AdvStPh20190219.pdf

AdvStPh20190221.pdf

AdvStPh20190226.pdf

AdvStPh20190228.pdf

AdvStPh20190305.pdf

AdvStPh20190307.pdf

AdvStPh20190311.pdf

AdvStPh20190314.pdf

AdvStPh20190319.pdf

AdvStPh20190321.pdf

AdvStPh20190326.pdf

AdvStPh20190328.pdf

AdvStPh20190402.pdf

AdvStPh20190404.pdf

AdvStPh20190409.pdf

AdvStPh20190411.pdf

AdvStPh20190416.pdf

AdvStPh20190425.pdf

AdvStPh20190430.pdf

AdvStPh20190502.pdf

AdvStPh20190507.pdf

AdvStPh20190509.pdf

AdvStPh20190514.pdf

AdvStPh20190516.pdf

1.1 Some books

- F.Reif: Fundamentals of Statistical and Thermal Physics
- L.E.Reichl: A Modern Course in Statistical Physics
- C.Kittel: Elementary Statistical Physics
- J.Sethna Statistical Mechanics,mEntropy, Order Parameters and Complexity (free on internet)
- R.K.Pathria: Statistical Mechanics
- Gould, Tobochnik: Thermal and Statistical Physics (free on internet)
- S.K.Ma: Statistical Physics
- MIT Open course project (http://ocw.mit.edu/OcwWeb/Physics/index.htm)
- J.M.Yeomans Statistical Mechanics of Phase Transitions

Probability density

$$p(a < x < b) = \int_a^b \varrho(x)dx$$

Mean value

$$\overline{f(x)} = \int f(x)\varrho(x)dx$$

Independent variables

$$\varrho(x, y) = \varrho_1(x)\varrho_2(y)$$

Marginal distribution

$$\varrho_1(x) = \int \varrho(x, y)dy$$

Function of random variable

$$y = f(x)$$

$$\overline{g(y)} = \int g(f(x))\varrho(x)dx = \int g(y)\delta(y - f(x))\varrho(x)dxdy =$$

$$= \int g(y)\varrho(f^{-1}(\xi))\delta(y-\xi)\frac{1}{f'(f^{-1}(\xi))}dyd\xi = \int g(y)\widetilde{\varrho}(y)dy$$

$$\widetilde{\varrho}(y) = \int \delta(y - f(x))\varrho(x)dx = \varrho(f^{-1}(y))\frac{1}{f'(f^{-1}(y))}$$

$$z = x + y$$

$$\varrho(z) = \int \delta(z - (x+y))\varrho_1(x)\varrho_2(y)dxdy = \int \varrho_1(x)\varrho_2(z-x)dx$$

It is a convolution

Characteristic function

$$f(k) = \langle \exp(ikx) \rangle = \int dx \exp(ikx) \varrho(x) = \sum \frac{(ik)^n}{n!} \langle x^n \rangle$$

$$\varrho(x) = \frac{1}{2\pi} \int dk \exp(-ikx) f(k)$$

$$\langle x^n \rangle = (-i)^n \frac{d^n f}{dk^n}$$

For the sum of two variables the density function is a convolution, so the characteristic function is a product, since the Fourier transform of a convolution is a the product of Fourier transforms.

$$\int \mathrm{e}^{ikz} \delta(z-(x+y)) \varrho_1(x) \varrho_2(y) dx dy dz = \int dx \mathrm{e}^{ikx} \varrho_1(x) \int dy \mathrm{e}^{iky} \varrho_2(y) = f_1(k) f_2(k)$$

Characteristic function of a function of a random variable

$$y = \varphi(x)$$

$$\varrho(y) = \int \varrho(x)\delta(y - \varphi(x))dx$$

$$f(k) = \int dy\mathrm{e}^{iky}\varrho(y) = \int dxdy\mathrm{e}^{iky}\varrho(x)\delta(y - \varphi(x)$$

$$f(k) = \int \mathrm{e}^{ik\varphi(x)}\varrho(x)dx$$

Cumulants

$$s(k) \equiv \ln(f(k)) = \sum \frac{(ik)^n}{n!} C_n$$

$$C_1 = \langle x \rangle$$

$$C_2 = \langle x^2 \rangle - \langle x \rangle^2 \equiv \langle x^2 \rangle_C$$

$$C_3 = \langle x^3 \rangle - 3\langle x \rangle \langle x^2 \rangle + 2\langle x^3 \rangle^3 \equiv \langle x^3 \rangle_C$$

$$C_n = (-i)^n s^{(n)}(k)|_{k=0}$$

$$\langle x^n \rangle = (-i)^n \frac{d^n f}{dk^n}$$

Cumulant expansion

$$\begin{aligned} \langle x_1 x_2 x_3 x_4 \rangle = \langle x_1 x_2 x_3 x_4 \rangle_C + \langle x_1 x_2 x_3 \rangle_C \langle x_4 \rangle_C + 3 \text{ terms} + \\ + \langle x_1 x_2 \rangle_C \langle x_3 x_4 \rangle_C + 2 \text{ terms} + \langle x_1 x_2 \rangle_C \langle x_3 \rangle_C \langle x_4 \rangle_C + 5 \text{ terms} + \\ + \langle x_1 \rangle_C \langle x_2 \rangle_C \langle x_3 \rangle_C \langle x_4 \rangle_C \end{aligned}$$

We now prove the cumulant expansion. Factor $(-i)^n$ is implicitely hidden in both moments and cumulants, so it gets canceled finally.

$$f(k) = \exp(s(k))$$

$$\langle x^n \rangle = f^{(n)}(k)|_{k=0} = \left(\frac{d^n}{dk^n} \exp(s(k)) \right)_{k=0}$$

$$1 = \langle x^0 \rangle = \exp(s(0)) = \exp(C_0) \quad \Rightarrow \quad C(0) = 0$$

$$\langle x^1 \rangle = \left(\frac{d}{dk} \exp(s(k)) \right)_{k=0} = \left(s^{(1)}(k) \exp(s(k)) \right)_{k=0} = C_1$$

$$\langle x^2 \rangle = \left(\frac{d}{dk} s^{(1)}(k) \exp(s(k)) \right)_{k=0} = \left((s^2(k) + s^1(k) s^1(k)) \exp(s(k)) \right)_{k=0} = C_2 + C_1 C_1$$

$$\langle x^3 \rangle = \left(\frac{d}{dk} \left(s^2(k) + s^1(k) s^1(k) \right) \exp(s(k))) \right)_{k=0} =$$

$$= \Big((s^{(3)}(k) + s^{(2)}(k) s^{(1)}(k) + s^{(1)}(k) s^{(2)}(k) +$$

$$+ s^{(2)}(k) s^{(1)}(k) + s^{(1)}(k) s^{(1)}(k) s^{(1)}(k)) \exp(s(k)) \Big)_{k=0}$$

Here the terms in the first line of the last equation originated by differentiating the terms in the bracket before the exponential, the terms in the second line of the last equation originated by differentiating the exponential.

$$
\begin{aligned}
\langle x^3\rangle &= \left(\frac{d}{dk}\left(s^2(k)+s^1(k)s^1(k)\right)\exp(s(k))\right)_{k=0} = \\
&= \Big((s^{(3)}(k)+s^{(2)}(k)s^{(1)}(k)+s^{(1)}(k)s^{(2)}(k)+ \\
&+ s^{(2)}(k)s^{(1)}(k)+s^{(1)}(k)s^{(1)}(k)s^{(1)}(k))\exp(s(k))\Big)_{k=0}
\end{aligned}
$$

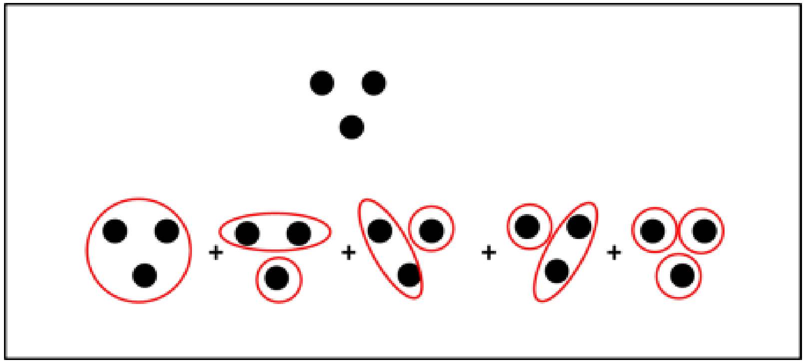


Figure 2.1: Third momentum into cumulants

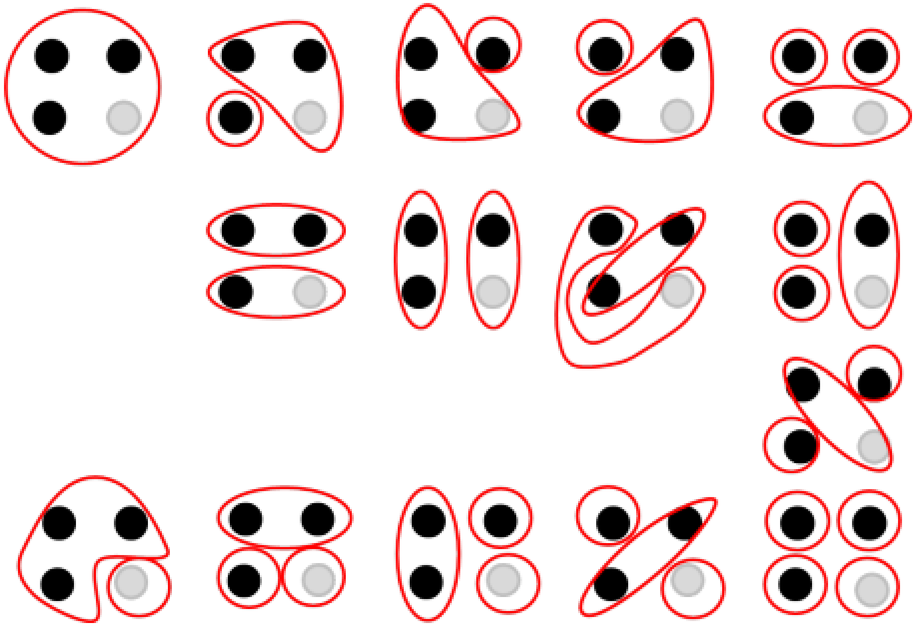


Figure 2.2: Fourth momentum into cumulants

Covariance

$$cov(x, y) = \langle xy \rangle - \langle x \rangle \langle y \rangle$$

Correlation function

$$cor(x, y) = \frac{cov(x, y)}{\sigma_x \sigma_y}$$

Normal distribution

$$\varrho(x) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp(-\frac{x^2}{2\sigma^2})$$

$$f(k) = \exp(-\frac{k^2\sigma^2}{2})$$

$$\ln(f(k)) = -\frac{k^2\sigma^2}{2}$$

only second cumulant non-zero, so all higher moments are expressible through the second cumulant.

Variance

$$\sigma^2 = \langle x^2 \rangle_C$$

Multivariant Gauss distribution

$$\varrho(x_1, \ldots, x_n) = \sqrt{\frac{\det(g)}{(2\pi)^n}} \exp(-\frac{1}{2} x^\top g x)$$

$$f(k_1, \ldots, k_n) = \langle \exp(i(k_1 x_1 + \ldots + k_n x_n)) \rangle$$

$$\begin{aligned} \langle x_1 x_2 x_3 x_4 \rangle &= (-i)^4 \frac{\partial}{\partial k_1} \frac{\partial}{\partial k_2} \frac{\partial}{\partial k_3} \frac{\partial}{\partial k_4} f(k_1, \ldots, k_n) \Big|_{k_i = 0} = \\ &= \langle x_1 x_2 \rangle \langle x_3 x_4 \rangle + \langle x_1 x_3 \rangle \langle x_2 x_4 \rangle + \langle x_1 x_4 \rangle \langle x_2 x_3 \rangle \end{aligned}$$

Wick theorem

Generating function

$$f(u) = \sqrt{\frac{A}{2\pi}} \int dx \exp(-\frac{1}{2}Ax^2 + ux)$$

to be compared with characteristic function with $u = ik$

Technique: shift of variables

$$x = x' - \frac{u}{A}$$

$$f(u) = \exp\left(\frac{u^2}{2A}\right)$$

Taylor series generates moments

Another technique to calculate moments: Euler Gamma function

$$\Gamma(z) = \int_0^\infty t^{z-1} \exp(-t) dt$$

$$\Gamma(n) = (n-1)!$$

$$\Gamma(z+1) = z\Gamma(z)$$

$$\Gamma(1/2) = \sqrt{\pi}$$

$$\Gamma(\frac{n}{2}) = \sqrt{\pi}\frac{(n-2)!!}{2^{(n-1)/2}} \quad \text{for odd } n$$

$$\int dx x^n \exp(-\frac{1}{2}Ax^2) = \left(\frac{A}{2}\right)^{-\frac{n+1}{2}} \int d\xi \xi^n \exp(-\xi^2) =$$

$$= \left(\frac{2}{A}\right)^{\frac{n+1}{2}} \int \frac{1}{2} t^{-1/2} t^{n/2} \exp(-t) = \frac{1}{2}\left(\frac{2}{A}\right)^{\frac{n+1}{2}} \Gamma(\frac{n+1}{2})$$

