1 Course Syllabus for Wroclaw, HGFei: May 2016

file: D:\conf\Conf16\Wroclaw16\WroclawCoursSum.pdf 24.04.2016, rev. 2023

2 Linear Algebra Recalled I

We have been discussing $m \times n$-matrices $\boldsymbol{A}$, which according to the MATLAB viewpoint consist of a collection of n column-vectors $\mathbf{a}_k, 1 \leq k \leq n$ in $\mathbb{C}^m$. The command `plot(A)` uses exactly this interpretation!

We recall that the mapping $T : \mathbf{x} \to \boldsymbol{A}{*}\mathbf{x}$ (matrix-vector-multiplication) corresponds to the formation of *linear combinations*, in fact $\boldsymbol{A}{*}\mathbf{x}$ gives exactly the linear combination $\sum_{k=1}^{n} x_k \mathbf{a}_k$ of the columns of the matrix $\boldsymbol{A}$ with the coefficients $x_1, \cdots, x_n \in \mathbb{C}$. Any linear mapping T is of this form (just put $\mathbf{a}_k = T(\mathbf{e}_k), 1 \leq k \leq n$). Hence the range of T is just the *column space* of $\boldsymbol{A}$, i.e. the linear space spanned by the column vectors, which is exactly the set of all possible linear combinations of the column vectors $\mathbf{a}_1, \cdots, \mathbf{a}_n$, because this *is a linear space* and obviously the smallest linear subspace of $\mathbb{C}^m$ containing all these column vectors! Similarly the range of $\boldsymbol{A}'$ (the transpose conjugate matrix, which is now an $n \times m$-matrix) is the *row space*, a subspace of $\mathbb{C}^n$ (it should be really better seen as column-space (!) of $\boldsymbol{A}'$!). Hence T is surjective if and only the sequence $(\mathbf{a}_k)_{k=1}^n$ is a *set of generators* of $\mathbb{C}^m$ (i.e. they span all of $\mathbb{C}^m$), while T is injective if and only if the columns of $\boldsymbol{A}$ are *linear independent*! The columns thus form a basis if both properties are valid, i.e. if the columns of the matrix form a linear independent generating system. Finally note that *linear algebra* tells us that this is possible if and only if $m = n$. For an argument one may invoke the good old Gauss elimination procedure.

Using the *pivot elements* in the Gauss Elimination Process, i.e. the leading ones in the row reduced echelon form of $\boldsymbol{A}$ we can find out that these two space have the same dimension, i.e. the number of pivot elements in the final stage of the row-reduced echelon form (in fact, the rows of `rref(a)` form a basis for the row space, while the columns of !original matrix $\boldsymbol{A}$ which correspond to those columns containing a leading ones in `rref(a)` are a basis for the column space. Due to the echelon form we see that there is thus at least one basis for each of the spaces with $r = rank(\boldsymbol{A}) =$ *number of pivot elements of* $\boldsymbol{A}$ elements.

Finally we have discussed that it is not difficult to show that the null-space of T, or equivalently *the subspace of all the vectors which are solutions to the homogeneous linear system* $\boldsymbol{A} * \mathbf{x} = \mathbf{0}$ is equal to the orthogonal complement of the row-space of $\boldsymbol{A}$ (meaning column-space of $\boldsymbol{A}'$).

The mapping T restricted to the column-space of $\boldsymbol{A}'$ is thus injective, but on the other hand has still the same range, so T, restricted to the row-space of $\boldsymbol{A}$ defines actually an isomorphism between the row-space of a and the column-space of $\boldsymbol{A}$.

The idea of the *pseudo-inverse* is now to invert the mapping at this level, while the rest (the two orthogonal complements) are ignored!

The realization of $PA = pinv(A)$ is realized with the help of the SVD (*singular value decomposition*). We only need so much: $\boldsymbol{A} = \mathbf{U} * \Sigma * \mathbf{V}'$, where $\mathbf{U}$ an $\mathbf{V}$ (hence also $\mathbf{V}'$) are two unitary matrices of dimension $m \times m$ and $n \times n$ respectively, and Σ is a *rectangular diagonal matrix* with exactly r *strictly positive* diagonal entries, the so-called *singular values*, sorted in decreasing order. The biggest singular value is in fact the *operator norm* of the matrix, i.e. of the linear mapping $\mathbf{x} \mapsto \boldsymbol{A} * \mathbf{x}$, or in other words $s_1 = \max_{\|\mathbf{x}\| \leq 1} \|\boldsymbol{A} * \mathbf{x}\|$. Note that in the case of an invertible 3×3 matrix $\boldsymbol{A}$ one has: $s_1 \geq s_2 \geq s_1 > 0$,and the image of the unit-sphere consisting of all $\mathbf{x} \in \mathbb{R}^3$ with $\|\mathbf{x}\| = 1$ is an ellipsoid, whose main axes have the directions $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ and the corresponding length are exactly the singular values. Thus, for example s_1/s_3 could be viewed as the eccentricity of this ellipsoid, but also

`pinv(A)` is then obtained by changing the order and inverting those parts of Σ which can be inverted. Of course, in practice, one often inverts only those singular values $\sigma_j, 1 \leq j \leq r$ which are not too small, while the very small ones are considered to be too close to zero and are thus ignored in the pseudo-inverse.

