Zusammenfassung der 14. Vorlesung (19.11.2010)

Zum Beweis des Satzes von Elliot Lieb war die Ungleichung

$$(S_1^{\frac{1}{2}}S_2^{\frac{1}{2}} + T_1^{\frac{1}{2}}T_2^{\frac{1}{2}}) \leq R_1^{\frac{1}{2}}R_2^{\frac{1}{2}}$$

hergeleitet worden. Die allgemeinere Ungleichung

$$(S_1^{\mu}S_2^{(1-\mu)} + T_1^{\mu}T_2^{(1-\mu)}) \leq R_1^{\mu}R_2^{(1-\mu)}$$

ist nach Voraussetzung trtivialer Weise für $\mu \in \{0,1\}$ erfüllt und nun auch für $\mu = \frac{1}{2}$. Wegen der Vertauschbarkeit der S_i und der der T_i gilt nun für $\nu, \sigma \in \{0, \frac{1}{2}, 1\}$ auch

$$\begin{aligned}
& S_1^{\frac{\nu+\sigma}{2}}S_2^{(1-\frac{\nu+\sigma}{2})} + T_1^{\frac{\nu+\sigma}{2}}T_2^{(1-\frac{\nu+\sigma}{2})} \\
= \quad & (S_1^{\nu}S_2^{(1-\nu)}S_1^{\sigma}S_2^{(1-\sigma)})^{\frac{1}{2}} + (T_1^{\nu}T_2^{(1-\nu)}T_1^{\sigma}T_2^{(1-\sigma)})^{\frac{1}{2}} \\
\leq \quad & (S_1^{\nu}S_2^{(1-\nu)} + T_1^{\nu}T_2^{(1-\nu)})^{\frac{1}{2}}(S_1^{\sigma}S_2^{(1-\sigma)} + T_1^{\sigma}T_2^{(1-\sigma)})^{\frac{1}{2}} \\
\leq \quad & (R_1^{\nu}R_2^{(1-\nu)})^{\frac{1}{2}}(R_1^{\sigma}R_2^{(1-\sigma)})^{\frac{1}{2}} = R_1^{\frac{\nu+\sigma}{2}}R_2^{(1-\frac{\nu+\sigma}{2})},
\end{aligned}$$

d. h. $\mu = \frac{\nu+\sigma}{2} \in \{0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1\}$ ist möglich. Wie eine Fortsetzung dieses Verfahrens zeigt, ist die allgemeinere Ungleichung auch für alle $\frac{K}{2^N}$, $N \in \mathbf{N}$, $(K = 0, 1, 2, \ldots, 2^N)$ erfüllt, die eine in $[0,1]$ dichte Menge bilden. Da für posittive Operatoren A die Funktion A^t auf $[0,1]$ stetig ist, gilt daher

$$S_1^tS_2^{(1-t)} + T_1^tT_2^{(1-t)} \leq R_1^tR_2^{(1-t)} \quad \text{für} \quad t \in [0,1].$$

Wir betrachten nun den M^2-dimensionalen linearen Raum der komplexen $M \times M$ Matrizen Z. Diese Matrizen bilden einen Hilbertraum mit dem inneren Produkt $\ll A, B \gg := \mathrm{tr}(A^+B)$, den wir mit $\mathcal{T}^2$ bezeichnen wollen. Man beachte, dass in $\mathcal{T}^2$ die Vektoren auch als Operatoren wirken. Seien nun $A_1, B_1 \geq 0$ und $\lambda \in [0,1]$. Dann betrachte man die Operatoren auf $\mathcal{T}^2$

$$S_1 : X \mapsto \lambda A_1 X, \qquad S_2 : X \mapsto \lambda X B_1.$$

Für alle $X, Y \in \mathcal{T}^2$ gelten

$$\begin{aligned}
&\ll Y, S_1X \gg = \lambda\mathrm{tr}(Y^+A_1X) = \lambda\mathrm{tr}((A_1Y)^+X) = \ll S_1Y, X \gg, \\
&\ll Y, S_2X \gg = \lambda\mathrm{tr}(Y^+XB_1) = \lambda\mathrm{tr}((YB_1)^+X) = \ll S_2Y, X \gg,
\end{aligned}$$

und

$$\ll X, S_1 X \gg = \lambda \mathrm{tr}(X^+ A_1 X) \geq 0, \qquad \ll X; S_2 X \gg = \lambda \mathrm{tr}(X^+ X B_1) \geq 0.$$

Die Operatoren S_1, S_2 sind also selbstadjungiert und positiv, $S_1, S_2 \geq 0$ auf $\mathcal{T}^2$. Überdies gilt $[S_1, S_2] = 0$, denn

$$S_1 S_2 = \lambda S_1 (X B_1) = \lambda^2 A_1 (X B_1) = \lambda^2 (A_1 X) B_1 = \lambda S_2 (A_1 X) = S_2 S_1,$$

so dass diese Operatoren auch vertauschbar sind. Für $A_2, B_2 \geq 0$ haben die Operatoren

$$T_1 : X \mapsto (1-\lambda) A_2 X, \qquad T_2 : X \mapsto (1-\lambda) X B_2.$$

die analogen Eigenschaften: $T_1, T_2 \geq 0$, $[T_1, T_2] = 0$. Schließlich gilt für