2.2 n-dimensional sphere

$$(\sqrt{\pi})^n = \int dx \exp(-\sum x_i^2) = \int \exp(-r^2) S_n(r) dr$$

from here one easily calculates the surface of the n-dimensional spehere $S_n(r)$ assuming the form

$$S_n(r) = C_n r^{n-1}$$

$$S_n(r) = \frac{2\pi^{n/2} r^{n-1}}{\Gamma(n/2)}$$

and then also the volume of an n-dimensional spehere $V_n(r)$ via

$$V_n(r) = \int_0^r S_n(r) dr$$

$$V_n(r) = \frac{\pi^{n/2} r^n}{\Gamma(n/2+1)}$$

2.3 Laplace method

Also called steepest descent method or saddle point method is an approximation scheme which is used to evaluate the integral I=

$$\int \exp(Mf(x))dx \qquad \text{for } M \to \infty$$

Let $f(x)$ has only one maximium at x_0, then the integral is dominated by this maximum

$$f(x) = f(x_0) + \frac{1}{2} f''(x_0)(x - x_0)^2$$

$$\int \exp(Mf(x))dx = \exp(Mf(x_0)) \int dx \exp(-\frac{M}{2}|f''(x_0)|(x - x_0)^2) =$$

$$= \exp(Mf(x_0)) \sqrt{\frac{2\pi}{M|f''(x_0)|}}$$

The logarithm is dominated by the maximum

$$\ln I = Mf(x_0)$$

Stirling formula

$$N! = \sqrt{2\pi N} N^N \mathrm{e}^{-N}$$

We denote $x = Nz$

$$N! = \Gamma(N+1) = \int x^N \mathrm{e}^{-x} dx = \int \mathrm{e}^{-Nz} (Nz)^N N dz =$$

$$= N^{N+1} \int dz \mathrm{e}^{(-Nz + N \ln z)} = N^{N+1} \int \mathrm{e}^{N(\ln z - z)} dz$$

Laplace method with

$$f(z) = \ln z - z$$

$$f'(z) = \frac{1}{z} - 1 \qquad \text{maximum for } \; z = 1$$

$$f''(z) = -\frac{1}{z^2}$$

$$N! = N^{N+1} \mathrm{e}^{-N} \sqrt{\frac{2\pi}{N}}$$

Central limit theorem

by the central limit theorem, according to which in the limit $n \to \infty$ the variable

$$X_n = \frac{\sum_{i=1}^{n}(x_i - \mu)}{\sqrt{n}\sigma}$$

is distributed according to the standard normal distribution $\mathcal{N}(0,1)$. Here $\mu < \infty$ is the mean of x_i, $\sigma < \infty$ is the root mean squared of $(x_i - \mu)$.

We define new random variables $y_i = (x_i - \mu)/\sigma$.These variables have the mean value 0 and variance 1.

Using this relation, we can write the Taylor series of the characteristic function of the probability density $f(x)$ with a mean 0 and variance 1 in the form

$$\varphi_f(t) = 1 - \frac{t^2}{2} + o(t^2) \qquad (t \to 0).$$

By convolution the characteristic functions get multiplied (they are the Fourie transforms). Since $X_n = \sum_i y_i/\sqrt{n}$, the characteristic function of its distribution $g(X_n)$ is the product of the characteristic functions of the distribution $f(y)$ and can be written as

$$\left[\varphi_g(t/\sqrt{n})\right]^n = \left[1 - \frac{t^2}{2n} + o(t^2/n)\right]^n \xrightarrow{n\to\infty} \mathrm{e}^{-t^2/2}.$$

The last function is the characteristic function of the standard normal distribution $\mathcal{N}(0,1)$, we have proved that the distribution $g(X_n)$ converges to the standard normal distribution $\mathcal{N}(0,1)$.

Characteristic function of a function of a random variable

$$y = \varphi(x)$$

$$\varrho(y) = \int \varrho(x)\delta(y - \varphi(x))dx$$

$$f(k) = \int dy \mathrm{e}^{iky}\varrho(y) = \int dx dy \mathrm{e}^{iky}\varrho(x)\delta(y - \varphi(x)$$

$$f(k) = \int \mathrm{e}^{ik\varphi(x)}\varrho(x)dx$$

Thebyshev inequality

$$\sigma^2 = \int_{-\infty}^{\infty} \big(x - \langle x\rangle\big)^2 f(x)\,\mathrm{d}x.$$

$$\sigma^2 \geq \int_{-\infty}^{\langle x\rangle-\epsilon} \big(x - \langle x\rangle\big)^2 f(x)\,\mathrm{d}x + \int_{\langle x\rangle+\epsilon}^{\infty} \big(x - \langle x\rangle\big)^2 f(x)\,\mathrm{d}x$$

$$\sigma^2 \geq \epsilon^2 \Big[\int_{-\infty}^{\langle x\rangle-\epsilon} f(x)\,\mathrm{d}x + \int_{\langle x\rangle+\epsilon}^{\infty} f(x)\,\mathrm{d}x \Big] = \epsilon^2\, P\big(|x - \langle x\rangle| \geq \epsilon\big)$$

$$P\big(|x - \langle x\rangle| \geq \epsilon\big) \leq \frac{\sigma^2}{\epsilon^2}$$

Thebyshev inequality

$$P\big(|x - \langle x\rangle| \geq \epsilon\big) \leq \frac{\sigma^2}{\epsilon^2}$$

Now let us consider the sum of N independent identically distributed random variables $y_N = (x_1 + \cdots + x_N)/N$. Clearly $\langle y\rangle = \langle x\rangle$. Since $\sigma_y^2 = \sigma_x^2/N$ (the reader can easily check this himself) the Tchebyshev inequality gives

$$P\big(|y_N - \langle x\rangle| \geq \epsilon\big) \leq \frac{\sigma_y^2}{\epsilon^2} = \frac{\sigma_x^2}{N\epsilon^2}.$$

Assuming $\sigma_x < \infty$ it is clear that for any chosen value ϵ we get

$$\lim_{N\to\infty} P\big(|y_N - \langle x\rangle| \geq \epsilon\big) = 0.$$

The probability that the arithmetic average y_N of independent identically distributed random variables x_i difference from $\langle x\rangle$, in the limit $N \to \infty$ declines to zero.

So far we have shown how to determine experimentally the mean value with arbitrary precision. Now let us do the same for a probability that the random variable value hits some predefined subset M. We have a random variable x with the probability density $\rho(x)$. Let us define a yes/no event

$$\text{yes: } x \in M \qquad\qquad \text{no: } x \notin M$$

Now we define a new random variable as a function $\chi(x)$ where ξ is the characteristic function of the set M, that is

$$\chi(x) = \begin{cases} 0 & x \notin M \\ 1 & x \in M \end{cases}$$

The mean value of χ is

$$\overline{\chi} = \int \chi(x)\rho(x)dx = \int_M \rho(x)dx = p(x \in M)$$

On the other hand the mean value χ can be determined experimentally with arbitrary precision as

$$\overline{\chi} = \frac{1}{N}(\chi_1 + \chi_2 + \ldots + \chi_n)$$

This fact justifies the experimental procedure to measure the probability as a ratio of the number favorable events to the total number of events. This ratio according to the law of large numbers converges to a certain value and this value we call the probability of the studied phenomenon.

Binomial distribution

Binomial distribution is the discrete probability distribution of the number of successes in a sequence of n independent yes/no experiments, each of which yields success with probability p.

$$p(k;n,p) = \binom{n}{k} p^k (1-p)^{n-k}$$

K successes from n trials

$$\overline{k} = \sum_0^n k \binom{n}{k} p^k q^{n-k} = p\frac{\partial}{\partial p}\sum_1^n \binom{n}{k} p^k q^{n-k} = p\frac{\partial}{\partial p}((p+q)^n - q^n) = np$$

$$\overline{k^2} = \sum_0^n k^2 \binom{n}{k} p^k q^{n-k} = p\frac{\partial}{\partial p} p\frac{\partial}{\partial p}\sum_1^n \binom{n}{k} p^k q^{n-k} =$$

$$= p\frac{\partial}{\partial p} p\frac{\partial}{\partial p}((p+q)^n - q^n) = p\frac{\partial}{\partial p} np(p+q)^{n-1} =$$

$$= np + n(n-1)p^2 = np(1-p) + n^2p^2$$

$$\overline{k^2} - \overline{k}^2 = np(1-p)$$

In the limit $n \to \infty$ assuming $pn = \text{const.} \equiv \lambda$, the binomial distribution approaches the Poisson distribution

$$f(k) = \frac{\lambda^k}{k!} \mathrm{e}^{-\lambda},$$

where λ is the mean value of the random variable k.

$$\overline{k^2} - \overline{k}^2 = np(1-p)$$

In the limit $\lambda \to \infty$ the Poisson distribution approaches the normal distribution $\mathcal{N}(\lambda, \sqrt{\lambda})$.

Random numbers

Generic random number generator

```
i:= (i*k) mod N
r:=i;
r:=r/N; //this is real number arithmetic
return(r)
```

The key is Monte Carlo integration

$$\overline{f} = \frac{1}{b-a}\int_a^b f(x)dx$$

The mean value can be determined ”experimentally”

$$\overline{f} = \frac{1}{N}\sum_1^N f(a + RND() * (b-a))$$

this is a default Monte Carlo integration algorithm.

We want to calculate by Monte Carlo the integral

$$\int_{-\infty}^{\infty} f(x)\varrho(x)dx$$

but we want to use random generator generating x distributed according to $\varrho(x)$
The key is the substitution to integral using ”inverse to primitive”

$$x = F^{-1}(\xi) \qquad \xi \in (0,1)$$

where

$$F(x) = \int_{-\infty}^{x} \varrho(x')dx'$$

The substitution is

$$dx = \frac{1}{F'}d\xi = \frac{1}{\varrho(F^{-1}(\xi))}d\xi$$

and we get

$$\int_{-\infty}^{\infty} f(x)\varrho(x)dx = \int_0^1 d\xi f(F^{-1}(\xi))\varrho(F^{-1}(\xi))\frac{1}{\varrho(F^{-1}(\xi))} = \int_0^1 d\xi f(F^{-1}(\xi))$$

Importance sampling

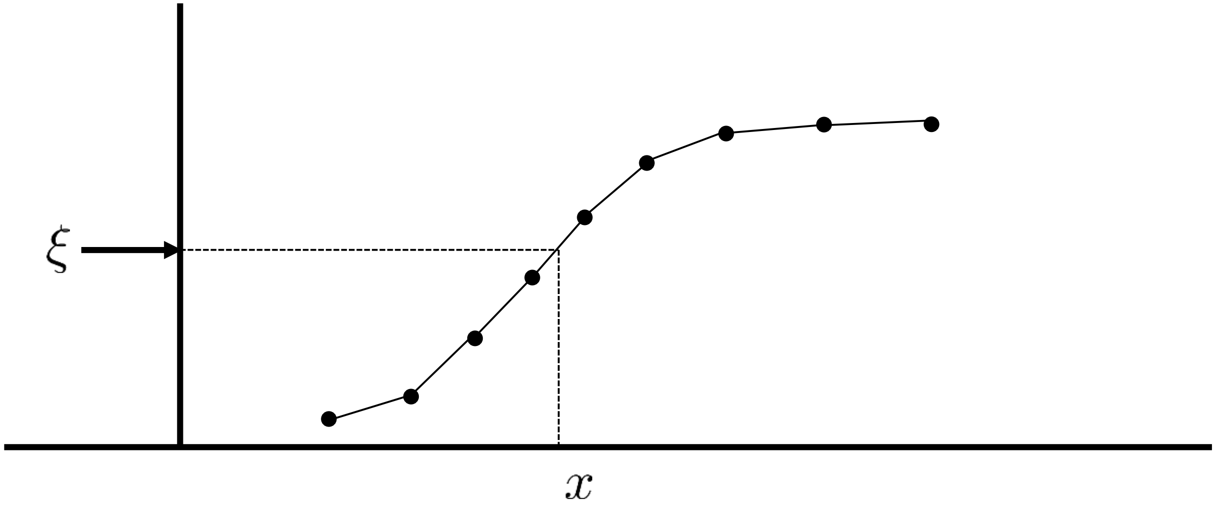
We see that the integral can be calculated by the following Monte Carlo

$$\int_{-\infty}^{\infty} f(x)\varrho(x)dx == \frac{1}{N}\sum_{1}^{N} f(F^{-1}(RND()))$$