There is also a connection of `pinv(A)` with the problem of finding the MNLSQ (the minimal norm least squares solution to $\boldsymbol{A} * \mathbf{x} = \mathbf{b}$. It is usually realized with the help of the so-called normal equations, i.e. by looking at the (necessary) solution of $(\boldsymbol{A}' * \boldsymbol{A}) * \mathbf{x} = \boldsymbol{A}' * \mathbf{b}$, which in the case of invertibility of $\boldsymbol{A}' * \boldsymbol{A}$, the so-called *Gram-Matrix*, and this is equivalent linear independence of the columns of $\boldsymbol{A}$, gives: [1] Thus the solution of the MNLSQ-problem is in this case just

$$\tilde{\mathbf{x}} = [(\boldsymbol{A}' * \boldsymbol{A})^{-1} * \boldsymbol{A}'] * \mathbf{b}. \tag{1}$$

This is in fact a useful way to compute $\boldsymbol{A}^+ = \mathrm{pinv}(\boldsymbol{A})$ for such matrices, namely as

$$\boldsymbol{A}^+ = [(\boldsymbol{A}' * \boldsymbol{A})^{-1} * \boldsymbol{A}']. \tag{2}$$

A very analogous question applies for the case that $\boldsymbol{A}'$ has linear independent columns, but since $\mathrm{null}(\boldsymbol{A}') = \mathrm{Col}(\boldsymbol{A})^\perp$ this the same as the assumption that the columns of $\boldsymbol{A}$ span all of $\mathbb{C}^m$, i.e. are a generating system for their range:

$$\boldsymbol{A}^+ = [\boldsymbol{A}' * (\boldsymbol{A} * \boldsymbol{A}')^{-1}]. \tag{3}$$

The unified version of these to statements is the following claim:
For any $m \times n$-matrix $\boldsymbol{A}$ of maximal rank $\mathrm{rank}(\boldsymbol{A}) = min(m,n)$ one has:

$$\boldsymbol{A}^+ = [\boldsymbol{A}' * \mathrm{pinv}(\boldsymbol{A} * \boldsymbol{A}')] = [\mathrm{pinv}(\boldsymbol{A}' * \boldsymbol{A}) * \boldsymbol{A}']. \tag{4}$$

where the "smaller of the two square matrices ($\boldsymbol{A}' * \boldsymbol{A}$ and $\boldsymbol{A} * \boldsymbol{A}'$ resp.) will be invertible and should be used.

Once more, in case of linear independence one has few column vectors in a larger space, test for linear independence is realized via the Gram matrix $\boldsymbol{A}' * \boldsymbol{A}$ (format $n \times n$!), while in the case of generating systems one looks at the so-called frame operator $\mathbf{S} := \boldsymbol{A} * \boldsymbol{A}'$, given (by its action on vectors) via

$$\mathbf{S}(\vec{x}) = \sum_{k=1}^{n} \langle \vec{x}, \vec{a}_k \rangle \vec{a}_k, \quad \vec{x} \in \mathbb{C}^n. \tag{5}$$

One can interpret this definition as follows: We *hope* that a generating system might be an orthonormal basis, and in this case we *would have* (for well known reasons) $\mathbf{S} = \mathrm{Id}_n$.

[1] Verify this claim by observing that $\mathrm{Null}(\boldsymbol{A}) = \mathrm{Null}(\boldsymbol{A}' * \boldsymbol{A})$, for any $m \times n$-matrix $\boldsymbol{A}$.

Of course it is almost the same good luck if we have $\mathbf{S} = \lambda \, \mathrm{Id}_n$, because then it is obvious that $\mathrm{Id}_n = \lambda^{-1}\mathbf{S}$ (which is the case for the FFT matrix $F = \mathrm{fft}(eye(N))$, with $\lambda = \sqrt{N}$). Such systems (of column vectors) in $\mathbb{C}^m$ are called *tight frames* for $\mathbb{C}^m$, and (as the example of the Fourier basis shows) equal length orthogonal systems are special cases (the non-redundant tight frames, i.e. the ones with exactly m elements).

From this geometric interpretation of the situation one can derive a couple of important things(writing $\boldsymbol{A}^+$ for $\mathrm{pinv}(\boldsymbol{A})$ such as:

$$\boldsymbol{A} * \boldsymbol{A}^+ = \mathbf{P}_{col}(\boldsymbol{A}), \quad \boldsymbol{A}^+ * \boldsymbol{A} = \mathbf{P}_{row}(\boldsymbol{A}), \quad \boldsymbol{A} * \boldsymbol{A}^+ * \boldsymbol{A} = \boldsymbol{A}, \quad \boldsymbol{A}^+ * \boldsymbol{A} * \boldsymbol{A}^+ = \boldsymbol{A}^+. \tag{6}$$

Finally we mention that we will also be able to derive tight frames, or for linear independent systems a symmetric form of orthonormalization, the so called Loewdin orthonormalization, using in the pinv-characterization not the inverse, but the inverse square root (which is the same as the square root of the inverse of a positive operator), namely $\mathbf{S}^{-1/2}$ in order to get a *tight frame*.

So the Loewdin orthogonalization creates from a given collection of n linear independent vectors an orthonormal system by forming suitable linear combinations of the given vectors (so that the original and the new orthonormal set span the same linear space) by using the coefficients showing up in the square root inverse (equal the inverse of the square root) of the (positive definite) Gram matrix:

$$\mathbf{W} = \boldsymbol{A} * (\boldsymbol{A}' * \boldsymbol{A})^{-1/2}. \tag{7}$$

It produces the orthonormal basis $\mathbf{W}$ for the column space of $\boldsymbol{A}$ which is closest to the given matrix $\boldsymbol{A}$ in the Frobenius sense, so in the sense of the Euclidean structure of $\mathbb{C}^{m \cdot n}$. It can be also described (alternatively) as $\mathbf{W} = \mathbf{U} * \mathbf{V}'$, where $\boldsymbol{A} = \mathbf{U} * \Sigma * \mathbf{V}'$ (via the interpretation as polar decomposition of operators, removing the dilation part Σ).

