Basics

Sylvester's inequality: If A is an $M\times N$ matrix and B is an $N\times K$ matrix, then $\text{rank}(A)+\text{rank}(B)-N\le\text{rank}(AB)$.
Cholesky decomposition: Any A hermitian PD matrix can be expressed as $A=LL^*$, where where L is an invertible lower triangular matrix with real and positive diagonal entries.
Vandermonde matrix: A Vandermonde matrix V of size $M\times N$ has entries of the form $V_{i,j}=t_i^j$ for $i=0,\dots,M-1$, $j=0,\dots,N-1$, and it is of full rank if $t_l\neq t_m$ when $l\neq m$ since $\det(V)=\prod_{0\le l<m\le N-1}(t_l-t_m)$.
Circulant matrix: Implement circular convolution. Its rows are circular rotations of a sequence. They are diagonalized bu the DFT matrix.
Toeplitz matrix; Constant coefficients along diagonals. Implements linear convolution. In a matrix representation of a linear and shift-invariant system, the matrix will be Toeplitz.
Unitary matrix: $U^{-1}=U^*$ and its eigenvalues satisfy $|\lambda_j|=1$. Unitary matrices are isometries, i.e. $\|U\mathbf{x}\|=\|\mathbf{x}\|$.
Normal matrices: $AA^*=A^*A$ (so need to be square).
Eigenvalue decomposition: $A\mathbf{x}=\lambda\mathbf{x}$ (characteristic equation), $\|\mathbf{x}\|=1$, $\lambda\in\overline{\mathbb{F}}$. λ verifies $\det(\lambda I-A)=0$. Square matrix diagonalizable if $\exists S$ $n\times n$ and Λ diagonal matrix s.t. $A=S\Lambda S^{-1}$. For all square matrices, $A=URU^*$ with R upper triangular matrix, U unitary (Schur decomposition). For normal matrices, $R=\Lambda$, with Λ diagonal.
Singular value decomposition: $A=USV^*$, U,V unitary, S (e.g. for tall matrix) hermitian diagonal block and 0 block, $U=[U_1\quad U_0]$. This leads to thin SVD $A=U_1S_1V^*$, U_1,U_0 form orthonormal basis for $\mathcal{R}(A),\mathcal{N}(A^*)$ resp. U,V are resp- eigenvectors of normal matrices AA^* and A^*A. $A^\dagger=VS^\dagger U^*$ ($S^\dagger$ is S^T with inverted singular values).
Determinant of a matrix: Oriented volume of the hyper-parallelepiped defined by column vectors of A. $|\det(U)|=1$ for unitary matrices (e.g. rotation and reflection).
Geometric series: For $r\neq1$, $\sum_{k=n_0}^{n}r^k=\frac{r^{n_0}-r^{n+1}}{1-r}$.
Pseudo-inverse: The pseudo-inverse of A denoted $A^\dagger$ satisfies all the four statements: $AA^\dagger A=A$, $A^\dagger AA^\dagger=A$, $(AA^\dagger)^*=AA^\dagger$ and $(A^\dagger A)^*=A^\dagger A$. If A has linearly independent columns (non-singular), A^*A invertible and $A^\dagger=(A^*A)^{-1}A^*$ is a left-inverse. If A has linearly independent rows (A^* non-singular), AA^* invertible and $A^\dagger=A^*(AA^*)^{-1}$ is a right-inverse.
Building projections: $AA^\dagger$ orthogonal projection into $\mathcal{R}(A)$. $A^\dagger A$ orthogonal projection into $\mathcal{R}(A^*)$ (so orthogonal projection into $\mathcal{N}(A)$ is $I-A^\dagger A$). For U orthonormal set of vectors, $U^\dagger=U^*$.
Spectral norm: $\|A\|:=\sup_{\mathbf{x}\neq\mathbf{0}}\frac{\|A\mathbf{x}\|}{\|\mathbf{x}\|}=\sup_{\|\mathbf{x}\|=1}\|A\mathbf{x}\|=\sigma_{\max}(A)=\sqrt{\lambda_{\max}(A^*A)}$. Spectral norm is submultiplicative: $\|AB\|\le\|A\|\,\|B\|$.
Change of coordinates: $\iint_C f(x,y)dxdy=\iint_P f(r,\phi)\begin{vmatrix}\frac{\partial x}{\partial r}&\frac{\partial x}{\partial\phi}\\ \frac{\partial y}{\partial r}&\frac{\partial y}{\partial\phi}\end{vmatrix}drd\phi$.
In case of switching from cartesian to polar, $x=r\cos(\phi)$, $y=r\sin(\phi)$, $r=\sqrt{x^2+y^2}$, $\phi=\arctan\left(\frac{y}{x}\right)$ and determinant becomes r.

Trigonometric identities

Even/odd: $\sin(-x)=-\sin(x)$, $\cos(-x)=\cos(x)$, $\tan(-x)=-\tan(x)$.
Cofunction: $\sin(\frac{\pi}{2}-x)=\cos(x)$, $\cos(\frac{\pi}{2}-x)=\sin(x)$.
Sum and difference of angles: $\sin(x+y)=\sin x\cos y+\cos x\sin y$, $\sin(x-y)=\sin x\cos y-\cos x\sin y$, $\cos(x+y)=\cos x\cos y-\sin x\sin y$, $\cos(x-y)=\cos x\cos y+\sin x\sin y$.
Double angles: $\sin(2x)=2\sin x\cos x$, $\cos(2x)=\cos^2x-\sin^2x=2\cos^2x-1=1-2\sin^2x$.
Product to sum: $\sin x\sin y=\frac12[\cos(x-y)-\cos(x+y)]$, $\cos x\cos y=\frac12[\cos(x-y)+\cos(x+y)]$, $\sin x\cos y=\frac12[\sin(x+y)+\sin(x-y)]$.
Derivatives: $\frac{d}{dx}\sin(x)=\cos(x)$, $\frac{d}{dx}\cos(x)=-\sin(x)$.

Ranges, nullspaces and invertibility

For matrix A of size $m\times n$.
Range: $\mathcal{R}(A)=\{y:y=Ax\}$ (linear subspace of $\mathbb{F}^m$).
Nullspace: $\mathcal{N}(A)=\{x:Ax=0\}$ (linear subspace of $\mathbb{F}^n$).
Rank: Cardinality of largest set of linearly independent columns (or rows) of A; $\text{rank}(A)=\dim(\mathcal{R}(A))=\dim(\mathcal{R}(A^*))\le\min(m,n)$ (if equality, full rank). $\text{rank}(A)=\text{rank}(A^*)$.
Rank-nullity theorem: $\dim(\mathcal{R}(A))+\dim(\mathcal{N}(A^*))=m$ (since $\mathcal{R}(A)=\mathcal{N}(A^*)^\perp$), $\dim(\mathcal{R}(A^*))+\dim(\mathcal{N}(A))=n$ (since $\mathcal{R}(A^*)=\mathcal{N}(A)^\perp$).
Invertibility (equivalences): For A square matrix, $\mathcal{N}(A)=\{0\}$ (trivial nullspace), full-rank matrix, no 0 eigenvalue, determinant equal to 0.