So random variable x distributed according tho the probability density $\varrho(x)$ is generated by the algorithm

$$x = F^{-1}(RND())$$

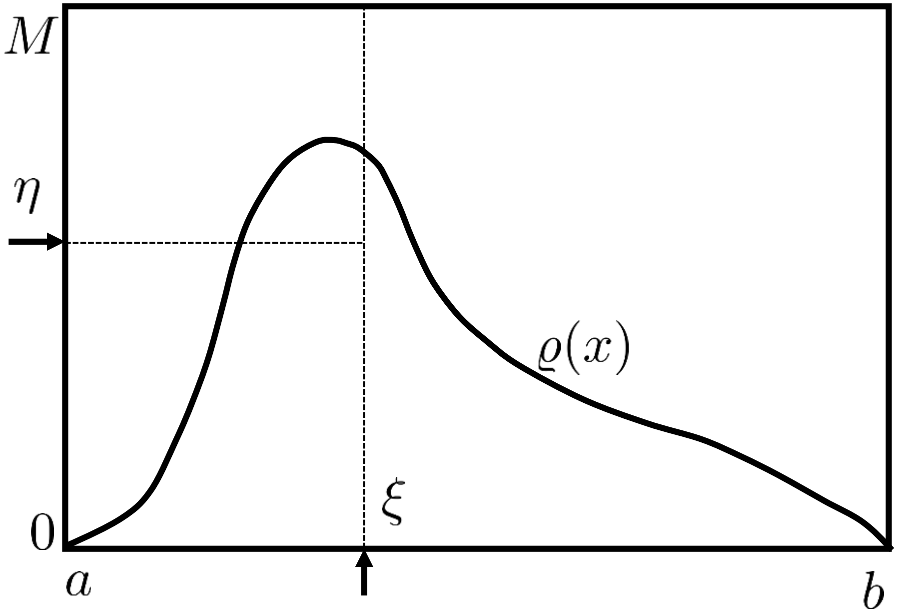
Numerically



Another method which can be used if $\varrho(x)$ is defined on the finite interval (a, b) if we know, that $\varrho(x)$ is majorized by some value M:

$$\forall x \in (a, b) : 0 < \varrho(x) < M$$

- we generate randomly a point (ξ, η) in the "bounding rectangle" that is $\xi = a + (b - a) \times RND()$ and $\eta = M \times RND()$
- if the generated point (ξ, η) is below the curve ϱ, that is if $\eta < \varrho(\xi)$ we accept value ξ as the generated value x
- otherwise we repeat the process by generating a new point (ξ, η)



accept - reject

Importance sampling for normal distribution

$$x = \sqrt{-2\sigma^2 \ln(\mathrm{RND}())}\,\cos\left(2\pi\ \mathrm{RND}()\right)$$

Selected items from statistics

$$\varrho(x) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{x^2}{2})$$

Now we form a new random variable

$$y = x^2$$

Its probability density is

$$\varrho_y(y) = \int \varrho(x)\delta(y{-}x^2)dx = \Theta(y)(\frac{1}{\sqrt{2\pi}} \exp(-\frac{y}{2})\frac{1}{2\sqrt{y}} + \text{term with the negative root})$$

$$\varrho_y(y) = \frac{1}{\sqrt{y}}\frac{1}{\sqrt{2\pi}} \exp(-\frac{y}{2})$$

This is a special case of a more general ”χ^2 distribution with n degrees of freedom”

$$\varrho_{\chi^2}(x) = \frac{1}{2^{n/2}\Gamma(\frac{n}{2})} x^{\frac{n}{2}-1} \exp(-\frac{x}{2})$$

If $x_1, x_2, \ldots, x_n$ are independent draws from a normal $N(0,1)$ distribution, then the random variable

$$y_n = \sum_i x_i^2$$

is distributed as χ^2 with n degrees of freedom (This is the definition.)

It is easy to derive the explicit formula for the χ^2 distribution.

$$\begin{aligned}
\rho_{\chi^2}(y) &= \int \frac{1}{(\sqrt{2\pi})^n} \exp(-\frac{1}{2}\sum x_i^2)\delta(y - \sum x_i^2) dx_1 \ldots dx_n = \\
&= \int \frac{1}{(2\pi)^{n/2}} \exp(-y/2)\delta(y - r^2) S_n(r) dr = \\
&= \int \frac{1}{(2\pi)^{n/2}} \exp(-y/2) \frac{1}{2\sqrt{y}} S_n(\sqrt{y}) dr \\
&= \frac{1}{2^{n/2}\Gamma(n/2)} y^{n/2-1} \exp(-y/2)
\end{aligned}$$

Suppose we have a container of random values x said to be distributed according to normal distribution $N(\mu, \sigma)$. We want to verify the fact experimentally. The procedure is

- We formulate a null-hypothesis according to which the assumed fact is true: that is we assume that the random values in the container are really distributed according to $N(\mu, \sigma)$.
- We perform an experiment: draw n independent samples $x_1, \ldots, x_n$.
- We calculate the statistics

$$\chi^2 = \sum_1^n \frac{(x-\mu)^2}{\sigma^2}$$

- The value of χ^2 obtained in the previous step is checked for probability. That is we evaluate the probability p that a chi-squared distribution can give value as observed or higher

$$p = \int_{\chi^2}^{\infty} \varrho_{\chi^2}(x) dx$$

- The value of χ^2 obtained in the previous step is checked for probability. That is we evaluate the probability p that a chi-squared distribution can give value as observed or higher

$$p = \int_{\chi^2}^{\infty} \varrho_{\chi^2}(x)dx$$

- It the obtained value p is very small (like $\lessapprox 0.001$) we say that it is very improbable that we have such a bad-luck that we observe such a big deviation form the expected null-hypothesis-behaviour. So we reject the null-hypothesis. If the value p is not small enough, we say that it may be that the value χ^2 we have observed is due to a fluctuation and we keep the null hypothesis as valid.

$$\chi^2 = \sum_{1}^{n} \frac{(x-\mu)^2}{\sigma^2}$$

Expected χ^2 value is of the order of n.

Another use of chi-squared distribution for statistical testing is given by the following Pearson theorem.

We have r histogram bins, and the expected probabilities for a random event to fall into the i-th bin are p_i. We observe n random events and the actual number of hits in the i-th bin are n_i. Then the statistics

$$\chi^2 = \sum_1^r \frac{(n_i - np_i)^2}{np_i}$$

is distributed as chi-squared with $(r-1)$ degrees of freedom.

- We assume we know the expected probabilities p_i and as the null hypothesis we assume they are true.
- We perform the experiment and observe the actual number of hits n_i and calculate the statistics χ^2.
- We calculate what is the probability p to observe the value χ^2 as observed or even higher.
- If p is reasonably small we reject the null hypothesis and say that the true probabilities are different from those expected.

W consider a a clinical study (experiment) the aim of which is to test which of the two available drugs (DRUG1, DRUG2) is better for treatment of a particular disease. The study is organized as follows. The patient having the disease are randomized (chosen randomly) for treatment either with DRUG1 or DRUG2. Then their health state is observed and classified into two categories RESPONSE (these are patients whose disease positively reacted to the treatment) and NO RESPONSE (those for whom the treatments did not have the required effect). Then a ”contingency table” is created where number of patients of each category is entered. The contingency table for our case is a 2×2 matrix and is defined by the white fields in the following table

	DRUG1	DRUG2	Totals
RESPONSE	15	8	23
NO RESPONSE	7	5	12
Totals	22	13	35

	DRUG1	DRUG2	Totals
RESPONSE	15	8	23
NO RESPONSE	7	5	12
Totals	22	13	35

There are four categories of patients in our table (number of rows $times$ number of columns). The actual number of entries in each category can be denoted as N_{ij}. So for example N_{21} is the observed number of patients in the second row and first column, that is it is the number of patients treated by the DRUG1 with status NO RESPONSE, so $N_{21} = 7$ in our table. The independence of features means that the probability for a patient to enter a particular row is independent on what column it enters and the probabilities to enter a particular cell factorize. If the null hypothesis holds, then a set of parameters r_1, r_2, c_1, c_2 must exist which give the probabilities to enter the rows 1 and 2 and columns 1 and 2, respectively. If there were no statistical fluctuations then the observed numbers of hits should be

$$N_{ij} = N r_i c_j$$

where N is the total number of patients. Now the task is twofold

- Estimate the optimal values of parameters r_1, r_2, c_1, c_2 which best describe the observed data provided the null hypothesis holds

- Test whether the optimal parameters really describe the observed data well or the description is so poor that we should abandon the null hypothesis.

To estimate the optimal parameters we introduce the following notation

$$N_{i+} = \sum_j N_{ij}$$

$$N_{+j} = \sum_i N_{ij}$$

These variables describe the number of hits of particular row (or column) and are entered as ”totals” in the above contingency table. For the total number of patients we get

$$N = \sum_{i,j} N_{ij} = \sum_i N_{i+} = \sum_j N_{+j}$$

Now it is intuitively clear[3] that

$$r_i = \frac{N_{i+}}{N} \quad c_j = \frac{N_{+j}}{N}$$

So the expected number of hits E_{ij} in the cell (i, j), provided the null hypothesis holds and fluctuations are zero, is

$$E_{ij} = N r_i c_j = \frac{N_{i+} N_{+j}}{N}$$

So we create another contingency table containing not the observed but rather the expected numbers of hits. Here it is

	DRUG1	DRUG2	Totals
RESPONSE	14.457	8.543	23
NO RESPONSE	7.543	4.457	12
Totals	22	13	35

$$\chi^2 = \sum_{i,j} \frac{(N_{ij} - E_{ij})^2}{E_{ij}}$$

	DRUG1	DRUG2	Totals
RESPONSE	15	8	23
NO RESPONSE	7	5	12
Totals	22	13	35

chi-squared with $(r - 1) \times (c - 1)$ degrees of freedom

Student test

The Student distribution with n degrees of freedom is defined by the probability density

$$\varrho_{\mathrm{t}}(x) = \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})\sqrt{\pi}\sqrt{n}} \left(1 + \frac{x^2}{n}\right)^{(n+1)/2}$$

It can be shown that if we have $(n+1)$ samples

$$x_0, x_1, x_2, \ldots, x_n$$

from the normal $N(0,1)$ distribution, then the statistics

$$t = \frac{x_0}{\sqrt{\frac{1}{n}\sum_1^n x_i^2}}$$

is distributed according to the Student distribution with n degrees fo freedom

The generic example looks as follows. We have two containers said to contain normally distributed samples of the same gaussian distributions $N(\mu, \sigma)$ with unknown but equal values of the parameters μ and σ. We make n draws from the first container, getting the sequence of random values

$$x_1, x_2, \ldots, x_n$$

and m draws from the second container getting the sequence

$$y_1, y_2, \ldots, y_m$$

The null hypothesis is that the two containers are truly equivalent. The question is can we rule out the null hypothesis on basis of the observed events?

- Calculate the sample means

$$\overline{x} = \frac{1}{n}\sum_i x_i \quad \bar{(}y) = \frac{1}{m}\sum_i y_i$$

- calculate sample variances

$$s_x^2 = \frac{1}{n-1}\sum_i (x_i - \overline{x})^2 \quad s_y^2 = \frac{1}{m-1}\sum_i (y_i - \overline{y})^2$$

- calculate the statistics t

$$t = \frac{\overline{x} - \overline{y}}{\sqrt{\frac{(n-1)s_x^2 + (m-1)s_y^2}{n+m-2}}\sqrt{\frac{1}{n} + \frac{1}{m}}}$$

- the random variable t should have the Student distribution with $(n+m-2)$ degrees of freedom. Usually it is $|t|$ which matters and we ask what is the probability that the absolute value of a student-distributed variable is greater than the value we have observed. So we have to integrate both the left and the right tail of the distribution to get the probability which we consider for estimating the significance.

Bayesian statistics

prior probability + data obtained ⟶ posterior probability

The prior probability here measures our complete knowledge obtained before the experiment considered, the posterior probability corresponds to our knowledge after the proces of the considered data evaluation. The scheme looks rather mysteriously. To understand it better we have to discuss its different parts.

For example we said that the prior probability measure our complete knowledge. How it is possible? If we understand the notion of probability as the ratio of the number favorable events to the number of all events, it cannot measure knowledge. As an alternative here we consider the probability to be the measure of our **confidence** of a realization of a particular event. The limiting values are 0 for the events we consider to be impossible and 1 for the events we consider to be certain.

The next reasoning is heavily based on the Bayes relation for the conditional probability, so we start by deriving it. Let A be some event and $\{B_i\}$ is a set of disjoint events such that $\cup B_i = \Omega$. The definition of conditional probability leads to the equality $P(A \cap B_i) = P(A)P(B_i|A)$ and in the same time $P(A \cap B_i) = P(B_i \cap A) = P(B_i)P(A|B_i)$. Comparing these two relations we get

$$P(B_i|A) = \frac{P(B_i)P(A|B_i)}{P(A)}.$$

$$P(B_i|A) = \frac{P(B_i)P(A|B_i)}{\sum_i P(B_i)P(A|B_i)}$$

Now we rename the entries in the Bayes formula to correspond more clearly to the context in which we shall use it

$$P(M_i|\mathcal{D}) = \frac{\pi(M_i)\, P(\mathcal{D}|M_i)}{\sum_i \pi(M_i)\, P(\mathcal{D}|M_i)}.$$

Here M_i denotes competing models, $\mathcal{D}$ the data obtained, $\pi(M_i)$ the prior probability of the model M_i and $P(\mathcal{D}|M_i)$ the probability to observe the data assuming validity of the M_i. The left side of the relation, $P(M_i|\mathcal{D})$, is the posterior probability. We shall demonstrate the practical use of the formula by several examples.

Let us investigate its use in the simple example of fitting the data by a straight line. This simple example will also manifest better the relation between the classical and Bayesian statistics.

We have results of some measurement, the set of data points $\{x_i, y_i\}$ for $i = 1, \ldots, N$. According to theory, the relation between x a y should be a linear one $y = Ax + B$. What values of A a B correspond best to the observed data?

If we measured absolutely exactly, we should obtain for each x_i the value $y_i^* = Ax_i + B$. Now let us assume, that the deviations from this value (measurement errors) are distributed normally with the standard deviation σ and the mean $\mu = 0$. This assumption is essential and is usually justified by the central limit theorem. The probability that for x_i we obtain the value y_i is

$$P(y_i|A, B) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(y_i - Ax_i - B)^2}{2\sigma^2}\right].$$

The probability to obtain the whole observed set of values $\{y_i\}$ is

$$P(\{y_i\}|A, B) = \frac{1}{(\sqrt{2\pi}\sigma)^N} \exp\left[-\sum_{i=1}^{N} \frac{(y_i - Ax_i - B)^2}{2\sigma^2}\right].$$

This is so called likelihood function.

