A fast, discrete Gabor transform via a partition of unity

Michael P. Lamoureux, Peter C. Gibson, Jeff P. Grossman, and Gary F. Margrave

ABSTRACT

A partition of unity on the d-dimensional integer lattice $\mathbb{Z}^d$ is used to create a generalized discrete Gabor transform, with analysis and synthesis windows of smooth, desirable characteristics. Factorizing the partition of unity allows for different choices of analysis and synthesis windows, a transformation on general lattices in the time and frequence domains is considered, and an approximate partition of unity via Gaussians gives an approximate inverse. Speed, implementation issues, and practical choices for partitions of unity useful in applications are discussed.

INTRODUCTION

The Gabor transform is a mathematical transformation of functions that provides a joint time/frequency representation of a given signal, as well as a means to recover the signal from this time/frequency representation. Although Denis Gabor's original motivation for introducing this transformation lay in quantum mechanics (see [5]), the technique which bears his name has proven to be useful in a variety of signal processing contexts, including image processing, acoustics, medical imaging, electrical engineering, and others. For instance, the authors of this paper have a particular interest in seismic data processing and have used Gabor transforms to good effect (see for example, references [8], [9], [10] and [12]).

Briefly, the Gabor transform is computed by comparing a signal with translations and modulations of a single analysis window, thereby deriving the Gabor coefficients via an inner product. The reconstruction of the signal is computed by summing certain translations and modulations of a fixed synthesis (dual) window, weighted by the previously computed Gabor coefficients. Since the windows are localized in space, the inner product only detects a localized portion of the original signal, and thus the Gabor coefficents depend only locally on the signal information. The Gabor transform may be thought of as a windowed, or short-time, Fourier transform.

Not only is the Gabor representation of a signal useful for analysing a signal, it also provides a convenient means to modify the signal in the time/frequency domain. For instance, adjusting the magnitude of the Gabor coefficients in a prescribed manner is a simple way to implement a time-varying filter; reconstruction of the modified signal using the inverse Gabor transform completes the implementation. A full exploration of these filtering techniues leads to the theory of Gabor multipliers (see for instance [2]), although we are not directly concerned with them in this paper. Another motivation for using the Gabor transform is that it provides a representation where the physics of the process being modeled is clear, and proper filters may bedesigned.

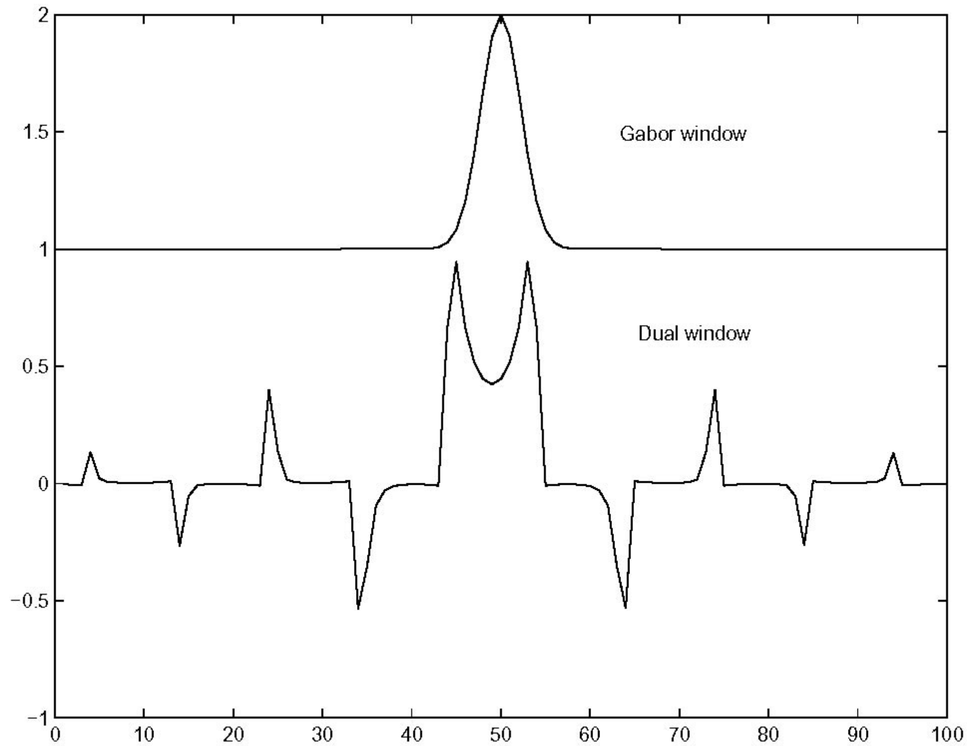

FIG.1. A Gabor window and its canonical dual.

In our seismic work, for instance, we have successfully implemented such time-variant filters to create commericially viable seismic data processing techniques, including deconvolution, forward and inverse Q-filtering, wavefield extrapolation, and wave propagation. In exploration seismology, data sets of a size on the order of tens of gigabytes are not uncommon; thus speed and robustness of any processing algorithm are critically important. This provides the motivation for the results presented in the present paper.

Here, we are concerned with the theoretcal issues that arise in a practical implementation of the Gabor transform. First of all, all our signals are sampled: that is, we are analysing functions in space $l^2(\mathbb{Z}^d)$ and so we are primarily interested in the discrete transform. Second, we want to make a choice of windows which will allow us to create a fast algorithm for the transform, with a fast inverse, while minimizing the appearance of artifacts in the final output. For instance, one often sees in the literature examples of Gabor windows and duals such as the pair illustrated in Figure 1; while the first (top) window is smooth and localized, the dual is objectionable both because of its many singular points, and because it is much less localized than the first window. In fact, we wish to be able to choose both analysis and synthesis windows so that we reduce artifacts, and so that the physics of the processes we are modeling in the time/frequency domain remain clear in this representatios. Often, this means we want both windows to be smooth and well-localized.

The principal new result of this paper is the observation that a simple choice of windows g and γ as factors of a partition of unity results in an exact reconstruction form for f; that is, g and γ so chosen gives directly the transformand inverse that recovers the original signal; the inverse frame operator is never required. Moreover, rather than translating a single window all over signal space, it can be preferable to start with a collection of windows $\{g_n\}$ and define a generalized Gabor transform; we still obtain exact inverses. Another major result of the paper is that we may translate and modulate over very general lattices in both the time and frequency domain, provided the partition of unity condition holds. Finally, if g and γ are factors of an approximate partition of unity, then the reconstruction is approximately exact.

The structure of the paper is as follows. In the following section, we review briefly the typical implementation of the discrete Gabor transform, using translations and modulations of a fixed window along a rectangular lattice. Section 3 moves away from translations of a single window, introducing a general collection of windows that cover signal space, and indicating the advantage of such an approach (namely, speed and reduction of processing artifacts). This is the starting point for the partition of unity approach to the Gabor transform. Section 4 moves away from a rectangular lattice in the frequency domain, again with the motivation being the removal of artifacts. Section 5 is the core of the paper, where we introduce the Gabor transform based on a partition of unity, and prove the key result on reconstruction of the signal from its Gabor transform. Section 6 discusses speed of the algorithm and its inverse, while Section 7 presents some common errors that can occur in implementation. Sections 8 and 9 discuss special cases of the partition of unity idea applied to the discrete Gabor transform on general lattices, and with Gaussian windows. Section 10 is a summary, and we have left to the appendices some examples of practical windows we have used in a variety of applications, with desirable properties for signal processing, and a proof of the modulation theorem in Section 4.

THE STANDARD GABOR TRANSFORM

The typical form of the Gabor transform is as follows: one begins with a Hilbert space $L^2(X)$ of square integrable functions on some measure space X, and a single function g in $L^2(X)$, then defines modulations and translations of g as

$$g_{mn} = M_m T_n g,$$

where T_n is translation by some amount indexed by n, and M_n is modulation, or multiplication by a complex exponential with some frequency indexed by m. The actual choice of index sets for m, n and indeed for the exact definition of translation and modulation, depends on the particular space X.

From this choice of g, for any function (signal) f in $L^2(X)$, one defines the Gabor coefficients of f, relative to g, as

$$V_{mn}(f) = \langle f, g_{mn} \rangle,$$

where $\langle f, g_{mn} \rangle$ is the standard inner product on $L^2(X)$. The frame operator S, for given window g, is defined as

$$Sf = \sum_{m,n} \langle f, g_{mn} \rangle g_{mn}$$

where the sum is not necessarily over all possible indices m,n, but perhaps some specified, select, subset of indices. When the frame operator is invertible, there is a canonical dual window given by $\gamma^c = S^{-1}g$, For instance, in Figure 1 we see a typical window and its canonical dual. With the dual, one also defines translations and dilations as

$$\gamma^c_{mn} = M_m T_n \gamma^c.$$

From these functions, one recovers f as the sum

$$f = \sum_{m,n} V_{mn}(f)\gamma^c_{mn},$$

where the sum is over the same subset as in the defining sum for S.

In general, any function $\gamma \in L^2(X)$ coud be used in an attempt to reconstruct f, so one can define a reconstruction operator $R = R^\gamma$ for any function γ as

$$Rf = \sum_{m,n} V_{mn}(f)\gamma_{mn}.$$

Once again, this is a sum over some subset of possible indices m,n. One challenge in Gabor theory is to find windows g and γ with suitable properties for signal processing, that also allow for perfect reconstructions, $Rf = f$.

In applications, the measure space X is usually a d-dimensional cube of points,

$$X = [0,1,...L_1 - 1] \times [0,1,...L_2 - 1] \times ... \times [0,1,...L_d - 1],$$

and the measure is just counting measure; the Hilbert space of square summable functions is usually denoted $l^2(X)$. Translations in X are just the operators $\{T_n : n \in X\}$ with

$$(T_n f)(x) = f(x-n),$$

where the difference $x-n$ is computed using modulo arithmetic, or by assuming f is zero-extended outside the cube. Modulations are given as operators $\{M_n : m \in X\}$, with

$$(M_m f)(x) = f(x)e^{2\pi i \frac{x_1 m_1}{L_1} + \frac{x_2 m_2}{L_2} + \ldots + \frac{x_d m_d}{L_d}}.$$

The Gabor coefficients

$$V_{mn}(f) = \langle f, g_{mn} \rangle$$

could be computed for all indices m, n in X, but in practice these are only computed for some fixed sublattice in X. The question of which sublattice leads directly to questions of over- and under-sampling.

