Finite-power spectral analytic framework for quantized sampled signals

Nguyen T. Thao
Dept. of Electrical Engineering
City College, CUNY, New York, NY 10031
Email: thao@ee.ccny.cuny.edu

***Abstract*—To be accurate, the theoretical spectral analysis of quantized sequences requires that the deterministic definition of power spectral density be used. We establish the functional space foundations for this analysis, which remarkably appear to be missing until now. With them, we then shed some new light on quantization error spectra in PCM and $\Sigma\Delta$ modulation.**

I. Introduction

The spectral analysis of quantized signals appears to miss clear functional foundations. In spite of their deterministic nature, quantized sequences are often theoretically described using a probabilistic definition of power spectral density. This however only leads to approximate statistical models that cannot predict quantization phenomena such as intermodulation, idle tones and limit cycles present in $\Sigma\Delta$ modulation for example. The first rigorous analysis of quantized signals in pulse code modulation (PCM) and $\Sigma\Delta$ modulation was performed by R. M. Gray [1], [2], [3] in the late 80's based on the deterministic time-averaged power function $\mathsf{M}(|x|^2)$ of a sequence $x[n]$, where

$$\mathsf{M}(x) := \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} x[n].$$

Similarly to the probabilistic case, power spectral density is obtained by taking the Fourier transform of the autocorrelation $r_x[m] := \mathsf{M}(x\,T^m x)$, where T is the shift sequence operator defined by $Tx[n] = x[n{+}1]$. From a functional space viewpoint, this appears at first sight as a mere extension of *energy spectral density* defined in the Hilbert space ℓ^2 of square-summable sequences. In this view, one would define the space of finite-power sequences as

$$\mathcal{P} := \big\{ x \in \mathbb{C}^{\mathbb{N}} : \mathsf{M}(|x|^2) \text{ exists} \big\}$$

with the inner-product

$$\langle x, y\rangle_{\mathcal{P}} := \mathsf{M}(x^* y).$$

This is however doomed to fail as $\mathcal{P}$ is not even a linear space as shown in this paper.

The goal of this article is to rigorously establish Hilbert space foundations to the spectral analysis of finite-power sequences. Based on this, standard theorems can be applied such as the spectral properties of unitary operators. We thus provide a functional space background to the work of R. M. Gray, explaining for example why mixed quantization spectra are to be expected. We also indicate some possible generalization to overloaded $\Sigma\Delta$ modulators. The detailed proofs of the claimed results are included in [4].

II. Hilbert spaces of finite-power sequences

The first obstacle to a direct analogy between ℓ^2 and $\mathcal{P}$ is that the function $\langle\cdot,\cdot\rangle_{\mathcal{P}}$ is not defined everywhere in $\mathcal{P}{\times}\mathcal{P}$. Consider for example the sequence $s[n] := (-1)^{\lfloor \log_2(n)\rfloor}$ for $n \geq 1$, which is clearly an element of $\mathcal{P}$. One easily shows that $\mathsf{M}(s)$ does not converge, which is equivalent to saying that $\langle s, 1\rangle_{\mathcal{P}}$ does not exist, although the constant sequence 1 is also in $\mathcal{P}$. This simultaneously shows that $\mathsf{M}(|s{+}1|^2)$ does not exist, otherwise we would obtain $2\langle s,1\rangle = \mathsf{M}(|s{+}1|^2) - \mathsf{M}(|s|^2) - \mathsf{M}(|1|^2)$. So $\mathcal{P}$ is not even a linear space.

To rigorously justify spectral analysis in the sense of finite power, one needs to explicitly build a Hilbert space within $\mathcal{P}$. One possible procedure is to come up with a known family of sequences $\{\varphi_k\}_{k\in K}$ of $\mathcal{P}$ such that $\langle \varphi_k, \varphi_{k'}\rangle_{\mathcal{P}}$ exists for all $k, k' \in K$ and is equal to $\delta_{k-k'}$ where δ is the Kronecker symbol. We say that $\langle \varphi_k, \varphi_{k'}\rangle_{\mathcal{P}}$ is orthonormal with respect to $\langle\cdot,\cdot\rangle_{\mathcal{P}}$. Up to some non-trivial theoretical considerations [4], it can be shown that the space of sequences of the form $x[n] = \sum_{k\in K} \alpha_k\, \varphi_k$ where $(\alpha_k)_{k\in K}$ is a family of complex coefficients whose nonzero values are in countable number and square summable, is a Hilbert space with respect to $\langle\cdot,\cdot\rangle_{\mathcal{P}}$ that is included in $\mathcal{P}$. We denote this space by