2.1 MATLAB facts and conventions for Gabor Analysis

We consider finite vectors as functions on the (finite Abelian) group $G = \mathbb{Z}_N$, which is $\mathbb{Z}$ *modulo* $n\mathbb{Z}$, or just the (multiplicative) group of unit roots of order N. In fact, there is the primitive unit-root $\omega_N = exp(2\pi i/N)$, and all the other elements are of the form $\omega_N^k, 0 \leq k \leq N-1$ (i.e. starting with $\omega^0 = 1$ and going clockwise through this discrete unit circle.

(NuHAG MATLAB command): `plotnum(uroots(12)); axis off; shg`
gives us the following picture for $N = 12$:

Viewing finite vectors in $\mathbb{C}^N$ as functions on the group $\mathbb{Z}_N$ is giving us the correct group-theoretical viewpoint. It is easy to characterize the *dual group* $\widehat{\mathbb{Z}_N}$, i.e. the set of all homomorphism from this finite Abelian group to $\mathbf{U} := \{z \mid \in \mathbb{C}, |z| = 1\}$ as the *pure frequencies* resp. the eigenvectors for the *cyclic shift operator*. Since $\mathbb{Z}_N$ is a *cyclic group* it suffices to check that such an homomorphism is essentially determined by its value at ω, and that there are only N different choices (because applying the cyclic shift operator N times we get the identity!), namely all the unit roots of order N again. If we look at this we see that this are exactly the entries in the second row of the FFT-matrix $F = fft(eye(N))$, which looks in a plot like this:

See also the presentation (in German)

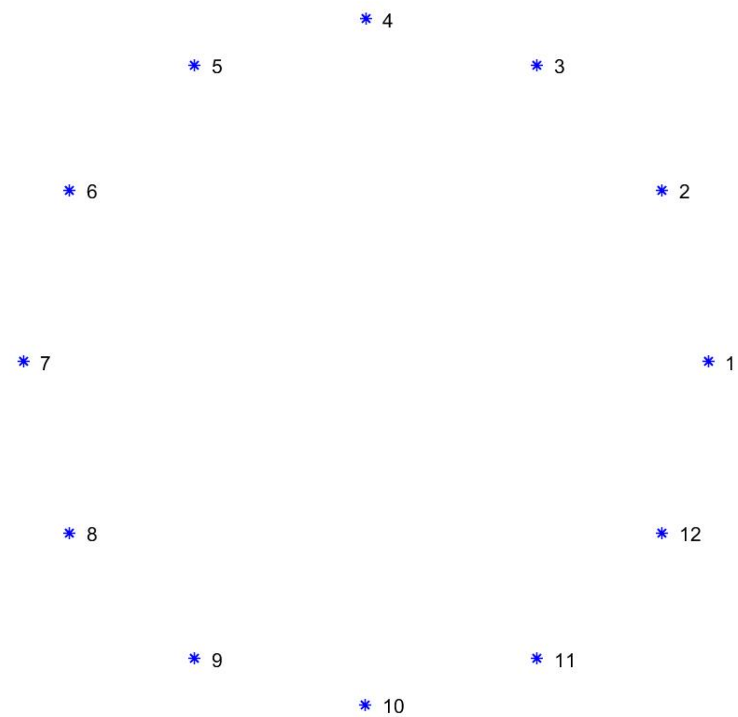

Abbildung 1:

`http://www.univie.ac.at/nuhag-php/program/talks_details.php?id=3113`

So we have N different frequency shifts, each of them being some power of the elementary frequency shift, which is just the cyclic rotation on the Fourier transform side, and each of them is just the pointwise multiplication of a column vector with the corresponding column of the DFT/FFT matrix $\mathcal{F}$. We call them M_l, for *modulation operators*, with $l = 0, \ldots, N-1$, because we think in terms of frequencies (the first column corresponds to frequency zero!).

Accordingly we have altogether N^2 time-frequency shifts on our signal space $\mathbb{C}^N = \boldsymbol{\ell}^2(\mathbb{Z}_N)$, namely

$$\pi(k,l) = M_l T_k, \quad k, l \leq N-1. \tag{8}$$

Since we have in this way N^2 particular matrices within the N^2-dimensional space of $n \times n$-matrices one may rise the question: Are they a basis for $\mathcal{M}_{N,N}(\mathbb{C})$?

One nice formulation about *TF-analysis (time-frequency) analysis* is to say:

"it is the part of analysis that one can do concentrating on this family of operators".

We will view the parameter space, which is just $\mathbb{Z}_N \times \mathbb{Z}_N$ in the following way: Since we story information about functions on the "TF-plane" in the form of a matrix, we have to set some conventions: The "first position" (i.e. the term within $\boldsymbol{A}$ usually

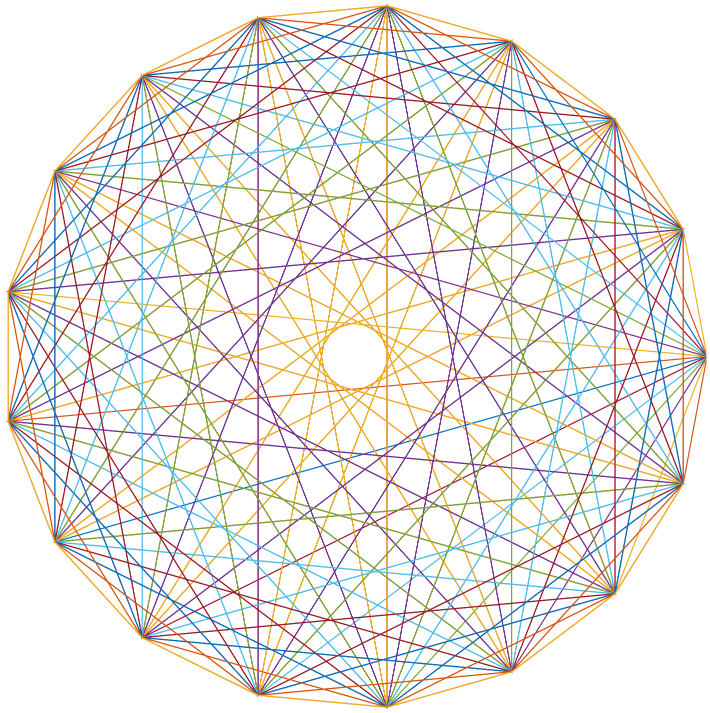

Abbildung 2:

