\[ I(X;Y) = h(X) - h(X|Y) \]
\[ = - \int f(x) \log f(x) \, dx + \int f(x,y) \log f(x|y) \, dx \, dy \]
\[ = \int f(x,y) \log \frac{f(x,y)}{f(x)} \, dx \, dy \]
\[ = \int f(x,y) \log \frac{f(x)}{f(x)f(y)} \, dx \, dy \]

also showing interchangeability of \( X \) and \( Y \).

Def. 2.5.6. The relative entropy or Kullback-Leibler divergence between two densities \( f \) and \( g \) is defined as
\[ D(f \| g) = \int f(x) \log \frac{f(x)}{g(x)} \, dx. \]

From (g) it follows that
\[ I(X;Y) = D(f(x,y) \| f(x)f(y)) \] (**)  

Th. 2.5.7. \( D(f \| g) \geq 0 \)

with equality iff \( f = g \) (almost everywhere, a.e.)

Proof. Let \( S = \{ x \mid f(x) > 0 \} \) the support of \( f \). Then
\[ -D(f \| g) = \int_S f(x) \log \frac{g}{f} \]
\[ \leq \log \int_S \frac{g}{f} \] (Jensen's inequality, La 2.1.6)
\[ = \log \int g \leq \log 1 = 0 \] (concave: \( f \) concave : \( \int f(x) \leq f(EX) \))

Equality holds iff \( f = g \) a.e.
**Corollary 2.5.8.**

a) $I(X;Y) \geq 0$ with equality iff $X$ and $Y$ are independent.

b) $h(X|Y) \leq h(X)$ with equality iff $X, Y$ are independent.

c) $-\sum f \log f \leq -\sum g \log g$

**Proof.**
a) follows from (**)**
b) $I(X;Y) = h(X) - h(X|Y) \geq 0$ by a)
c) by Def. of $D(f||g)$.

**Th. 2.5.9.** (Chain rule for diff. entropy)

$h(X_1, \ldots, X_n) = \sum_{i=1}^{n} h(X_i; X_{i+1} \ldots X_n)$

**Proof.** From the definition it follows that $h(X, Y) = h(X) + h(Y|X)$

This implies

$h(X_1, \ldots, X_i) = h(X_1, \ldots, X_{i-1}) + h(X_i; X_{i+1} \ldots X_n)$

The assertion follows by induction.

**Corollary 2.5.10.**

$h(X_1, \ldots, X_n) \leq \sum_{i=1}^{n} h(X_i)$

with equality iff $X_1, \ldots, X_n$ are stoch. independent.
Thu, 25.11. Let $X \in \mathbb{R}^n$ with density $f(x)$, $A \in \mathbb{R}^{m \times n}$ or full rank, $b \in \mathbb{R}^n$. Then

$$\ln(Ax + b) = \ln(X) + \log|A|.$$ 

(Note that $|A| = 1 \det(A)$)

Proof. If $X \sim f(x)$, then $Y = AX + b \sim \frac{1}{|A|} f(A^{-1}(y - b))$, $x, y \in \mathbb{R}^n$. 

$$-\int \frac{1}{|A|} f(A^{-1}(y - b)) \log \left( \frac{1}{|A|} f(A^{-1}(y - b)) \right) dy$$

$$= -\log \frac{1}{|A|} - \int \frac{1}{|A|} f(A^{-1}y) \log(f(A^{-1}y)) dy$$

$$= -\log \frac{1}{|A|} - \int f(x) \log f(x) dx$$

$$= \log |A| + \ln(X).$$
Th. 2.5.12. Let $X \in \mathbb{R}^n$ abs. cont. with density $p(x)$ and $\text{Cov}(X) = C$, $C$ pos. definite. Then
\[ h(X) \leq \frac{1}{2} \ln \left( (2\pi e)^n |C| \right), \]
c.i., $N_n(0, C)$ has largest entropy amongst all r.v. with pos. def. covariance matrix $C$.

Proof: W.l.o.g. assume $EX = 0$ (see Th. 2.5.11)
Let $q(x) = \frac{1}{(2\pi)^{n/2} |C|^{1/2}} \exp\left\{ -\frac{1}{2} x^T C^{-1} x \right\}$ density of $N_n(0, C)$.
Let $X \sim q(x)$, $EX = 0$, $\text{Cov}(X) = E(XX^T) = \int x x^T p(x) \, dx$.

\[ h(X) = -\int p(x) \ln p(x) \, dx \]
\[ \leq -\int p(x) \ln q(x) \, dx \quad \text{(Cor. 2.5.8 a)} \]
\[ = -\int p(x) \ln \left[ \frac{1}{(2\pi)^{n/2} |C|^{1/2}} \exp\left\{ -\frac{1}{2} x^T C^{-1} x \right\} \right] \, dx \]
\[ = \ln \left( (2\pi)^{n/2} |C|^{1/2} \right) + \frac{1}{2} \int x^T C^{-1} x \, p(x) \, dx \]
\[ = \ln ((2\pi)^{n/2} |C|^{1/2}) + \frac{1}{2} \int \text{tr}(C^{-1} x x^T) \, p(x) \, dx \]
\[ = \ln ((2\pi)^{n/2} |C|^{1/2}) + \frac{1}{2} \int \text{tr}(C^{-1} x x^T) \, p(x) \, dx \]
\[ = \ln ((2\pi)^{n/2} |C|^{1/2}) + \frac{1}{2} \ln |C|^{-1} \int x x^T \, p(x) \, dx \]
\[ = \ln ((2\pi)^{n/2} |C|^{1/2}) + \frac{1}{2} \ln |C|^{-1} \left( E(x x^T) \right) = \ln ((2\pi e)^n |C|^{1/2}) \]
\[ = \ln ((2\pi e)^n |C|^{1/2}) \]