FUNCTIONAL ANALYSIS LECTURE NOTES:
ADJOINTS IN BANACH SPACES

CHRISTOPHER HEIL

1. Adjoints in Banach Spaces

If $H$, $K$ are Hilbert spaces and $A \in \mathcal{B}(H, K)$, then we know that there exists an adjoint operator $A^* \in \mathcal{B}(K, H)$, which is uniquely defined by the condition

$$\forall x \in H, \quad \forall y \in H, \quad \langle Ax, y \rangle_K = \langle x, A^*y \rangle_H. \quad (1.1)$$

Now we will consider the case where $X$, $Y$ are Banach spaces and $A \in \mathcal{B}(X, Y)$. We will see that there exists a unique adjoint $A^* \in \mathcal{B}(Y^*, X^*)$ which is defined by an equation that generalizes equation (1.1) to the setting of Banach spaces. Note, however, that while we have

$$A: X \to Y,$$

the adjoint will be a map

$$A^*: Y^* \to X^*.$$

In particular, unlike the Hilbert space case, we cannot consider compositions of $A$ with $A^*$.

**Exercise 1.1.** Let $X$, $Y$ be Banach spaces, and let $A \in \mathcal{B}(X, Y)$ be fixed. Show that there exists a unique operator $A^* \in \mathcal{B}(Y^*, X^*)$ that satisfies

$$\forall x \in X, \quad \forall \mu \in Y^*, \quad \langle Ax, \mu \rangle = \langle x, A^*\mu \rangle. \quad (1.2)$$

Further, show that

$$\|A^*\mu\| \leq \|A\| \|\mu\|, \quad \mu \in Y^*. \quad (1.3)$$

Hint: Fix $\mu \in Y^*$. Then define $A^*\mu: X \to \mathbb{F}$ so that (1.2) is satisfied, i.e., set $\langle x, A^*\mu \rangle = \langle Ax, \mu \rangle$ for $x \in X$. Show that $A^*\mu$ defined in this way satisfies (1.3), and conclude that $A^*\mu \in X^*$, and also that $A^*$ is bounded. Finally, show there is no other map $B: Y^* \to X^*$ that satisfies (1.2).

The preceding exercise defines $A^*$ as a map from $Y^*$ to $X^*$. Furthermore, equation (1.3) shows that $\|A^*\| \leq \|A\|$. The next exercise will show that equality holds.

**Exercise 1.2.** Let $X$, $Y$ be Banach spaces, and choose $A \in \mathcal{B}(X, Y)$. This exercise will show that $A^* \in \mathcal{B}(Y^*, X^*)$ satisfies

$$\|A^*\| = \|A\|.$$

Since we already know that $\|A^*\| \leq \|A\|$, we need only show the opposite inequality.

a. Show that if $Y = \{0\}$, then $A = 0$ and $A^* = 0$.  

1
b. Therefore, assume $Y \neq \{0\}$. Show that $Y^* \neq \{0\}$. Hint: Hahn–Banach.

c. Now choose any $x \in X$ with $\|x\| = 1$. Suppose that $Ax \neq 0$, and set $d = \|Ax\|$. Show that there exists a $\mu \in Y^*$ such that

$$\langle Ax, \mu \rangle = \|Ax\|, \quad \|\mu\| = 1. \quad (1.4)$$

d. Show that if $Ax = 0$, then there still exists a $\mu \in Y^*$ such that (1.4) holds. Hint: Use part b.

e. Show that $\|Ax\| \leq \|A^*\|$. Since this is true for every unit vector $x \in X$, conclude that $\|A\| \leq \|A^*\|$. 

**Exercise 1.3.** Let $X, Y$ be Banach spaces, and let $A, B \in \mathcal{B}(X, Y)$ be given. Prove the following properties of the adjoint maps.

a. $(A + B)^* = A^* + B^*$.

b. $(cA)^* = \bar{c}A^*$ for $c \in \mathbb{F}$.

c. If $X, Y$ are reflexive, then $A^{**} = A$.

**Remark 1.4.** Our choice of notation for linear functionals, i.e., using the sesquilinear form

$$\mu(x) = \langle x, \mu \rangle, \quad x \in X, \ \mu \in X^*,$$

that is linear in $x$ and anti-linear in $\mu$, allows us to create a definition of the adjoint that directly generalizes the definition of the adjoint of an operator on Hilbert spaces. In other words, if $H$ and $K$ are Hilbert spaces and $A \in \mathcal{B}(H, K)$ then there is no difference between the adjoint $A^*$ defined by considering $H$ and $K$ to be Hilbert spaces, and the adjoint $A^*$ defined by considering $H$ and $K$ to be Banach spaces.

However, we have seen that there are good reasons for considering other notations for linear functionals. To illustrate this, consider the specific example of the space $\ell^p$ with $1 \leq p < \infty$. We know that each element $y \in \ell^p$ determines a linear functional $\mu_y \in (\ell^p)^*$. However, there are at least three natural ways to define the identification of $y$ with $\mu_y$:

$$\langle x, \mu_y^1 \rangle = \sum_{k=1}^{\infty} x_k \bar{y}_k, \quad \langle x, \mu_y^2 \rangle = \sum_{k=1}^{\infty} x_k y_k, \quad \langle x, \mu_y^3 \rangle = \sum_{k=1}^{\infty} \bar{x}_k y_k.$$

**Notation 1.** This notation has the advantage of directly extending the inner product on the Hilbert space $\ell^2$. That is, if $p = 2$, then $\langle x, \mu_y^1 \rangle$ is exactly equal to the inner product $\langle x, y \rangle$.