$$\mathcal{H} := \overline{\mathrm{span}}\{\varphi_k\}_{k\in K}.$$

The most basic example of such a construction is obtained with the exponential sequences $e_\xi[n] := e^{\mathrm{i}2\pi\xi n}$. Note that $\langle e_\xi, e_{\xi'}\rangle_{\mathcal{P}} = \mathsf{M}(e_{\xi'-\xi}) = \delta_{\xi'-\xi}$ when $\xi, \xi' \in [0,1)$, where δ is the Kronecker symbol. Thus, $\mathcal{B} := \overline{\mathrm{span}}\{e_\xi\}_{\xi\in[0,1)}$ is a Hilbert space included in $\mathcal{P}$. This is called the space of *almost periodic sequences in the sense of Besicovitch* (Besicovitch-AP sequences) [5]. Its elements can be presented in the form

$$x[n] = \textstyle\sum_{k\in\mathbb{Z}} \alpha_k\, e_{\xi_k}[n] \tag{1}$$

where $(\alpha_k)_{k\in\mathbb{Z}}$ is square summable and $(\xi_k)_{k\in\mathbb{Z}}$ are distinct values in $[0,1)$.

III. Finite power by Weyl's criterion

Weyl's equidistribution criterion [6] states that a real sequence $s[n]$ is uniformly distributed modulo 1 (i.e., its fractional part is uniformly distributed in $[0,1)$) if and only if

$\mathsf{M}(e^{\mathrm{i}2\pi ks}) = 0$ for all nonzero integers k, where $e^{\mathrm{i}2\pi ks}$ designates the sequence $(e^{\mathrm{i}2\pi ks[n]})_{n\geq 1}$. Noticing the relation $\langle e^{\mathrm{i}2\pi k\cdot s}, e^{\mathrm{i}2\pi k'\cdot s}\rangle_{\mathcal{P}} = \mathsf{M}(e^{\mathrm{i}2\pi l\cdot s})$ where $l := k-k'$, the uniform distribution of $s[n]$ is then equivalent to the orthonormality of the family of sequences $\{e^{\mathrm{i}2\pi ks}\}_{k\in\mathbb{Z}}$. We state below the multi-dimensional version [6] of this result.

Proposition 3.1: Let $\boldsymbol{s}[n]$ be a sequence of vectors in $\mathbb{R}^d$. Then $\boldsymbol{s}[n]$ is uniformly distributed modulo 1 (u.d. mod 1) *if and only if* $\{e^{\mathrm{i}2\pi \boldsymbol{k}\cdot\boldsymbol{s}}\}_{\boldsymbol{k}\in\mathbb{Z}^d}$ is an orthonormal family with respect to $\langle\cdot,\cdot\rangle_{\mathcal{P}}$ (where $\boldsymbol{k}\cdot\boldsymbol{s}$ is the dot product of $\boldsymbol{k}$ and $\boldsymbol{s}$ in $\mathbb{R}^d$).

As soon as a u.d. mod 1 sequence $\boldsymbol{s}[n]$ is found, one therefore generates a (separable) Hilbert space in $\mathcal{P}$ by forming the space

$$\mathcal{H}_{\boldsymbol{s}} := \overline{\mathrm{span}}\{e^{\mathrm{i}2\pi\boldsymbol{k}\cdot\boldsymbol{s}}\}_{\boldsymbol{k}\in\mathbb{Z}^d}.$$

The next proposition characterizes a useful subspace of $\mathcal{H}_{\boldsymbol{s}}$.

Proposition 3.2: Let $\boldsymbol{s}[n]$ be a sequence of vectors in $\mathbb{R}^d$ that is u.d. mod 1. Then, for any d-variable 1-periodic *Riemann integrable* function $h(\boldsymbol{u})$, the sequence $h(\boldsymbol{s}[n])$ belongs to $\mathcal{H}_{\boldsymbol{s}} \subset \mathcal{P}$ and yields the orthogonal expansion

$$h(\boldsymbol{s}[n]) = \sum_{\boldsymbol{k}\in\mathbb{Z}^d} \hat{h}_{\boldsymbol{k}}\, e^{\mathrm{i}2\pi\boldsymbol{k}\cdot\boldsymbol{s}[n]} \tag{2}$$

where $(\hat{h}_{\boldsymbol{k}})_{\boldsymbol{k}\in\mathbb{Z}^d}$ are the Fourier coefficients of $h(\boldsymbol{u})$.