denoted by $a_{1,1}$ corresponds to the case $l = 0 = k$, so no shift und multiplication with constant 1, i.e. this is the identity operator. The position $(5, 3)$ means that we move down within the matrix by 4 units, so we should think of M_4, while the horizontal shift is just 2 to the right (compared to the first column), so T_2, so anything related to the pair $(5, 3)$ should have to do in this interpretation with $\pi(4, 2) = M_4 \circ T_2$, because we speak about *time-frequency* shifts, meaning that we are applying the time-shift operator first, and the modulation (or frequency shift) operator afterwards. This is in contrast to the usual description of matrix entries, which gives first the row number k and then the column number j if we look at $a_{k,j}$. Overall there is again a match (up to the index shift by one): The operator M_4T_2 is associated with the matrix position $(5, 3)$ (which is $(1, 1) + (4, 2)$!). Note that on the other hand it is plausible to assign the time-shift operator to the horizontal shift (left and right), while the frequency direction (up and down) is related to frequency shifts (like in a musical score, where higher notes are in the sense of higher frequency are depicted higher up on the score-sheet!)

The most important object defined over the TF-plane is the STFT (the short-time Fourier transform, or Sliding-Window Fourier Transform). We think of the "signal" (piece of musical recording, from an audio-file) is called x while the window g is assumed to be quite smooth, like e.g. a plateau-functions. Following Gabor ([10]) one should use a Gauss-function $g_0(t) = exp(-\pi|t|^2)$ for $G = \mathbb{R}$, so we often use a discrete variant of this function (routine-name `gaussnk.m`), which is (up to the unavoidable scaling factor of $\sqrt{N}$ effectively FFT-invariant. In any case it makes sense to assume that g is real-valued and symmetric around the origin, so that one case say that we see in column k the spectrum (the frequency distribution) of the localized version $x \cdot T_{k-1}g$, which is essentially the part of the (long) signal concentrated around $k-1$. The function/vector g is called a *Gabor window* or *Gabor atom*.

Definition 1.

$$\mathrm{STFT}(x,g)(j,k) = \langle x, M_{j-1}T_{k-1}g\rangle, \quad 1 \le j,k \le N.$$

We speak of the STFT of the *signal* $\mathbf{x}$ with respect to the window g.This is the definition of a full STFT, which produces from a signal of length N an image of size $N \times N$. We have seen in the experiment that for $g0 = gaussnk(n)$ (e.g. $n = 480$) the function $STg0g0 = stft(g0, g0)$ looks like (and is essentially) a $2D$-Gaussian (at least in term of absolute values), but with larger broadness compared to the ordinary $2D$-Gauss function which would be of the form $g0(:) * g0(:)'$ (tensor product of functions, or table-multiplication!).

Of course one may vary in different ways. The Fourier transform is understood as the Fourier transform over the full group. So the signal length of $\mathbf{x}$ might be huge (in fact ongoing), while the window length might be much smaller. Now we can easily verify that taking a short FFT of the localized signal $x\dot{T}_{k-1}g$ (signal length determined by the length of the *window!*) amounts to the same as sampling the full/long FFT of that localized signal (at least of we choose as FFT-length some divisor of N!)

FOR OUR PRACTICAL DISCUSSIONS WE USE OVERALL SIGNAL LENGTH (LOWER CASE) n! SORRY FOR THAT.

Typically - for the theoretical considerations, we assume (aside from questions of efficient implementation) that we have the full STFT and sample it, typically on a regular grid, which means a product lattice, with horizontal shift-parameter usually called a and vertical shift-parameter b. Both should divide the overall signal length N so that we have n/a positions in time which are marked and n/b frequencies, so overall n^2/ab as the number of sampling points. This compared to n is just red $:= n/ab$, the redundancy of the *Gabor system*.

3 Going to the continuous setting, first $G = \mathbb{R}^d$

Although the theory of time-frequency analysis can be carried out without big problems over LCA (locally compact Abelian) groups we will concentrate here mostly on the case of $G = \mathbb{R}^d$, and in the concrete explanations probably mostly on the case $G = \mathbb{R}$ even. We will define Gabor families, Gabor Riesz sequences and Gabor frames accordingly.

The concept of a frame is arguable the "correct" generalization of "generating system" in a finite dimensional vector space. It brings in so-to-say the "quality of a generating system". If it is good it should not be close to a singular system, which in turn is geometrically related to the fact that some vectors are only hard to represent, or are *almost orthogonal* to the linear span. And it is more or less the SVD (singular value decomposition) of a matrix $\boldsymbol{A}$ which plays the crucial role, in fact mostly the distribution of singular values, which are (up to taking square roots) the eigenvalues of the positive (in fact in this case positive definite!) matrix $\boldsymbol{A} * \boldsymbol{A}'$.