Hilbert spaces and projection operators

Vector space: Set of vectors ($\mathbb{R}^N$, functions,...), field of scalars (real, complex), vector addition, scalar multiplication. Satisfy $x+y=y+x$ (commutativity), $(x+y)+z=x+(y+z)$ and $(\alpha\beta)x=\alpha(\beta x)$ (associativity), $\alpha(x+y)=\alpha x+\alpha y$ and $(\alpha+\beta)x=\alpha x+\beta x$ (distributivity), $\exists\mathbf{0}$ s.t. $x+\mathbf{0}=x$ (additive identity), $\exists-x$ s.t. $x+(-x)=\mathbf{0}$ (additive inverse), $1x=x$ (multiplicative identity).
Subspace: $S\subseteq V$, V vector space and also S itself. $S\neq\emptyset$, closed under vector addition ($x+y\in S$ $\forall x,y\in S$) and scalar multiplication ($\alpha x\in S$ $\forall x\in S,\alpha\in\mathbb{F}$, so $\{\mathbf{0}\}\in S$).
span(S): $\{\sum_{k=0}^N\alpha_k\phi_k:\alpha_k\in\mathbb{F},\phi_k\in S,N\in\mathbb{N}\setminus\{\infty\}\}$. Smallest vector space containing the set of vectors S. For S infinite, $\overline{\text{span}}(S)$ (o.w. some vectors cannot be represented with **finite** linear combiation of ϕ_k). Always a subspace.
Inner product: $\langle\mathbf{x}+\mathbf{y},\mathbf{z}\rangle=\langle\mathbf{x},\mathbf{z}\rangle+\langle\mathbf{y},\mathbf{z}\rangle$ (distributivity), $\langle\alpha\mathbf{x},\mathbf{y}\rangle=\alpha\langle\mathbf{x},\mathbf{y}\rangle$ (linearity in the first argument), $\langle\mathbf{x},\mathbf{y}\rangle^*=\langle\mathbf{y},\mathbf{x}\rangle$ (hermitian symmetry), $\langle\mathbf{x},\mathbf{x}\rangle\ge0$ equality iff $\mathbf{x}=\mathbf{0}$ (PD).
On $\mathbb{C}^\mathbb{R}$, $\langle x,y\rangle=\int_{-\infty}^{\infty}x(t)y^*(t)dt$.
Norm: $\|\mathbf{x}\|\ge0$ equality iff $\mathbf{x}=\mathbf{0}$ (PD), $\|\alpha\mathbf{x}\|=|\alpha|\,\|\mathbf{x}\|$ (positive scalability), $\|\mathbf{x}+\mathbf{y}\|\le\|\mathbf{x}\|+\|\mathbf{y}\|$ equality iff $\mathbf{y}=\alpha\mathbf{x}$ (triangle inequality). $\langle\mathbf{x},\mathbf{y}\rangle=\|\mathbf{x}\|\,\|\mathbf{y}\|\cos\alpha$.
Pythagorean theorem: For $\{\mathbf{x}_k\}$ orthogonal, $\left\|\sum_k\mathbf{x}_k\right\|^2=\sum_k\|\mathbf{x}_k\|^2$.
Cauchy-Schwarz: $|\langle\mathbf{x},\mathbf{y}\rangle|\le\|\mathbf{x}\|\,\|\mathbf{y}\|$. Equality when $\mathbf{x}=\alpha\mathbf{y}$ (collinear).
Parallelogram law: $\|x+y\|^2+\|x-y\|^2=2(\|x\|^2+\|y\|^2)$.
Linear operator: $A(x+y)=Ax+Ay$ (additivity), $A(\alpha x)=\alpha(Ax)$ (scalability). Operator norm $\|A\|=\sup_{\|x\|=1}\|Ax\|$.
Adjoint: For $A:\mathcal{H}_0\to\mathcal{H}_1$, $\langle Ax,y\rangle_{\mathcal{H}_1}=\langle x,A^*y\rangle_{\mathcal{H}_0}$. Exists and is unique. $\|A^*\|=\|A\|$. Operator is unitary iff $A^{-1}=A^*$. If a matrix is self-adjoint (hermitian), its eigenvectors define an orthogonal basis. If A invertible, $(A^{-1})^*=(A^*)^{-1}$.
Projection: Bounded linear operator ($\|P\|<\infty$; linear operators with finite-dimensional domains are always bounded) that is idempotent $P^2=P$ (in such case $\|P\|\ge1$). If self-adjoint $P^*=P$, orthogonal projection and all eigenvalues are real valued (o.w. oblique).Bounded linear operator P satisfies $\langle x-Px,Py\rangle=0\forall x,y\in H$ iff P orthogonal projection. If S,T closed subspaces s.t. $H=S\oplus T$ $\exists$projection P on H s.t. $S=\mathcal{R}(P),T=\mathcal{N}(P)$. In an orthogonal projection, all eigenvalues are either 0 or 1, and we have that $\|P\mathbf{x}\|\le\|\mathbf{x}\|$ (orthogonal projection is a contraction). Moreover, if range space is not trivial, $\|P\|=1$.
Projection theorem: For S closed subspace of Hilbert space H and $x\in H$, $\|x-\hat{x}\|\le\|x-s\|$ $\forall s\in S$ iff $x-\hat{x}\perp S$, $\hat{x}=Px$, P orthogonal projection.
Basis: $\Phi=\{\phi_k\}_{k\in\mathcal{K}}\subset V$, $V=\overline{\text{span}}(\Phi)$ so any $x\in V$ can be expressed as $x=\sum_{k\in\mathcal{K}}\alpha_k\phi_k$ and expansion coefficients α_k are unique.
Linear independence: $\sum_{k\in\mathcal{K}}\alpha_k\phi_k=0$ iff $\alpha_k=0$ $\forall k$.
Riesz basis: Φ basis of Hilbert space H, $\exists0<\lambda_{\min}\le\lambda_{\max}<\infty$ s.t. $\forall x\in H$, $\lambda_{\min}\|x\|^2\le\sum_{k\in\mathcal{K}}|\alpha_k|^2\le\lambda_{\max}\|x\|^2$. Let $G=\Phi^*\Phi$ the **Gram matrix**. $\lambda_{\max}(G)=1/\lambda_{\min}$, $\lambda_{\min}(G)=1/\lambda_{\max}$.
Orthogonal projection: $\alpha_k=\langle x,\phi_k\rangle$, i.e. $\alpha=\Phi^*x$. This gives orthogonal projection to $\overline{\text{span}}(\{\phi_k\})$ (also for orthonormal set, not necessarily basis).
Gram matrix: $G=\Phi^*\Phi$, $G_{i,k}=\langle\phi_k,\phi_i\rangle$.
Biorthogonal pairs of bases: $\Phi,\tilde{\Phi}$ both bases for H and biorthogonal: $\left\langle\phi_i,\tilde{\phi}_k\right\rangle=\delta_{i-k}$. In this case $\alpha_k=\left\langle x,\tilde{\phi}_k\right\rangle$, i.e. $\alpha=\tilde{\Phi}^*x$. Residual of projection on biorthogonal pairs of sets, satisfies $x-Px\perp\overline{\text{span}}(\{\tilde{\phi}\}_{k\in\mathcal{I}})$ $\tilde{\Phi}=\Phi G^{-1}$ if Φ Riesz basis.
Parseval equality: For Φ orthonormal basis for H, $\langle x,y\rangle=\langle\Phi^*x,\Phi^*y\rangle=\langle\alpha,\beta\rangle$ since $\Phi^*\Phi=I$ on $\ell^2(\mathcal{K})$, $\Phi\Phi^*=I$ on H (unitary operator). For biorthogonal pairs of bases, $\langle x,y\rangle=\langle\tilde{\alpha},\beta\rangle$ since $\tilde{\Phi}^*\Phi=I$ on $\ell^2(\mathcal{K})$, $\Phi\tilde{Phi}^*=I$ on H.
Bessel's inequality: For orthonormal sets, $\Phi^*\Phi=I$ but $\Phi\Phi^*\neq I$ in general. $\|x\|^2\ge\|\Phi_{\mathcal{I}}^*x\|^2$ with equality when Φ is a basis.
Inverse problem: Find x s.t. $y=Ax$. Solution $\hat{x}$ is consistent with measurements if $\{\hat{x}:\hat{x}=x+\tilde{x},\tilde{x}\in\mathcal{N}(A)\}$, i.e. $A\tilde{x}=y$.
Dual basis; $\tilde{\Phi}=\Phi(\Phi\Phi^*)^{-1}=\Phi G^{-1}$, G the Gram matrix.
Change of basis: $x,y\in H$, $y=Ax$, $x=\Phi\alpha$, $y=\Psi\beta$, $\beta=\Gamma\beta$. $\Gamma_{i,j}=\left\langle A\phi_j,\tilde{\psi}_i\right\rangle$.
Gram-Schmidt procedure: Want to convert original set $\{s^{(k)}\}$ onto orthonormal set $\{u^{(k)}\}$. At each step k, $p^{(k)}=s^{(k)}-\sum_{n=0}^{k-1}\left\langle u^{(n)},s^{(k)}\right\rangle u^{(n)}$. $u^{(k)}=p^{(k)}/\left\|p^{(k)}\right\|$.