The 1-periodicity of h implies that $h(\boldsymbol{u}+\boldsymbol{k}) = h(\boldsymbol{u})$ for all $\boldsymbol{u}\in\mathbb{R}^d$ and $\boldsymbol{k}\in\mathbb{Z}^d$. The proof of this proposition uses the more general equivalent criterion for equidistribution implying that $\mathsf{M}\big(f(\boldsymbol{s}[n])\big) = \int_{[0,1)^d} f(\boldsymbol{u})\,\mathrm{d}\boldsymbol{u}$ for any d-variable 1-periodic Riemann integrable function $f(\boldsymbol{u})$ [6]. The conceptual difficulty of (2) is that the summation does not necessarily converge pointwise, although it converges in the sense of the norm $\|\cdot\|_{\mathcal{P}}$.

A simple case of interest is when $\boldsymbol{s}[n] = n\boldsymbol{\zeta}$ where $\boldsymbol{\zeta} = (\zeta_1,\cdots,\zeta_d)\in\mathbb{R}^d$. It is known that $\boldsymbol{s}[n] = n\boldsymbol{\zeta}$ is uniformly distributed modulo 1 if and only if $\zeta_1,\cdots,\zeta_d,1$ are rationally independent (i.e., no rational combination of $\zeta_1,\cdots,\zeta_d,1$ other than the zero combination is equal to zero) [6]. Assuming that this condition is realized and that h has the required property, Proposition 3.2 implies that $h(n\boldsymbol{\zeta})$ is a finite-power sequence that belongs to the closed space $\mathcal{H}_{\boldsymbol{s}}$ spanned by the orthonormal family $\{e_{\xi_{\boldsymbol{k}}}\}_{\boldsymbol{k}\in\mathbb{Z}^d}$ where

$$\xi_{\boldsymbol{k}} := \boldsymbol{k}\cdot\boldsymbol{\zeta} = k_1\,\zeta_1 + \cdots + k_d\,\zeta_d, \tag{3}$$

since $e^{\mathrm{i}2\pi\boldsymbol{k}\cdot\boldsymbol{s}[n]} = e^{\mathrm{i}2\pi\boldsymbol{k}\cdot(n\boldsymbol{\zeta})} = e^{\mathrm{i}2\pi\xi_{\boldsymbol{k}}n}$. In this case, $\mathcal{H}_{\boldsymbol{s}} \subset \mathcal{B}$. Hence, $h(n\boldsymbol{\zeta})$ is a Besicovitch-AP sequence. Its orthogonal expansion in $\mathcal{B}$ is explicitly

$$h(n\boldsymbol{\zeta}) = \textstyle\sum_{\boldsymbol{k}\in\mathbb{Z}^d} \hat{h}_{\boldsymbol{k}}\, e_{\xi_{\boldsymbol{k}}}[n]. \tag{4}$$

IV. Basics of Spectral Analysis

A. Power spectral measure

Similarly to the probabilistic approach, the power spectral density of a sequence $x[n]$ would be the Fourier transform of the autocorrelation

$$r_x[m] := \langle x, T^m x\rangle_{\mathcal{P}} \tag{5}$$

and T is the shift sequence operator defined in the introduction. The sequence $r_x[m]$ is guaranteed to exist if $x[n]$ is shown to be in some Hilbert space $\mathcal{H}\subset\mathcal{P}$ that is invariant by T. Assume that this is the case. Then, T is a *unitary* operator of $\mathcal{H}$ as it preserves the inner-product and is invertible. Now rigorously speaking, $r_x[m]$ may not have a Fourier transform. Given the positive definite property of $r_x[m]$, it is however shown to be at least the Fourier coefficients of a positive measure μ_x [7], which we call the *power spectral measure* of $x[n]$. By Lebesgue's decomposition theorem, μ_x has in general *three* components, a pure-point part (purely discrete measure composed of Dirac masses), an absolutely-continuous part (the actual and only part that yields a power spectral density by Radon-Nikodym derivative) and a singular-continuous part.

B. Spectral decomposition

A key to analyzing the measure structure of μ_x is to decompose $\mathcal{H}$ as an orthogonal sum

$$\mathcal{H} = \textstyle\bigoplus_{k\in K} \mathcal{I}_k \tag{6}$$

of Hilbert subspaces $\mathcal{I}_k$ that are invariant by T. Writing $x = \sum_{k\in K} x_k$ where x_k is the orthogonal projection of x onto $\mathcal{I}_k$, it is clear that $\langle x_k, T^m x_{k'}\rangle = 0$ when $k\neq k'$. This leads to the decompositions

$$r_x[m] = \textstyle\sum_{k\in K} r_{x_k}[m] \qquad \text{and} \qquad \mu_x = \sum_{k\in K}\mu_{x_k}.$$

In this reduction process, we will encounter in this paper two types of measure μ_{x_k}.