We point out three important limitations in this standard implementation. First, only one window g gets translated to cover signal space; why not use several different windows? (Section 3) Second, a rectangular lattice specifes the collection of vectors used to sample in the frequency domains; one can use a more general lattice (Section 4). Finally, there may be many possible choices for the dual window that give an accurate reconstruction of f. We create a method toproduce "nice" windows by a factorization of a partition of unity (Section 5).

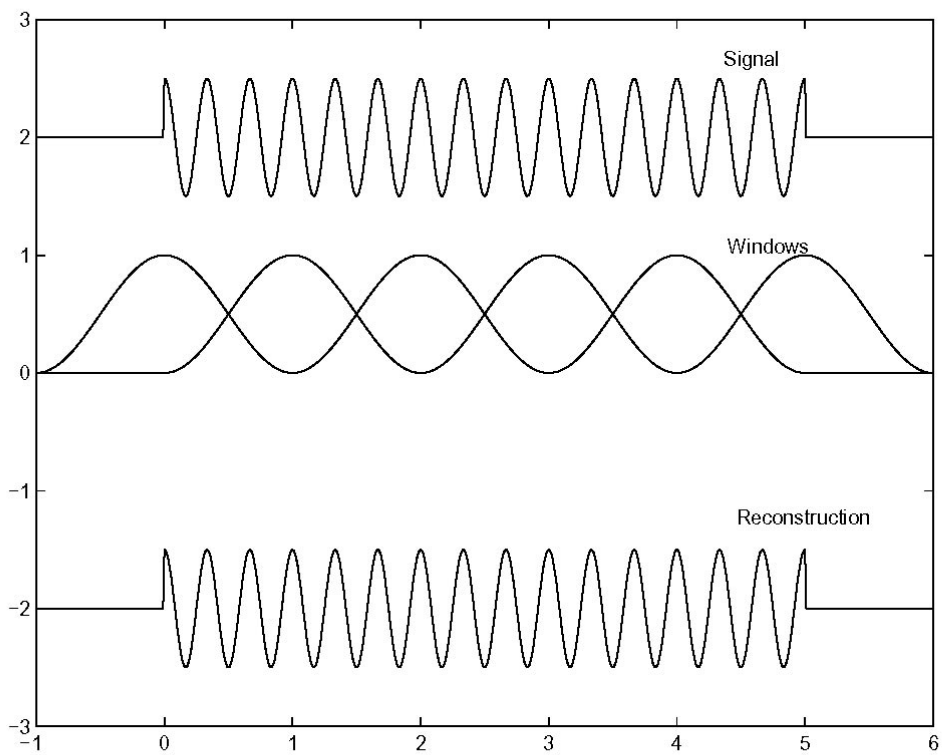

FIG. 2. Regular window translated along signal space.

GENERAL WINDOWS

The standard implementation of the discrete Gabor transform uses a single window function g, defined on discrete points, which is translated along the length of the signal, as in Figure 2.

In practice, there are a few problems with this regular structure. First of all, in certain areas the signal may be changing rapidly, while in other areas it may have a more uniform character. It would be useful to use short windows in the first case, to accurately track those changes, and use larger windows in the second case. (Grossman's work on the adaptive Gabor transform demonstrates this approach, see [10].) The second problem is that when applying a time-varying filter, typically we get edge effects near the sides of any window, and regularly spaced windows will give regularly spaced artifacts. It is important to point out, of course, that artifacts never appear if one only computes Gabor coefficients, and then inverts, as the numerical inversion is essentially exact. However, when modifying the Gabor coefficients before inverting (as when implementing a time-varying filter, or applying a Gabor multiplier), there will be changes in the output signal; unwanted changes we loosely describe as artifacts. Any artifacts, particularly regularly-spaced artifacts, can be confusing or annoying to the person viewing the Gabor coefficients and can obfuscate the processed data. For instance, in image processing, one often observes artifacts that lie on a rectangular boundary (the jpeg effect, see [14]). Perhaps a more optimal tiling of the image space could be useful; for instance, the hexagonal tiling shown in Figure 3 corresponds to an optimal packing of discs in the plane. Even a non-periodic tiling might prove to be useful. In our seismic experiments, it is numerically troublesome to have artifacts that accumulate along the direction of wave propagation, so again, something other than regular, rectangular division of the image space is useful.

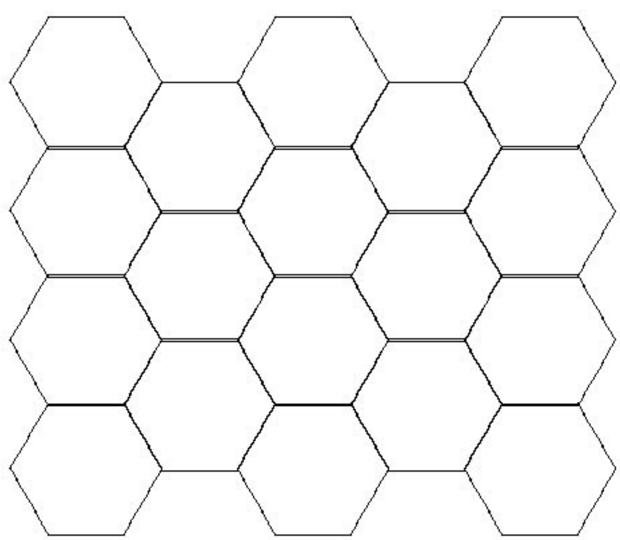

FIG. 3. A hexagonal tiling of the plane.

Thus, we propose generating a Gabor transform using a collection of windows $\{g_n\}$, and duals $\{\gamma_n\}$ which appropriately cover signal space, as in Figure 4. Although the functions g_n are not related to each other by translation, we may still define the modulation appropriately, as

$$g_{mn} = M_m g_n$$

and form Gabor coefficients of signal f as

$$V_{mn}(f) = \langle f, g_{mn} \rangle,$$

and similarly for the reconstruction operator, using the general collection of dual windows $\{\gamma_n\}$ rather than a collection of translates.

The next section shows how we avoid rectangular lattices in the frequency domain, and Section 5 pulls this all together to create a general Gabor transform.

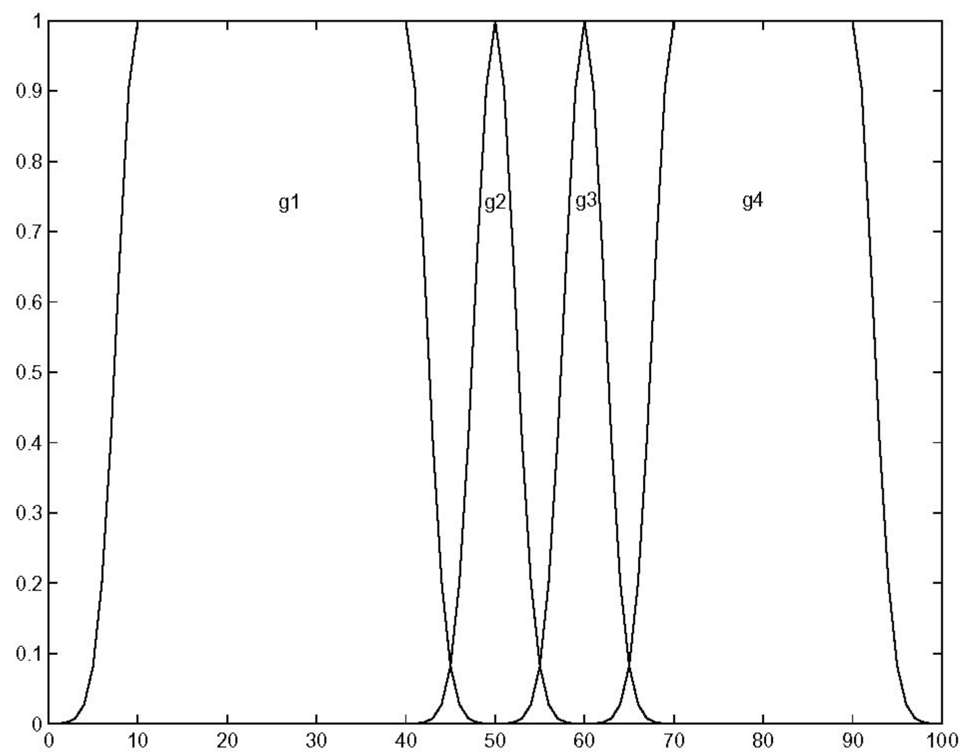

FIG. 4. A collection of four non-uniform windows covering signal space.

MODULATION FUNCTIONS AND INTEGER LATTICES

In the Gabor transform, the window functions are modified by modulation functions which, in one dimension, are simply complex exponential functions of the form

$$P_\alpha(j) = e^{2\pi i \alpha j} \qquad \text{for all } j \in \mathbb{Z},$$

for some fixed parameter α. In applications, it is common to choose only periodic modulation functions, obtained by setting parameter α to be a rational number, say $\alpha = m/M$. In higher dimensions, one may take a product of several one dimensional modulaton functions, and obtain functions

$$P_m(j) = e^{2\pi i m_1 j_1/M_1} e^{2\pi i m_2 j_2/M_2} \ldots e^{2\pi i m_d j_d/M_d}$$

for all $j = (j_1, j_2, \ldots j_d)$ in $\mathbb{Z}^d$, where $m = (m_1, m_2, \ldots m_d)$ is an index in $\mathbb{Z}^d$ and $M_1 \ldots M_d$ is some fixed choice of integer denominators. It is convenient to fix a diagonal matrix B with integer entries

$$B = \begin{bmatrix} M_1 & 0 & \ldots & 0 \\ 0 & M_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & M_d \end{bmatrix},$$

and then express the above modulation function via the standard inner product on $\mathbb{Z}^d$ as

$$P_m(j) = e^{2\pi i (B^{-1}m)\cdot j} \qquad \text{for all } j \in \mathbb{Z}^d.$$

In this form, it is clear that each modulation function is periodic in each component of j, and the "vector of frequencies" represented by function P_m is the vector $B^{-1}m$. Less obvious, but also true, is that there are only finitely many different functions P_m (as two different indices m, m' can give the functions $P_m = P_{m'}$), and a sum of the different P_m gives a delta function.