Definition 2. A family $(f_i)_{i\in I}$ in a Hilbert space $\mathcal{H}$ is called a *frame* if there exist positive constants $A, B > 0$ such that for all $f \in \mathcal{H}$

$$A\|f\|^2 \leq \sum_{i\in I} |\langle f, f_i\rangle|^2 \leq B\|f\|^2 \tag{9}$$

One shows that condition (9) is satisfied if and only if the so-called *frame operator*

Definition 3.

$$S(f) := \sum_{i\in I} \langle f, f_i\rangle f_i, \quad \text{for} \quad f \in \mathcal{H},$$

is invertible. In fact, the double equation (9) is fully equivalent (!Exercise) to the ordering relationship

$$A \cdot \mathrm{Id}_{\mathcal{H}} \preceq \mathbf{S} \preceq B \cdot \mathrm{Id}_{\mathcal{H}}, \tag{10}$$

remembering that the ordering for positive operators is given through their quadratic forms, i.e. we have by convention

$$\boldsymbol{B}_1 \preceq \boldsymbol{B}_2 \quad \text{if and only if} \quad \langle \boldsymbol{B}_1 * \mathbf{x}, \mathbf{x}\rangle \leq \langle \boldsymbol{B}_2 * \mathbf{x}, \mathbf{x}\rangle, \quad \forall \mathbf{x} \in \mathcal{H},$$

one can find that the appropriate quality description can be put in the following form.

The obvious fact $S \circ S^{-1} = Id = S^{-1} \circ S$ implies that the (canonical) *dual frame* $(\tilde{f}_i)_{i\in I}$, defined by $\tilde{f}_i := S^{-1}(f_i)$ has the property that one has for $f \in \boldsymbol{H}$:

Definition 4.

$$f = \sum_{i\in I} \langle f, \tilde{f}_i\rangle f_i = \sum_{i\in I} \langle f, f_i\rangle \tilde{f}_i \tag{11}$$

Since S is *positive definite* in this case we can also get to a more symmetric expression by defining $h_i = S^{-1/2} g_i$. In this case one has

$$f = \sum_{i\in I} \langle f, h_i\rangle h_i \quad \text{for all} f \in \boldsymbol{H}. \tag{12}$$

The family $(h_i)_{i\in I}$ defined in this way is called the *canonical tight frame* associated to the given family $(g_i)_{i\in I}$.

LITERATURE: Book by Ole Christensen: [1]

If we consider $\mathbf{A}$ as a collection of column vectors $(f_i)_{i\in I}$ in some Hilbert space $\mathcal{H}$, then the role of $\mathbf{A}'$ is that of a coefficient mapping: $f \mapsto (\langle f, f_i\rangle)$.

$$\begin{tikzcd}
 & \ell^2(I) \arrow[dl, "\mathcal{R}"'] \arrow[d, "\mathbf{P}"] \\
\mathcal{H} \arrow[r, shift right, "\mathcal{C}"'] & \mathcal{C}(\mathcal{H}) \arrow[l, shift right, "\mathcal{R}"']
\end{tikzcd}$$

This diagram is **fully equivalent** to the frame inequalities.The *reconstruction mapping* is the pseudo-inverse of the matrix, and the commutativity of the diagram is just the fact that the (bounded linear) mapping $\mathcal{R}$ is just the left inverse to the coefficient mapping $\mathcal{C}$, i.e. $\mathcal{R} \circ \mathcal{C} = \mathrm{Id}_{\mathcal{H}}$. As a consequence (and in fact equivalently formulated) we have: The range of $\mathcal{C}$, i.e. $\mathcal{C}(\mathcal{H})$ a closed subspace within the prototypical Hilbert space $(\boldsymbol{\ell}^2(I), \|\cdot\|_2)$. We also recognize that the orthogonal projection $\mathbf{P}$ from $\boldsymbol{\ell}^2(I)$ onto $\mathcal{C}(\mathcal{H})$ is given by $\mathcal{C} \circ \mathcal{R}$, which is the analogue of our formula $\mathbf{P}_{col}(\boldsymbol{A}) = \boldsymbol{A}^+ * \boldsymbol{A}'$.

This diagram also indicated how to generalize the notion of frames to the Banach spaces setting. For now only in words: A *Banach frame* is an embedding, more precisely an isomorphism between a given Banach space $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ and a corresponding Banach space $(\boldsymbol{Y}, \|\cdot\|_{\boldsymbol{Y}})$ of sequences taking the role of $\boldsymbol{\ell}^2(I)$, as a partner, in fact a closed and complemented subspace of $\boldsymbol{Y}$, to make the diagram work. *In my opinion* it should not be just a sequence space, but a so-called Banach lattice of *solid* sequence space. So if the index set is (for simplicity, motivated by many practical examples) just the index set I, we assume that every $\mathbf{y} \in \boldsymbol{Y}$ is a complex-valued sequence of the form $\mathbf{y} = (y_i)_{i \in I}$. Solidity then means: if one has a sequence $\mathbf{z} = (z_i)_{i \in I}$ with $|z_i| \leq |y_i|, \forall i \in I$, then this should imply $\|\mathbf{z}\|_{\boldsymbol{Y}} \leq \|\mathbf{y}\|_{\boldsymbol{Y}}$

In the definition of a *Riesz basic sequence* the roles of the sequence space and of the "general function space" (typically a space of distributions, or measurable functions) are reversed, so the sequence space is embedded in the same, complemented way into the Banach space, and again this is expressed in the Hilbert space setting easily by just an inequality, but in the more general context better by a diagram.