1) Simple discrete spectrum: This is the case where $\mathcal{I}_k$ is spanned by a single vector φ which we can assume of norm 1. By T-invariance of $\mathcal{I}_k$, φ must be an eigenfunction of T. Since T is unitary, it is known that the eigenvalue of φ must be of the form $e^{\mathrm{i}2\pi\xi}$ where $\xi\in[0,1)$. Since $T^m\varphi = e^{\mathrm{i}2\pi\xi m}\varphi$, the sequence x_k which is of the form $a\,\varphi$ yields the autocorrelation $r_{x_k}[m] = \langle a\varphi, aT^m\varphi\rangle_{\mathcal{P}} = |a|^2 e^{\mathrm{i}2\pi\xi m}$. The corresponding spectral measure μ_{x_k} is then discrete and equal to $|a|^2\delta_\xi$ where δ_ξ denotes the Dirac mass at frequency ξ.

2) Simple absolutely-continuous spectrum: This is the case where $\mathcal{I}_k$ is the closed span of an orthonormal family of the form $\{T^n\varphi\}_{n\in\mathbb{Z}}$. The sequence x_k is then of the form $\sum_{n\in\mathbb{Z}} a_n\, T^n\varphi$, where a_n is a square-summable sequence. By orthonormality of $\{T^n\varphi\}_{n\in\mathbb{Z}}$, one finds that

$$r_{x_k}[m] = \langle x_k, T^m x_k\rangle_{\mathcal{P}} = \textstyle\sum_{n\in\mathbb{Z}} a^*_{n+m}\, a_n.$$

This is precisely the *finite-energy* autocorrelation of a_{-n}. In this case, $r_{x_k}[m]$ yields a Fourier transform $R_{x_k}(\xi) = |A(-\xi)|^2$ where $A(\xi)$ is the Fourier transform of a_n in $L^2([0,1))$, making $R_{x_k}(\xi)$ a function in $L^1([0,1))$. This makes the measure μ_{x_k} absolutely-continuous.

C. Besicovitch almost-periodic sequences

The most straightforward example of spectral decomposition is achieved in the space $\mathcal{B}$. Every function of its orthonormal basis $\{e_\xi\}_{\xi\in[0,1)}$ turns out to be an eigenfunction of T since

$$Te_\xi = e^{\mathrm{i}2\pi\xi}\, e_\xi.$$

Once a Besicovitch-AP sequence $x[n]$ is written in the form (1), it can be presented as element of a space sum (6) with $K := \mathbb{Z}$ and $\mathcal{I}_k := \text{span}\{e_{\xi_k}\}$. This falls in the case of Section IV-B1. One then obtains the discrete power spectral measure $\mu_x = \sum_{k\in\mathbb{Z}} |\alpha_k|^2\, \delta_{\xi_k}$.

V. PCM WITH TRIGONOMETRIC POLYNOMIAL INPUT

We show that the quantizer error sequence $\epsilon[n]$ from the pulse code modulation (PCM) of a finite sum of sinusoids is Besicovitch almost-periodic, thus yielding a purely discrete power spectral measure. PCM consists in transforming every sample of a sequence $x[n]$ individually by a nonlinear memoryless scalar function $Q(\cdot)$ that is basically piecewise constant. This results in an error sequence $\epsilon[n] := x[n]{-}Q(x[n])$[1]. An input sequence $x[n]$ that is a finite sum of sinusoids can always be expanded as a trigonometric polynomial[2]. $x[n] = \sum_{k=1}^{p} \alpha_k\, e^{\mathrm{i}2\pi\zeta_k n}$ where ζ_k are distinct frequency values in $[0,1)$. By defining the p-variable 1-periodic function $\mathrm{x}(u_1,\cdots,u_p) = \sum_{k=1}^{p} \alpha_k\, e^{\mathrm{i}2\pi u_k}$, one can write $x[n] = \mathrm{x}(n\boldsymbol{\zeta})$ where $\boldsymbol{\zeta} := (\zeta_1,\cdots,\zeta_p)$. Thus, the error sequence is of the form

$$\epsilon[n] = h(n\boldsymbol{\zeta}) \qquad \text{where} \qquad h(\boldsymbol{u}) := Q(\mathrm{x}(\boldsymbol{u})){-}\mathrm{x}(\boldsymbol{u}). \quad (7)$$

The function $h(\boldsymbol{u})$ is p-variable, 1-periodic and Riemann integrable since Q is piecewise constant and $\mathrm{x}(\boldsymbol{u})$ is continuous. Under the condition that $\zeta_1,\cdots,\zeta_p,1$ are rationally independent, we know from Section III that $\epsilon[n]$ is a Besicovitch-AP sequence of orthogonal expansion (4). From Section IV-C, we conclude that the power spectral measure of $\epsilon[n]$ is purely discrete with the autocorrelation expansion

$$r_\epsilon[m] = \sum_{\boldsymbol{k}\in\mathbb{Z}^r} |\hat{h}_{\boldsymbol{k}}|^2\, e^{\mathrm{i}2\pi\xi_{\boldsymbol{k}} m}.$$

The discrete frequencies of the spectrum of $\epsilon[n]$ are the values $\xi_{\boldsymbol{k}}$ given by (3) and are nothing but the *intermodulation products* of the fundamental input frequencies $\zeta_1,\cdots,\zeta_p$, as seen in (3).

