From last time:

$$H(P) = -Tr(P\log(P))$$

$$D(P||Q) = Tr(P\log(P)) - Tr(P\log(Q))$$

Klein's inequality

$$\rho, \sigma \in D() \rightarrow D(\rho||\sigma \geq 0 \quad \text{with equality if and only if } \rho = \sigma$$

Suppose $\rho \in D()$ and that $dim() = n$. It holds that $H(\rho) \leq \log(n)$ with equality if and ontly if $\rho = \nVdash/n$.

We know that $0 \leq D(\rho||\mathbb{1}/n) = -H(\rho) - Tr(\rho\log(\mathbb{1}/n))$.

In general, if you want to know $\log(\alpha P) = \sum\limits_k \log(\alpha\lambda_k(P))x_k x_k^*$ and

$$\sum_k (\log(\alpha) + \log(\lambda_k(P)))x_k x_k^* = \log(P) + \log(\alpha)\mathbb{1}$$

Plugging this expression in to our earlier expression for $D(\rho||\mathbb{1}/n)$.

$$\begin{aligned} D(\rho||\mathbb{1}/n) &= -H(\rho) - Tr(\rho\log(\mathbb{1}) + \log(\frac{1}{n})\mathbb{1}) \\ &= -H(\rho) - \log(1/n) \\ &= -H(\rho) + \log(n) \end{aligned}$$

Thus:

$$H(\rho) \leq \log(n)$$

Let's demonstrate the subadditivity of the quantum entropy. If X is a register in some state $\rho \in \mathcal{X}$, we write $H(\mathcal{X})$ to mean $H(\rho)$. Consider two registers X and Y in the state $\rho \in D(\mathcal{X} \otimes D(Y)$. $H(X) = H(Tr_y(\rho))$ and $H(X,Y) =$ "$H((X,Y))''$.

Proposition: for any state $\rho \in D(\mathcal{X} \otimes \mathcal{Y})$ it holds that $H(X) + H(Y) \geq H(X,Y)$. This is a statement of subadditivity.

More notation: If $\rho \in D(\mathcal{X} \otimes \mathcal{Y})$ then we write $\rho[X] = Tr_y\rho$ and $\rho[Y] = Tr_x\rho$. This is often shown as ρ^X or ρ_X.

Subaddivity means that

$$H(\rho[X]) + H(\rho[Y]) \geq H(\rho).$$

It also means that

$$0 \leq D(\rho||\rho[X] \otimes \rho[Y])$$

Suppose that you have $\log(P \otimes Q)$. What should that be? If they were scalars we'd get the sum. But, with a tensor product we get $\log(P \otimes Q) = \log(P) \otimes \nVdash + \nVdash \otimes \log(Q)$. This generalizes the log for tensor products of operators.

Going back to subadditivity:

$$
\begin{aligned}
0 &\leq D(\rho||\rho[X] \otimes \rho[Y]) \\
&= -H(\rho) - Tr(\rho \log(\rho[X] \otimes \rho[Y])) \\
&= -H(\rho) - Tr(\rho(\log(\rho[X]) \otimes \mathbb{1})) - Tr(\rho(\mathbb{1} \otimes \log(\rho[Y]))) \\
&= -H(\rho) - Tr(\rho[X] \log(\rho[X])) \\
&= -H(\rho) + H(\rho[X]) + H(\rho[Y])
\end{aligned}
$$

Let's next demonstrate the concavity of the von Neumann entropy. Consider $P, Q \in Pd(\mathcal{X})$ (they are positive definite).