In the terminology introduced in XXX this means that $\mathcal{R}$ is injective, but not surjective, and $\mathcal{C}$ is a left inverse of $\mathcal{R}$. Thus we have the following commutative diagram.

$$\begin{tikzcd}
\boldsymbol{X} \arrow[d, "\mathbf{P}"'] \arrow[dr, "\mathcal{C}"] & \\
\boldsymbol{X}_0 \arrow[r, shift left, "\mathcal{C}"] & \boldsymbol{Y} \arrow[l, shift left, "\mathcal{R}"]
\end{tikzcd}$$

Definition 5. A sequence (h_k) in a separable Hilbert space $\mathcal{H}$ is a *Riesz basis* for its closed linear span (sometimes also called a *Riesz basic sequence*) if for two constants $0 < D_1 \leq D_2 < \infty$,

$$D_1 \|c\|_{\boldsymbol{\ell}^2}^2 \leq \Big\| \sum_k c_k h_k \Big\|_{\mathcal{H}}^2 \leq D_2 \|c\|_{\boldsymbol{\ell}^2}^2, \qquad \forall c \in \boldsymbol{\ell}^2 \tag{13}$$

A detail description of the concept of *Riesz basis* can be found in ([12]) where the more general concept of Riesz projection bases is explained.

In particular a sequence (h_k) is a *Riesz basis* if and only if the corresponding *Gram matrix*, whose entries are the scalar products $(\langle h_k, h'_k \rangle)_{k,k'}$ is invertible on the corresponding $\boldsymbol{\ell}^2$-space.

A good example are the *spline-type spaces* which have a Riesz basis of translates of a given function (along a lattice). For example, it can be shown, that for any lattice $\Lambda \lhd \mathbb{R}^d$ the set of translates $(T_\lambda g_0)_{\lambda \in \Lambda}$ of the Gauss function are always a Riesz basis, and the Riesz bounds C, D are not too bad if the lattice is not too fine.

MENTION CUBIC SPLINES ETC.

Prototypical questions about frames arise in the context of RKHS (*reproducing kernel Hilbert spaces*), i.e. in situations, where the (abstract) Hilbert space consists of continuous functions on some domain, and where the point evaluations are continuous linear functionals.

Again one can refer to the finite dimensional case. Assume we have the 4-dimensional space of cubic polynomials over some interval $[a, b]$, or over $\mathbb{R}$ or even $\mathbb{C}$. Assume further that we have a finite number (at least 4) measurements, say 7 measurements of a cubic polynomials. Can we reconstruct the polynomial? What if we have imprecise measurements? What can we still say. The answer is in fact relatively easy and natural to formulate in words: Yes: there is a best fitting polynomial $p(x)$ (even a uniquely determined one), which may not pass exactly through the data, but minimizes the Euclidean error between the data vector, say $\vec{d} \in \mathbb{R}^7$ and the actual values of $p(x)$ at those given 7 points.

So what we have to look at is the 7×4 matrix which represents the linear mapping from the coefficients of $p(x)$ (MATLAB convention `polyval(a,x)`:

$$p(x) = p_{\mathbf{a}}(x) := a_1x^3 + a_2x^2 + a_3x + a_4, \quad \text{with } \vec{a} \in \mathbb{R}^4$$

which represents the linear (!) mapping $\vec{a} \to (p(x_j))_{j=1}^7$, but this is simply a part of the corresponding Vandermonde Matrix, namely the last 4 rows.

Just for illustration, let us take $\vec{x} = -3 : 3$ or $\vec{x} = [-3, -2 - 1, 0, 1, 2, 3]$; and $V = vander(-3 : 3)$, the we get this matrix:

$$\begin{pmatrix} 729 & -243 & 81 & -27 & 9 & -3 & 1 \\ 64 & -32 & 16 & -8 & 4 & -2 & 1 \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 64 & 32 & 16 & 8 & 4 & 2 & 1 \\ 729 & 243 & 81 & 27 & 9 & 3 & 1 \end{pmatrix} \tag{14}$$

Obviously we do not need the contributions to the monomials $x^k, k \geq 4$, so our system matrix $V47 = V(:, 4 : 7)$, or

$$\begin{pmatrix} -27 & 9 & -3 & 1 \\ -8 & 4 & -2 & 1 \\ -1 & 1 & -1 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 \\ 8 & 4 & 2 & 1 \\ 27 & 9 & 3 & 1 \end{pmatrix} \tag{15}$$

Given a data vector $\vec{d} \in \mathbb{R}^7$, which pretends to provide measured (i.e. approximate) values of $p(x_j), 1 \leq j \leq 7$ we cannot expect that this vector is in the (in fact) 4-dimensional subspace of all the tuples which *are in fact* samples of the form $(p(x_j))_{j=1}^7$. So we may want to find (in the spirit of the MNLSQ solution, using `pinv`) the best approximation of $\vec{d}$ from this four-dimensional range space, and then (since the mapping from the coefficients resp. cubic polynomials to the values is obviously injective, since we have more than 4 different points!) find the unique polynomial (or coefficients of that polynomial) which passes through the modified/projected version of $\vec{d}$, i.e. through $\mathbf{P}(\vec{d})$, where $\mathbf{P}$ is the projection on that four-dim. space.

Following our earlier recipes this is achieved by the one-line command

$$\vec{a} = \mathrm{pinv}(V47) * \vec{d}. \tag{16}$$

In the context of the frame we view this matrix multiplication with pinv($V47$) as the matrix multiplication forming the linear combinations of 7 vectors in $\mathbb{R}^4$, which are in fact the canonical dual frame to the linear functionals, which are point evaluations. But we know the matrix, which represents the transition from $\vec{a}$ (resp. the space of cubic polynomial functions with the monomial basis taken in decreasing order, i.e. with $\mathscr{M}^3$ it is just our starting point) to the values $(p(x_j))_{j=1}^7$, it is the (reduced) Vandermonde matrix $V47$.