VI. QUANTIZATION ERROR IN IDEAL $\Sigma\Delta$ MODULATION

In an ideal $\Sigma\Delta$ modulator with a polynomial trigonometric input, we show that the errors due to quantization can be presented as output of a system of the form

$$\left\{\begin{array}{l} \boldsymbol{s}[n] = \boldsymbol{M}\boldsymbol{s}[n{-}1] + \boldsymbol{\tau} \\ \epsilon[n] = h(\boldsymbol{s}[n]) \end{array}\right. \quad (8)$$

where h is a d-variable 1-periodic Riemann-integrable function and $\boldsymbol{M}$ is a square matrix that is *unimodular* (i.e., invertible with integer entries) and *unipotent* (i.e., with all eigenvalues equal to 1).

A. General equations

The general diagram of a $\Sigma\Delta$ modulator is shown in Figure 1 and defines the signal notation we will use. In this section,

[1]The usual convention for a system error is $e[n] = Q(x[n]){-}x[n]$. Working with the sequence $\epsilon[n] := -e[n]$ will prove more convenient from a dynamical system viewpoint.

[2]This is in the largest sense of sequences of the form $x[n] := \sum_{k=1}^{N} \alpha\, e^{\mathrm{i}2\pi\xi_k n}$, where the ξ_k's are not necessarily harmonics of a single frequency.

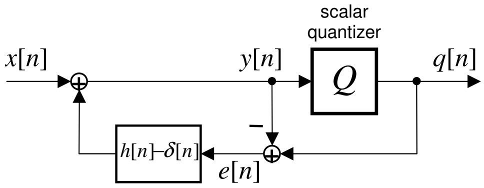

Fig. 1. General diagram of a $\Sigma\Delta$ modulator

the modulator is assumed to be ideal, i.e., $h[n]$ is a purely differentiating sequence of z-transform $H(z) = (1{-}z^{-1})^r$ and the quantizer is not overloaded. Like in PCM, we define the quantizer error to be $\epsilon[n] = y[n]{-}q[n] = -e[n]$. Using the vector sequence $\boldsymbol{v}[n] = (v_1[n],\cdots,v_r[n])$ such that $v_i[n]$ is the $(r{-}i)$th order differentiation of $\epsilon[n]$, one can show the following system of equation

$$\left\{\begin{array}{l} \boldsymbol{v}[n] = \boldsymbol{L}\,\boldsymbol{v}[n{-}1] + \mathbf{1}(x[n]{-}q[n]) \\ \epsilon[n] = \boldsymbol{j}\cdot\boldsymbol{v}[n] \end{array}\right. \quad (9)$$

where $\boldsymbol{L}$ is the lower-triangular matrix of 1's and size r, $\mathbf{1} := (1,\cdots,1) \in \mathbb{R}^r$ and $\boldsymbol{j} := (0,\cdots,0,1) \in \mathbb{R}^r$. The outstanding property of this state model is that $\boldsymbol{L}$, $\mathbf{1}$, $\boldsymbol{j}$ and $q[n]{+}\frac{1}{2}$ all have entries or coefficients that are integers. So, if we recursively construct the sequence

$$\boldsymbol{u}[n] = \boldsymbol{L}\,\boldsymbol{u}[n{-}1] + \mathbf{1}\,x'[n] \quad (10)$$

where $x'[n] := x[n]{+}\frac{1}{2}$ and $\boldsymbol{u}[0] = \boldsymbol{v}[0]$, then the relation $\boldsymbol{v}[n]{-}\boldsymbol{u}[n] \in \mathbb{Z}^r$ is maintained for all $n \geq 0$. Then,

$$\epsilon[n] \equiv \boldsymbol{j}\cdot\boldsymbol{u}[n] \mod 1. \quad (11)$$

B. Trigonometric polynomial inputs

Assuming that $x[n]$ is a trigonometric polynomial, we show that $\epsilon[n]$ can be at least determined modulo 1 via an *autonomous* system of the type

$$\left\{\begin{array}{l} \boldsymbol{s}[n] = \boldsymbol{M}\boldsymbol{s}[n{-}1] + \boldsymbol{\tau} \\ \epsilon[n] \equiv g(\boldsymbol{s}[n]) \mod 1 \end{array}\right. \quad (12)$$