This is so called likelihood function. Now let be non-Bayesian (frequentist) for a moment. Then the likelihood function is all we need and we estimate the best values for A a B by maximizing the likelihood function $P(\text{data}|\text{parameters})$. The above expression shows that the maximum is achieved when the sum of the squares of deviations is minimal. The estimates $\hat{A}$ a $\hat{B}$ are given as

$$\hat{A}, \hat{B} = \underset{A,\ B}{\operatorname{argmin}} \left\{ \sum_{i=1}^{N} \frac{(y_i - Ax_i - B)^2}{\sigma^2} \right\}.$$

Within the Bayesian spirit, however, we

should also consider the prior probability $\pi(A, B)$. Our main interest is not the likelihood function, but the posterior probability of the values of A and B after observing $\{y_i\}$. According to the Bayes formula it is given as

$$f\left(A, B|\{y_i\}\right) = \frac{\pi(A, B)\, P\big(\{y_i\}|A, B\big)}{\int \pi(A, B)\, P\big(\{y_i\}|A, B\big)\, \mathrm{d}A\mathrm{d}B} = \\ = \frac{\pi(A, B)}{C(\sqrt{2\pi}\sigma)^N} \exp\left[- \sum_{i=1}^{N} \frac{(y_i - Ax_i - B)^2}{2\sigma^2} \right].$$

The denominator of the fraction plays no role for the further consideration and is denoted simply by C. The maximum of the posterior probability differs in general from the maximum of the likelihood function.

Competition of two models

In physics we often meet situations when we have to choose between two alternative models. For example for the set of data points $\{x_i, y_i\}$ we have to decide whether the data support the two-parameters linear model

$$y_i = A + Bx_i$$

or an alternative three-parameters quadratic model

$$y_i = A + Bx_i + Cx_i^2$$

This is a question from the methodology of physics and within the Popper-like spirit it seems to be wrongly stated. Our basic belief is that in physics we cannot prove hypotheses, we can just disprove (reject) them. Here we have two hypotheses, linear and quadratic. If one of them can be rejected (by suitable statistical test) then we favor the other hypothesis. But what if neither of them can be disproved?

Then we usually use ”The Occam razor”.

So the ”standard folklore” (non-Bayesian, frequentist) is the following.

- Choose the simplest model (here the linear one). Adjust its parameters (A, B) by maximizing the likelihood
- Having fixed the parameters of the model perform a statistical test. For example you can calculate the χ^2 statistics

  $$\chi^2 = \sum \frac{(y_i - A - Bx_i)^2}{\sigma^2}$$

  and calculate the probability to observe the value obtained if the null hypothesis (linear model is true) holds
- If you get a reasonable probability, accept the model until new analysis is forced by new data.
- If you get too small a probability, reject the model and repeat the analysis with a more complicated model

Now, the Bayesian approach to the same problem would be:

- Estimate the prior model probabilities $\pi(M_1)$ and $\pi(M_2)$. (M_1 is the linear model, M_2 is the quadratic model.)The model prior probabilities might depend on the model parameters.
- Express the posterior model probabilities through the model parameters for both models
- Maximize both posterior probabilities choosing the optimal values of the model parameters.
- Compare the two posterior model parameters calculating their ratio (so that the normalization factor from the denominator drops out)

$$\frac{P(M_1|\text{data})}{P(M_2|\text{data})} = \frac{\pi(M_1)P(\text{data}|M_1)}{\pi(M_2)P(\text{data}|M_2)}$$

- Accept the model for which the posterior probability is higher

Model as data compression

$$010101010101010101010101010101010101010101010101\ldots,$$

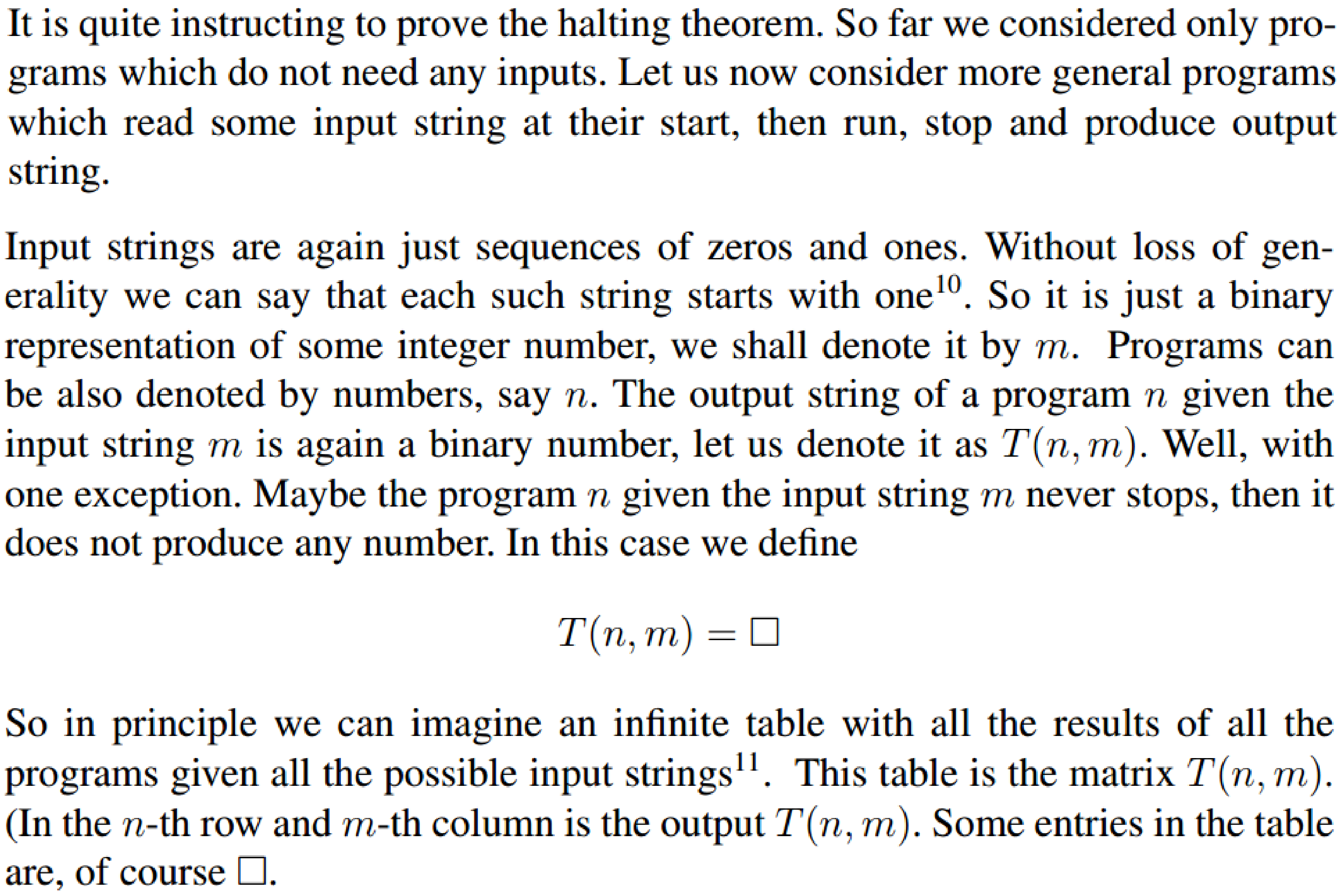
Can be reproduced by a very short program

```
for(i=0;i<10^9;i++) print("01");
```

- Programs are just finite strings of zeros and ones. This is a countable set. This set can be even ordered according to the length of the string representing the program. So the set of all programs can be numbered giving number 1 to the shortest program, 2 to the next one etc.
- Begin with program No 1. Run it. Wait until it stops. If it prints the desired data string, you have found the shortest program, the ”theory”. If not, continue with program No.2. And so on.
- Some program eventually prints the desired data string, at least the ”copy program” must do so.
- The first program (found in this way) which prints the data string is the shortest one doing so.

THEORY

The argumentation was wrong. Halting theorem!



It is quite instructing to prove the halting theorem. So far we considered only programs which do not need any inputs. Let us now consider more general programs which read some input string at their start, then run, stop and produce output string.

Input strings are again just sequences of zeros and ones. Without loss of generality we can say that each such string starts with one[10]. So it is just a binary representation of some integer number, we shall denote it by m. Programs can be also denoted by numbers, say n. The output string of a program n given the input string m is again a binary number, let us denote it as $T(n, m)$. Well, with one exception. Maybe the program n given the input string m never stops, then it does not produce any number. In this case we define

$$T(n, m) = \Box$$

So in principle we can imagine an infinite table with all the results of all the programs given all the possible input strings[11]. This table is the matrix $T(n, m)$. (In the n-th row and m-th column is the output $T(n, m)$. Some entries in the table are, of course $\Box$.

We said ”in principle” because we are not sure whether the matrix can really be constructed (computed). Well, the matrix is infinite, so it can never be constructed (computed) in totality. But we can define the matrix be computable if we can compute the entry in any given (n, m) position. Can we?

Yes if the halting problem can be solved. We need the program symbolically denoted as H which if given two[12]numbers (n, m) (that is some program and its data) always halts and outputs 1 if the considered program n with data string m stops and outputs 0 if the program n with data m does not stop. [13]

With the help of the magic program H one can compute the matrix $T(n, m)$. Indeed: given the two numbers (n, m) we first run the H program with (n, m) as input. If it writes 0, we know $T(n, m) = \Box$ and that is all. If the H program ends with 1, it is safe to run the program n with data m and we get the result $T(n, m)$. So if the program H exists, the table $T(n, m)$ is computable.

Now we are going to show, that the table $T(n,m)$ cannot be computable. Suppose it is. Then its diagonal $T(n,n)$ is also computable. Then a vector $V(n)$ defined[14]as $V(n) = T(n,n) + 1$ is also computable. It means there exists a program which computes this vector. There must be a number n_0 such that $T(n_0,m) = V(m)$ but then

$$T(n_0,n_0) = V(n_0) = T(n_0,n_0) + 1 \neq T(n_0,n_0)$$

We got contradiction. So the program H does not exist. The halting theorem is proven.

Coding theory

In this section we start to investigate deep relation between probability and information. We shall first need elements of the coding theory.

It is possible to construct codes for letters such that no sequence of symbols which is a true code for some symbol appears as a prefix in the sequence of some other letter. Then a sequence of symbols without any gaps can be cut into pieces representing individual letters (the gaps can be introduced automatically).

Let us denote the lengths of codes of letter of an m-letter alphabet as $s_1, \ldots, s_m$. Let the corresponding code is prefix-free. Then the following (Kraft inequality) holds

$$\sum_i 2^{-s_i} \leq 1.$$

The inverse is also true, that is if the lengths of m sequences $s_1, \ldots, s_m$ fulfill the Kraft inequality, then a prefix-free code exists for an m-letter alphabet with the code sequences of the corresponding lengths.

Let us prove the theorem. Without a loose of generality we assume that the lengths are ordered and s_m ix maximal. Any given prefix code can be represented by an binary tree of depth s_m where the branches from each node correspond code symbols (0 and 1) and each codeword is represented by a path to a leaf at depth s_i. This guarantees that no codeword is a prefix of another. For each leaf in such a code tree, consider the set of descendents A_i the leaf considered would have at depth s_m in a full binary tree. Then

$$A_i \bigcap A_j = \varnothing, \quad i \neq j$$

and

$$|A_i| = 2^{s_m - s_i}$$

Thus, given that the total number of nodes at depth s_m is 2^{s_m} ,

$$|\bigcup_{i=1}^{m} A_i| = \sum_{i=1}^{n} 2^{s_m - s_i} \leqq 2^{s_m}$$

$$\sum_i 2^{-s_i} \leq 1.$$

It is useful to demand that the message should have the total length as small as possible. Let us denote the length of the code assigned to the letter x by $L(x)$ and the alphabet of all letter by $\mathcal{A}$. The mean length assigned to a letter is $\langle L\rangle = \sum P(x)L(x)$. We assume here that the probabilities $P(x)$ are known beforehand.

$$\mathcal{L} = \sum_{x\in\mathcal{A}} P(x)L(x) - \lambda\Big(\sum_{x\in\mathcal{A}} 2^{-L(x)} - 1\Big).$$

$$\begin{aligned}
\frac{\partial}{\partial L(x)}\left(\sum_x P(x)L(x) + \lambda\Big(\sum_x 2^{-L(x)} - 1\Big)\right) &= \\
= P(x) - \lambda\ln(2)\exp(-L(x)\ln(2)) &= 0 \\
P(x) &= \lambda\ln(2)\exp(-L(x)\ln(2)) \\
\frac{P(x)}{\lambda\ln(2)} &= 2^{-L(x)} \\
\sum_x 2^{-L(x)} &= 1 \\
\frac{\sum_x P(x)}{\lambda\ln(2)} = \frac{1}{\lambda\ln(2)} &= 1 \\
\lambda &= \frac{1}{\ln(2)} \\
P(x) &= 2^{-L(x)}
\end{aligned}$$

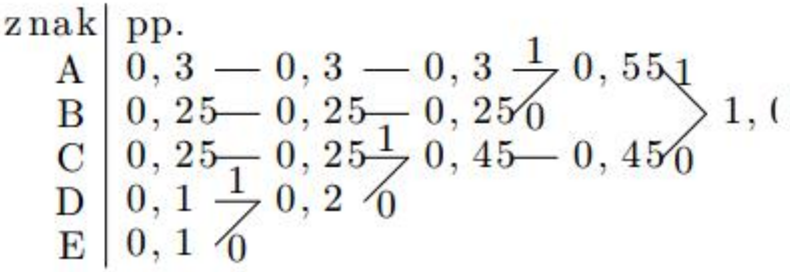
$$\boxed{L(x) = -\log_2 P(x)}$$

The solution is $L(x) = -\log_2 P(x)$. The optimal mean code length per letter is

$$H = -\sum_{x \in \mathcal{A}} P(x) \log_2 P(x)$$

This expression is called the entropy of the information source (when coding in the binary code). Since $\log_2 x = \ln 2 \cdot \ln x$, H differs only by a multiplicative factor from the expression for entropy $S = -k \sum p_i \ln p_i$ known in statistical physics.