TEST:

```
>> norm(pinv(pinv(V47))- V47)
ans =    9.5978e-14
```

So we may try to see what we can do in a similar, infinite-dimensional Kontext! And it is frame theory which allows us to think/act/compute in a quite similar way, of course with additional questions coming about (about boundedness, convergence etc.)

A verbal summary of the situation in the case of the STFT (the continuous Gabor transform or Short-Time Fourier Transform is, due to Balian-Low): There is no ideal = critical sampling density which on the one hand allows represent every signal (enough building blocks for stable expansion, e.g. $\mathscr{G}(g_0, a, b)$ with $a \cdot b < 1$) or enough linear independence in order to ensure that representations are unique (this is on the other hand valid even for $\vec{c} \in \boldsymbol{\ell}^\infty(I)$.

In a similar way the low-redundant case with $a \cdot b > 1$ provides Riesz basic sequences.

But there is a symmetry effect (the Ron-Shen principle) which implies that a family $\mathscr{G}(g, a, b)$ forms a (Gabor) frame if and only if $\mathscr{G}(g, 1/b, 1/a)$ forms a (Gaborian) Riesz basic sequence, in fact with equal condition number of the corresponding operators, i.e. the condition number of the frame operator of the first system coincides with the condition number of the system of the second system.

Qualitatively the situation is as follows: as $(a, b) \to 0$ the frame is getting more and more tight ($B/A \to 1$), while obviously $|(1/b, 1/a)| \to \infty$ which implies that for large values of $1/b$ and $1/a$ the Gabor system is only a mildly perturbed orthonormal basis. The nearest orthonormal basis is in fact automatically of Gaborian type as well (and can be obtained using the Loewdin orthonormalization, resp. the best tight frame is obtained by applying $\mathbf{S}^{-1/2}$ to the Gabor atom g.

4 Linear Algebra Recalled II

There are some further aspects of the theory of frames and Riesz basic sequences which will be better first formulated in a linear algebra context:

Given a bunch of vectors in the Hilbert space $\mathbb{C}^m$, say n vectors (which we put into a the $m \times n$-matrix $\boldsymbol{A}$, just for convenience of storage), in fact a bunch of linear functionals, realized by these vectors (due to the Riesz Representation Theorem we could formulate the same problem in the context of an abstract Hilbert space $\mathcal{H}$. So we have n linear measurements. What can we say about $f \in \mathcal{H}$?

Let us denote by $\mathcal{H}_0$ the (closed, finite-dimensional) subspace spanned by the vectors in $\mathcal{H}$ (representing these linear functions, via $f \mapsto \langle f, h_j \rangle, 1 \leq j \leq n$).

Then it is clear that one does not have any information about the part of f which is in $\mathcal{H}_0^\perp$, the orthogonal complement of $\mathcal{H}_0$. But can we recover at least the projection of f onto $\mathcal{H}_0$, which we denote for simplicity by $\mathbf{P}_0(f)$? Since we discuss the case that the vectors are linear independent we know that the Gramian matrix is invertible and should allow us to correct for the deviation from the perfect case, the case of orthogonality (which is the same as to say that the Gram matrix is the identify matrix of size $n \times n$!). Again, if we had an orthonormal basis from the very beginning, we could use the vectors (representing the linear measurements) for synthesis and get $\mathbf{P}_0(f) = \sum_{k=1}^n \langle f, h_j \rangle f_j$. But we are in general not in this lucky position.

So in matrix notation we start from $\boldsymbol{A}' * \mathbf{x}$ for some unknown vector $\mathbf{x} \in \mathbb{C}^m$ (replacing now h_j by $\mathbf{a}_j, 1 \leq j \leq n$). We want to get the projection onto the column space which is of the form $\boldsymbol{A} * \boldsymbol{A}^+$. But using now our formula ... (to be continued, April 2017, TUM, HGFei)

5 Comments on Coorbit Theory

The core of *coorbit theory* as outlined in [3,4] is to generate, from a given group representation π of some locally compact group $\mathscr{G}$ on a Hilbert space $\mathcal{H}$ a whole family of spaces $\boldsymbol{Co}(\boldsymbol{Y})$, with atomic decompositions spaces, providing coefficients in solid BK-spaces $\boldsymbol{Y}_d$ naturally associated with the coorbit space.

The first step is to define the so-called voice-transform V_g, which maps (suitably normalized) $\mathcal{H}$ isometrically into $\big(\boldsymbol{L}^2(\mathscr{G}), \|\cdot\|_2\big)$. For $f, g \in \mathcal{H}$ we set $V_g(f)(x) = \langle f, \pi(x) g \rangle, x \in \mathscr{G}$. Given any translation invariant, solid BF-space $(\boldsymbol{Y}, \|\cdot\|_{\boldsymbol{Y}})$ on $\mathscr{G}$ with $\boldsymbol{L}^2(\mathscr{G}) \cap \boldsymbol{Y}$ being a proper subspace of one can define the coorbit space $(\boldsymbol{Co}(\boldsymbol{Y}), \|\cdot\|_{\boldsymbol{Co}(\boldsymbol{Y})})$ as the space $\boldsymbol{Co}(\boldsymbol{Y}) := \{f \,|\, f \in \mathcal{H}, V_g(f) \in \boldsymbol{Y}\}$, with the norm $\|f\|_{\boldsymbol{Co}(\boldsymbol{Y})} := \|V_g(f)\|_{\boldsymbol{Y}}$.