where $\boldsymbol{M}$ is unimodular and unipotent, and g is a continuous function of $\mathbb{R}^d$ such that $g(\boldsymbol{s}){-}g(\boldsymbol{s}') \in \mathbb{Z}^d$ when $\boldsymbol{s}{-}\boldsymbol{s}' \in \mathbb{Z}^d$. When $x'[n]$ is equal to a constant $\bar{x}$, this is easily achieved from (10) and (11) by taking $\boldsymbol{s}[n] = \boldsymbol{u}[n]$, $\boldsymbol{M} = \boldsymbol{L}$, $\boldsymbol{\tau} = \mathbf{1}\bar{x}$ and $h(\boldsymbol{u}) = \langle \boldsymbol{j}\cdot\boldsymbol{u}\rangle_I$. When $x'[n]$ is not constant, one goes from (10-11) to (12) by the technique of *skew-product* first used in $\Sigma\Delta$ modulation in [8]. This requires the following preliminary result.

Proposition 6.1: Let $\mathrm{x}(\boldsymbol{u}) = \sum_{k=1}^{p} \alpha_k\, e^{\mathrm{i}2\pi u_k}$ and $\boldsymbol{\zeta} \in (0,1)^p$. An explicit solution to the equation

$$\boldsymbol{w}[n] = \boldsymbol{L}\,\boldsymbol{w}[n{-}1] + \mathbf{1}\,\mathrm{x}(n\boldsymbol{\zeta}) \quad (13)$$

is $\boldsymbol{w}[n] = \mathbf{x}(n\boldsymbol{\zeta})$, where $\mathbf{x}(\boldsymbol{u}) := \sum_{k=1}^{p} \alpha_k\, e^{\mathrm{i}2\pi u_k}\, \boldsymbol{x}_{\zeta_k}$ and $\boldsymbol{x}_\zeta := (x_\zeta, x_\zeta^2,\cdots,x_\zeta^r)$ with $x_\zeta := (1{-}e^{\mathrm{i}2\pi\zeta})^{-1}$.

Calling $\bar{x}$ the constant component of $x'[n]$, we express the "AC-component" $x'[n]{-}\bar{x}$ in the form $\mathrm{x}(n\boldsymbol{\zeta})$ as was done in Section V with the difference that $\boldsymbol{\zeta} \in (0,1)^p$. With the resulting function $\mathbf{x}(\boldsymbol{u})$ as obtained in the above proposition, we obtain the following result.

Proposition 6.2: The sequence $\boldsymbol{s}[n] := \big(n\boldsymbol{\zeta}, \boldsymbol{u}[n]-\mathbf{x}(n\boldsymbol{\zeta})\big)$ achieves the system (12), with

$$\boldsymbol{M} := \begin{bmatrix} \boldsymbol{I} & 0 \\ 0 & \boldsymbol{L} \end{bmatrix}, \quad \boldsymbol{\tau} = (\boldsymbol{\zeta}, \mathbf{1}\bar{x}) \quad \text{and} \quad g(\boldsymbol{\theta}, \bar{\boldsymbol{u}}) := \boldsymbol{j} \cdot (\bar{\boldsymbol{u}}+\mathbf{x}(\boldsymbol{\theta}))$$

where $\boldsymbol{I}$ is the identity matrix of size p.

The central argument of the proof is that the second component $\bar{\boldsymbol{u}}[n] := \boldsymbol{u}[n]-\mathbf{x}(n\boldsymbol{\zeta})$ of $\boldsymbol{s}[n]$ satisfies the recursion $\bar{\boldsymbol{u}}[n] = \boldsymbol{L}\,\bar{\boldsymbol{u}}[n-1] + \mathbf{1}\,\bar{x}$.

C. Non-overloaded quantizer

The fact that the quantizer is not overloaded implies that $\epsilon[n]$ remains in the interval $I := [-\frac{1}{2}, \frac{1}{2})$. Let $\langle\cdot\rangle_I$ be the unique 1-periodic function that is identity in I (explicitly equal to $\langle\cdot+\frac{1}{2}\rangle - \frac{1}{2}$ where $\langle\cdot\rangle$ is the fractional part function). Since $\epsilon[n] = \langle\epsilon[n]\rangle_I$ and $\epsilon[n] \equiv g(\boldsymbol{s}[n]) \bmod 1$, then $\epsilon[n] = h(\boldsymbol{s}[n])$ where $h(\boldsymbol{s}) := \langle g(\boldsymbol{s})\rangle_I$. One easily verifies that h is 1-periodic and Riemann integrable. The system (8) is thus fully achieved.