An optimal solution (coding individual letters) is the Huffman code[16]. We start with two least probable letters and assign to them (as final symbols for their code sequences) 0 and 1. We sum their probabilities and count them together as if one letter the number of "effective letters" of the alphabet decreases by one. We iterate the procedure until all the letters have their code sequences assigned (see the figure).



The code sequences for the example in the above figure are $A \mapsto 11, B \mapsto 10, C \mapsto 01, D \mapsto 001, E \mapsto 000$. Huffman code is used in many applications like zip, mp3 a mpeg.

Supplements to the foundations of statistical physics

Density matrix

$$\overline{A} = \sum_x p(x)\langle x|\,\hat{A}|x\rangle$$

$$\hat{\rho} = \sum_x |x\rangle p(x)\langle x|$$

$$\overline{A} = Tr(\hat{\rho}\hat{A})$$

$$\begin{aligned}
Tr(\rho\hat{A}) &= \sum_n \langle n|\,\hat{\rho}\hat{A}|n\rangle = \sum_n \langle n|\,\sum_x |x\rangle p(x)\langle x|\,\hat{A}|n\rangle = \\
&= \sum_n \sum_x \langle n \mid x\rangle\, p(x)\langle x|\,\hat{A}|n\rangle = \\
&= \sum_n \sum_x p(x)\langle x|\,\hat{A}|n\rangle\, \langle n \mid x\rangle = \\
&= \sum_x p(x)\langle x|\,\hat{A}|x\rangle = \overline{A}
\end{aligned}$$

The density matrix is evidently a hermitian operator, it can be expressed in any base (in general it would not be diagonal) as

$$\hat{\rho} = \sum_{m,n} |m\rangle \rho_{mn} \langle n|$$

The time evolution of a macrostate will be described by a time dependent density matrix

$$\begin{aligned} \hat{\rho}(t) &= \sum_{x} |x(t)\rangle p(x) \langle x(t)| = \\ &= \sum_{x} \exp(-\frac{i}{\hbar}\hat{H}t)|x\rangle p(x)\langle x| \exp(\frac{i}{\hbar}\hat{H}t) \end{aligned}$$

Differentiating with respect to time we get the equation of motion

$$i\hbar \frac{\partial}{\partial t}\hat{\rho}(t) = [\hat{H}, \hat{\rho}]$$

If the macrostate is stationary, then it will be described by a time independent density matrix. According to the equation of motion we see that such a density matrix commutes with the Hamiltonian. Therefore there exists a base of stationary states in which the density matrix will be diagonal having the form

$$\hat{\rho} = \sum_n |n\rangle \rho_{nn} \langle n|$$

This is the reason why in statistical physics of stationary macrostates we can limit ourselves to the ensembles composed of stationary states.

In the case of a canonical ensemble we get for the density matrix

$$\hat{\rho} = \sum_n |n\rangle \frac{1}{Z} \exp(-\frac{E_n}{kT}) \langle n|$$

which can be formally written as

$$\hat{\rho} = \frac{1}{Z} \exp(-\frac{\hat{H}}{kT})$$

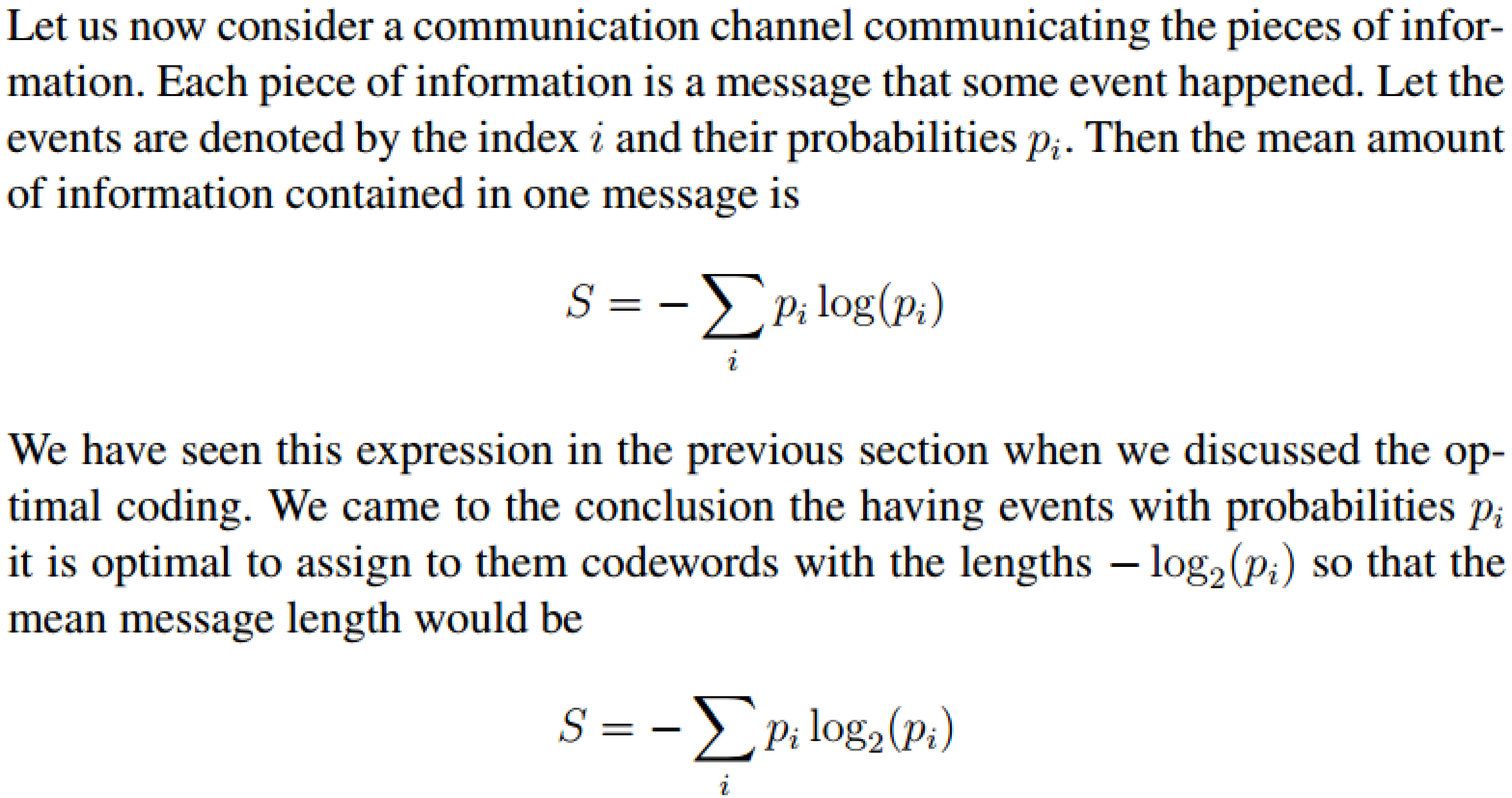
Entropy

How to measure the quantity of information?

The starting point is the requirement that the message saying an improbable event happened carries more information then a message reporting that a probable event happened. The quantity of information is therefore related to the probability of the message carrying the information. Information about the event is related to the probability of that event before it happened.

Now we require that the information can be communicated in parts. For example I can first announce that the unknown number is odd, and then only the number itself. It is natural to require that the information from partial messages should be added to get the total information. Having in mind that the probability of independent events combine multiplicatively, but the corresponding independent pieces of information should be combined additively it is natural to assume that the amount of information is given by the logarithm of the corresponding probability.

$$I = -\log(p)$$



Let us now consider a communication channel communicating the pieces of information. Each piece of information is a message that some event happened. Let the events are denoted by the index i and their probabilities p_i. Then the mean amount of information contained in one message is

$$S = -\sum_i p_i \log(p_i)$$

We have seen this expression in the previous section when we discussed the optimal coding. We came to the conclusion the having events with probabilities p_i it is optimal to assign to them codewords with the lengths $-\log_2(p_i)$ so that the mean message length would be

$$S = -\sum_i p_i \log_2(p_i)$$

Information is, in fact, measured by the price of telex carrying that information. If telex is priced by some amount of money per letter.

Using the logarithms of different bases just changes the units in which we measure the amount of information. (Changing the base just multiplicatively renormalizes the logarithm.) The unit corresponding to binary logarithm is called **bit**, the unit corresponding to natural logarithm is called **nat**.

Now we prove that the expression for entropy has the following important optimality property. Considering arbitrary sequences of positive numbers q_i satisfying the condition $\sum_i q_i = 1$ the following inequality holds

$$-\sum_i p_i \log(q_i) \geq -\sum_i p_i \log(p_i)$$

$$\begin{aligned}
\log(x) &\leq x - 1 \\
\log(\frac{q_i}{p_i}) &\leq \frac{q_i}{p_i} - 1 \\
p_i \log(q_i) - p_i \log(p_i) &\leq q_i - p_i \\
\sum_i (p_i \log(q_i) - p_i \log(p_i)) &\leq \sum_i (q_i - p_i) = 0 \\
-\sum_i p_i \log(p_i) &\leq -\sum_i p_i \log(q_i)
\end{aligned}$$

By the way we have proved again the optimal coding theorem, that is that one should use the codewords of the length log(q_i) = log(p_i).

Now, why all this is relevant to statistical physics?

A macrostate in statistical physics is represented by a statistical ensemble of microstates i, whose probabilities are p_i. Statistical entropy (of the macrostate) is defined as

$$S = -\sum_i p_i \log(p_i)$$

So we see that the statistical entropy can be interpreted in the following way.
Let us imagine that somebody gives us a sample of some macrostate. A macrostate is a virtual notion, he must actually give us some specific microstate. He just does not tell us which specific microstate from the corresponding ensemble he gave us.

So our knowledge about the system considered just corresponds to its macrostate, we are completely unaware of the microstate actually delivered. Now imagine that someone tells us which particular microstate was actually delivered. Then our original unawareness is changed to a complete knowledge. The amount of information contained in the message was $-\log(p_i)$ where p_i is the probability assigned in the statistical ensemble to the microstate actually delivered.

So the mean amount of information needed to complete our knowledge from macrostate to microstate level is

$$S = -\sum_i p_i \log(p_i)$$

$$S = -k \sum_i p_i \log(p_i)$$

In thermodynamic:

$$\Delta S = \int \frac{\delta Q}{T}$$

therefore thermodynamical entropy is measured in J/K

Variation principles

We have already met the inequality

$$-\sum_i p_i \log(q_i) \geq -\sum_i p_i \log(p_i)$$

in terms of the density matrix it can be written in the following form

$$-Sp(\rho \ln(\rho)) \leq -Sp(\rho \ln(\rho'))$$

This inequality holds for arbitrary density matrixρ'.

The proof is not completely trivial, since in general the two matrices ρ, ρ' cannot be diagonalized concurrently. Let us suppose that both matrices are strictly positive. Let ρ is diagonal in base $|m\rangle$ and ρ' in base $|q\rangle$. Then we get

$$Sp(\rho \ln \rho') - Sp(\rho \ln \rho) = \sum_{m,q} \rho_m \langle m \mid q\rangle \ln \rho'_q \langle q \mid m\rangle - \sum_m \rho_m \ln \rho_m =$$

$$= \sum_{m,q} \rho_m \langle m \mid q\rangle \ln \rho'_q \langle q \mid m\rangle - \sum_{m,q} \langle m \mid q\rangle \langle q \mid m\rangle \rho_m \ln \rho_m =$$

$$= \sum_{m,q} | \langle m \mid q\rangle |^2 \rho_m \ln \frac{\rho'_q}{\rho_m} \leq \sum_{m,q} | \langle m \mid q\rangle |^2 \rho_m (\frac{\rho'_q}{\rho_m} - 1) =$$

$$= Sp(\rho') - Sp(\rho) = 0$$

We have proven the above inequality for strictly positive matrices. By continuity it is true also for positive semidefinite matrices.