Discrete Systems

Delta function: $\int_{-\infty}^{\infty}\delta(t)dt=1$, $\delta(t)=0$ $\forall t\neq0$, $\int_{-\infty}^{\infty}x(t)\delta(t)dt=x(0)$, $\int_{-\infty}^{\infty}x(\tau)\delta(t-\tau)d\tau=x(t)$. $\delta(at)=\delta(t)/|a|$.
Impulse response: Let $y_k(t)=A(xk(t))$. If A linear ($A(\sum_k\alpha_kx_k(t))=\sum_k\alpha_ky_k(t)$) shift-invariant ($A(x_k(t-\tau))=y_k(t-\tau)$) system, can express A as convolution with impulse response $h(t)=A(\delta(t))$; $A(x(t))=(x*h)(t)$.
Filter as projection: A filter is said to be a projection if $h_n=(h*h)_n$, or equivalently $H(e^{j\omega})=H^2(e^{j\omega})$. In case of being a projection, orthogonal projection if $h_n=h_{-n}^*$, or equivalently $H(e^{j\omega})=H^*(e^{j\omega})$.

Transforms

	Time domain	Frequency domain
Fourier Transform	Continuous aperiodic	Continuous aperiodic
Fourier Series	Continuous periodic	Discrete aperiodic
DTFT	Discrete aperiodic	Continuous aperiodic
DFT	Discrete periodic	Discrete periodic

Fourier Transform

Definition: $X(\omega)=\int_{-\infty}^{\infty}x(t)e^{-j\omega t}dt$, $x(t)=\frac{1}{2\pi}\int_{-\infty}^{\infty}X(\omega)e^{j\omega t}$.
Continuous-time convolution: $(x*y)(t)=\int_{-\infty}^{\infty}x(\tau)y(t-\tau)d\tau$. Eigenfunctions $e^{j\omega t}$.
Parseval: $\langle x(t),y(t)\rangle=\frac{1}{2\pi}\langle X(\omega),Y(\omega)\rangle$.
Properties: $x(\alpha t)\leftrightarrow\frac{1}{\alpha}X\left(\frac{\omega}{\alpha}\right)$, $(x*y)(t)\leftrightarrow X(\omega)Y(\omega)$, $x(t)y(t)\leftrightarrow\frac{1}{2\pi}(X*Y)(\omega)$, $x(t-\tau)\leftrightarrow e^{-j\omega\tau}X(\omega)$, $e^{j\omega_0t}x(t)\leftrightarrow X(\omega-\omega_0)$, $x(-t)\leftrightarrow X(-\omega)$, $x^*(t)\leftrightarrow X^*(-\omega)$, $\frac{d^nx(t)}{dt^n}\leftrightarrow(j\omega)^nX(\omega)$, $(-jt)^nx(t)\leftrightarrow\frac{d^nX(\omega)}{d\omega^n}$, $\int_{-\infty}^{t}x(\tau)d\tau\leftrightarrow\frac{X(\omega)}{j\omega}$, $X(0)=0$.
Poisson sum formula: $\sum_{n\in\mathbb{Z}}x(t-nT)=\frac{1}{T}\sum_{m\in\mathbb{Z}}X(\frac{2\pi m}{T})e^{j\frac{2\pi mt}{T}}$. Proof: LHS is $T-$periodic so compute Fourier Series and anti-transform result.
Common transforms: $\text{sinc}(\omega_0t)=\frac{\sin\omega_0t}{\omega_0t}\leftrightarrow\frac{\pi}{\omega_0}\mathbb{1}\{\omega\in[-\omega_0,\omega_0]\}$. $\sum_{n\in\mathbb{Z}}\delta(t-nT)\leftrightarrow\frac{2\pi}{T}\sum_{k\in\mathbb{Z}}\delta(\omega-\frac{2k\pi}{T})$. $\frac{1}{t_0}\mathbb{1}\{t\in[-t_0/2,t_0/2]\}\leftrightarrow\text{sinc}(\frac{t_0\omega}{2})$. $1-|t|$, $|t|<1\leftrightarrow\text{sinc}^2(\frac{\omega}{2})$. $1\leftrightarrow2\pi\delta(\omega)$. $\delta(t)\leftrightarrow1$.