$$\begin{aligned} R_1 &: X \mapsto (S_1 + T_1) X = (\lambda A_1 + (1-\lambda) A_2) X =: FX \\ R_2 &: X \mapsto (S_2 + T_2) X = X(\lambda B_1 + (1-\lambda) B_2) =: XG \end{aligned}$$

auch

$$R_1 R_2 = F(XG) = (FX)G = R_2 R_1,$$

also $[R_!, R_2]$. R_i, S_i und T_i, $(i = 1, 2)$ erfüllen damit alle Voraussetzungen für die Gültigkeit von

$$S_1^t S_1^{(1-t)_2 + T^t} T_2^{(1-t)} \leq R_1^t R_2^{(1-t)} \quad \text{für} \quad t \in [0,1].$$

in $\mathcal{T}^2$. Für beliebiges $X \in \mathcal{T}^2$ gilt also

$$\ll X, (S_1^t S_2^{(1-t)} + T_1^t T_2^{(1-t)}) X \gg \leq \ll X, R_1^t R_2^{(1-t)} X \gg \quad \text{für} \quad t \in [0,1].$$

Einsetzen der expliziten Definitionen von R_i, S_i und T_i liefert für die linke Seite der Ungleichung

$$\begin{aligned} & \lambda \mathrm{tr}(X^+ A_1^t X B_1^{(1-t)}) + (1-\lambda) \mathrm{tr}(X^+ A_2^t X B_2^{(1-t)}) \\ = \quad & \lambda f(A_1, B_1) + (1-\lambda) f(A_2, B_2), \end{aligned}$$

und für die rechte Seite

$$\begin{aligned} & \mathrm{tr}(X^+ (\lambda A_1 + (1-\lambda) A_2)^t X (\lambda B_1 + (1-\lambda) B_2)^{(1-t)} \\ = \quad & f(\lambda A_1 + (1-\lambda) A_2, \lambda B_1 + (1-\lambda) B_2). \end{aligned}$$

Damit ist

$$\lambda f(A_1,B_1)+(1-\lambda)f(A_2,B_2)\leq f(\lambda A_1+(1-\lambda)A_2,\lambda B_1+(1-\lambda)B_2),$$

was zu beweisen war.

Auf dem Satz von Elliot Lieb beruht der Beweis des folgenden Satzes:

Satz: $S(\rho\|\sigma)$ ist in $(\rho\|\sigma)$ kombiniert konvex.

Beweis: Dazu betrachtet man $\mathcal{H}\oplus\mathcal{H}\cong\mathbf{C}^2\otimes\mathcal{H}$ und stellt fest, dass mit

$$A=\begin{pmatrix}\rho & 0\\ 0 & \sigma\end{pmatrix},\qquad X=\begin{pmatrix}0 & 0\\ 1 & 0\end{pmatrix}$$

die relative Entropie in der Form

$$S(\rho\|\sigma)=\mathrm{tr}(X^+X(\log_2 A)A-X^+(\log_2 A)XA)$$

geschrieben werden kann. Aus

$$A^t=2^{t\log_2 A}=e^{t\ln 2\log_2 A}\Rightarrow\frac{d}{dt}A^t=(\ln 2\log_2 A)A^t,$$

also ist

$$\frac{d}{dt}A^t|_0=\ln 2\log_2 A\quad\text{und}\quad\frac{d}{dt}A^{(1-t)}|_0=-(\ln 2\log_2 A)A,$$

und damit

$$\begin{aligned}S(\rho\|\sigma)&=-\frac{1}{\ln 2}\frac{d}{dt}\mathrm{tr}(X^+A^tXA^{(1-t)})|_0\\&=-\frac{1}{\ln 2}\frac{d}{dt}\mathrm{tr}(X^+A^tXA^{(1-t)}-X^+XA)|_0.\end{aligned}$$

Nach dem Satz von Elliot Lieb ist $f(A,B)=\mathrm{tr}(X^+A^tXB^{(1-t)})$ kombiniert konkav in (A,B) und damit ist $\mathrm{tr}(X^+A^tXA^{(1-t)})$ konkav in A. Ebenso ist $g_t(A):=\frac{1}{\ln 2}\mathrm{tr}(X^+A^tXA^{(1-t)}-X^+XA)$ in A konkav, hat aber überdies den Vorteil, dass $g_0(A)=0$ ist. Es gilt deshalb für $h>0$ und $\lambda\in[0,1]$

$$\frac{1}{h}(g_h(\lambda A_1+(1-\lambda)A_2)\geq\frac{1}{h}(\lambda g_h(A_1)+(1-\lambda)g_h(A_2)),$$

und im Limes $h \downarrow 0$ ergibt sich wegen $g_0(A) = 0$

$$\frac{d}{dt}(g_t(\lambda A_1 + (1-\lambda)A_2)|_0 \geq \lambda \frac{d}{dt} g_t(A_1)|_0 + (1-\lambda)\frac{d}{dt} g_t(A_2)_0.$$

Weil $S(\rho\|\sigma) = -\frac{d}{dt}(g_t(A)|_0$ ist, folgt die Behauptung des zu beweisenden Satzes:

$$S(\lambda\rho_1 + (1-\lambda)\rho_2\|\lambda\sigma_1 + (1-\lambda)\sigma_2) \leq \lambda S(\rho_1\|\sigma_1) + (1-\lambda)S(\rho_2\|\sigma_2).$$