However, this is too restrictive a class of modulation functions, as essentially it restricts us to sampling in the frequency domain with a simple rectangular lattice, in this case the set of points $\frac{1}{M_1}\mathbb{Z} \times \frac{1}{M_2}\mathbb{Z} \times \cdots \times \frac{1}{M_d}\mathbb{Z}$. Just as in the time domain, we wished to moved away from rectangular sample lattices, so here too we may do so in the frequency domain.[1] The key is using a more general form for matrix B and considering the lattices these matrices generate.

A lattice is, roughly speaking, a regularly spaced collection of points in Euclidean space $\mathbb{R}^d$; more precisely, it is a discrete subgroup of $\mathbb{R}^d$ under the operation of vector

[1]Our approach to lattices is quite a bit different than the symplectic case as considered by Grochenig, [7].

addition. The standard example of a discrete lattice is the subset $\mathbb{Z}^d$ of points in Euclidean space with integer coordinates. For the Gabor transform, we are interested in more general lattices that "fill out space," which is to say, are not confined to some hyperplane in $\mathbb{R}^d$, yet the points don't "bunch up" into a dense set; such lattices may always be represented as the image of the standard lattice $\mathbb{Z}^d$ under a linear transformation. That is, there is an invertible $d \times d$ matrix $A \in M_d(\mathbb{R})$ with the lattice given as the set

$$A\mathbb{Z}^d = \{Az : z \in \mathbb{Z}^d\}.$$

Figure 5 shows a portion of a typical lattice in the plane.

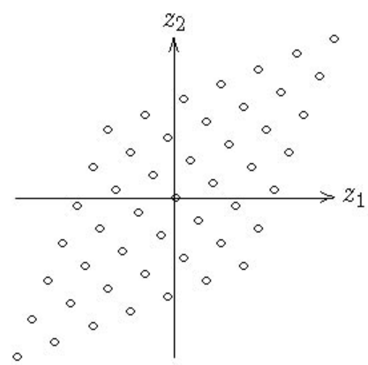

FIG. 5. A non-rectangular lattice in the plane.

In the discussion on modulation functions, the diagonal matrix B defines a lattice of frequencies, namely the set

$$B^{-1}\mathbb{Z}^d = \{B^{-1}z : z \in \mathbb{Z}^d\}.$$

The matrix B also defines a lattice of periodicity for the modulation functions,

$$B\mathbb{Z}^d = \{Bz : z \in \mathbb{Z}^d\},$$

which (since B is integer), is a subset of the standard lattice $\mathbb{Z}^d$ in $\mathbb{R}^d$. By lattice of periodicity, we simply mean that if indices j and j' differ by some integer multiple of B (i.e. $j' - j \in B\mathbb{Z}^d$), then P_m takes the same values there, $P_m(j') = P_m(j)$.

To generalize the frequency sampling to non-rectangular lattices, fix B to be any invertible $d \times d$ matrix with integer entries, and for any index $m \in \mathbb{Z}^d$, define modulation function P_m by

$$P_m(j) = e^{2\pi i (B^{-1}m)\cdot j}, \qquad \text{for all } j \in \mathbb{Z}^d.$$

We also define two equivalence relations on the lattice $\mathbb{Z}^d$ by

$$m \sim m' \qquad \text{iff} \qquad m - m' \in B\mathbb{Z}^d,$$

and

$$m \sim^* m' \qquad \text{iff} \qquad m - m' \in B^*\mathbb{Z}^d,$$

where B^* is the transpose of matrix B.

The following proposition is easily proved in the case where B diagonal; the proof for general B is deferred to an appendix.

Proposition 1. Let B be an invertible $d \times d$ matrix with integer entries, defining equivalence relations $\sim$ and $\sim^*$, and modulation functions $\{P_m : m \in \mathbb{Z}^d\}$ as above. Then

1. each function P_m is periodic, with $j \sim^* j' \Rightarrow P_m(j) = P_m(j')$;
2. $P_m = P_{m'}$ if and only if $m \sim m'$;
3. there are exactly $|\det B|$ different modulation functions P_m, which are uniquely indexed by some finite cube
$$M = [0,...,M_1 - 1] \times [0,...,M_2 - 1] \times ... \times [0,...,M_d - 1] \subset \mathbb{Z}^d;$$
4. the finite sum of distinct modulation functions
$$P = \sum_{m \in M} P_m$$
is a delta function, with
$$P(j) = \begin{cases} |\det(B)| & \text{if } j \in B^*\mathbb{Z}^d \\ 0 & \text{otherwise,} \end{cases}$$

In the following sections, there may be only one fixed modulation matrix B, but in general, the matrix may vary with the choice of window function.

THE DISCRETE GABOR TRANSFORM ON A PARTITION OF UNITY

A *partition of unity* is a collection of functions which sum to the constant function one; that is, a collection of functions $\{w_n\}$ which sum as

$$\sum_n w_n(j) = 1 \qquad \text{for all } j \in \mathbb{Z}^d.$$

For typical signal processing applications, it will be useful to choose the w_n with particular characteristics; say smooth, non-negative, rapidly decreasing, and/or with a well-behaved Fourier transform. In Appendix A, we list some examples of practical windows that we have used in a variety of signal processing applications. In the following, however, the only restriction we might make on each w_n is that it have compact (i.e. finite) support. Even this restriction may be dropped: see the remark at the end of the section.

Given a partition of unity $\{w_n\}$ on the set $\mathbb{Z}^d$, with each function w_n of compact support, choose for each index n an invertible integer matrix B_n so that the lattice $B_m^*\mathbb{Z}^d$ intersects the difference set [2]

$$\mathrm{supp}(w_n) - \mathrm{supp}(w_n) \equiv \{j-k : j,k \in \mathrm{supp}(w_n)\}$$

only at the origin. That is, choose B_n with entries large enough so that the non-zero entries in the lattice $B_m^*\mathbb{Z}^d$ are far enough from the (finite) set of differences of pairs of elements in $\mathrm{supp}(w_n)$.

Then, choose functions g_n and γ_n to form what we call a weighted factorization of the partition of unity; that is, g_n and γ_n are chosen with

i) $\mathrm{supp}(g_n) = \mathrm{supp}(\gamma_n) = \mathrm{supp}(w_n)$;

ii) $\gamma_n \overline{g_n} \, |\det B_n| = w_n$ on the set $\mathbb{Z}^d$.

For instance, with a non-negative partition of unity, one may choose

$$g_n = \left(\frac{w_n}{|\det B_n|}\right)^p$$

$$\gamma_n = \left(\frac{w_n}{|\det B_n|}\right)^{1-p}$$

for some real parameter $0 \le p \le 1$. (In applications, the choice of p is important in controlling the rate of roll-off for a window.)

For each window index n, and for each integer vector $m \in \mathbb{Z}^d$, the modulated version of the window functions are denoted by g_{mn} and γ_{mn} and are defined as

$$g_{mn}(j) = g_n(j) e^{2\pi i (B_n^{-1} m)\cdot j}$$

$$\gamma_{mn}(j) = \gamma_n(j) e^{2\pi i (B_n^{-1} m)\cdot j}$$

for all vectors $j \in \mathbb{Z}^d$. The Gabor transform of function $f \in l^2(\mathbb{Z}^d)$ is defined as

[2]the Minkowski difference

$$V_{mn}(f)=\langle f,g_{mn}\rangle$$

and the inverse tranform given by

$$\tilde{f}=\sum_{m,n}V_{mn}(f)\gamma_{mn}.$$

By this careful choice of window functions, we have the following result.

Proposition 2. With this choice of modulated window functions g_{mn} and γ_{mn} construction from a partition of unity, we have the exact reconstruction formula

$$f=\sum_{m,n}\langle f,g_{mn}\rangle\gamma_{mn}$$

for all functions $f\in l^2(\mathbb{Z}^d)$. That is, the Gabor transform with windows g_n has an inverse transform with windows γ_n.

Proof. We compute

$$\begin{aligned}\sum_{m,n}\langle f,g_{mn}\rangle\gamma_{mn}(j)&=\sum_{m,n,k}f(k)\overline{g_n}(k)e^{-2\pi i(B_n^{-1}m)\cdot k}\gamma_n(j)e^{2\pi i(B_n^{-1}m)\cdot j}\\&=\sum_{n,k}f(k)\overline{g_n}(k)\gamma_n(j)\sum_m e^{2\pi i(B_n^{-1}m)\cdot(j-k)}\end{aligned}$$

where the inner sum (over m) we recognize as the sum of modulation functons, which is non-zero only when $j-k$ is in the lattice $B^*\mathbb{Z}^d$. The term $\overline{g_n}(k)\gamma_n(j)$ is non-zero only when $j-k$ is in the finite difference set

$$\text{supp}(\gamma_n)-\text{supp}(g_n)$$

which, by construction, is the same set as $\text{supp}(w_n)-\text{supp}(w_n)$. But the intersection

$$B^*\mathbb{Z}^d\cap\left(\text{supp}(w_n)-\text{supp}(w_n)\right)$$

is just the origin, and so the above sum collapses to the non-zero term with $j=k$, and becomes

$$\begin{aligned}&=\sum_n f(j)\overline{g_n}(j)\gamma_n(j)\,|\det B_n|\\&=f(j)\sum_n w_n(j)\\&=f(j),\end{aligned}$$

since the w_n sum to one.

Having chosen the analysis windows g_n with small support relative to matrix B_n, the frame operator is also in a particularly simply form, as indicated in the following.