Fourier Series (signal of period T)

Definition: $X_k=\frac{1}{T}\int_{-\frac{T}{2}}^{\frac{T}{2}}x(t)e^{-j\frac{2\pi kt}{T}}dt$, $x(t)=\sum_{k\in\mathbb{Z}}X_ke^{j\frac{2\pi kt}{T}}$.
Circular continuous-time convolution: For x,y $T-$periodic. $(x\circledast y)(t)=\int_{-\frac{T}{2}}^{\frac{T}{2}}x(\tau)y(t-\tau)d\tau$. Eigenfunctions $e^{j\frac{2\pi kt}{T}}$.
Parseval: $\langle x(t),y(t)\rangle=T\langle X_k,Y_k\rangle$.
Properties: $x(t-t_0)\leftrightarrow e^{-j\frac{2\pi kt_0}{T}}X_k$, $e^{j\frac{2\pi k_0t}{T}}x(t)\leftrightarrow X(k-k_0)$, $(h\circledast x)(t)\leftrightarrow TH_kX_k$, $h(t)x(t)\leftrightarrow(H*X)_k$.
Common transforms: $\sum_{n\in\mathbb{Z}}\delta(t-nT)\leftrightarrow1/T$.

DFTF

Definition: $X(e^{j\omega})=\sum_{n\in\mathbb{Z}}x_ne^{-j\omega n}$, $x_n=\frac{1}{2\pi}\int_{-\pi}^{\pi}X(e^{j\omega})e^{j\omega n}d\omega$.
Discrete convolution: $(x*y)_n=\sum_{k\in\mathbb{Z}}x_ky_{n-k}$. Eigenfunctions $e^{j\omega n}$.
Properties: Transform is $2\pi-$periodic. $(h*x)_n\leftrightarrow H(e^{j\omega})X(e^{j\omega})$, $h_nx_n\leftrightarrow\frac{1}{2\pi}(H\circledast X)(e^{j\omega})$.

DFT

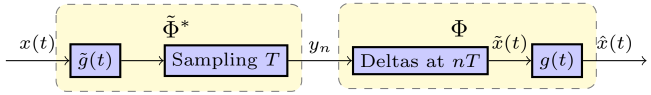
Definition: $X_k = \sum_{n=0}^{N-1} x_n W_N^{kn}$, $x_n = \sum_{k=0}^{N-1} X_k W_N^{-kn}$, $W_N^{kn} = e^{-j\frac{2\pi kn}{N}}$.

Discrete circular convolution: For x, y sequences of length N. $(x \circledast y)_n = \sum_{k=0}^{N-1} x_k y_{n-k}$. Eigenfunctions $e^{j\frac{2\pi kn}{N}}$. Same as discrete convolution if x has support M, y has support L and $N \geq L + M - 1$.

Properties: $x_{(n-n_0) \mod N} \leftrightarrow W_N^{kn_0} X_k$, $W_n^{-k_0 n} \leftrightarrow X_{(k-k_0) \mod N}$, $(h \circledast x)_n \leftrightarrow H_k X_k$, $h_n x_n \leftrightarrow \frac{1}{N}(H \circledast X)_k$, $x_n^* = X_{-k \mod N}^*$, $x_{-n \mod N}^* \leftrightarrow X_k^*$.

Sampling and Interpolation

Interpolation $\tilde{\Phi}$, sampling $\tilde{\Phi}^*$. $\tilde{S} = \mathcal{N}(\tilde{\Phi}^*)^\perp = \mathcal{R}(\tilde{\Phi})$ (what can be measured). $S = \mathcal{R}(\Phi)$ (what can be reproduced). $P = \Phi\tilde{\Phi}^*$ projection (with range S and $x - \hat{x} \perp \tilde{S}$, where $\hat{x} = Px$) iff $\tilde{\Phi}^*\Phi = I$ (or equivalently $\langle g(t-kT), \tilde{g}^*(nT-T)\rangle_t = \delta_{k-n}$). In this case, we say that sampling and interpolation are consistent (and $\hat{x}$ is the best least-squares approximation of x in S). When $\Phi = \left(\tilde{\Phi}^*\right)^\dagger = \tilde{\Phi}(\tilde{\Phi}^*\tilde{\Phi})^{-1}$ (for orthogonal vectors, pseudo-inverse is the adjoint), they form a biorthogonal pair of bases for S and hence $S = \tilde{S}$ and we say that operators are ideally matched (orthogonal projection). To prove $S = \tilde{S}$ when $\Phi = \left(\tilde{\Phi}^*\right)^\dagger$, note that Φ is a linear combination of columns of $\tilde{\Phi}$ with coefficients given by the corresponding column of $(\tilde{\Phi}^*\tilde{\Phi})^{-1}$ (full rank matrix). Hence, columns of $\Phi, \tilde{\Phi}$ must span the same space.



$y_n = \int_{-\infty}^{\infty} x(\tau) g^*(\tau - t) d\tau \Big|_{t=nT}$; $\tilde{x}(t) = \sum_{n\in\mathbb{Z}} y_n \delta(t - nT)$

$\hat{x}(t) = \sum_{k\in\mathbb{Z}} y_k g(t-kT) = \int_{-\infty}^{\infty} \tilde{x}(\tau) g(t-\tau) d\tau = \Phi\tilde{\Phi}^* x = Px$.

In the orthogonal case, $\tilde{\Phi}^* = \Phi^*$ and $\tilde{g}(t) = g^*(-t)$

Shift-invariant subspace: A subspace $S \subset \mathcal{L}^2(\mathbb{R})$ is a shift-invariant subspace with respect to shift $T \in \mathbb{R}^+$ if $x(t) \in S \implies x(t-kT) \in S\ \forall k \in \mathbb{Z}$. $s \in \mathcal{L}^2(\mathbb{R})$ is called a generator of S when $S = \overline{\text{span}}(\{s(t-kT)\}_{k\in\mathbb{Z}})$. To check for latter, have to show that $\forall x \in S$, $\exists\{\alpha_k\}_{k\in\mathbb{Z}}$ unique s.t. $x(t) = \sum_{k\in\mathbb{Z}} \alpha_k s(t-kT)$.

Subspace of bandlimited functions: A function $x(t) \in \mathcal{L}^2(\mathbb{R})$ has bandwidth $\omega_0 \in [0, \infty)$ if its FT satisfies $X(\omega) = 0\ \forall|\omega| > \frac{\omega_0}{2}$. This defines the space of ω_0- bandlimited functions, $\text{BL}[-\omega_0/2, \omega_0/2]$ which is a shift-invariant subspace (proof by delay property of FT).

Sampling theorem: If $x \in \text{BL}[-\pi/T, \pi/T]$, $x(t) = \sum_{n\in\mathbb{Z}} x(nT)\text{sinc}\left(\frac{\pi}{T}(t-nT)\right)$. This means that sinc is a generator for $\frac{2\pi}{T}-$band limited functions. In this case, choose $g(t) = \frac{1}{\sqrt{T}}\text{sinc}\left(\frac{\pi t}{T}\right) = g^*(-t)$ since it s a generator with shit T of $\text{BL}[-\pi/T, \pi/T]$ and $\{g(t-kT)\}_k$ are orthonormal. If $x \in \text{BL}[-\omega_0/2, \omega_0/2]$, we need $T < 2\pi/\omega_0$ (Nyquist interval). $\omega_0/2\pi$ is called the Nyequist rate.