$$
D\left(\begin{pmatrix} P & 0 \\ 0 & Q \end{pmatrix} || \begin{pmatrix} \frac{P+Q}{2} & 0 \\ 0 & \frac{P+Q}{2} \end{pmatrix}\right) \tag{1}
$$

$$
= Tr(P \log(P)) - Tr(P \log(\frac{P+Q}{2})) \tag{2}
$$

$$
= Tr(Q \log(Q)) - Tr(Q \log(\frac{P+Q}{2})) \tag{3}
$$

$$
= Tr(P \log(P)) + Tr(Q \log(Q)) - 2Tr(\frac{P+Q}{2} \log(\frac{P+Q}{2})) \tag{4}
$$

$$
0 \leq \frac{1}{2} Tr(P \log(P)) + \frac{1}{2} Tr(Q \log(Q)) - Tr(\frac{P+Q}{2} \log(\frac{P+Q}{2})) \tag{5}
$$

$$
= H(\frac{P+Q}{2} - \frac{1}{2} H(P) - \frac{1}{2} H(Q) \tag{6}
$$

So,

$$
\frac{1}{2}\left(H(P) + H(Q)\right) \leq H(\frac{P+Q}{2})
$$

This states that the entropy is "midpoint concave". Since the entropy is continuous, though, and we know that it's midpoint concave we can say that it's concave. For any continuous midpoint concave function it is concave.

Let's consider strong subadditivity. Let X,Y and Z be registers and $\rho \in D(\mathcal{X} \otimes \mathcal{Y}\mathcal{Z})$ be any state. It can be shown that $H(X,Y,Z) + H(Z) \leq H(X,Z) + H(Y,Z)$.

Let's prove this (it will take a while)

$$
\begin{aligned}
&\rho_0, \rho_1, \sigma_0, \sigma_1 \in D(\mathcal{X}) \\
&\lambda \in [0,1] \\
D(\lambda\rho_0 + (1-\lambda)\rho_1 || \lambda\sigma_0 + (1-\lambda)\sigma_1) &\leq \lambda D(\rho_0||\sigma_0) + (1-\lambda) D(\rho_1||\sigma_1)
\end{aligned}
$$

Step 1 for the proof is going to be to get rid of the logarithms. So consider $\rho, \sigma \in D(\mathcal{X})$. Let's define a funnction $f_{\rho,\sigma} : \mathbb{R} \mapsto \mathbb{R}$ as $f_{\rho,\sigma}(\alpha) = Tr(\sigma^\alpha \rho^{1-\alpha})$.

This function is continous and, in fact, differentiable. So, let's calculate its derivative.

$$f^{'}(\alpha) = Tr(\sigma^{\alpha}\rho^{1-\alpha}(\ln(\sigma) - \ln(\rho)))$$

This function is interesting for one reason because of its value at "$\alpha = 0''$: f'(0) =Tr($\rho\ln(\sigma) - \rho\ln(\rho)) = -\ln(2)D(\rho||\sigma)$. This new function is nice because it has no logarithms but it has relationships to things we want to calculated.

There is another way that we can think about this function:

$$D(\rho||\sigma) = \frac{1}{\ln(2)} \lim_{\alpha^{+}\to 0} \frac{Tr\left(\sigma^{\alpha}\rho^{1-\alpha}\right) - 1}{\alpha}$$

Imagine we could prove this theorem. This theorem is known as the "Lieb concavity theorem".

$$\rho_0, \rho_1, \sigma_0, \sigma_1 \in D(\mathcal{X}) \tag{7}$$

$$\alpha, \lambda \in [0, 1] \tag{8}$$

$$Tr\left((\lambda\sigma_0 + (1-)\,\sigma_1)^{\alpha}\left(\lambda\rho_0 + (1-\lambda)\rho_1\right)^{1-\alpha}\right) \geq \lambda Tr\left(\sigma_0^{\alpha}\rho_0^{1-\alpha}\right) + (1-\lambda)\,Tr(\sigma_1^{\alpha}\rho_1^{1-\alpha}) \tag{9}$$

