathcal{G} \right)}{\partial \boldsymbol{z}^*}(\boldsymbol{S}) \right) \tag{36}$$

Combining (34), (35) and (36) then gives:

$$\frac{\partial \tilde{\mathcal{C}}}{\partial \boldsymbol{z}^*}(\boldsymbol{s}) = \mathcal{F}^{-1} \left(\boldsymbol{W}.\mathcal{F}\left(\frac{\partial \mathcal{C}}{\partial \boldsymbol{z}^*}(\boldsymbol{b}) \right) \right) \tag{37}$$

Using this relation we can now convert the W^*-gradient at $\boldsymbol{b}$ expressed in (32) to a W^*-gradient at $\boldsymbol{s}$, the base signal.

In addition to this, since $\boldsymbol{s}$ is a purely real-valued signal it is also straightforward to make use of the previous result to obtain the more common real gradient of $\tilde{\mathcal{C}}$ at $\boldsymbol{s}$. Indeed, from the definition of the W^*-derivative in (8) and since $\tilde{\mathcal{C}}$ is real-valued, we have that if $\boldsymbol{s} \in \mathbb{R}^M$:

$$\frac{\partial \tilde{\mathcal{C}}}{\partial \boldsymbol{x}}(\boldsymbol{s}) = 2\Re \left\{ \frac{\partial \tilde{\mathcal{C}}}{\partial \boldsymbol{z}^*}(\boldsymbol{s}) \right\} \tag{38}$$

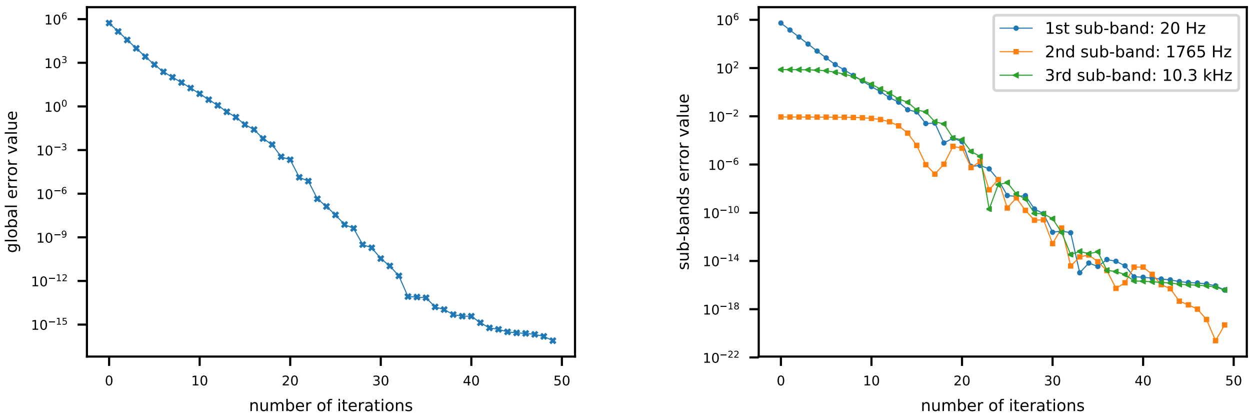
Which results in the final expression of the real gradient of the cost function with respect to the real base time signal:

$$\frac{\partial \tilde{\mathcal{C}}}{\partial \boldsymbol{x}}(\boldsymbol{s}) = 6 \left[\left(\sum_{k \in [0, M-1]} |b_k|^3 \right) - \gamma \right] \Re \left\{ \mathcal{F}^{-1} \left(\boldsymbol{W}.\mathcal{F}\left(\boldsymbol{b}.\,|\boldsymbol{b}| \right) \right) \right\} \tag{39}$$

We now have all the tools required to use a gradient descent algorithm over the base signal $\boldsymbol{s}$ in order to minimize the cost function $\mathcal{C}$ and thus impose a given statistic over a selected sub-band of it, without having to first perform it over the sub-band and then resort to a potentially damaging reconstruction.

Additionally, it is now possible to impose a given set of statistics to several sub-bands at once. Indeed, if we define a general cost function as the sum of several sub-bands cost functions, such as the one expressed in (31), then the linearity of derivation gives that the global cost gradient with respect to the base time signal is simply the sum of the sub-band cost gradients which we established in (39): using this global cost in the gradient descent will then tend to lower the sub-band cost functions, thus imposing the desired values to the statistics of each sub-bands at once.

So as to act both as an example and an experimental validation, such a simultaneous imposition was made over a 10 seconds-long white noise signal sampled at 48 kHz. The statistics chosen as goals are the raw third moments extracted from a rain sound texture over 3 sub-bands arbitrarily chosen at 20 Hz, 1765 Hz and 10.3 kHz, while the algorithm used is a conjugate gradient descent. The evolution over the iterations of the gradient descent of the cost functions, both global and band-wise, is plotted in Figure 1. Because the band-wise cost functions are decreasing during the gradient descent, the optimization successfully outputs a time signal possessing the exact statistics we meant to impose on it, without requiring a separate imposition for each sub-band nor a possibly harmful reconstruction of the time signal from the sub-bands signals.



(a) global error over the sub-bands

(b) separate errors over each sub-band

Figure 1: Value of the cost functions over the number of iterations made during the gradient descent. The total gradient is obtained by summing the gradients of 3 arbitrary sub-bands (20 Hz, 1765 Hz and 10.3 kHz) as expressed in (39)

5. CONCLUSION

By making use of Wirtinger formalism it is possible to link the gradient of a real-valued function with respect to a signal to the gradient with respect to its spectrum and transition between the two simply by use of a digital Fourier Transform and its inverse: this allows any gradient descent algorithm to be performed equivalently in time or frequency domain while avoiding any complication that may come from the complex non-differentiability of cost functions.

Put to use in sound texture synthesis this leads to a fully coherent approach to imposing arbitrary values to the statistics of the sub-bands of a signal, which is currently being investigated in order to establish an efficient sound texture synthesis algorithm.

6. REFERENCES

[1] Stefan Bilbao, Alberto Torin, and Vasileios Chatziioannou, "Numerical modeling of collisions in musical instruments," *Acta Acustica united with Acustica*, vol. 101, no. 1, pp. 155–173, 2015.

[2] Josh H McDermott and Eero P Simoncelli, "Sound texture perception via statistics of the auditory periphery: Evidence from sound synthesis," *Neuron*, vol. 71, no. 5, pp. 926–940, 2011.

[3] Wei-Hsiang Liao, *Modelling and transformation of sound textures and environmental sounds*, Ph.D. thesis, Université Pierre et Marie Curie, 2015.

[4] Wilhelm Wirtinger, "Zur formalen theorie der funktionen von mehr komplexen veränderlichen," *Mathematische Annalen*, vol. 97, no. 1, pp. 357–375, 1927.

[5] DH Brandwood, "A complex gradient operator and its application in adaptive array theory," in *IEE Proceedings F-Communications, Radar and Signal Processing*. IET, 1983, vol. 130, pp. 11–16.

[6] P Bouboulis, "Wirtinger's calculus in general hilbert spaces," *arXiv preprint arXiv:1005.5170*, 2010.

[7] Robert FH Fischer, *Precoding and signal shaping for digital transmission*, John Wiley & Sons, 2005.