Entropy is maximal

Let us consider the Hilbert space of some physical system and a subspace of this space corresponding to vectors with a given fixed energy. Let this subspace have the dimension N. This subspace corresponds to microcanonical ensemble with the density matrix given as

$$\rho = \frac{1}{N}\sum_n |n\rangle\langle n|$$

$$-Sp(\rho \ln(\rho)) \leq -Sp(\rho \ln(\rho'))$$

For arbitrary density matrix defined in the energy subspace we get (according to the above inequality)

$$S' = -Sp(\rho' \ln(\rho')) \leq -Sp(\rho' \ln(\rho)) = -Sp(\rho' \ln(\frac{\widehat{1}}{N})) = \ln N$$

Free energy is minimal

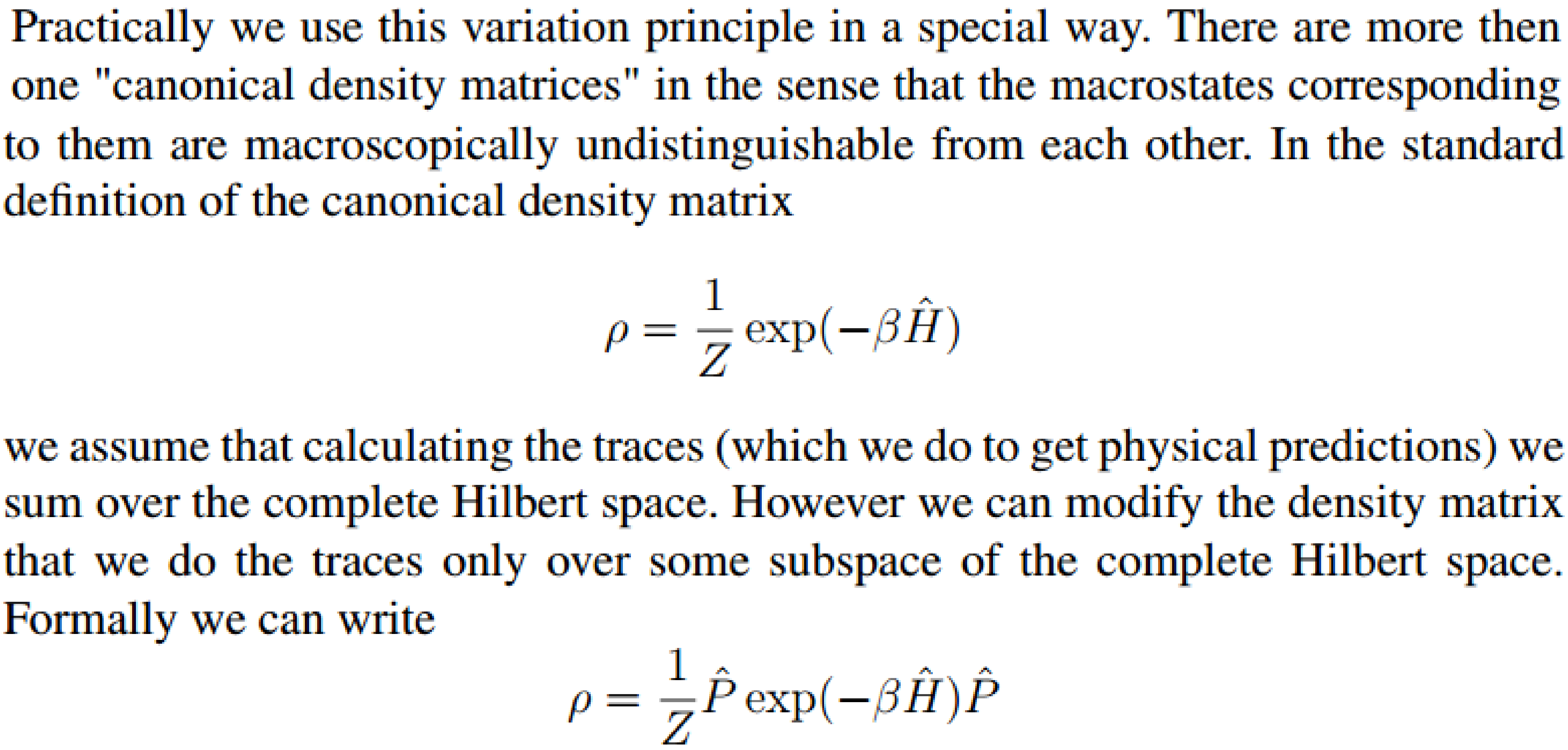
$$\rho = \frac{1}{Z}\exp(-\beta\hat{H})$$

and an arbitrary density matrix ρ' (defined on the whole Hilbert space). We get the inequality

$$\begin{aligned} S' = -Sp(\rho' \ln(\rho')) &\leq -Sp(\rho' ln(\rho)) \\ S' &\leq -Sp(\rho'(-\ln(Z) - \frac{\hat{H}}{T}) = +\ln(Z) + \frac{1}{T}E' \\ E - TS &\leq E' - TS' \\ F &\leq F' \end{aligned}$$

F' is different from canonical, so it is non-equilibrium free energy

How to calculate a non-equilibrium free energy at a given temperature



Practically we use this variation principle in a special way. There are more then one "canonical density matrices" in the sense that the macrostates corresponding to them are macroscopically undistinguishable from each other. In the standard definition of the canonical density matrix

$$\rho = \frac{1}{Z}\exp(-\beta\hat{H})$$

we assume that calculating the traces (which we do to get physical predictions) we sum over the complete Hilbert space. However we can modify the density matrix that we do the traces only over some subspace of the complete Hilbert space. Formally we can write

$$\rho = \frac{1}{Z}\hat{P}\exp(-\beta\hat{H})\hat{P}$$

Formally we can write

$$\rho = \frac{1}{Z}\hat{P}\exp(-\beta\hat{H})\hat{P}$$

where $\hat{P}$ is the projection operator projecting to the subspace considered. Usually we choose the subspace as the subspace on which all the states have specific value A of some (macroscopically measurable) variable $\hat{A}$. We denote the corresponding projection operators symbolically as

$$\hat{P}_{\hat{A}=A}$$

and we get

$$\rho_A = \frac{1}{Z_A}\hat{P}_{\hat{A}=A}\exp(-\beta\hat{H})\hat{P}_{\hat{A}=A}$$

we mean by this notation that, for example

$$Z_A = \sum_{i,A_i=A} exp(-\beta H_i)$$

We know that for a macroscopical system all sums are saturated by the subspace on which the value of some variable is equal to the macroscopical equilibrium value of that variable. In other words ρ is equivalent to ρ_A if $A = \langle A \rangle = Tr(\rho A)$.

Now we can use the value A as the variation parameter in the above described variation principle. So among the subspaces having specific fixed value A of the variable $\hat{A}$ we look for the subspace for which the corresponding free energy is minimal. The value A_{min} for which the minimum is achieved is the equilibrium value of the quantity $\hat{A}$

$$A_{min} = \langle A \rangle = Sp(\rho A)$$

This is the way how equilibrium values are often calculated in statistical physics. We minimize the non-equilibrium free energy corresponding to a general non-equilibrium value of the quantity of interest. The value of this quantity for which the minimum is achieved is its equilibrium value.

So far we have considered a variation procedure based on the inequality saying that the non-equilibrium free energy is greater then the equilibrium free energy **of the same system**.

Now we formulate an alternative variation principle based on comparing two **equilibrium** free energies. We start from the previous inequality but now as a trial density matrix we shall choose canonical (equilibrium) matrix corresponding, however, to different Hamiltonian $\hat{H}'$. We keep the notation

$$E' = Sp(\rho' \hat{H})$$

so that E' is the mean energy corresponding to the original Hamiltonian $\hat{H}$ in the trial state ρ'. Now we introduce a new quantity

$$E'' = Sp(\rho' \hat{H}')$$

it is the mean energy corresponding to the trial Hamiltonian $\hat{H}'$ in the trial state ρ'. Adding and subtracting the same term E'' to the previous inequality we get

$$\begin{aligned} E - TS &\leq E' - E'' + E'' - TS' \\ F &\leq (E' - E'') + F'' \end{aligned}$$

where F'' is the canonical (equilibrium) free energy of a trial system with the Hamiltonian $\hat{H}'$ Variation technology proceeds in a standard way. We introduce a free parameter α into the trial Hamiltonian $\hat{H}'$ and we get a variable upper estimate

$$F \leq (E'(\alpha) - E''(\alpha)) + F''(\alpha)$$

Classical and quantum statistics

Problem: how to define microcanonical ensemble in classical physics, which is continuous. For continuous system we generally cannot define uniform probability, since this notion does not survive a change of continuous variables.

Solution: only canonical transformations are allowed.

Let us have canonical coordinates p, q. We shall omit indexes everywhere, so p, q can represent many (10^{23}) variables with indexes. We have Hamiltonian $H(p, q)$ and the equations

$$\dot{p} = -\frac{\partial H}{\partial q} \qquad \dot{q} = \frac{\partial H}{\partial p}$$

We introduce new variables P, Q by transformations $P(p, q), Q(p, q)$. The new coordinates P, Q are called canonical, if

$$\dot{P} = -\frac{\partial \widetilde{H}}{\partial Q} \qquad \dot{Q} = \frac{\partial \widetilde{H}}{\partial p}$$

where

$$\widetilde{H}(P, Q) = H(p(P, Q), q(P, Q))$$

We get from definition of P

$$\dot{P} = \left(\frac{\partial P}{\partial q}\right)_p \dot{q} + \left(\frac{\partial P}{\partial p}\right)_q \dot{p} = \left(\frac{\partial P}{\partial q}\right)_p \left(\frac{\partial H}{\partial p}\right)_q - \left(\frac{\partial P}{\partial p}\right)_q \left(\frac{\partial H}{\partial q}\right)_p$$

We should get

$$\dot{P} = -\frac{\partial \widetilde{H}}{\partial Q} = -\left(\frac{\partial H}{\partial p}\right)_q \left(\frac{\partial p}{\partial Q}\right)_P - \left(\frac{\partial H}{\partial q}\right)_p \left(\frac{\partial q}{\partial Q}\right)_P$$

Comparing the two expressions we get two conditions of canonicity

$$\left(\frac{\partial P}{\partial q}\right)_p = -\left(\frac{\partial p}{\partial Q}\right)_P$$

$$\left(\frac{\partial P}{\partial p}\right)_q = \left(\frac{\partial q}{\partial Q}\right)_P$$

The two other conditions of canonicity can be get similarly considering $\dot{Q}$.

Now the Jacobian can be calculated by dividing the total canonical transformation into two subsequent canonical transformations

$$J = \frac{\partial(Q,P)}{\partial(q,p)} = \frac{\partial(Q,P)}{\partial(q,P)}\frac{\partial(q,P)}{\partial(q,p)}$$

$$J = \left(\frac{\partial Q}{\partial q}\right)_P \left(\frac{\partial P}{\partial p}\right)_q$$

Substituting the second condition of canonicity we get

$$J = \left(\frac{\partial Q}{\partial q}\right)_P \left(\frac{\partial q}{\partial Q}\right)_P = 1$$

We shall show now that the volume of the phase space is also invariant with respect to time development according to the equations of motion. So let we have initial point p, q, and the final point p', q' which is reached after an infinitesimal time development during the time dt.

$$q' = q + \frac{\partial H}{\partial p} dt \qquad p' = p - \frac{\partial H}{\partial q} dt$$

The Jacobian will be

$$J = \begin{vmatrix} \frac{\partial p'}{\partial q} & \frac{\partial p'}{\partial p} \\ \frac{\partial q'}{\partial q} & \frac{\partial q'}{\partial p} \end{vmatrix} = \begin{vmatrix} -\frac{\partial^2 H}{\partial q^2} dt & 1 - \frac{\partial^2 H}{\partial p \partial q} dt \\ 1 + \frac{\partial^2 H}{\partial q \partial p} dt & \frac{\partial^2 H}{\partial p^2} dt \end{vmatrix}$$

$$|J| = 1 + O(dt^2)$$

And so

$$\frac{dJ}{dt} = 0 \Rightarrow |J| = 1 = \text{const}$$

Now suppose we define a probability density in the phase space, which generally depends explicitly on time

$$\varrho(t,p,q)$$

Since probability has to be conserved, the continuity equation should be satisfied

$$\frac{\partial\varrho}{\partial t}+\frac{\partial}{\partial p}(\varrho\dot{p})+\frac{\partial}{\partial q}(\varrho\dot{q})=0$$

Substituting from the equations of motion we get

$$\frac{\partial\varrho}{\partial t}-\frac{\partial}{\partial p}(\varrho\frac{\partial H}{\partial q})+\frac{\partial}{\partial q}(\varrho\frac{\partial H}{\partial p})=0$$

The mixed derivatives drop out and we get

$$\frac{\partial\varrho}{\partial t}+\frac{\partial}{\partial p}\dot{p}+\frac{\partial}{\partial q}\dot{q}=0$$

That means that if we insert into ϱ the solutions of the equations of motion

$$\varrho(t,p(t),q(t))$$

we get the total time derivative is zero

$$\frac{d}{dt}\varrho(t)=0$$

Entropy for a classical system

If we just speak about probabilities, the choice of canonical coordinates makes things unique. If we want to introduce entropy we arrive at a problem of choosing the unit for the volume of the phase space. The discrete formula for the entropy

$$S = -\sum_i \ln(p_i) p_i$$

can be generalized to the case of continuous space and probability density as

$$S = -\int \ln(\rho)\rho d\Phi$$

where $d\Phi$ denotes the element of the phase space volume. However ρ is the probability density, that means probability divided by the phase space volume. So the normalization of the probability density depends on the choice of a unit for the phase space volume.

phase space volume. In the integrals defining probabilities we get something like

$$P(subset) = \int_{subset} \rho d\Phi$$

and the unit of the volume of the space, being in the numerator in $d\Phi$ and in the denominator in ρ cancels. In the integral defining the entropy, the phase space volume unit is present also in ρ under the logarithm $\ln(\rho)$ and does not cancel in the integral. Moreover, the phase space volume is not dimensionless, in one particle mechanics it has the dimension of distance $times$ momentum what in the usual units is Js. The logarithm is, however, well defined only for dimensionless quantities. So we have to express the phase space volume in some prescribed unit and take for ρ just the numerical factor. Changing the unit of the phase space volume does not change the values of probabilities, but it does change the value of entropy. Since changing the volume unit renormalizes the probability density ρ multiplicatively, and ρ appears under logarithm, the entropy gets additively renormalized. So classical entropy is well defined up to an arbitrary additive constant.