Continuous-time convoluton via DSP: For $x \in \text{BL}[-\pi/T, \pi/T]$, $y = h*x$ can be computed by sampling, filtering the resulting sequence and interpolating the result of the convolution. $\tilde{h}_n = \langle h(t), \text{sinc}(\frac{\pi}{T}(t-nT))\rangle$ without pre-filter (first filter in sampling is multiplying by $\sqrt{T}$).

Computational Tomography

Parametrize line with angle $\theta \in [0, \pi)$ of the line's normal vector and signed distance $t \in \mathbb{R}$ from the origin. Cannot use $y = mx + n$ since vertical lines cannot be described.

Explicit: x, y parametrized by $s \in \mathbb{R}$, where $L_{\theta,t} = \{(x(s), y(s))\}$. $x(s) = t\cos\theta - s\sin\theta$, $y(s) = t\sin\theta + s\cos\theta$

Implicit: $L_{\theta,t} = \{(x,y) : x\cos\theta + y\sin\theta = t\}$.

Radon transform: Gives sinogram. $\mathcal{R}[f](\theta, t) = p(\theta, t) = \int_{L_{\theta,t}} f(x,y)ds = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\delta(x\cos\theta + y\sin\theta - t)dxdy$.

Laminogram or Backprojection Summation: Adjoint of Radon transform. Assign every point in the image along $L_{\theta,t}$ the value $p(\theta, t)$. Gives blurred reconstruction. $\mathcal{B}[p](x,y) = f_b(x,y) = \int_0^\pi p(\theta, x\cos\theta + y\sin\theta)d\theta$.

Fourier Slice Theorem: The Fourier transform of a parallel projection of an image $f(x,y)$ taken at angle θ gives a slice of its two-dimensional Fourier transform, $F(u,v)$, that subtends an angle θ with the u-axis. $P(\theta, \omega) = F(u,v)|_{u=\omega\cos\theta, v=\omega\sin\theta} = F(\omega\cos\theta, \omega\sin\theta)$, where $P(\theta,\omega)$ is the 1D-FT of $p(\theta, t)$ w.r.t t.

Fourier Reconstruction Method: Take the 1D Fourier transform of each projection (w.r.t. t). Insert the results in the appropriate slices in the (u,v)-plane. Resample on a rectangular grid in the (u,v)-plane. Take the 2D IFT of the formed 2D spectrum.

Filtered Backprojection (FBP)

Gives images very close to the original. FBP does not work easily with constraints. With limited data and/or non-uniform distribution of projection angles, reconstruction with FBP contains artifacts. In theory, $f(x,y) = \mathcal{B}[q](x,y)$ where $q(\theta, t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} P(\theta,\omega)\frac{|\omega|}{2\pi}e^{j\omega t}d\omega$. To proof this, do inverse 2D-IFT of $F(u,v)$. Change Cartesian to polar coordinates with $\omega \in \mathbb{R}_+, \theta \in [0, 2\pi)$ (Recall that $dudv = \omega d\theta d\omega$). Split integral in θ from 0 to π for θ and $\theta + \pi$. Note that $\tilde{F}(\omega, \theta+\pi) = \tilde{F}(-\omega, \theta)$ since $\cos(\theta+\pi) = -\cos\theta$ and $\sin(\theta+\pi) = -\sin\theta$. Apply Fourier Slice Theorem.

Convolution backprojection: Can also write $f(x,y) = \int_{-\infty}^{\infty} p(\theta,t)h(x\cos\theta + y\sin\theta - t)dtd\theta$, where $h(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{|\omega|}{\pi}e^{j\omega t}d\omega$. Given that $|\omega| \notin \mathcal{L}^2(\mathbb{R})$, its FT doesn't exist, so in practice $H(\omega) = \frac{|\omega|}{2\pi}W(\omega)$ for W a suitable window. IFT of ramp filter $\frac{|\omega|}{2\pi}$ is $\text{sinc} - \text{sinc}^2$ (rectangle - triangle).

Reconstruct image from sinogram: Take 1D FT of projections (w.r.t. t) apply filter $H(\omega)$ and take 1D IFT (if $h(t)$ short, do convolution in spatial domain). Backproject filtered projections and sum backprojected images.

Algebraic Reconstruction

Suppose $f(x,y) = \sum_{i=0}^{N-1} f_i\phi_i(x,y)$, where ϕ_k are the basis functions. Beam shapes are h_i. Hence, $b_i = \langle \mathbf{f}, h_i\rangle$. With spline surface model, $\mathbf{b}$ (the measurement) results from $\mathbf{b} = A\mathbf{f}$, where A is the measurement matrix or forward operator with $A_{i,k} = \langle\phi_k, h_i\rangle$ of size $M \times N$ and $\mathbf{f}$ the unknown pixel weight vector. Usually, number of non-zero coefficients in each row of A is $\mathcal{O}(\sqrt{N})$, so A sparse (so avoid computing pseudo-inverse, usually dense). Can view $A\mathbf{f}$ as forward projection, $A^T\mathbf{b}$ as backprojection. In presence of noise system can be inconsistent: Usually $\hat{\mathbf{f}} = \arg\min_{\mathbf{f}}\|\mathbf{b} - A\mathbf{f}\|_2^2$. If $\text{rank}(A) = N$, $\hat{\mathbf{f}} = (A^TA)^{-1}A^T\mathbf{b}$ unique. If $\text{rank}(A) = M < N$, take solution with minimum norm $\hat{\mathbf{f}} = A^T(AA^T)^{-1}\mathbf{b}$ (both interpreted as FBP). With this approach, handle constraints of scanning topology.

Kaczmarz's algorithm

Let $\{\mathbf{r}_0^T, \ldots, \mathbf{r}_{M-1}^T\}$ the rows of A. Hence $\langle\mathbf{f}, \mathbf{r}_i\rangle = b_i$, which define affine hyperplanes in $\mathbb{R}^N$. If solution to $A\mathbf{f} = \mathbf{b}$ exists and is unique, it's intersection of M affine hyperplanes. Kaczmarz's satisfies constraints iteratively starting from initial guess $\mathbf{f}^{(-1)}$. $\mathbf{f}^{(k)} = \mathbf{f}^{(k-1)} + \frac{b_n - \left\langle\mathbf{f}^{(k-1)}, \mathbf{r}_n\right\rangle}{\|\mathbf{r}_n\|_2^2}\mathbf{r}_n$ for $n = k \mod M$. That is, apply orthogonal projection of $\mathbf{f}^{(k-1)}$ onto $\mathcal{H}_n = \{\mathbf{f}^{(-1)} : \left\langle\mathbf{f}^{(k-1)}, \mathbf{r}_n\right\rangle = b_n\}$. If there's a unique solution, found with $k \to \infty$. Ordering of rows influence convergence rate (can order to increase angles between consecutive rows or select randomly with e.g. $p_n = \frac{\|\mathbf{r}_n\|_2^2}{\|A\|_F^2}$). Can incorporate constraints such as box constraint $(0 \leq f_i \leq 1)$ by projecting each step onto this convex set.