Proposition 3. With window functions g_m chosen as above, the frame operator

$$Sf = \sum_{m,n} \langle f, g_{mn} \rangle g_{mn}$$

is a multiplication operator on $l^2(\mathbb{Z}^d)$, by values

$$s(j) = \sum_n | g_n(j) |^2 | \det B_n |.$$

Proof. We compute the pointwise values

$$\begin{aligned}(Sf)(j) &= \sum_{m,n} \langle f, g_{mn} \rangle g_{mn}(j) \\ &= \sum_{n,k} f(k)\overline{g_n}(k) g_n(j) \sum_m e^{2\pi i (B_n^{-1} m)\cdot(j-k)}\end{aligned}$$

which collapses, as in the previous proof, to the sum

$$\begin{aligned}&= \sum_n f(j)\overline{g_n}(j) g_n(j) | \det B_n | \\ &= f(j) s(j)\end{aligned}$$

with $s(j) = \sum_n | g_n(j) |^2 | \det B_n |$.

There is a well-developed theory for the expansion of functions by non-orthogonal families of "basis" functions, namely frame theory, which is beyond the scope of this article (see for instance [1]). However, it is useful to insert thefollowing observation, which is relevant to frame theory.

Corollary 4. A tight frame is obtained if and only if the functions w'_n defined by

$$w'_n(j) = | g_n(j) |^2 | \det B_n |$$

form a partition of unity on $\mathbb{Z}^d$, times a fixed constant.

Proof. This is a consequence of the observation that the frame is tight if and only if the frame operator S is a scalar multiple of the identity; equivalently, that the multiplication operator S with values $s(j)$ is a constant function. Since $s(j) = \sum_n w'_n(j)$, tightness is equivalent to the w'_n forming a partition of unity, up to scaling by some constant.

It is worth observing that if the original partition of unity $\{w_n\}$ is non-negative, the choice of analysis window

$$g_n = \left(\frac{w_n}{|\det B_n|} \right)^{1/2}$$

will give a tight frame. It is also interesting to observe that once an analysis window g_n is chosen with suitably small support relative to $B_n^* \mathbb{Z}^d$, then *any* choice of dual window with small support (including the canonical dual $\gamma_n^c = S^{-1} g_n$) gives rise to a partition of unity, as noted in the following proposition.

Proposition 5. Suppose analysis windows g_n and matrix B_n are chosen such that $\mathrm{supp}(g_n)$ is finite, and the intersection

$$B_n^* \mathbb{Z}^d \cap \left(\mathrm{supp}(g_n) - \mathrm{supp}(g_n) \right)$$

contains only the origin. If γ'_n is *any* choice of dual windows, with support the same as the corresponding g_n, that satisfies the reconstruction

$$f = \sum_{m,n} \langle f, g_{mn} \rangle \gamma'_{mn} \qquad \text{for all } f \in l^2(\mathbb{Z}^d),$$

then the functions

$$w'_n = \gamma' \overline{g_n} \, |\det B_n|$$

form a partition of unity for the set $\mathbb{Z}^d$.

Proof. As in the calculation above, we have

$$\begin{aligned} f(j) &= \sum_{m,n} \langle f, g_{mn} \rangle \gamma'_{mn}(j) \\ &= f(j) \sum_n \gamma'_n(j) \overline{g_n}(j) \, |\det B_n| \\ &= f(j) \sum_n w'_n(j). \end{aligned}$$

The only way this equality can hold for all f is if the w'_n sum to one.

For the canonical dual $\gamma_n^c = S^{-1} g_n$, since operator S is simply an invertible multiplication operator, the support of γ_n^c is the same as the support of g_n, and thus the canonical dual times $\overline{g_n} \, |\det B_n|$ forms a partition of unity $\{w'_n\}$. It is important to notice this is usually not the well-designed partition of unity $\{w_n\}$ that we started out with, and the canonical dual does not necessarily have the nice properties, such as smoothness, that we designed in the factorization. Again, Figure 1 is a typical example of a poor canonical dual.

Thus, in standard Gabor theory, there often are partitions of unity lurking around in the background. Our approach in this paper can be summarized as saying that we begin with a well-designed partition of unity, and create well-behaved windows from this partition.

Remark: In certain applications, it is convenient to chose some windows with non-finite support; for instance, in some filtering applications, it is useful to set $g_n = w_n$ and $\gamma_n \equiv 1$. Technically, this doesn't quite work, as one recovers (in the reconstruction) a periodization of the original signal. However, all is not lost: by a good choice of the frequency lattice, one can ensure the reconstruction is exact on the support of signal f, and just ignore the periodization that occurs outside the support. More precisely, in many applications, there is some reasonable finite set $F \subset \mathbb{Z}^d$ such that every signal f of interest has support in F; this set F may be used to truncate the window γ_n to finite support. Givena partition of unity $\{w_n\}$, we wish to chose a factorization with

$$w_n = \gamma_n \overline{g_n} \, |\det B_n| \qquad \text{on set } F,$$

where for the moment, matrix B_n is unspecified. One can choose supports for the windows to lie in the finite set F, and also

$$\operatorname{supp}(\gamma_n) \cap \operatorname{supp}(g_n) = \operatorname{supp}(w_n) \cap F.$$

Now choose matrix B_n so that

$$B_n^* \mathbb{Z}^d \cap \left(\operatorname{supp}(\gamma_n) - \operatorname{supp}(g_n)\right) = \{0\}.$$

All the result above apply. Thus larger support are possible for γ_n or g_n than specified by the partition of unity. However, the price paid is that matrix B_n may have a large determinant, which means many modulation functions are required inanalysis and reconstruction.

SPEED OF INVERSE

The discrete Gabor transform may be implemented in a straightforward manner using the Fast Fourier Transform. The Gabor coefficients

$$\begin{aligned} V_{mn}(f) &= \langle f, g_{mn} \rangle \\ &= \left\langle f \overline{g_n}, P_m \right\rangle \end{aligned}$$

are obtained by fixing n, and taking a discrete Fourier transform of the windowed signal $f\overline{g_n}$, over the frequencies determined by matrix $(B_n)^{-1}$. Of course, for the fastest FFT implementation, B_n should be chosen to be a diagonal matrix with powers of 2 along the diagonal. With window function g_n of finite support, or more generally with signal f of finite support, this truly gives a discrete, finite Fourier transform.

Other implementations of the forward Gabor transform use periodization techniques to speed up the transform computations. A MATLAB routine provided by Feichtinger [4] is particularly effective. Figure 6 shows a comparision of Feichtinger's frame-based implementation to the direct method via the fast Fourier transform. Clearly, Feichtinger's implementation is much faster than the direct method.

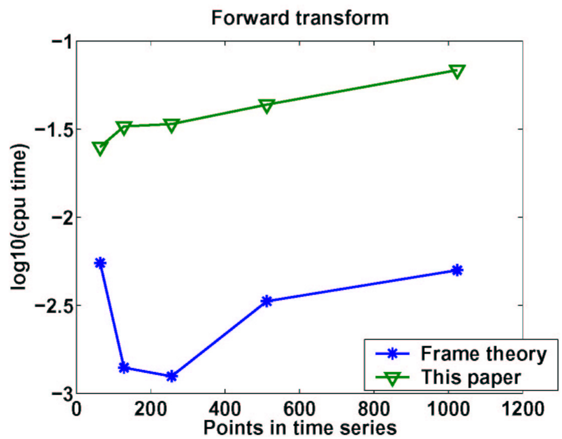

FIG. 6. Comparison of speed: frame-based implementation, and direct method.

For the inverse transform arising from a partition of unity, a direct implementation by the FFT is also possible. The reconstruction of signal f is obtained from the Gabor coefficients V_{mn} as

$$\begin{aligned} f(j) &= \sum_{m,n} V_{mn}\gamma_{mn}(j) \\ &= \sum_{n} \gamma_n(j)\left[\sum_{m} V_{mn} e^{2\pi i (B^{-1}m)\cdot j}\right] \\ &= \sum_{n} \gamma_n(j)\mathrm{iFFT}_m(V_{mn})(j); \end{aligned}$$

that is, for each index n, we take the inverse FFT of the coefficients V_{mn} (transforming over index m), then take the pointwise product with γ_n to obtained a windowed slice of the output. Summing over each slice (indexed by n) recovers the signal f. (Keep in mind, the index set for m is some finite cube as determined by matrix B_n, and in practice there are only finitely many windows as indexed by n, so this is all a finite calculation.)

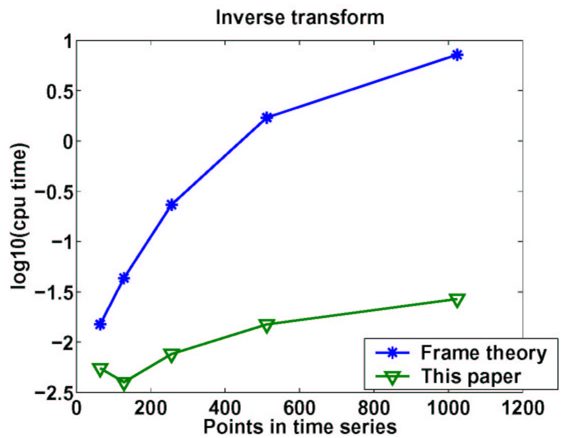

FIG. 7. Inverse speed: frame-based implementation, and our method.

This direct FFT implementation of the inverse transform is quite fast. Figure 7 compares this method with code we obtained from Feichtinger. In this case, our POU technique allows for a significant speedup for computing the inverse transform; in particular, we never need to use the frame operator S explicitly, and never need to compute its inverse. The total time for our method, combining both the forward and inverse transforms, is generally much faster, especially for long time series.

Although we have not done a detailed analysis of why our method is faster, or whether there is room for improvement, it has allowed us to implement a practical method for Gabor deconvolution, a technique using Gabor multipliers to enhange the resolution of images in seismic data processing [13].