VII. Spectral analysis in "unipotent" dynamical system

In this section, we perform the spectral analysis of sequences $\epsilon[n]$ output by systems of the type (8) with a unimodular and unipotent matrix $\boldsymbol{M}$.

A. State equidistribution

Proposition 7.1: Let $\boldsymbol{s}[n]$ be a sequence satisfying (8) where $\boldsymbol{M}$ is unimodular and unipotent. Then, $\boldsymbol{s}[n]$ is u.d. mod 1 *if and only if* $\boldsymbol{k}\cdot\boldsymbol{\tau} \notin \mathbb{Z}$ for all $\boldsymbol{k} \in \mathbb{Z}^d\backslash\{\mathbf{0}\}$ such that $\boldsymbol{M}^\top\boldsymbol{k} = \boldsymbol{k}$.

The basic ingredients of the proof are as follows. Due to the unipotent property of $\boldsymbol{M}$, the components of $\boldsymbol{s}[n]$ are shown to be polynomial sequences. Next, one uses the known fact that a polynomial sequence is u.d. mod 1 if and only if at least one of the coefficients of its non-constant terms is irrational [6]. In the setting of ideal $\Sigma\Delta$ modulation of the previous section, this proposition implies that $\bar{x}, \zeta_1, \cdots, \zeta_r, 1$ must be rationally independent for $\boldsymbol{s}[n]$ to be u.d. mod 1. In the constant input case, this reduces to the condition that $\bar{x}$ be irrational.

B. Spectral analysis

Assume that the condition of Proposition 7.1 is realized. Proposition 3.2 then implies that $\epsilon[n]$ is a finite-power sequence with the orthogonal expansion

$$\epsilon = \textstyle\sum_{\boldsymbol{k}\in\mathbb{Z}^d} \hat{h}_{\boldsymbol{k}}\,\varphi_{\boldsymbol{k}} \qquad \text{where} \qquad \varphi_{\boldsymbol{k}}[n] := e^{\mathrm{i}2\pi\boldsymbol{k}\cdot\boldsymbol{s}[n]}.$$

To derive the autocorrelation $r_\epsilon[m] := \langle\epsilon, T^m\epsilon\rangle_{\mathcal{P}}$, one needs to apply T^m on the basis vectors $\varphi_{\boldsymbol{k}}$. From the mere relation $\boldsymbol{s}[n+1] = \boldsymbol{M}\boldsymbol{s}[n]+\boldsymbol{\tau}$, one finds that

$$T\varphi_{\boldsymbol{k}} = e^{\mathrm{i}2\pi\xi_{\boldsymbol{k}}}\,\varphi_{\boldsymbol{k}'} \qquad \text{where} \qquad \xi_{\boldsymbol{k}} := \boldsymbol{k}\cdot\boldsymbol{\tau} \tag{14}$$

and $\boldsymbol{k}' := \boldsymbol{M}^\top\boldsymbol{k} \in \mathbb{Z}^d$. This first implies that the space $\mathcal{H} := \overline{\mathrm{span}}\{\varphi_{\boldsymbol{k}}\}_{\boldsymbol{k}\in\mathbb{Z}^d}$ is T-invariant. But as $\boldsymbol{M}$ is unimodular, it defines a permutation of $\mathbb{Z}^d$. Hence, the action of T on $\{\varphi_{\boldsymbol{k}}\}_{\boldsymbol{k}\in\mathbb{Z}^d}$ amounts to a permutation of the basis vectors plus some phase shift. One then obtains an orthogonal decomposition of $\mathcal{H}$ of the type (6) with the following definitions: $\mathcal{I}_{\boldsymbol{k}} := \overline{\mathrm{span}}\{\varphi_{\boldsymbol{l}}\}_{\boldsymbol{l}\in\mathcal{O}(\boldsymbol{k})}$, $\mathcal{O}(\boldsymbol{k}) := \big\{\boldsymbol{k}_n : n \in \mathbb{Z}\big\}$, $\boldsymbol{k}_n := {\boldsymbol{M}^\top}^n\boldsymbol{k}$ and K equal to a subset of $\mathbb{Z}^r$ such that $\{\mathcal{O}(\boldsymbol{k})\}_{\boldsymbol{k}\in K}$ is a partition of $\mathbb{Z}^r$. We know from Section IV-B that $\mu_\epsilon = \sum_{\boldsymbol{k}\in K}\mu_{\epsilon_{\boldsymbol{k}}}$ where $\epsilon_{\boldsymbol{k}}$ is the orthogonal projection of ϵ onto $\mathcal{I}_{\boldsymbol{k}}$.