Cimmino's method

Instead of updating one row at a time, update once per sweep with average of all projections. $\mathbf{f}^{(k)} = \mathbf{f}^{(k-1)} + A^TW^{-1}(\mathbf{b} - A\mathbf{f}^{(k-1)})$ where $W = \text{diag}(M\|\mathbf{r}_0\|^2, \ldots, M\|\mathbf{r}_{M-1}\|^2)$.

Stochastic Processes

Covariance $\Sigma_x = \mathbb{E}[(x-\mu_x)(x-\mu_x)^*]$, which is PSD since $u^*\Sigma_x u = \mathbb{E}[u^*(x-\mu_x)(x-\mu_x)^*u] = \mathbb{E}[|u^*(x-\mu_x)|^2]$. $\langle x, y\rangle = \sum_{n=0}^{N-1} E[x_n y_n^*]$. Autocorrelation $a_{x,n,k} = \mathbb{E}[x_n x_{n-k}^*]$, crosscorrelation $c_{x,y,n,k} = E[x_n y_{n-k}^*]$. For i.i.d. process $a_{x,n,k} = |\mu_x|^2 + \sigma_x^2\delta_k$.

Stationary process: Joint distribution of $(x_{n_0}, \ldots, x_{n_L})$ and $((x_{n_0+k}, \ldots, x_{n_L+k}))$ are identical $\forall\{n_0, \ldots, n_L\} \subset \mathbb{Z}$, $\forall k \in \mathbb{Z}$ and L finite.

WSS: $\mu_{x,n} = \mu_x$, $a_{x,n,k} = a_{x,k}$ $n, k \in \mathbb{Z}$. x and y jointly WSS if each is WSS and $c_{x,y,n,k} = c_{x,y,k}$. With WSS, we have $\sigma_{x,n}^2 = a_{x,0} - |\mu_x|^2 = \sigma_x^2$, $a_{x,k} = a_{x,-k}^*$. With joint WSS, $c_{y,x,k} = c_{y,x,-k}^*$.

White noise: $\mu_{x,n} = 0, \sigma_{x,n}^2 = \sigma_x^2$ and $a_{x,k} = \sigma_x^2\delta_k$ (uncorrelated). Gaussian rv's uncorrelated iff independent, so white Gaussian process is i.i.d.

Independent vs. uncorrelated: If X, Y independent, $\mathbb{E}[XY] = \mathbb{E}[X]\,\mathbb{E}[Y]$, which implies that they are uncorrelated, i.e. $\mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])] = 0$. Converse is only true for zero-mean random variables and jointly Gaussian random variables.

Whitening or decorrelation: Processing that results in white noise process. Diagonalization of covariance matrix.

Filtering WSS processes: $y = x*h$ with h the impulse response of BIBO-stable LSI system (BIBO means L1 norm of h finite). $\mu_{y,n} = \mu_x H(e^{j0}) = \mu_y$, $a_{y,n,k} = \sum_{p\in\mathbb{Z}} a_{h,p}a_{xk-p} = a_{y,k}$ so if x WSS, y WSS and x and y jointly WSS. $C_{x,y}(e^{j\omega}) = H^*(e^{j\omega})A_x(e^{j\omega})$ and $C_{y,w}(e^{j\omega}) = H(e^{j\omega})A_x(e^{j\omega})$.

Power spectral density: x WSS. DTFT of its autocorrelation (FT in case of continuous time). $A_x(e^{j\omega}) = \sum_{k\in\mathbb{Z}} a_{x,k}e^{-j\omega k}$. $A_y(e^{j\omega}) = |H(e^{j\omega})|^2 A_x(e^{j\omega})$.

Power: $P_x = \frac{1}{2\pi}\int_{-\pi}^{\pi} A_x(e^{j\omega}) = a_{x,0}$.

Orthogonal Stochastic Processes: $c_{x,y,k,n} = \mathbb{E}[x_k y_{k-n}^*] = 0\ \forall k, n \in \mathbb{Z}$. For jointly WSS processes, equivalent to $C_{x,y}(e^{j\omega}) = 0\ \forall\omega \in \mathbb{R}$.

Wiener filtering: $\hat{x} = h*y$, $h = \arg\min_h \mathbb{E}[|e_n := x_n - \hat{x}_n|^2]$. Assume x, w uncorrelated, WSS, zero-mean. $\hat{x} \in \mathcal{S} := \overline{\text{span}}(\{y_{n-k}\}_{k\in\mathbb{Z}})$, so best estimator $e \perp \mathcal{S}$. This gives $H(e^{j\omega}) = \frac{C_{x,y}(e^{j\omega})}{A_y(e^{j\omega})}$.

Beamforming

In narrowband beamforming, output is $y_n = \sum_{k=0}^{M-1} h_k^* x_{k,n}$, M number of array elements. **Array response vector:** $\mathbf{a}(\theta) = [H_0(\omega)e^{-j\omega\tau_0(\theta)} \cdots H_{M-1}(\omega)e^{-j\omega\tau_{M-1}(\theta)}]^T$, H_k the transfer function of the $k-$th sensor, $\tau_k(\theta)$ the time needed for the wave to travel from reference point to sensor k. If sensors ideal ($H_k(\omega) = 1\ \forall k$) and sensor 0 taken as reference, $\mathbf{a}(\theta) = [1 \quad e^{-j\omega\tau_1(\theta)} \cdots (\omega)e^{-j\omega\tau_{M-1}(\theta)}]^T$. $x_n = \mathbf{a}(\theta)s_n + e_n$ For multiple sources $\mathbf{x} = [\mathbf{a}(\theta_0)\cdots\mathbf{a}(\theta_{N-1})][\mathbf{s}_0\cdots\mathbf{s}_{N-1}]^T + \mathbf{e} = A\mathbf{s} + \mathbf{e}$.

Uniform linear array: Equispaced sensors in same line. Under plane-wave assumption, $\tau_k = k\frac{d\sin\theta}{c}$, $\theta \in [-\frac{\pi}{2}, \frac{\pi}{2})$ (front-back ambiguity), c velocity of wave propagation, d distance between consecutive sensors.

Beamformer's response: $r(\theta) = \mathbf{h}^*\mathbf{a}(\theta)$. Beampattern is $|r(\theta)|^2$. If $\mathbf{a}(\theta_1) = \mathbf{a}(\theta_2)$ and $\theta_1 \neq \theta_2$, spatial aliasing, i.e. spacing between sensors is too large.

Signal model: For narrowband source with DOA θ and power σ_s^2, if there's no noise $\mathbf{x}_n = \mathbf{a}(\theta)s_n$, so $R_x = \sigma_s^2\mathbf{a}(\theta)\mathbf{a}^*(\theta)$.

Data independent beamforming

Phased array (delay-and-sum) beamformer: $\mathbf{h} = \mathbf{a}(\theta_0)$, where signal comes from single location θ_0. Can control beam width of main lobe and height of side lobes with tapering function $\mathbf{h} = T\mathbf{a}(\theta_0)$, with T diagonal $M \times M$ with tapering weights.