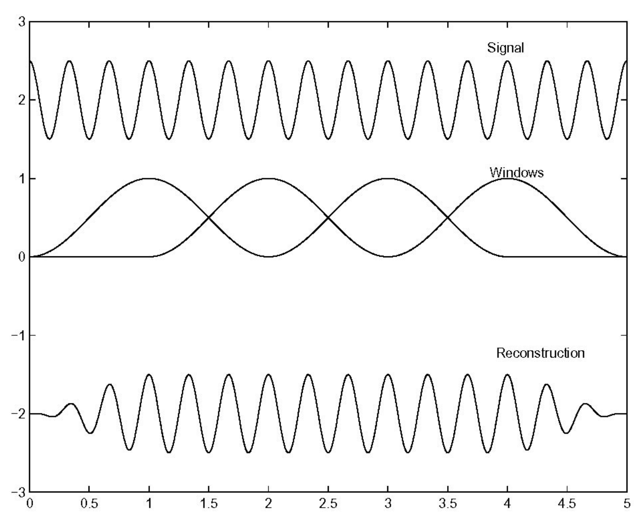

FIG. 8. Too few windows give a poor reconstruction.

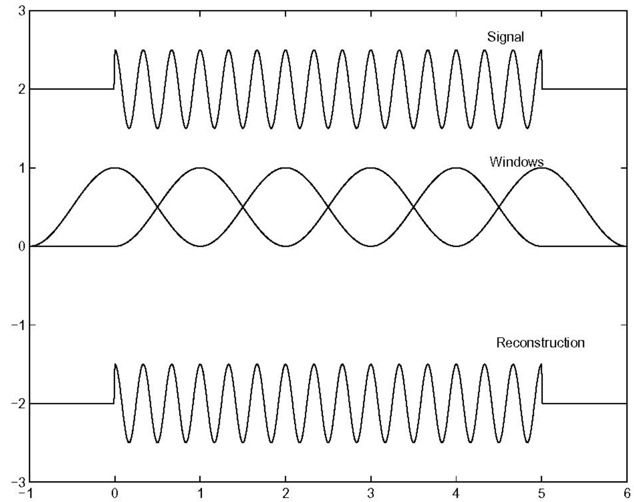

FIG. 9. Windows across the edge give a proper reconstruction.

IMPLEMENTATION ERRORS

In many applications, the signal f to be analyzed has compact support; that is, for sampled signals, only finitely many sample values are nonzero. It is convenient to choose windows with compact support, appropriately spread out, so that only finitely many Gabor coefficients $\langle f, g_{mn} \rangle$ are non-zero. Thus, complete information about the signal in Gabor space may be stored in finite number of coefficients – namely, coefficients for those (translated) windows whose support intersects with the support of the original signal.

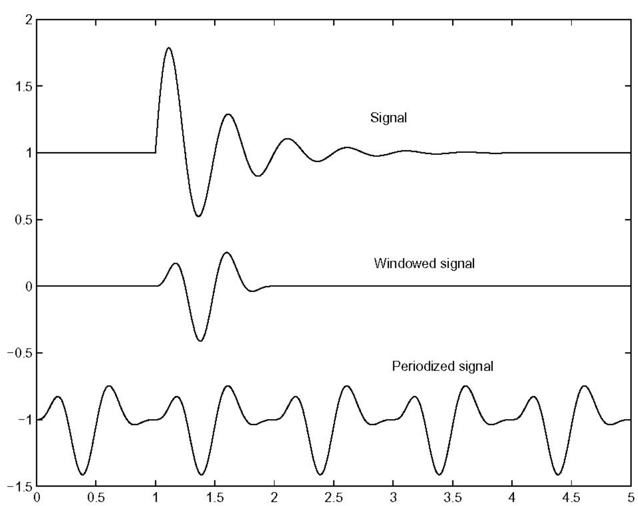

FIG. 10. Effective periodization of a signal in Gabor domain.

However, a common implementation error is to use coefficients for a smaller set of windows, often those windows whose support is a subset of the support of the signal –

this is often for practical considerations in programming, where it is difficult to increase the size of the data array where the signal lives. This almost always leads to an "edge effect" type of error, when the signal is not properly reconstructed near the ends of the signal. Figure 8 shows a typical example, where a signal supported on interval $[0,5]$ is analyzed, and re-synthesized with raised cosine windows each of length 2. By not properly dealing with the edges, an error in reconstruction occurs. This type of error can occur even when no filtering is done.

In Figure 9, two additional windows are included, and this gives the proper reconstruction of the signal. In general, the key to a successful implementation, including in higher dimensions, is to use all windows translated via the lattice whose support intersects with the support of the signal. For non-compactly supported windows, such as Gaussians, every window may intersect with the signal and thus any finite implementation introduces some error; with Gaussians, this error can be made arbitrarily small by including sufficiently many.

Another potential source of error (or "edge effects") again caused by a programmer's insistence on not letting data arrays "get bigger than they need to be" is caused by the implicit periodization of a signal in the Gabor domain. When represented in the Gabor domain, a single windowed signal is equivalent to periodic version of the signal, repeating out to infinity. Performing a smoothing operation, such as a convolution, can cause a portion of the periodic signal to "leak" back into the windowed area; reconstruction would then include these leakages. Figure 10 shows a conceptualization of this effect. This type of error only occurs when filtering (or multiplication in the Gabor domain) has been implemented.

The solution to this is, of course, to chose matrix B with even larger entries; effectively, this increases the length of the period, and so leakage errors have further to travel before mapping back into the windowed area. More precisely, given the windowed signal in some interval $[m,n]$, one creates a zero pad by increasing the interval to $[m-a,n+a]$, and perform the FFT on this larger interval. (More data points corresponds to larger B.) Indeed, in some applications, it can be useful to use the whole support of signal f as the analysis interval for each windowed signal — wraparound errors are almost never a problem then.

THE DISCRETE GABOR TRANSFORM ON A LATTICE

Notwithstanding our interest in general partitions of unity, it is useful to restrict to the case where the partition arises as a collection of translations of a single function. More precisely, we wish to analyze functions f in $l^2(\mathbb{Z}^d)$, via translations and modulations of a given window function g along some lattice, and resynthesize it with a dual function γ. To this end, fix g and γ as bounded functions on $hbbR^d$, translation matrix $A \in M_d(\mathrm{R})$ an invertible $d \times d$ matrix with real entries, and inverse frequency matrix $B \in M_d(\mathbb{Z})$ an invertible $d \times d$ matrix with integer entries.

We point out explicitly that here, the analysis and synthesis windows are more generally functions on Euclidean space (for instance, Gaussians), and one is permitted to translate them by arbitrary real vectors – typically, these windows and their translates are computed on the fly, and there is no need to restrict the translations to the standard lattice. Indeed, in many applications, it is advantageous to use these general forms for the window functions. In constrast, the function f is defined on the regular lattice $\mathbb{Z}^d$, as it usually comes from sampled data in real applications. One must require some moderate conditions on the decay of g,γ to ensure the sum of their translates converge; piecewise continuous with compact support, or integrable (in $\mathbb{R}^d$) is sufficient for our purposes here. Also note the inverse frequency matrix B need not be diagonal, as again there may be advantages to sampling on a non-standard frequency lattice. Integer entries for B, however, are required for the reconstruction theorem.

The set of points $\{An : n \in \mathbb{Z}^d\}$ forms a discrete lattice in $\mathbb{R}^d$, while the set $\{Bm : m \in \mathbb{Z}^d\}$ forms a sublattice of the discrete lattice $\mathbb{Z}^d$. A translate of function g along the lattice $A\mathbb{Z}^d$ is defined by

$$g_n(x) = g(x - An), \text{ for all } x \in \mathbb{R}^d,$$

while the modulation of g_n by frequencies $B^{-1}m$ is defined as

$$g_{mn}(x) = g_n(x)e^{2\pi i(B^{-1}m)\cdot x}, \text{ for all } x \in \mathbb{R}^d.$$

The translations and modulation of the dual window γ are defined similarly. (The order of operation of translation and modulation is important. We have fixed it here with translations first, but observe that in some applications, it is relevant whether one is measuring phase relative to the signal, or relative to the translated window.)

The discrete Gabor transform (for given A, B, g and γ) is defined as a map of functions $f \in l^2(\mathbb{Z}^d)$ to $V(f) \in l^\infty(\mathbb{Z}^d \times \mathbb{Z}^d)$ using the Gabor coefficients

$$V_{mn}(f) = \langle f, g_{mn} \rangle = \sum_k f(k)\overline{g}(k - An)e^{-2\pi i(B^{-1}m)\cdot k},$$

where $\langle , \rangle$ denotes the usual inner product on $l^2(\mathbb{Z}^d)$. Since the inverse of matrix B occurs in the modulation, it is immediately clear that $V_{mn}(f)$ is periodic in m, with period B; that is $V_{mn}(f) = V_{m'n}(f)$ when $m \sim m'$ (modulo $B\mathbb{Z}^d$). Thus the Gabor transform is completely determined on the finite quotient set $\mathbb{Z}^d / B\mathbb{Z}^d$, and the reconstruction of f may be defined by the sum

$$\tilde{f} = \sum_{\substack{m \in \mathbb{Z}^d / B\mathbb{Z}^d, \\ n \in \mathbb{Z}^d}} V_{mn}(f) \gamma_{mn}.$$

That is, the reconstruction of f is obtained by taking a linear combination of translations and modulation γ_{mn} of the dual window function γ, using weights $V_{mn}(f)$, which come from the Gabor transform of f. The remarkable result is that this reconstruction depends only sparsely on the original function f; that is, the matrix representing this linear transformation from f to $\tilde{f}$ has many zeroes — we only need sum over equivalence classes in the sublattice $B^*\mathbb{Z}^d$.