Using this result:

$$\begin{aligned}
&(\lambda\rho_0 + (1-\lambda)\rho_1||\lambda\sigma_0 + (1-\lambda)\sigma_1) \\
&= -\frac{1}{\ln(2)} \lim_{\alpha^{+}\to 0} \left(\frac{Tr((\lambda\sigma_0 + (1-\lambda)\sigma_1)^{\alpha}(\lambda\rho_0 + (1-\lambda)\rho_1)^{1-\alpha}) - 1}{\alpha}\right) \\
&\leq \frac{1}{\ln 2} \lim_{\alpha^{+}\to 0} \frac{1}{\alpha}\left(\lambda Tr(\sigma_0^{\alpha}\rho_0^{1-\alpha}) + (1-\lambda)Tr(\sigma_1^{\alpha}\rho_1^{1-\alpha}) - \lambda(1-\lambda)\right)\big) \\
&= \lambda D(\rho_0||\sigma_0) + (1-\lambda)D(\rho_1||\sigma_1)
\end{aligned}$$

QED

Consider another theorem: $A_0, A_1 \in Pd(\mathcal{X})$ and $B_0, B_1 \in Pd(\mathcal{Y})$ and $\alpha \in [0, 1]$. We can show that $(A_0 + A_1)^{\alpha} \otimes (B_0 + B_1)^{1-\alpha} \geq A_0^{\alpha} \otimes B_0^{1-\alpha} + A_1^{\alpha} \otimes B_1^{1-\alpha}$. To get Lieb, take $A_0 = \lambda\sigma_0, A_1 = (1-\lambda)\sigma_1$ and $B_0 = \lambda\rho_0^T, B_1 = (1-\lambda)\rho_1^T$.

Writing out all of this yields: $(\lambda\sigma_0 + (1-\lambda)\sigma_1)^{\alpha} \otimes (\lambda\rho_0^T + (1-\lambda)\rho_1^T)^{1-\alpha} \geq \lambda\sigma_0^{\alpha} \otimes (\rho_0^T)^{1-\alpha} + (1-\lambda)\sigma_1^{\alpha} \otimes (\rho_1^T)^{1-\alpha}$. Now, we make a "sandwich" where the "bread" is $vec(\mathbb{1})$.

$$vec(\mathbb{1}^{*}\sigma_0^{\alpha} \otimes (\rho_0^T)^{1-\alpha} vec(\mathbb{1}) + (1-\lambda)vec\mathbb{1}^{*}\sigma_1^{\alpha} \otimes (\rho_1^T)^{1-\alpha} vec(\mathbb{1})$$

Now we use the fact that $vec(\mathbb{1}^{*}(X \otimes Y)vec(\mathbb{1}) = Tr(XY)$. Using this fact we can show that the previous expression implies Lieb's concavity theorem.

Now let us prove Ando's version of Lieb's concavity theorem. We will use the following lemma: Suppose that we have $P, Q \in Pd(\mathcal{X})$. We'll also assume that $[P, Q] = 0$. Let's also assume that $H \in Herm(\mathcal{X})$. If

$$\begin{pmatrix} P & H \\ H & Q \end{pmatrix} \geq 0$$

then $H \leq \sqrt{P}\sqrt{Q}$.

From before we had the condition that if and only if $\begin{pmatrix} P & X \\ X^* & Q \end{pmatrix} \geq 0$ then $X = \sqrt{P}X\sqrt{Q}$ for $||K|| \leq 1$. So, $H = \sqrt{P}K\sqrt{Q}$ for $||K|| \leq 1$. So $||P^{-\frac{1}{2}}HQ^{-\frac{1}{2}}|| \leq 1$.

If λ is an eigenvalue of $P^{-\frac{1}{2}}HQ^{-\frac{1}{2}}$, then $|\lambda| \leq 1$.

In general, if $X, Y \in L(\mathcal{X})$, then the eigenvalues of XY and YX must always be the same.

Considering more operator stamements: $\lambda_1 \left(P^{-1/4}Q^{-1/4}HQ^{-1/4}P^{-1/4}\right) \leq 1$. This implies that $P^{-1/4}Q^{-1/4}HQ^{-1/4}P^{-1/4} \leq \mathbb{1}$. So, $H \leq \sqrt{P}\sqrt{Q}$. Bhatia's Matrix Analysis has a lot of these relatonships. It's a good reference text.