To choose the unit for the phase space volume we look for something with the dimension Js. The obvious candidate is the Planck constant $\hbar = 1.055 \times 10^{-34}$Js.

In quantum statistics we have discrete space of stationary states, their number can be counted in a unique way and it is a dimensionless quantity, so the quantum entropy is absolutely defined, without any additive arbitrariness[2].

We can ask whether we can choose the phase space volume unit in such a way that we get for the entropy in the classical case the same number as in the classical limit of quantum statistical physics. The answer is positive, the matching unit of the phase space can be found. We demonstrate it on the example of ideal gas.

In the classical limit of the quantum statistics of the ideal gas we get for entropy the well known Sackur-Tetrode formula

$$S_q = kN \ln \frac{V}{N} + kN\frac{3}{2} \ln kT + kN\frac{3}{2} \ln \frac{m}{2\pi\hbar^2} + \frac{5}{2}kN$$

In classical statistical physics we get the expression[3]

$$S_c = -k \int \frac{1}{N!} \prod_i d^3x_i d^3p_i \frac{N!}{V^N} C^{3N/2} \exp(-\sum_i \frac{p_i^2}{2mkT}) \ln(\frac{N!}{V^N} C^{3N/2} \exp(-\sum_i \frac{p_i^2}{2mkT}))$$

where

$$C = (\frac{1}{2\pi mkT})$$

$$S_c = -kN \ln N + kN + kN \ln V + \frac{3kN}{2} \ln(2\pi mkT) + \frac{3kN}{2}$$

$$S_c = kN \ln \frac{V}{N} + kN\frac{3}{2} \ln kT + kN\frac{3}{2} \ln(2\pi m) + \frac{5}{2}kN$$

The classical calculation and the classical limit of the quantum calculation differ by an additive term

$$S_q = S_c - 3kN \ln(2\pi\hbar)$$

The results will be identical if we write the integration measure as

$$\int \prod_i \frac{d^3x_i d^3p_i}{(2\pi\hbar)^3}$$

Metropolis algorithm

$$\bar{A} = \sum_i A_i p(i)$$

$$p(i) = \frac{1}{Z} \exp\left(-\frac{E_i}{kT}\right)$$

Blind Monte Carlo

$$\bar{A} = \sum_{i=\mathrm{rg}()} A_i p(i)$$

Importance sampling needed

Nature:

$$\bar{A} = \sum_{events} A_i$$

```
repeat{
$\xi_{\text{trial}} := \text{Generate\_trial\_state\_near\_to\_the\_actual\_state}(\xi_{\text{current}})$
if $(E(\xi_{\text{trial}}) < E(\xi_{\text{current}}))$
    then $\xi_{\text{current}} := \xi_{\text{trial}}$
    elseif $(\text{rnd}() < \exp(-\frac{E(\xi_{\text{trial}})-E(\xi_{\text{current}})}{T}))$
        then $\xi_{\text{current}} := \xi_{\text{trial}}$
print$(\xi_{\text{current}})$
}
```

The Metropolis algorithm is a special case of a Markov process. This is a name given to a random process where the probability distribution relevant for the n-th step depends only on the value drawn in the $(n-1)$-th step, and (generally) on the value n. If the probability does not depend on n, we speak about a homogenous Markov process.

The dynamics of a homogenous Markov process in discrete space is completely given by a transfer matrix

$$P_{ij}$$

The matrix element P_{ij} denotes[1] the probability that the actual state j is (in the actual step) changes into the state i. It is the probability of the transition $j \to i$.

The matrix elements P_{ij} are non-negative and satisfy the normalization condition

$$\sum_i P_{ij} = 1$$

The ensemble probability that in the n-th step the i state will be generated we denote by $p_n i$. Then from the definition of the transfer matrix it is clear that the following relation holds.

$$p_n(i) = \sum_j P_{ij} p_{n-1}(j)$$

This recurrence relation can be explicitly solved and ensemble probability after n steps can be written in terms of the ensemble probability at the beginning[2].

$$p_n(i) = \sum_j (P^n)_{ij} p_0(j)$$

As usual in matrix linear algebra, investigation of the properties of the expression is based on the technique of eigenstates and eigenvalues. We shall not do a general analysis, we just demonstrate the procedure on the simple two-dimensional case. So let us have two states ”1” and ”2” with energies E_1 a E_2, where the notation is chosen so that$E_1 < E_2$. Then the Metropolis algorithm is given by the probabilities

$$\begin{aligned}
p(2 \to 1) &= 1 = P_{12} \\
p(2 \to 2) &= 0 = P_{22} \\
p(1 \to 2) &= p = \exp(-\frac{E_2 - E_1}{kT}) = P_{21} \\
p(1 \to 1) &= 1 - p = P_{11}
\end{aligned}$$

The corresponding transfer matrix is

$$P = \begin{pmatrix} 1-p & 1 \\ p & 0 \end{pmatrix}$$

Notice that the matrix is not symmetric, therefore the left eigenvactors are not equal to the right eigenvectors. We clearly need the right eigenvectors, they are easily found in this simple case. They are

$$\begin{pmatrix} \frac{1}{1+p} \\ \frac{p}{1+p} \end{pmatrix} \quad \text{with eigenvalue } 1$$

and

$$\begin{pmatrix} 1 \\ -1 \end{pmatrix} \quad \text{with eigenvalue } -p$$

Notice, that the first eigenvector has positive components and we have normalized it to the sum 1. So this vector can be interpreted as the ensemble probability vector. The second eigenvector does not have non-negative components, their sum is equal to zero. It is a ”non-physical” vector, its components cannot be interpreted as ensemble probabilities. However, the vector is useful, since it is linearly independent from the first one and they together form a base in the space of ensemble probabilities. Each ”physical” vector of ensemble probabilities can be written as a linear combination of the two eigenvectors.

$$\begin{pmatrix} p(1) \\ p(2) \end{pmatrix} = c_1 \begin{pmatrix} \frac{p}{1+p} \\ \frac{1}{1+p} \end{pmatrix} + c_2 \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

$$\begin{pmatrix} p(1) \\ p(2) \end{pmatrix} = c_1 \begin{pmatrix} \frac{p}{1+p} \\ \frac{1}{1+p} \end{pmatrix} + c_2 \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

The coefficients are easily found

$$\begin{aligned} c_1 &= 1 \\ c_2 &= \frac{p(1) - pp(2)}{1+p} \end{aligned}$$

So we express the initial ensemble probability in the base of the eigenvectors and after n steps we get

$$\begin{pmatrix} p_n(1) \\ p_n(2) \end{pmatrix} = 1^n \begin{pmatrix} \frac{1}{1+p} \\ \frac{p}{1+p} \end{pmatrix} + c_2(-p)^n \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

In the limit $n \to \infty$ we get

$$\begin{pmatrix} p_\infty(1) \\ p_\infty(2) \end{pmatrix} = \begin{pmatrix} \frac{1}{1+p} \\ \frac{p}{1+p} \end{pmatrix} \qquad p = \exp(-\frac{E_2 - E_1}{kT})$$

Supplements to thermodynamics

Basic notions and laws of thermodynamics

First postulate: an isolated system spontaneously changes its (macro)state until it arrives at a steady state which we call the equilibrium state. The typical time needed to arrive to the equilibrium state is called the relaxation time. In equilibrium state all the (macroscopic)parameters of the system are constant with time. Sometimes the equilibrium time to reach the true equilibrium state is unrealistically high, the system might for a very long time stay in a metastable state (like glass).

We have the notion of external parameters (like volume or external magnetic field). Intuitively these are parameters which can be externally set by the experimentalist.

Going beyond the thermodynamics these are the parameters which set the energy levels of the (quantum) system. The energy of the system in the microstate i depends on the value of some external parameter V like $E_i(V, \ldots)$. Then the conjugate force is

$$p_i = -\frac{\partial E_i}{\partial V}$$

and the corresponding macroscopic parameter is

$$p = \langle p_i \rangle$$

Staying at the level of phenomenological thermodynamics we have to identify the macroscopic parameter conjugate to the external parameter V for example by observing it as a force to be applied to the system from outside to keep the external parameter V constant at the set value. Then the work done by an external agent changing the set value of V by dV is

$$\delta A' = -p\,dV$$

and the work done by the system upon the external agent is

$$\delta A = p\,dV$$

Now we need the notion of thermal contact. It is a way of interaction of two systems such that non of their (individual) external parameter is changed (but they still can exchange energy. We shall discuss energy in detail when we shall speak on the first law of thermodynamics.

Then we need the notion of ”two systems being in thermal equilibrium with each other”. We say the two systems are in thermal equilibrium with each other when, after bringing them to thermal contact, nothing macroscopic happens, that is the compound system is immediately in thermal equilibrium.

Zeroth law of thermodynamics

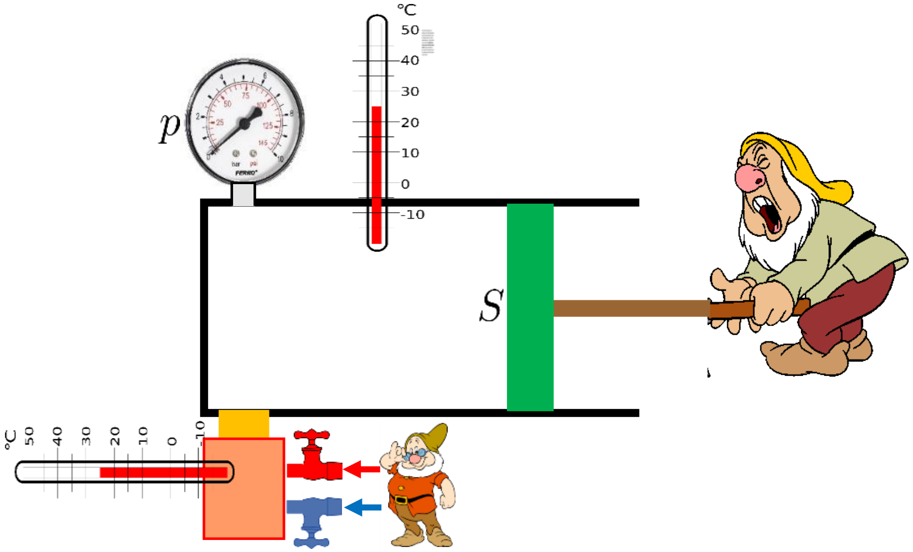
Then there is the zeroth law of thermodynamics, definition of temperature. We experimentally find and then postulate: The notion of of ”being in thermal equilibrium with each other” is transitive.

The properties of reflexivity and symmetry are automatic by definition, so the zeroth law of thermodynamics effectively says ”being in thermal equilibrium is equivalence (in mathematical sense). So all the systems can be classified into (not-overlapping) classes of equivalence. And the classes can be labeled. The label is called ”the temperature”. Temperature defined in this way is not unique: the only requirement is that different classes have different labels. We, however, usually require more, we need property of continuity in labeling the classes. We need to find some practical (experimental method) to label the classes.

To do that we had to choose a suitable ”thermoscope”. By that we mean some system which can be find in many equivalence classes (in each class the system is, of course, in a different (macro)state. Ideally the system should be found in states having the same values of external parameters, differing by the value of some internal parameter (like pressure) in different equivalence classes. The system (and its chosen parameter) can be used as thermoscope: bringing it to thermal contact with any other system and reading the value of pressure we can recognize to which equivalence class the ”measured” system belongs.

the readings of our thermoscope to label the classes: so we use it effectively as a thermometer. Most often we first choose some (arbitrary) calibration to define our temperature scale. Like this: we take the thermoscope of a defined prescribed size, put it into contact with a melt of ice and water and label the pressure reading by zero. Then we put it to contact with boiling water and label the pressure reading as 100. The other pressure readings are simply linearly interpolated or extrapolated. The temperature defined in this way is rather arbitrary but can serve our purposes and is ”continuously defined”. Topology is introduced into the space of equivalence classes: the two classes are considered to be close to each other if the thermoscope readings are close to each other.

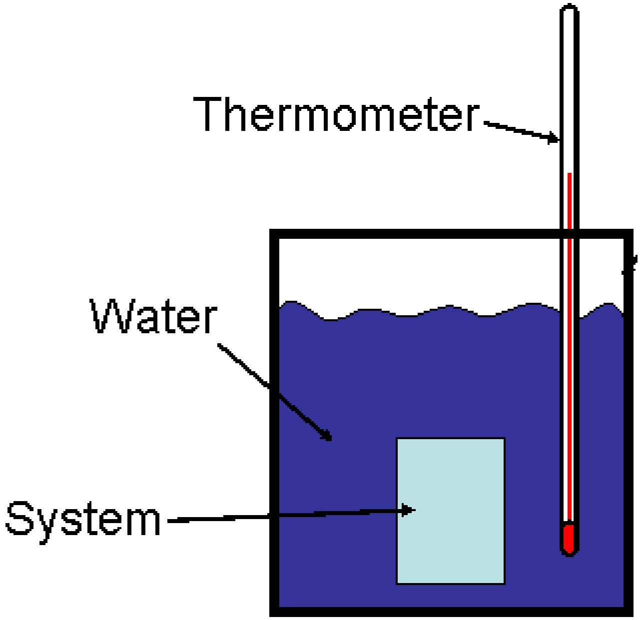
Reversible processes in gas – experimental setup



The first dwarf (**piston-pusher**) can set any V value (in the vicinity of the current value) by moving carefully the piston by small amount. He has to act by a proper force on the piston: a bit higher then corresponding to the gas pressure (if he wants to decrease V) or a bit lower then corresponding to gas pressure (if he wants to increase V).