Proposition 6. The reconstruction $\tilde{f}$ satisfies

$$\tilde{f}(j) = |\det(B)| \sum_{j' \sim^* j} c(j, j') f(j')$$

for all $j \in \mathbb{Z}^d$, where the sum is over all indices j' equivalent to j modulo $B^*\mathbb{Z}^d$, and c is the correlation function between the two functions γ, g over the lattice generated by A, given as

$$c(j, j') = \sum_n \gamma(j - An) \overline{g}(j' - An).$$

Proof. We compute:

$$\begin{aligned}
\tilde{f}(j) &= \sum_{m,n} V_{mn}(f) \gamma_{mn}(j) \\
&= \sum_{m,n} V_{mn}(f) \gamma(j - An) e^{2\pi i (B^{-1}m) \cdot j} \\
&= \sum_{m,n} \Big(\sum_{j'} f(j') \overline{g}(j' - An) e^{-2\pi i (B^{-1}n) \cdot j'} \Big) \gamma(j - An) e^{2\pi i (B^{-1}m) \cdot j} \\
&= \sum_{n,j'} f(j') \overline{g}(j' - An) \gamma(j - An) \left[\sum_m e^{2\pi i (B^{-1}m) \cdot (j - j')} \right] \\
&= \sum_{n,j'} f(j') \overline{g}(j' - An) \gamma(j - An) P(j - j')
\end{aligned}$$

where we recognize the last sum over $m \in M$ as a sum of modulation functions as in Proposition 1, and thus this sum is equal to $|\det B|$ when $j - j' \in B^*\mathbb{Z}^d$, and zero otherwise. Thus, the sum over j' collapses to a sum over those j' equivalent to j modulo $B^*\mathbb{Z}^d$, and we write

$$\begin{aligned}\tilde{f}(j) &= |\det B| \sum_{n,j'\sim^* j} f(j')\overline{g}(j'-An)\gamma(j-An) \\ &= |\det B| \sum_{j'\sim^* j}\left[\sum_n \overline{g}(j'-An)\gamma(j-An)\right] f(j') \\ &= |\det B| \sum_{j'\sim^* j} c(j,j')f(j')\end{aligned}$$

where $c(j,j') = \sum_n \gamma(j-An)\overline{g}(j'-An)$ is the correlation function.

Combining with the results in Section 5, we obtain the following.

Corollary 7. If invertible integer matrix B is chosen so that

$$\left(\mathrm{supp}(\gamma) - \mathrm{supp}(g)\right) \cap B^*\mathbb{Z}^d = \{0\}$$

and the functions $w_n = \gamma_n g_n$ form a partition of unity on $\mathbb{Z}^d$, then the reconstruction above satisfies

$$\tilde{f} = |\det B|\, f \qquad \text{for all } f \in l^2(\mathbb{Z}^d).$$

Proof. As in the previous section, the correlation function

$$c(j,j') = \sum_n \gamma(j-An)\overline{g}(j'-An)$$

is non-zero only when $j-j'$ is in the set $\mathrm{supp}(\gamma)-\mathrm{supp}(g)$. In the sum for the reconstruction formula, we only get non-zero terms when $j-j'$ is in the lattice $B^*\mathbb{Z}^d$; combined with the observation about where c is non-zero, we conclude the sum collapses, and thus

$$\tilde{f}(j) = |\det B|\, c(j,j) f(j).$$

With $\gamma_n g_n$ a partition of unity, these diagonal entries $c(j,j)$ are exactly one, and thus

$$\tilde{f}(j) = |\det B|\, f(j).$$

One easy way to obtain a partition of unity over lattice translations is simply to fix a non-negative function $v : \mathbb{R}^d \to \mathbb{R}$ and symmetrize over the lattice $A\mathbb{Z}^d$. That is, we define

$$w(x) = \frac{v(x)}{\sum_{n\in Z^d} v(x-An)} \qquad \text{for all } xin\mathbb{R}^d.$$

Provided the denominator is never zero, this gives rise to a partition of unity by translating along $A\mathbb{Z}^d$. A fatorization of the partition of unity may be obtained by setting

$$g(x) = w(x)^p$$
$$\gamma = w(x)^{1-p}$$

for any real parameter $0 \le p \le 1$.

The constructions the appendix show other ways to obtain practical partitions of unity.

GAUSSIANS AND THE GABOR TRANSFORM

The original transform of Denis Gabor was defined using Gaussian windows on Euclidean space, for analysizing functions in $L^2(\mathbb{R}^d)$. These Gaussians, restricted to the discrete lattice $\mathbb{Z}^d$, are useful windows for implementing the discrete Gabor transform. For instance, the corresponding discrete correlation function is almost constant on the diagonal, and vanishes rapidly off the diagonal. Thus, while not exact, the reconstruction formula gives an approximate rescaling of the original signal f. Thefollowing provides a precise statement of these facts.

Proposition 8. Suppose functions γ and g are Gaussians on $\mathbb{R}^d$; that is, there are constants $a, a' > 0$ with

$$\gamma(x) = e^{-ax\cdot x}$$
$$g(x) = e^{-a'x\cdot x}$$

for all $x \in \mathbb{R}^d$. For a, a' sufficiently small, the correlation function is nearly constant on the diagonal.

Proof. More precisely, we will prove for any $\varepsilon > 0$, we may choose a, a' sufficiently small that

$$C(1-\varepsilon) \le c_{\gamma,g}(j,j) \le C(1+\varepsilon), \text{ for all } j \in \mathbb{Z}^d,$$

where C is some normalizing constant. This result follows from the observation that the sum of integer translates of a one-dimensional Gaussian e^{-ax^2} is an even periodic function with cosine expansion

$$\sum_{n\in\mathbb{Z}} e^{-a(x-n)^2} = C(1 + 2e^{-1/a\pi^2}\cos(2\pi x) + 2e^{-4/a\pi^2}\cos(4\pi x) + \ldots);$$

that is, the LHS is simply the convolution of a Gaussian with a comb of delta functions, so its Fourier transform over the real line is a dual Gaussian multiplied by the same comb of delta functions, which we can express as a sum of cosines with weights given by the dual Gaussian evaluated at integer locations. In this one dimensional case, given $\varepsilon < 1$, simply take a small enough that $3e^{-1/a\pi^2}\cos(2\pi x) < \varepsilon$ to force almost constant. (Note: two times the exponential is enough to dominate the first non-constant term in the expansion, while the three is more than enough to dominate the first term and all the rest.) In the d-dimensional case on lattice $A\mathbb{Z}^d$, diagonalize matrix $A^T A$, and reduce to a product of one dimensional cases. The correlation function is simply a product of scaled versions of the above cosine expansion.

Proposition 9. Suppose window functions γ, g are Gaussians on $\mathbb{R}^d$; that is, there are constants $a, a' > 0$ with

$$\gamma(x) = e^{-ax\cdot x}$$
$$g(x) = e^{-a'x\cdot x}$$

for all $x \in \mathbb{R}^d$. Then the correlation function $c(j, j')$ for the windows goes to zero rapidly as $j - j'$ gets large. That is, there exists constant C' with

$$c(j, j') \le C' e^{-\frac{aa'}{a+a'}|j-j'|^2}, \text{ for all } j, j' \in \mathbb{Z}^d.$$

Proof. With j, j' fixed, let $j'' = (aj + a'j')/(a + a')$ denote the weighted average of j, j' and an easy computation shows

$$a(j-x)\cdot(j-x) + a'(j'-x)\cdot(j'-x) = (a+a')(j''-x)\cdot(j''-x) + \frac{aa'}{a+a'}(j-j')\cdot(j-j').$$

Thus

$$\gamma(j - An)\overline{g}(j' - An) = e^{-(a+a')(j''-An)\cdot(j''-An)} e^{-\frac{aa'}{a+a'}(j-j')\cdot(j-j')}$$

and summing over all n gives the estimate

$$c(j, j') \le C' e^{-\frac{aa'}{a+a'}|j-j'|^2}$$

where $C' = \sum_{n\in\mathbb{Z}^d} e^{-(a+a')(An)\cdot(An)}$.

Note, we can estimate C' above by replacing the sum with an integral, and find that

$$C' \approx \left(\frac{\pi}{a+a'}\right)^{d/2} \frac{1}{\det(A)}.$$

We may now choose a, a' small enough that the correlation function is nearly constant on the diagonal, then choose integer matrix B "large enough" so that the term $Ce^{-\frac{aa'}{a+a'}|j-j'|^2}$ is sufficiently small on the non-zero points in the sublattice $B\mathbb{Z}^d$.

Proposition 10. Suppose the functions γ and g are Gaussians on $\mathbb{R}^d$; that is, there are constants $a, a' > 0$ with

$$\gamma(x) = e^{-ax\cdot x}$$
$$g(x) = e^{-a'x\cdot x}$$

for all $x \in \mathbb{R}^d$. Then for a, a' sufficiently small, and integer matrix B sufficiently large, the reconstruction satisfies

$$\tilde{f} \approx |\det B| C \cdot f,$$

for some constant C. That is, we have an approximate reconstruction.

Proof. With a, a' sufficiently small, the diagonal terms in the expansion

$$\tilde{(}f) = |\det B| \sum_{j' \sim_{B^*} j} c(j, j') f(j')$$

are nearly constant, and introduce an error no larger than $(|\det B| C\varepsilon)$, times the size of f, as indicated in Proposition 6. The off-diagonal terms are bounded by $C' e^{-\frac{aa'}{a+a'}|j-j'|^2}$, with $j - j' \in B^* \mathbb{Z}^d$, so with B large enough, these terms are small.

SUMMARY

We have shown that by using a partition of unity on the integer lattice $\mathbb{Z}^d$, we may define a generalized Gabor transform that provides a non-uniform time/frequency decomposition of a signal, which comes with an associated inverse transform. The Gabor coefficients depend only locally on the signal, as determined by the size and location of the support of the functions in the partition, while the size of the frequency lattice must be chosen complimentarily to the size of the support. By factoring the functions in the partition, we obtain directly the Gabor synthesis and analysis windows; with a good choice of functions in the partition (eg. smooth, non-negative, compact support), we obtain Gabor windows with similar characteristics, which are useful for practical signal processing tasks.

We have used the partition of unity technique to extend the standard discrete Gabor transform to more general lattices in Euclidean space, and to the approximate case with Gaussian windows. The inverse transform is shown to be fast, and thus practical, and we have examined a number of other practical issues that arise in a real implementation of this generalized Gabor transform, in particular the reduction of artifacts.

ACKNOWLEDGMENTS

We thank NSERC, MITACS, the sponsors of POTSI (Pseudodifferential Operator Theory in Seismic Imaging - a joint project between the departments of Geophysics and Mathematics at University of Calgary) and the CREWES Project for their generous financial support. This work evolved from a seminar on Gabor Analysis, led by POTSI, attended by the authors over the past two years at University of Calgary.