Response design: To find $\mathbf{h}$ giving response similar to $r_d(\theta)$, pose as overdetermined problem $\arg\min_{\mathbf{h}}\|A^*\mathbf{h} - \mathbf{r}_d\|_2^2$, $A = [\mathbf{a}(\theta_0)\cdots\mathbf{a}(\theta_{P-1})]$, $\mathbf{r}_d = [r_d(\theta_0)\cdots r_d(\theta_{P-1})]^*$. If A full rank $h = (AA^*)^{-1}A\mathbf{r}_d$.

White noise gain: Output power due to white noise of unit power, i.e. $\mathbf{h}^*\mathbf{h}$. If this is high, beamformer output could have poor SNR. Control it by low-rank approximation of A or solving previous problem with regularization $+\lambda\|\mathbf{h}\|_2^2$, which gives $h = (AA^* + \lambda I)^{-1}A\mathbf{r}_d$.

Data dependent beamforming

Output of beamformer approximates desired $\mathbf{y}_d$. Usually pose as $\arg\min_h \mathbb{E}[|\mathbf{y} - \mathbf{y}_d|^2]$, which gives $\mathbf{h} = R_x^{-1}r_{x,d}$.

LCMV: Constrain so that signals from desired directions have specified gain with $C^*\mathbf{h} = \mathbf{f}$. Solve $\min_{\mathbf{h}}\mathbf{h}^*R_x\mathbf{h}$ (minimize power at beamformer's output) subject to constraints. This gives $\mathbf{h} = R_x^{-1}C(C^*R_x^{-1}C)^{-1}\mathbf{f}$.

GSC: Transform LCMV to unconstrained. Decompose $\mathbf{h} = \mathbf{h}_0 - \mathbf{g}$, $\mathbf{h}_0 \in \mathcal{R}(C), \mathbf{g} \in \mathcal{N}(C^*)$. $\mathbf{h}_0 = C(C^*C)^{-1}\mathbf{f}$, the min norm solution satisfying constraints (data independent), $\mathbf{g} = C_n \mathbf{h}_n$, where C_n basis of $\mathcal{N}(C^*)$ (has no contribution of satisfaction of constraint bt allows to minimize objective). Solve $\min_{\mathbf{h}_n} (\mathbf{h}_0 - C_n \mathbf{h}_n)^* R_x (\mathbf{h}_0 - C_n \mathbf{h}_n)$, which gives $\mathbf{h}_n = (C_n^* R_x C_n)^{-1} C_n^* R_x \mathbf{h}_0$. The data-dependent beamformer vector has nulls in the directions of the constraints, which is ensured by the signal blocking matrix C_n.

Approximation Theory

Polynomials (finite interval)

Approximate $x(t)$ in finite interval $[a,b]$ by polynomial of order K $p_k(t) = \sum_{k=0}^K a_k t^k$. Approximation error $e_K(t) = x(t) - p_K(t)$, $t \in [a,b]$. Smooth approximation. Can approximate continuous functions arbitrarily well over finite intervals (Weierstrass theorem). Polynomials are infinitely differentiable. Approximating continuous functions with high degree polynomials tends to be problematic (e.g. at ending points of interval). Cannot approximate discontinuous functions or over infinite intervals well.

Least square minimization

Minimize $\|e_K\|_2^2 = \int_a^b |x(t) - p_K(t)|^2 dt$. Since $\mathcal{P}_K([a,b]) = \text{span}(\{1, t, \dots, t^K\}) \subset \mathcal{L}^2([a,b])$, solution (by projection theorem) is $p_K(t) = \sum_{k=0}^K \langle x, \phi_k \rangle \phi_k(t)$, $\{\phi_k\}_{k=0}^K$ orthonormal basis of $\mathcal{P}_K([a,b])$. These inner products with basis functions are not always easy to obtain (e.g. if we only have samples). **Gibbs phenomenon:** No matter how large K is, absolute error stays the same at the ripple near to the boundary.

Legendre polynomials: Orthonormal basis of $\mathcal{P}_K([-1,1])$. The Legendre polynomial of degree n, $L_n(t) = \frac{1}{2^n n!} \frac{d^n}{dt^n}(t^2-1)^n$, has n real distinct zeros in the interior of the interval $[-1,1]$

Lagrange interpolation

Observe function $x(t)$ at points (nodes) $t_0, \dots, t_K$. Constrain $p_K(t_i) = x(t_i) \forall i$. Leads to Vandermonde system, invertible iff $\{t_i\}$ distinct with solution $p_K(t) = \sum_{k=0}^K x(t_k) \prod i = 0, i \neq k^K \frac{t-t_i}{t_k - t_i}$.

Error: For $x(t) \in CK+1([a,b])$ and $\{t_i\}$ distinct, $|e_K(t)| \leq \frac{\prod_{k=0}^K |t-t_k|}{(K+1)!} \max_{\eta \in [a,b]} |x^{(K+1)}(\eta)|$. So error increases at the boundaries. Last term means error higher for wigglier functions.

Taylor series expansion

Assume $x(t) \in C^K([a,b])$ and find degree K polynomial with matching derivatives at $t_0 \in [a,b]$. Solution $p_K(t) = \sum_{k=0}^K x(t_k) \frac{(t-t_0)^k}{k!} x^{(k)}(t_0)$.

Error: For $x(t) \in CK+1([a,b])$, $|e_K(t)| \leq \frac{|t-t_0|^{K+1}}{(K+1)!} \max_{\eta \in [a,b]} |x^{(K+1)}(\eta)|$.

Minimax approximation

Minimize $\|e_K\|_\infty = \max_{t \in [a,b]} |e_K(t)|$. Non trivial since not Hilbert space. With polynomial of degree $\leq K$, minimax approximation p_K unique and determined by at least $K+2$ points $a \leq s_0 < s_1 < \dots < s_{K+1} \leq b$ for which $e_K(s_k) = \pm 1 (-1)^k \|e_K\|_\infty$ (Chebyshev equioscillation theorem). So expect to reach maximum error in $K+2$ points with alternating sign. Nearly solve by taking nodes minimizing maximum error for Lagrange interpolation. These optimal $K+1$ nodes given by roots of $K+1-$degree Chebyshev polynomial (its scaled version $2^{-K} T_{K+1}$ have the minimum ℓ_∞ norm among all $K+1-$degree polynomials): $t_k = \cos\left(\frac{2k+1}{2(K+1)}\pi\right)$.

Splines

Can approximate discontinuous functions well and over the entire real line. Spline of degree K with knots τ (countable strictly-increasing sequence) is a polynomial of degree $\leq K$ on $[\tau_n, \tau_{n+1})$ and its derivatives of order $0, \dots, K-1$ are continuous. $S_{K,\tau} \subset \mathcal{L}^2(\mathbb{R})$ is the spline space of degree K with knots τ. When τ evenly spaced and doubly infinite, spline space called uniform. For spline of degree K with $L+1$ knots, $K-1$ degrees of freedom. In practice for $K=3$, specifies derivatives at end-points or make sure 3rd derivative continuous at second and penultimate knots (not-a-knot condition).