Since one has to define a replacement for the set of Schwartz functions (the classical function spaces are usually defined as a subset of the Schwartz space $\mathscr{S}'(\mathbb{R}^d)$ of tempered distributions) we introduce the space $\boldsymbol{H}_w^1 := \{f \in \mathcal{H} \,|\, V_g(f) \in \boldsymbol{L}_w^1\}$, where w is a submultiplicative function on $\mathscr{G}$ (satisfying $w(xy) \leq w(x) \cdot w(y)$, and $\boldsymbol{L}_w^1 = \{f \,|\, fw \in \boldsymbol{L}^1(\mathscr{G})\}$).

Once we have a replacement for the space of test functions, namely $\boldsymbol{H}_w^1 = \boldsymbol{Co}(\boldsymbol{L}_w^1)$ we can define general coorbit spaces as subspace of the (anti)dual space of $\boldsymbol{H}_w^1$, we use the symbol $(\mathcal{H}_w^1)^\angle$.

Due to the π-invariance of $\boldsymbol{H}_w^1$ we can define for any $f \in (\mathcal{H}_w^1)^\angle$ the (generalized) *voice-transform* with respect to $g \in \boldsymbol{H}_w^1$:

$$V_g(f) = \langle f, \pi(\lambda)g \rangle, \quad x \in \mathscr{G},$$

where the pairing is understood as the natural extension of the usual pairing via the given sesquilinear-form on $\mathcal{H}$. Having now the appropriate general setting we can define for any *solid, translation invariant* BF-space $(\boldsymbol{Y}, \|\cdot\|_{\boldsymbol{Y}})$ the corresponding coorbit space

$$\boldsymbol{\mathcal{C}o}(\boldsymbol{Y}) := \{f \in (\mathcal{H}_w^1)^\angle \,|\, V_g(f) \in \boldsymbol{Y}\},$$

with the usual natural norm.

One of the main results of coorbit theory is the fact that in such a situation for any $g \in \boldsymbol{H}_w^1$ (or a slightly smaller space) one can say the following:

Theorem 1. *For any $g \in \boldsymbol{H}_w^1$ there exists a neighborhood of the identity U of $e \in \mathscr{G}$ such that the following is true: Any well-separated[2] which is U-dense discrete family $(x_i)_{i\in I}$ in $\mathscr{G}$ defines a Banach frame for $\boldsymbol{\mathcal{C}o}(\boldsymbol{Y})$.*

More precisely: There exists a solid BK-space $\boldsymbol{Y}_d$, consisting of complex-valued sequences over the index set I, and a bounded linear mappingf $\mapsto (c_i)_{i\in I}$ from $\boldsymbol{\mathcal{C}o}(\boldsymbol{Y})$ to $\boldsymbol{Y}_d$, such that

$$f = \sum_{i\in I} c_i \pi(x_i) g, \quad f \in \boldsymbol{\mathcal{C}o}(\boldsymbol{Y}),$$

with (unconditional) convergence in $(\boldsymbol{\mathcal{C}o}(\boldsymbol{Y}), \|\cdot\|_{\boldsymbol{\mathcal{C}o}(\boldsymbol{Y})})$.

One may observe a variety of similarities and differences compared to the original approach by Frazier/Jawerth:

6 References

[8, 9, 11]

A recent summary is given in ([7]). A draft version of this contribution to the Encyclopedia for Applied and Computational Mathematics is given in

http://www.univie.ac.at/nuhag-php/bibtex/open_files/13217_ENCAFeiLuef7.pdf

Concerning Gabor Analysis over finite Abelian groups (no convergence problem, just the algebraic structure) ([5]).

Dealing with the problems in the continuous setting: see ([2, 6]).

Literatur

[1] O. Christensen. *An Introduction to Frames and Riesz Bases.* Applied and Numerical Harmonic Analysis. Birkhäuser, Boston, 2003.

[2] This means that it is the finite union of uniformly separated sets in $\mathscr{G}$.

[2] E. Cordero, H. G. Feichtinger, and F. Luef. Banach Gelfand triples for Gabor analysis. In *Pseudo-differential Operators*, volume 1949 of *Lect. Notes Math.*, pages 1–33. Springer, Berlin, 2008.

[3] H. G. Feichtinger and K. Gröchenig. Banach spaces related to integrable group representations and their atomic decompositions, I. *J. Funct. Anal.*, 86(2):307–340, 1989.

[4] H. G. Feichtinger and K. Gröchenig. Banach spaces related to integrable group representations and their atomic decompositions, II. *Monatsh. Math.*, 108(2-3):129–148, 1989.

[5] H. G. Feichtinger, W. Kozek, and F. Luef. Gabor Analysis over finite Abelian groups. *Appl. Comput. Harmon. Anal.*, 26(2):230–248, 2009.

[6] H. G. Feichtinger and F. Luef. Wiener amalgam spaces for the Fundamental Identity of Gabor Analysis. *Collect. Math.*, 57(Extra Volume (2006)):233–253, 2006.

[7] H. G. Feichtinger and F. Luef. Gabor Analysis and Algorithms. In B. Engquist, editor, *Encyclopedia of Applied and Computational Mathematics*, pages 575–579. Springer Berlin Heidelberg, 2015.

[8] H. G. Feichtinger and T. Strohmer. *Gabor Analysis and Algorithms. Theory and Applications.* Birkhäuser, Boston, 1998.

[9] H. G. Feichtinger and T. Strohmer. *Advances in Gabor Analysis.* Birkhäuser, Basel, 2003.

[10] D. Gabor. Theory of communication. *J. IEE*, 93(26):429–457, 1946.

[11] K. Gröchenig. *Foundations of Time-Frequency Analysis.* Appl. Numer. Harmon. Anal. Birkhäuser, Boston, MA, 2001.

[12] G. Zimmermann. *Projective Multiresolution Analysis and Generalized Sampling.* PhD thesis, University of Maryland, College Park, December 1994.

STATE of May 18th, 2017, checked June 2023