**Basic information: The Moore-Penrose inverse**
A linear inverse problem consists in solving an equation

$$Tx = y \tag{1}$$

for $x \in X$, where $X, Y$ ´ Hilbert spaces, $y \in Y$ and $T \in L(X, Y)$ (the space of linear continuous functions from $X$ to $Y$) are given.
The Moore-Penrose inverse (also generalized or pseudo inverse) is defined as follows. Set $\widetilde{T}\colon \ker(T)^\perp \to \mathrm{Rg}(T)$ (from the orthogonal of the kernel of $T$ onto the range of $T$) with $\widetilde{T}x = Tx$ for $x \in \ker(T)^\perp$, which is obviously a linear bijective function that possesses an inverse $\widetilde{T}^{-1}\colon \mathrm{Rg}(T) \to \ker(T)^\perp$. When decomposing $y \in \mathrm{Rg}(T) \oplus \mathrm{Rg}(T)^\perp := \mathrm{Dom}(T^\dagger)$ into $y = y_1 + y_2$ with $y_1 \in \mathrm{Rg}(T)$, $y_2 \in \mathrm{Rg}(T)^\perp$ one can define the Moore-Penrose inverse $T^\dagger\colon \mathrm{Dom}(T^\dagger) \to \ker(T)^\perp$ with $T^\dagger y = \widetilde{T}^{-1} y_1$. Note that neither is $\mathrm{Dom}(T^\dagger)$ necessarily the entirety of $Y$ nor is $T^\dagger$ necessarily continuous.
We recall a few basic results from the functional analysis: On a Hilbert space, any bounded sequence possesses a weakly convergent subsequence and conversely, a weakly convergent sequence is bounded (uniform boundedness principle). Moreover, note that in a Hilbert space $H$, the unit ball is compact if and only if $H$ is finite-dimensional. We call an operator $T \in L(X, Y)$ compact (and write $T \in \mathcal{K}(X, Y)$), if for any bounded set $A \subset X$ the image $T(A)$ is relatively compact (i.e., every sequence has a convergent subsequence).

**Example 1.1) [Compact operators]**
Let $X$ and $Y$ be Hilbert spaces. Show the following claims concerning compact operators:

**a)** Let $T \in L(X, Y)$. The function $T$ is compact if and only if $T$ is completely continuous, i.e., $x^n \rightharpoonup x$ in $X$ (weak convergence in $X$) implies $\|Tx^n - Tx\|_Y \to 0$.

**b)** Let $T \in \mathcal{K}(X, Y)$. $T$ has closed range if and only if $\mathrm{Rg}(T)$ is finite-dimensional.

**Remark.** *Compact operators are in a sense the closest analog to finite dimensional linear functions in Hilbert spaces, and in particular, offer easy insight due to being described by countable singular systems. Statement a) can be quite useful in verifying compactness of an operator. Statement b) on the other hand, as will be shown in the lecture, implies that inverse problems with compact operators are always ill-posed (do not possess continuous $T^\dagger$).*

**Example 1.2) [Computation of Moore-Penrose Inverse]**
Compute for the following $T \in L(X, Y)$ the Moore-Penrose inverse of $T$ and check whether $\mathrm{Dom}(T^\dagger) = Y$.

**a)** For Hilbert spaces $X, Y$ and fixed $u \in X$ with $u \neq 0$ and $v \in Y$ with $v \neq 0$ let $T\colon X \to Y$ be defined according to $Tx = v\langle u, x\rangle$ (where $\langle \cdot, \cdot \rangle$ denote the scalar product on $X$).

**b)** Let $T\colon L^2([0, 1]) \to L^2([0, 1])$ according to

$$[Tx](t) = \int_0^t x(s)\,\mathrm{d}s \qquad \text{for almost every } t \in [0, 1]. \tag{2}$$

**Remark.** *The Moore-Penrose inverse is a natural generalization of inverse functions in case said function is not bijective, but can lack desirable analytical properties, thus requiring regularization methods, see the lecture.*

**Example 1.3) [Deconvolution via Fourier transform]**
The Fourier transform $\mathcal{F}_{L^2}\colon L^2(\mathbb{R},\mathbb{C}) \to L^2(\mathbb{R},\mathbb{C})$ is a linear unitary (isometric and surjective) operator, while $\mathcal{F}_{L^1}\colon L^1(\mathbb{R},\mathbb{C}) \to \mathcal{C}_0(\mathbb{R},\mathbb{C})$ (the decaying continuous functions endowed with the supremum norm) is linear and continuous (they coincide on $L^1(\mathbb{R},\mathbb{C}) \cap L^2(\mathbb{R},\mathbb{C})$). Let $k \in L^1(\mathbb{R},\mathbb{C})$, $k \neq 0$ and define (the linear and continuous) $K\colon L^2(\mathbb{R},\mathbb{C}) \to L^2(\mathbb{R},\mathbb{C})$ with

$$[Kf](x) = [k * f](x) = \int_{\mathbb{R}} k(x-y)f(y)\,\mathrm{d}y \qquad \text{for almost ever } x \in \mathbb{R}. \tag{3}$$

Moreover, the Fourier transform is connected to the convolution in the following sense: for $u \in L^2(\mathbb{R},\mathbb{C})$ and $v \in L^1(\mathbb{R},\mathbb{C})$

$$\mathcal{F}_{L^2}[u * v] = \sqrt{2\pi}[\mathcal{F}_{L^2}u] \cdot [\mathcal{F}_{L^1}v], \tag{4}$$

where $\cdot$ represents pointwise multiplication.

- **a)** Characterize $\mathrm{Rg}(K)$ and $\ker(K)$ via the properties of their elements concerning the Fourier transform.
- **b)** Find $K^\dagger$ and conclude that $K^\dagger$ is not continuous.

**Hint.** *Many analytical properties such as orthogonality, convergence or closedness are invariant under application of unitary operators.*

**Remark.** *Deconvolution is a classical example of an ill-posed inverse problem, and as we have just seen, the problem is not difficult to solve analytically, however, the solution process is not continuous and is thus only of limited use. Nonetheless, the identity of $K^\dagger$ can be used to find regular approximations of $K^{-1}$, see Example 1.4.*

**Example 1.4) [Numerical deconvolution]**
In Exercise_1_4.m you find a Matlab/Octave code segment performing deconvolution in a discrete setting. Note that the scaling of the discrete Fourier transform is different from the analytical one, and the padding of data with zeros is necessary since the discrete analog to (4) holds for periodic convolution (and with the help of padding, it coincides with the classical convolution). You may adapt the code for your purposes to answer the following questions, in particular, create suitable figures to support your findings: Please also upload your code to Moodle.

- **a)** What role does *epsilon* play in this code, and why is it necessary? Try the code with different epsilon in the range of $10^{-16}$ to 1 as well as zero, what do you observe?
- **b)** Adapt the code by computing $W$ via Fourier transforms of $k$ and $y$ (which is not possible in real applications since $y$ is unknown). Check that this new $W$ coincides with the previous $W$ to a reasonable degree. Repeat your considerations for different *epsilon* with this new $W$. What do you observe and why does the behavior differ?