REFERENCES

[1] Daubechies, I., Grossmann, A., and Meyer, Y. (1986). Painless nonorthogonal expansions: *J. Math. Phys.,* **27** No. 5, 1271–1283.

[2] Feichtinger, H. G. and Nowak, K., (2003) A first survey of Gabor multipliers, *Advances in Gabor analysis, 99–128, Appl. Numer. Hamon. Anal.* Birkhasuer Boston.

[3] Feichtinger, H. G. and Strohmer, T. (1998). Gabor analysis and algorithms, theory and applications, *Birkhauser.*

[4] Feichtinger, H. G. (2003). MATLAB code, in personal communication.

[5] Gabor, D. (1946). Theory of communication, *J. IEEE,* **93**, 429–457.

[6] Gibson, P. and Lamoureux, M., (2003). Maximally symmetric, minimally redundant partitions of unity in the plane, *preprint*.

[7] Grochenig, K. (2001). Foundations of time-frequency analysis, *Birkhauser.*

[8] Grossman, J. P., Margrave, G. F., Lamoureux, M. P., and Aggarwala, R., (2001). Constant-Q wavelet estimation via a nonstationary Gabor spectral model, *CREWES Research Report 2001, Vol. 13*.

[9] Grossman, J. P., Margrave, G. F., and Lamoureux, M. P., (2002). Fast wavefield extrapolation by phase-shift in the nonuniform Gabor domain, *CREWES Research Report 2002, Vol. 14*.

[10] Grossman, J. P., Margrave, G. F., and Lamoureux, M. P., (2002). Constructing adaptive nonuniform Gabor frames from partitions of unity, *CREWES Research Report 2002, Vol. 14*.

[11] Hoffman, H. and Kunze, R. (1971). Linear Algebra, *Prentice-Hall.*

[12] Margrave, G. F. and Lamoureux, M. P., (2001). Gabor deconvolution, *CREWES Research Report 2001, Vol. 13*.

[13] Margrave, G. F., Henley D. C., Lamoureux, M. P., Iliescu, V., and Grossman, J. P., (2002). An update on Gabor deconvolution, *CREWES Research Report 2002, Vol. 14*.

[14] Strang, G. and Nguyen, T., (1997). Wavelets and Filter Banks, *Wellesley-Cambridge Press*.

A. APPENDIX: EXAMPLES OF PARTITIONS OF UNITY

As discussed in the main section of this paper, a partition of unity is a collection of functions on $\mathbb{Z}^d$ which sum to the constant function one; that is, a collection of functions w_n with

$$\sum_n w_n(z) = 1, \qquad \text{for all } z \in \mathbb{Z}^d.$$

It is worth pointing out that there is a general construction to obtain a partition of unity with useful characteristics: given any family of functions ϕ_n (with n running over some index set, possibly infinite), we set

$$w_n(z) = \frac{\phi_n(z)}{\sum_{n'} \phi_{n'}(z)}.$$

Provided the denominator is never zero, nor infinite, the resulting functions w_n form a partition of unity. If the ϕ_n are chosen non-negative, then so are the w_n; if the ϕ_n each have compact support, then so do the w_n; and so on. That is, by choosing the ϕ_n with suitable characteristics, often the partition of unity has simliar characteristics, which can be important for typical signal processing applications. It is sometimes convenient to begin with one basic function ϕ_0 defined on Euclidean space $\mathbb{R}^d$ and define a family of functions by translation:

$$\phi_n(z) = \phi_0(z - An),$$

where A is a fixed $d \times d$ matrix and $n \in \mathbb{Z}^d$ the index vector; with the support of ϕ_0 compatible with the norm of matrix A, a partition of unity is obtained by the normalization method mentioned above.

There are of course many examples of partitions of unity built via simple constructions. In dimension one, translates of the boxcar function of length a give the simplest partition, with

$$w_n(z) = \begin{cases} 1 & \text{for } (n - \tfrac{1}{2})a \le z < (n + \tfrac{1}{2})a \\ 0 & \text{otherwise.} \end{cases}$$

More generally, any doubly infinite, increasing sequence $\dots k_{-1}, k_0, k_1, \dots$ of integers can be used to define a partition of unity with boxcar functions

$$w_n(z) = \begin{cases} 1 & \text{for } k_n \le z < k_{n+1} \\ 0 & \text{otherwise.} \end{cases}$$

A smoother choice of partition is obtained via translates of the raised cosine window, yielding functions

$$w_n(z) = \begin{cases} \frac{1}{2}(1+\cos(\pi z/a)) & \text{for } (n-1)a < z < (n+1)a \\ 0 & \text{otherwise,} \end{cases}$$

which also forms partition of unity on the set $\mathbb{Z}$; that is, translates of the cosine hump sum uniformly to one. (In fact, they sum to one even when extended to all real points $\mathbb{R}$.) Each member in this family of functions has compact support, and is the restriction (to the integers) of a continuous function with continuous derivative, which has some advantages as a window function for signal processing.

More generally, one may design a partition of unity using translates of a spline of class $C^k(\mathbb{R})$; that is, a smooth function on $\mathbb{R}$ supported on some compact interval $[-a,a]$, constant one on the subinterval $[-b,b]$ and chosen so the overlapping edges of translates sum to one. For instance, begin with the unique odd polynomial $p(x)$ of degree $2k+1$ that takes value $p(1)=1$, with derivatives $p^{(k')}(1)=0$ for $k'=1,2,...k$. With parameters $a>b>0$ fixed, define the basic window function by

$$w_0(z) = \begin{cases} \frac{1}{2}(1+p(\frac{b+a+2z}{a-b})) & \text{if } -a \le z \le -b \\ 1 & \text{if } -b \le x \le b \\ \frac{1}{2}(1+p(\frac{b+a-2z}{a-b})) & \text{if } b \le z \le a. \end{cases}$$

Translates of this window by multiples of $a+b$ forms a partition of unity for $\mathbb{Z}$, and indeed for $\mathbb{R}$ as well. (Our research team has named these “Lamoureux windows” although they are properly called splines.) A C^∞ window may be formed by replacing the polynomial $p(x)$ in the formulas above with a rescaled version of the hyperbolic tangent function, $p(x)=\tanh(\frac{x}{x^2-1})$. Figure 11 gives an example of a degree five Lamoureux window, with continuous second derivatives.

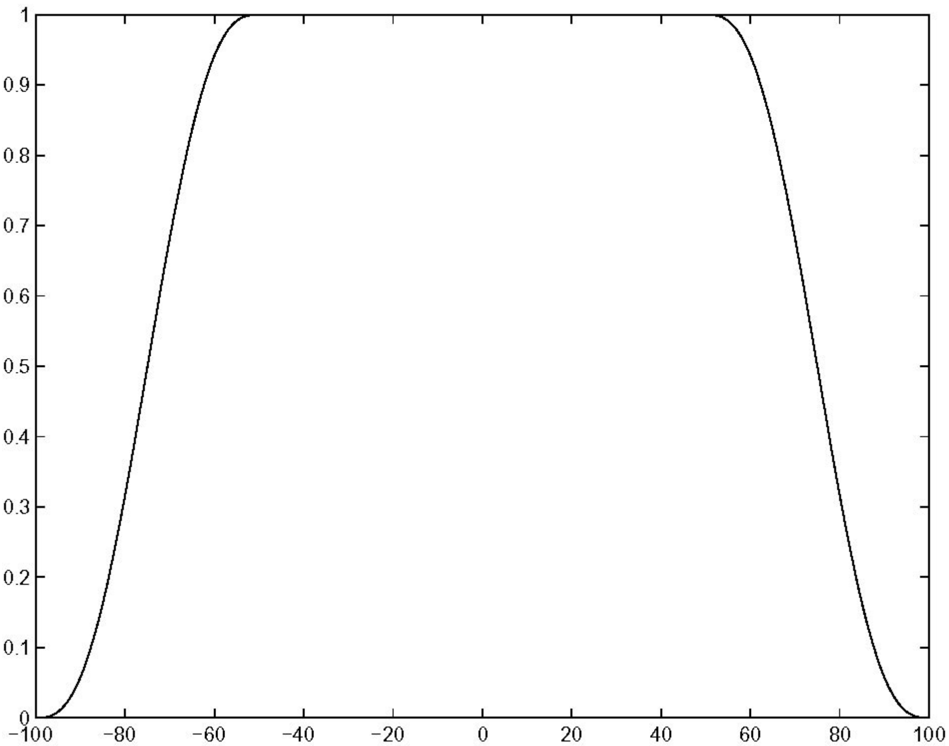

FIG. 11. A C^2 Lamoureux window, with $p(x)=(15x-10x^3+3x^5)/8$, $a,b=100,50$.

In dimension d, a direct product of one-dimensional partitions over $a_1\mathbb{Z},...,a_d\mathbb{Z}$, such as

$$w_d(x)=w(x_1)w(x_2)...w(x_d),$$

forms a discrete partition of unity over the d-dimensional lattice $a_1\mathbb{Z}\times...\times a_d\mathbb{Z}$. Other partitions, not of this form, may be defined: for instance, the hexagonal function shown in Figure 12 is given by symmetrizing afunction ϕ of compact support in $\mathbb{R}$ along a hexagonal lattice $\mathcal{L}_6$ to obtain a basic window function w_0 as

$$w_0(z)=\frac{\varphi(z\cdot z)}{\sum_{l\in L_6}\varphi((z-l)\cdot(z-l))}, \qquad \text{for all } z\in\mathbb{Z}^2.$$

Translates of w_0 over the lattice $\mathcal{L}_6$ gives a partition of unity.