With this notation, $(\ell^p)^*$ is the normed linear space containing all linear functionals $\mu_y^1$ for $y \in \ell^p$. The disadvantage is that the mapping $y \mapsto \mu_y^1$ is antilinear, so $(\ell^p)^*$ and $\ell^p$ are antilinearly isomorphic under this identification of $y$ with $\mu_y^1$.

A great advantage of this notation becomes apparent when consider unitary operators on Hilbert spaces and especially notation for the Fourier transform and other distributional operations.
Notation 2. This notation has the very nice property that \((\ell^p)^*\) is the normed linear space containing all the linear functionals \(\mu^2_y\) for \(y \in \ell^p\), and that the mapping \(y \mapsto \mu^2_y\) is linear. Thus \((\ell^p)^*\) and \(\ell^p\) are linearly isomorphic under this identification of \(y\) with \(\mu^2_y\).

The disadvantage of this notation is that for \(p = 2\), we do not have that the inner product \(\langle x, y \rangle\) equals \(\langle x, \mu^2_y \rangle\). One consequence of this is that it makes a difference whether we define adjoints of operators on the Hilbert space \(\ell^2\) using the Hilbert space definition and the Banach space definition (compare Example 1.5 below).

Another disadvantage involves unitary operators on the Hilbert space \(\ell^2\). If \(A : \ell^2 \to \ell^2\) is unitary, then we have that \(\langle x, y \rangle = \langle Ax, Ay \rangle\) for all \(x, y \in \ell^2\). However, \(\langle Ax, \mu^2Ay \rangle\) is not the same as \(\langle Ax, Ay \rangle\), and so notational difficulties arise. In particular, this results in very unpleasant notation when we consider extensions of operations to distributional settings.

Notation 3. This notation has the advantage that the mapping \(y \mapsto \mu^3_y\) is a linear isomorphism of \(\ell^p\) onto \((\ell^p)^*\). One disadvantage is that the elements \(\mu^3_y\) of \((\ell^p)^*\) are antilinear functionals on \(\ell^p\), i.e., \((\ell^p)^*\) is the space of bounded antilinear functionals on \(\ell^p\). There is also the issue of incompatibility with the inner product on \(\ell^2\).

Example 1.5. Let us consider \(H = \mathbb{C}^n\) and \(K = \mathbb{C}^m\).

Notation 1. The inner product on \(\mathbb{C}^d\) is the sesquilinear form given by

\[
\langle x, y \rangle_{\mathbb{C}^d} = x \cdot y = x^T \bar{y} = \sum_{k=1}^d x_k \bar{y}_k.
\]

If \(A\) is an \(m \times n\) matrix, then \(x \mapsto Ax\) is a linear map \(\mathbb{C}^n \to \mathbb{C}^m\). The adjoint map \(A^* : \mathbb{C}^m \to \mathbb{C}^n\) is determined by the condition that

\[
\langle x, A^*y \rangle_{\mathbb{C}^n} = \langle Ax, y \rangle_{\mathbb{C}^m}, \quad x \in \mathbb{C}^m, \; y \in \mathbb{C}^m.
\]

Let \(A^H = \overline{A^T}\) be the Hermitian, or conjugate-transpose, of \(A\). Then we have

\[
x \cdot A^*y = Ax \cdot y = (Ax)^T \bar{y} = x^T A^T \bar{y} = x^T \overline{A^T y} = \langle x, A^H y \rangle.
\]

Hence \(A^H = A^*\) is the adjoint of \(A\). With this notation, we have that

\[
(\alpha A + \beta B)^* = \bar{\alpha} A^* + \bar{\beta} B^*.
\]

Notation 2. Now let us consider the definition of the adjoint using Notation 2. We change notation and define a bilinear form

\[
\langle x, y \rangle_{\mathbb{C}^d} = x^T y = \sum_{k=1}^d x_k y_k.
\]

The adjoint of an \(m \times n\) matrix \(A\) is now the matrix \(A^*\) that satisfies

\[
\langle x, A^*y \rangle_{\mathbb{C}^n} = \langle Ax, y \rangle_{\mathbb{C}^m}, \quad x \in \mathbb{C}^n, \; y \in \mathbb{C}^m,
\]

with respect to the bilinear form given in (1.5). We then have

\[
x \cdot A^*y = Ax \cdot y = (Ax)^T y = x^T A^T y = \langle x, A^T y \rangle_{\mathbb{C}^n}.
\]
Hence, with respect to this bilinear form, the adjoint of \( A \) is \( A^* = A^T \), the transpose of \( A \) instead of the Hermitian of \( A \). With this notation, we have that

\[
(\alpha A + \beta B)^* = \alpha A^* + \beta B^*.
\]

In the end, we simply have to make a choice of whether the form \( \langle x, \mu \rangle \) on \( X \times X^* \) is sesquilinear or bilinear in \( \mu \). In the former case, if \( X \) is a Hilbert space then the definition of the adjoint coincides with the definition of adjoint corresponding to the inner product on \( X \), whereas if we choose the form to be bilinear then we obtain a different definition of the adjoint compared to the Hilbert space definition. However, this is really just a technical problem, not an ideological obstruction—it just requires us to be careful about our notation.

**Exercise 1.6.** Let \( X, Y \) be Banach spaces. Assume that \( T: X \to Y \) is an isometric isomorphism (a linear isometric bijection). Prove that the adjoint \( T^*: Y^* \to X^* \) is an isometric isomorphism.