B-Splines

Elementary B-Spline of degree 0 is $\beta^{(0)}(t) = 1$ for $t \in [-0.5, 0.5)$ (0 o.w.) and of degree K, $\beta^{(K)} = \beta^{(K-1)} * \beta^{(0)}$ (supported on $[-\frac{K+1}{2}, \frac{K+1}{2})$). Shifts of this are called B-splines of degree K. FT is $B^{(K)}(\omega) = \text{sinc}^{K+1}(\omega/2)$. For even K, $\beta^{(K)}$ smooth on $(z-0.5, z+0.5)$ and if odd on $(z, z+1)$ $z \in \mathbb{Z}$.
Causal B-Splines: Causal elementary B-Spline of degree 0 is $\beta_+^{(0)}(t) = 1$ for $t \in [0,1)$ (0 o.w.) and of degree K, $\beta_+^{(K)} = \beta^{(K)}(t - 0.5(K+1))$. $\beta_+^{(K)}$ is a generator of shift-invariant subspace $S_{K,\mathbb{Z}}$, in fact the one with shortest support. $\overline{\text{span}}(\{\beta^{(K)}(t-k)\}_{k\in\mathbb{Z}}) = S_{K,\mathbb{Z}}$ for odd K and $S_{K,\mathbb{Z}+0.5}$ for even K.
Differentiation: $\frac{d}{dt}x(t) = \sum_{k\in\mathbb{Z}} \alpha'_k \beta_+^{(K-1)}(t-k)$, $\alpha'_k = \alpha_k - \alpha_{k-1}$.
Integration: $\int_{-\infty}^\tau x(\tau) d\tau = \sum_{k\in\mathbb{Z}} \alpha_k^{(1)} \beta_+^{(K+1)}(t-k)$, $\alpha_k^{(1)} = \sum_{m=-\infty}^k \alpha_m$.
Canonical Dual Spline Basis: For dual basis of $\{\beta_+^{(1)}(t-k)\}_{k\in\mathbb{Z}}$, need $\left\langle \tilde{\beta}_+^{(1)}(t-i), \beta_+^{(1)}(t-k) \right\rangle = \delta_{i-k}$. For canonical dual, need $\tilde{\beta}_+^{(1)} \in S_{1,\mathbb{Z}}$. $\hat{x}(t) = \sum_{k\in\mathbb{Z}} \left\langle x(t), \tilde{\beta}_+^{(1)}(t-k) \right\rangle \beta_+^{(1)}(t-k)$. Dual spline of degree 1 has infinite support but decays exponentially.

Polynomial reproduction (Strang-Fix theorem)

If $\int_{-\infty}^\infty (1+|t|^K)|\phi(t)| dt < \infty$ for $k \in \mathbb{N}$ and ϕ function with FT Φ, following are equivalent:
(i) $p_K(t) = \sum_{k\in\mathbb{Z}} \alpha_k \phi(t-k)$ for p_K a polynomial of degree $\leq K$.
(ii) Φ and its first K derivatives satisfy $\Phi(0) \neq 0$ and $\Phi^{(k)}(2\phi l) = 0$ for $k = 1, \dots, K$, $l \in \mathbb{Z} \setminus \{0\}$.
Partition of unity (case of Strang-Fix): $\phi_1(t) = \sum_{n\in\mathbb{Z}} \phi(t-n) = 1$ (periodized version with period 1 of $\phi \in \mathcal{L}^1(\mathbb{R})$) iff $\Phi(2\phi k) = \delta_k$ $k \in \mathbb{Z}$.

Series Truncation

Given orthonormal expansion in infinite Hilbert space $x = \sum_{k\in\mathbb{Z}} c_k \phi_k$, $c_k = \langle x, \phi_k \rangle$. Cannot store all coefficients. **Linear approximation:** Retain coefficients with a priori fixed set of indices. Don't depend on x. Linear. Non optimal in error. **Nonlinear approximation:** Retain M largest coefficients in absolute value. Depend on x. Non linear. Store index of coefficients. Optimal in error. To proof latter select indices in $\mathcal{I}_M$ and apply Parseval: $\|x - \hat{x}\|^2 = \left\| \sum_{m\in\mathbb{Z}\setminus\mathcal{I}_M} c_m \phi_m \right\| = \sum_{m\in\mathbb{Z}\setminus\mathcal{I}_M} |c_m|^2$.

Uncertainty Principles

$\mu_t := \int_{-\infty}^\infty t \frac{|x(t)|^2}{\|x\|^2} dt$, $\Delta_t^2 := \int_{-\infty}^\infty (t-\mu_t)^2 \frac{|x(t)|^2}{\|x\|^2} dt$. $\mu_f := \int_{-\infty}^\infty \omega \frac{|X(\omega)|^2}{2\pi\|x\|^2} d\omega$, $\Delta_f^2 := \int_{-\infty}^\infty (\omega - \mu_t)^2 \frac{|X(\omega)|^2}{2\pi\|x\|^2} d\omega$.

Heisenberg principle

For $x \in \mathcal{L}^2(\mathbb{R})$ $\Delta_t^2 \Delta_f^2 \geq \frac{1}{4}$ with equality for Gaussian functions $x(t) = \gamma e^{-\alpha t^2}$, $\alpha > 0$. To proof, suppose w.l.o.g. $\mu_t = \mu_f = 0$, use Cauchy-Schwarz, Parseval, integration by parts with $(|x(t)|^2)$ and $\lim_{t\to\infty} tx^2(t) = 0$ (since $x \in \mathcal{L}^2(\mathbb{R})$, i.e. decays faster than $1/t$).
Shifts and scalings: Shifting changes μ in domain of shift. $\sqrt{a}x(at) \leftrightarrow X(\omega/a)/\sqrt{a}$ gives $\mu_t/a, \Delta_t/a, a\mu_f, a\Delta_f$.

Heisenberg principle for infinite sequences

Same definitions with discrete sum in time and DTFT (so integrals in $[-\pi,\pi]$). For $x \in \ell^2(\mathbb{Z})$ and $X(e^{j\pi}) = 0$ (necessary extra condition), $\Delta_n^2 \Delta_f^2 > \frac{1}{4}$ and lower bound cannot be achieved.
Shifts and scalings: Shifting in frequency gives $\mu_f + \omega_0$ if signal still $X(e^{j\pi}) = 0$ ("signal not splitted"). Upsample and postfilter gives $N\mu_n, N\Delta_n, \mu_f/N, \Delta_f/N$. Prefilter and downsample $\mu_n/N, \Delta_n/N, N\mu_f, N\Delta_f$ if $x \in \text{BL}[-\pi/N, \pi/N]$.

Uncertainty principle for finite-length sequences

$x_n \in \mathbb{C}^N$ and X_k its DFT with N_n and N_k nonzero coefficients respectively. $N_n N_k \geq N$. So cannot be sparse in both domains.