Proposition 7.2: For any $\boldsymbol{k} \in \mathbb{Z}^d$, the following statements are equivalent: (i) $\mathcal{O}(\boldsymbol{k})$ is a finite orbit, (ii) $\mathcal{O}(\boldsymbol{k}) = \{\boldsymbol{k}\}$, (iii) $\boldsymbol{k}$ belongs to the set $J_{\boldsymbol{M}} := \big\{\boldsymbol{l} \in \mathbb{Z}^d : \boldsymbol{M}^\top\boldsymbol{l} = \boldsymbol{l}\big\}$.

While $(ii) \Leftrightarrow (iii)$ is trivial, $(i) \Leftrightarrow (iii)$ uses the unipotent property of $\boldsymbol{M}$. When $\boldsymbol{k} \in J_{\boldsymbol{M}}$, $\mathcal{I}_{\boldsymbol{k}} = \mathrm{span}\{\varphi_{\boldsymbol{k}}\}$, so $\mu_{\epsilon_{\boldsymbol{k}}}$ is a single Dirac mass according to Section IV-B1. When $\boldsymbol{k} \in K\backslash J_{\boldsymbol{M}}$, one easily sees from (14) that $\{T^n\varphi_{\boldsymbol{k}}\}_{n\in\mathbb{Z}}$ is equal to $\{\varphi_{\boldsymbol{k}_n}\}_{n\in\mathbb{Z}}$ up to some phase shifts, and is therefore an orthonormal basis of $\mathcal{I}_{\boldsymbol{k}}$. So $\mu_{\epsilon_{\boldsymbol{k}}}$ is absolutely continuous according to Section IV-B2. We conclude that the power spectral measure of $\epsilon[n]$ is *a priori* mixed, with a pure-point part and an absolutely continuous-part equal to $\mu_{\mathring{\epsilon}}$ and $\mu_{\bar{\epsilon}}$, respectively, where $\mathring{\epsilon} := \sum_{\boldsymbol{k}\in J_{\boldsymbol{M}}}\epsilon_{\boldsymbol{k}}$ and $\bar{\epsilon} := \sum_{\boldsymbol{k}\in K\backslash J_{\boldsymbol{M}}}\epsilon_{\boldsymbol{k}}$, but *no* singular-continuous part.

VIII. Discussion and extensions

In his work [1], [2], [3], R. M. Gray found with ideal $\Sigma\Delta$ modulators the more precise result that μ_ϵ is either purely discrete or uniform (white noise), which is a particular case of absolutely-continuous spectral measure. The reason for this special result is particular to the specific function $h(\boldsymbol{s}) = \langle g(\boldsymbol{s})\rangle_I$ found in Section VI-C. It is however shown in [4] that systems of the type (8) are achieved with a class of $\Sigma\Delta$ modulators that are more representative of practical configurations [9] including quantizer overloading, and yield truly mixed spectra. Finally, although this paper constantly assumed the uniform distribution of $\boldsymbol{s}[n]$ modulo 1, similar results can be obtained with absolutely no condition on $\boldsymbol{s}[n]$. This uses the result that the closure of the set of points $\boldsymbol{s}[n]$ modulo 1 is a compact group and the generalized notion of uniform distribution in a compact group [6].

Acknowledgment

A large number of the mathematical results have been proved in collaboration with Sinan Güntürk [4].

References

[1] R. M. Gray, "Spectral analysis of quantization noise in a single loop sigma-delta modulator with dc input," *IEEE Trans. Commun.*, vol. 37, pp. 588-599, June 1989.

[2] R. M. Gray, "Quantization noise spectra," *IEEE Trans. Inform. Th.*, vol. 36, pp. 1220-1244, Nov. 1990.

[3] P.-W. Wong and R. M. Gray, "Two-stage sigma-delta modulation," *IEEE Trans. Acoust., Speech, Signal Proc.*, vol. 38, pp. 1937-1952, Nov. 1990.

[4] C. S. Güntürk and N. T. Thao, "Signal-analytic theory of $\Sigma\Delta$ quantization," *in preparation*.

[5] C. Corduneanu, *Almost Periodic Oscillations and Waves*, Springer, 2009.

[6] L. Kuipers and H. Niederreiter, *Uniform Distribution of Sequences*, Wiley, 1974.

[7] M. Queffélec, *Substitution Dynamical Systems - Spectral Analysis*, 2nd ed., Springer, 2010.

[8] A. Teplinsky, E. Condon and O. Feely, "Driven interval shift dynamics in sigma-delta modulators and phase-locked loops," IEEE Trans. Circuits and Systems I, vol. 52, pp. 1224-1235, June 2005.

[9] R. Schreier, "An empirical study of high-order single-bit delta-sigma modulators," IEEE Trans. Circuits and Systems II, vol. 40, pp. 461-466, Aug. 1993.