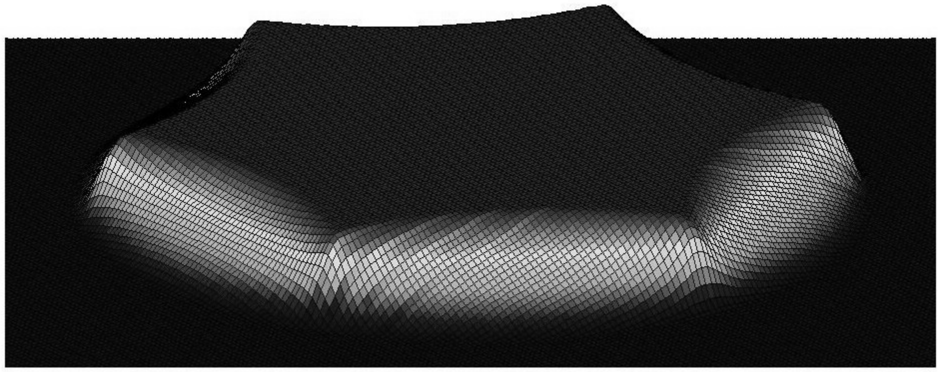

FIG. 12. A partition of unity with hexagonal symmetry.

Each of these examples is a partition of unity with compact support.

B. APPENDIX: PROOF OF PROPOSITION 1

In Section 4, we introduced the notion of frequency sampling on a non-rectangular lattice. Here we prove the key results leading to Proposition 1. Recall, B is an invertible $d \times d$ matrix with integer entries, for any index $m \in \mathbb{Z}^d$, we define modulation function P_m by

$$P_m(j) = e^{2\pi i (B^{-1}m)\cdot j}, \qquad \text{for all } j \in \mathbb{Z}^d,$$

and two equivalence relations are defined on the lattice $\mathbb{Z}^d$ by

$$m \sim m' \qquad \text{iff} \qquad m - m' \in B\mathbb{Z}^d,$$

and

$$m \sim^* m' \qquad \text{iff} \qquad m - m' \in B^*\mathbb{Z}^d,$$

where B^* is the transpose of matrix B.

As is well-known from group theory, the equivalence $\sim$ divides up the group $\mathbb{Z}^d$ into equivalence classes, and these classes form elements of the so-called quotient group, denoted $\mathbb{Z}^d/B\mathbb{Z}^d$. Since B is invertible, the quotient is a finite abelian group, and thus a direct product of cyclic groups. In fact, one may determine this quotient group precisely as in the following:

Proposition 11. Suppose $B \in M_d(\mathbb{Z})$ is an invertible, $d \times d$ matrix with integer entries. Then B factors uniquely as $B = UZ$, where U, Z are integer matrices, U is upper triangular with positive diagonal, and Z has determinant ± 1. With

$$U = \begin{bmatrix} M_1 & * & \dots & * \\ 0 & M_2 & \dots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & M_d \end{bmatrix},$$

where $M_1, M_2, \dots M_d$ are positive integers, then the cube

$$M = [0,1,\dots M_1 - 1] \times [0,1,\dots M_2 - 1] \times \dots [0,1,\dots M_d - 1] \subset \mathbb{Z}^d$$

maps bijectively onto the quotient group $\mathbb{Z}^d / B\mathbb{Z}^d$.

Proof. Column reduction of matrix B over the ring of integers gives the factorization above; forcing the diagonal to be positive gives uniqueness. The cardinality of the quotient group $\mathbb{Z}^d / B\mathbb{Z}^d$ is given by a volume form, which is just $|\det(B)| = M_1 \cdot M_2 \cdot \dots M_d$, and agrees with the cardinality of the cube M. Thus, we need only verify the quotient map $\mathbb{Z}^d \to \mathbb{Z}^d / B\mathbb{Z}^d$ is injective on M. Since this map is a group homomorphism, it suffices to show the kernel of the quotient map, intersect M, is just the zero element. So, if $m \in M$ is equivalent to $0 \mod B\mathbb{Z}^d$, then $m = Bz$ for some element $z \in \mathbb{Z}^d$. By the factorization,

$$m = \begin{bmatrix} M_1 & * & \dots & * \\ 0 & M_2 & \dots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & M_d \end{bmatrix} Zz$$

for some integer vector z. Starting at the last component in this equation, one sees that $m_d = M_d z_d$, for some integer z_d. Since $0 \le m_d < M_d$, one concludes that $m_d = z_d = 0$. For the next-to-last component, then $m_{d-1} = M_{d-1} z_{d-1} + (*) z_d$; but since z_d was zero, this simplifies to $m_{d-1} = M_{d-1} z_{d-1}$. Once again, since $0 \le m_{d-1} < M_{d-1}$, one obtains $m_{d-1} = z_{d-1} = 0$. Working back in this manner, one concludes all the components of m are zero, so $m = 0$, as desired.

The cube $M = [0,\dots,M_1 - 1] \times [0,\dots,M_2 - 1] \times \dots \times [0,\dots,M_d - 1]$ above is precisely the cube that appears in Proposition 1. With this factorization, parts 1, 2, and 3 of Proposition 1 become clear (the periodicity comes from observing $j \sim^* j'$ introduces a difference of an integer offset in the complex exponential $e^{2\pi i}$, which exponentiate to no difference; similar for $m \sim m'$). Part 4 is proved in the following.

Proposition 12. Define function P as

$$P(j)=\sum_{m\in\mathbb{Z}^d/B\mathbb{Z}^d} e^{2\pi i(B^{-1}m)\cdot j}, \qquad \text{for vectors } j\in\mathbb{Z}^d,$$

which the sum of modulation functions $P_m(j)$. Then

$$P(j)=\begin{cases}|\det(B)| & \text{if } j\in B^*\mathbb{Z}^d\\ 0 & \text{otherwise,}\end{cases}$$

where $B^*\mathbb{Z}^d$ is the discrete sublattice generated by the integer matrix B^* obtained by transposing matrix B.

Proof. As noted above, the sum over the quotient is well-defined, since the modulation functions $P_m, P_{m'}$ are the same for vectors m, m' equivalent mod $B\mathbb{Z}^d$. Also, observe that the function $P(j)$ is constant on equivalence classes of j mod $B^*\mathbb{Z}^d$, for if $j\sim^* j'$, then $j'=j+B^*z$ for some integer vector z and thus the corresponding exponentials agree, for

$$e^{2\pi i(B^{-1}m)\cdot j'}=e^{2\pi i(B^{-1}m)\cdot(j+B^*z)}=e^{2\pi i(B^{-1}m)\cdot j}e^{(BB^{-1}m)\cdot(z)}=e^{2\pi i(B^{-1}m)\cdot j};$$

Hence $P(j)=P(j')$.

Now clearly, $P(0)=\sum 1$, where the sum of these ones is just the cardinality of the quotient group, which is exactly $|\det(B)|$; by periodicity, $P(j)=|\det(B)|$ for all $j\in B^*\mathbb{Z}^d$.

To show the other j evaluate to zero, we will show that $P(j)\neq 0$ implies $j\in B^*\mathbb{Z}^d$. Fix j with $P(j)\neq 0$, and observe for all vectors m, by the factorization of B as $B=UZ$, we have

$$(B^{-1}m)\cdot j=(Z^{-1}U^{-1}m)\cdot j=(U^{-1}m)\cdot(Z^*)^{-1}j=(U^{-1}m)\cdot z,$$

where we set $z=(Z^*)^{-1}j$, an integer vector. Now, sum exponentials over vectors m in the cube M defined in the previous proposition, using the factorization for B, and expanding each m as a linear combination of standard basis vectors $e_1,\ldots e_d$, we observe

$$P(j) = \sum_{m \in M} e^{2\pi i (B^{-1}m)\cdot j}$$

$$\begin{aligned} P(j) &= \sum_{m \in M} e^{2\pi i (U^{-1}m)\cdot z} \\ &= \sum_{m \in M} \prod_{k=1}^{d} e^{2\pi i ((U^{-1}e_k)\cdot z) m_k} \\ &= \prod_{k=1}^{d} \left(\sum_{0 \le m_k < M_k} e^{2\pi i ((U^{-1}e_k)\cdot z) m_k} \right) \end{aligned}$$

Now, from the upper triangular form for U, we know U^{-1} is also upper triangular, with diagonal entries $1/M_1, 1/M_2, ..., 1/M_d$. Looking at the $k = 1$ factor in the above product, we see $(U^{-1}e_1) \cdot z = z_1/M_1$, the first component of vector z, divided by the integer M_1 on the diagonal of U. Since the factor

$$\sum_{0 \le m_1 < M_1} e^{2\pi i ((U^{-1}e_1)\cdot z) m_1} = \sum_{0 \le m_1 < M_1} e^{2\pi i (\frac{z_1}{M_1}) m_1}$$

is non-zero, the M_1-root of unity $e^{2\pi i (\frac{z_1}{M_1})}$ must the the trivial root (i.e. one), so z_1 is a multiple of M_1. Thus

$$z = \begin{bmatrix} aM_1 \\ z_2 \\ z_3 \\ \vdots \\ z_d \end{bmatrix}.$$

Now, define $j' = j - B^* \begin{bmatrix} a \\ 0 \\ \vdots \\ 0 \end{bmatrix}$, so j' is equivalent to j mod $B^*\mathbb{Z}^d$, and hence $P(j')$ is non-zero as well. The corresponding vector z' is thus

$$z' = (Z^*)^{-1} j' = z - U^* \begin{bmatrix} a \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ z'_2 \\ z'_3 \\ \vdots \\ z'_d \end{bmatrix}.$$

That is, this new vector z' has a zero in the first component. As in the argument above, we now examine the factor for $k = 2$, with

$$\sum_{0 \le m_2 < M_2} e^{2\pi i((U^{-1}e_2)\cdot z')m_2} = \sum_{0 \le m_2 < M_2} e^{2\pi i(\frac{z'_2}{M_2})m_2}$$

and since it is non-zero, we now conclude z'_2 is a multiple of M_2, and thus

$$z' = \begin{bmatrix} 0 \\ bM_2 \\ z'_3 \\ \vdots \\ z'_d \end{bmatrix}.$$

Continuing for each factor in the product, we eventually construct vector $z^{(d)}$, all of whose components are zero, and $j^{(d)} = (Z^*)z^{(d)} = 0$ which is equivalent to the original j (mod $B^*\mathbb{Z}^d$), so we have that $j \in B^*\mathbb{Z}^d$, as desired.