The second dwarf (**boiler attendant**) can set any pressure p (in the vicinity of the current value) indirectly by increasing or decreasing the current temperature of the gas. To control temperature in a reversible manner is a tricky task. Here it is how it can be done. The boiler attendant control roughly the water temperature (by manipulating the valves with hot and cold water) in an external boiler which serves as a heat exchanger. The boiler is in thermal contact with the gas through an interface with low coefficient of heat transfer. If the boiler water is a bit hotter then the gas, energy slowly flows from the boiler to the gas thus increasing its temperature. If the boiler is colder then the gas, energy slowly flows from the gas into the boiler thus decreasing slowly the gas temperature. The **energy exchange** between the boiler and the gas is **an irreversible process** (the boiler and the gas are not in the state of mutual equilibrium), **but the gas remains in internal equilibribrium** because the energy transfer is slow.

Calorimetry



Quantitative experiments about heating or cooling were done using calorimeters.
A calorimeter is a thermally isolated vessel. Inside the vessel we bring into contact two physical objects having initially different temperatures. Typically one object is a liquid like water, the other may be some solid body. After some time the thermal equilibrium is established, both object having the same temperature which we measure.

So the experiment looks like this
Water: mass m_1, initial temperature t_1. Solid body: mass m_2, initial temperature $t_2 > t_1$. The final temperature is t.
Empirical facts show that there is a material constant for water c_1 and for the body c_2, so that a following calorimetric equation holds

$$\boxed{m_1 c_1 (t - t_1) = m_2 c_2 (t_2 - t)}$$

The material constants c_1, c_2 are called specific heat constants.

Calorimetric equation as a conservation law

Calorimetric equation can be rewritten as: $m_1c_1t_1 + m_2c_2t_2 = m_1c_1t + m_2c_2t$

Inspecting that equation it looks like a conservation law. Some quantity calculated like mct is the same at the beginning an at the end of the process. This something looks like to be hidden inside the objects but distributed differently between the two objects during the process, the total sum being conserved.

We certainly cannot resist the temptation to introduce a new physical quantity: **the heat**. The heat is hidden inside the objects, can be transferred between the objects, the total amount of heat is conserved. More heat inside the object means higher temperature. It lead to terminology like **heat transfer, loss of heat, "heat is transferred through conduction, flow or radiation"**. **All this was found to be wrong!**

The crucial point was, that the first calorimetric measurements were done using fluids or solid bodies both having the property that they change their volumes with temperature only slightly. The situation is dramatically different with gases.

We have seen during a qualitative discussion of an adiabatic process **that we can change the temperature of some gas just by changing its volume by the "piston pusher" who is acting by fore on a moving piston therefore performing mechanical work.**

So the lesson is simple, to change the temperature, we do not need to increase the "heat content" of the body, we **can just perform mechanical work**.

Phenomenological units of heat

Historically, when people defined heat as “something that is conserved according to the calorimetric equation

$$m_1c_1(t-t_1)=m_2c_2(t_2-t)$$

it was obvious, that there is not any “natural” physical unit of heat, The content of heat in a body defined as mct is not well defined until we define the physical unit for the heat capacity.

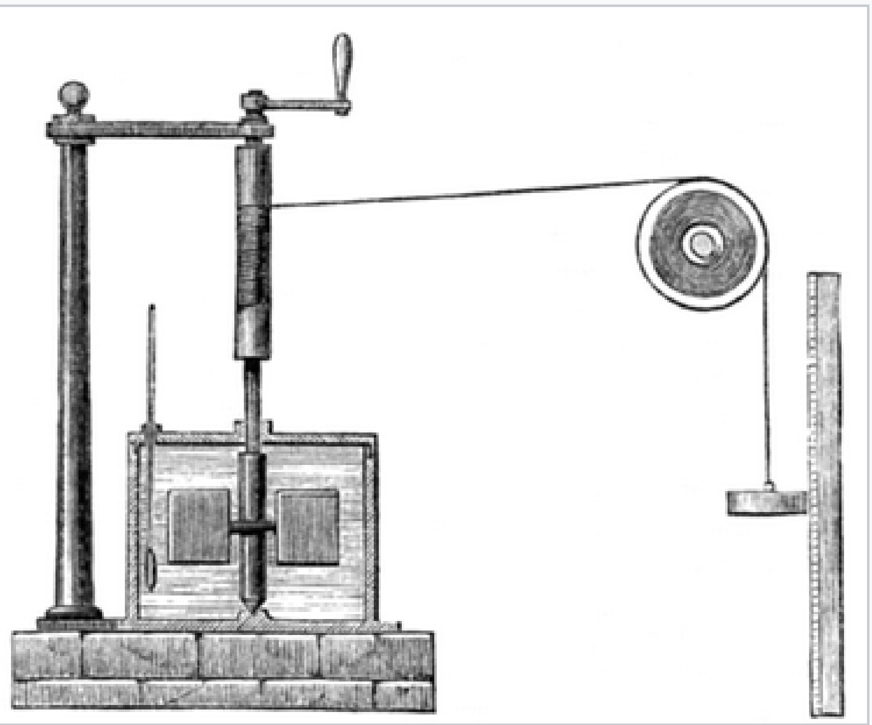
It is obvious, that if we multiply the heat capacities of all materials by the same constant, the calorimetric equation would still hold, That means the unit for heat capacity can be arbitrary, since the heat capacity and/or heat itself is (at this level of physics knowledge) not directly connected to any other physical quantity (with already defined physical unit).

So the unit of heat was chosen arbitrarily to be one calorie “cal” defined as the amount of heat needed to increase the temperature of one gram of water by one degree.

Therefore the specific heat of water was defined as 1 cal/g/K. The specific heat of any other material can then be measured by a suitable calorimetric measurement.

Mechanical equivalent of heat

Benjamin Thompson, Count Rumford, had observed the frictional heat generated by boring cannon at the arsenal in Munich, Germany circa 1797. More prcecise measurements were done by James Prescott Joule in the 1840s.



Joule's apparatus for measuring the mechanical equivalent of heat in which the "work" of the falling weight is converted into the "heat" of agitation in the water.

Finally it was found that the same increase of temperature of some object as achieved by “transfer of heat” of **1 cal can be achieved by performing a mechanical work of 4186 J**.

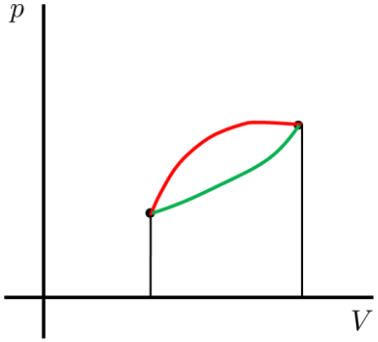
So it was found that the “amount od heat” can be more naturally measured by the units of work, Joules. In this way a “mechanical equivalent of heat” was found. **Calories can be universally converted to Joules by the conversion constant 4186 J/cal.**

Work and heat for reversible processes in gas

We have found that the work performed by the “piston-pusher” during a reversible process changing the state of gas from the state “1” to the state “2” can be calculated as

$$A = \int_1^2 p dV$$

and this work depends on “the trajectory” of the process between the initial and final states, as it is clearly seen from the figure below.



The obvious question is how much heat is to be provided by the boiler attendant on a particular process (trajectory). A careful investigation of this problem has lead to precise specification of the concept of energy formulating the law called “first law of thermodynamics”.

First law of thermodynamics

For any process between the two states "1" and "2" we can calculate the work performed by the piston pusher

$$A = \int_1^2 p dV$$

and we can measure the heat "transferred" to the gas by calorimetric consideration of the actions performed by the boiler attendant. The following empirical law was established.

The sum of mechanical work and heat (converted to Joules) depends only on the initial and final states of the system and so this sum is the same for any (reversible) process (trajectory) between those two states.

Let us stress again: the work by itself and the heat by itself depend on the specific trajectory, their sum does not.

This means that the sum of work and heat must be calculable from the characteristics of the initial and final states only.

This means that we can define a physical quantity of state in such a way, that the difference of this quantity between the two states is equal to the sum of work and heat performed during (any reversible) process between the two states.

First law of thermodynamics

Repeating: the first law of thermodynamics says, that there must be a physical quantity of state such that the sum of work and heat is calculable as a difference of this quantity between the final and initial state.

This quantity was given a name: **energy**. More precisely, the phenomenological thermodynamics used the name **internal energy**. The reason perhaps was, that it was not obvious that the state function "energy" as found by thermodynamical considerations has anything common with the quantity "energy" as found in the studies of Newtonian mechanics of particles.

Of course, there was a connection: the internal energy can be changed at the expense of mechanical work, the same work as found in classical mechanics.

Only after molecules were discovered, it was clear that the "internal energy" is just a "standard" mechanical energy of molecules, just macroscopically "not directly visible".

In what follows we shall just use the terminology "energy", without the attribute internal.

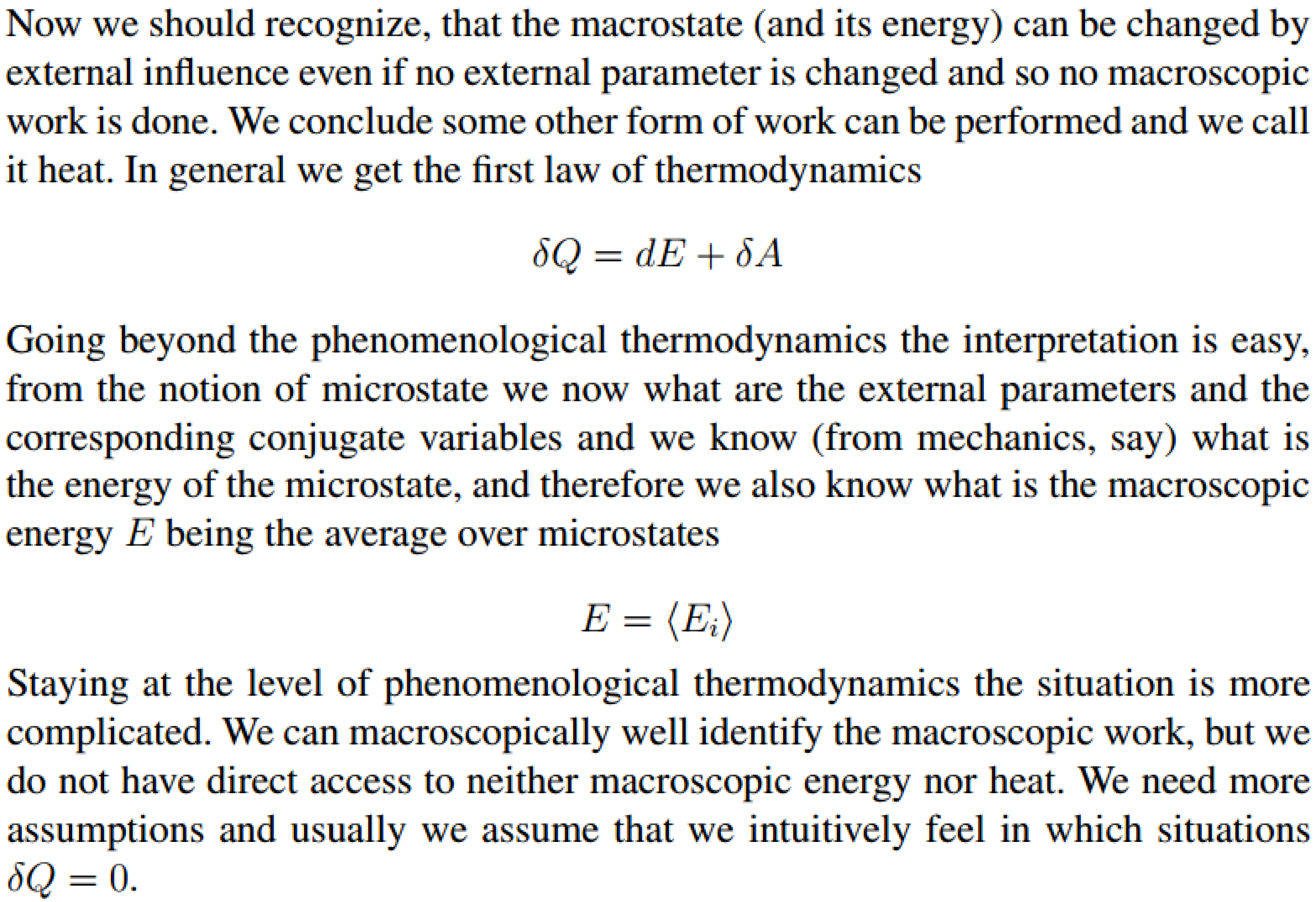
First law of thermodynamics

The obvious task for a phenomenologist is to find the formula for energy as a function of quantities defining the macrostate.

Obviously, the first law of thermodynamics can determine only the difference of energies of twos states, so energy in thermodynamics is defined up to an arbitrary additive constant.

The phenomenologist has to define some arbitrary reference macrostate whose energy is set to 0 by definition. Then he has to measure the sum of work and heat going from the reference state to an arbitrary macrostate. Based on experimental data he has to “guess” a formula for energy.

A theoretician, who already knows that behind the macrostate there is some microstate of molecules can in principle calculate the (mechanical) energy of molecules in the chosen representing microstate and**, based on statistical considerations, express the energy of the microstate through the macroscopic quantities defining the macrostate considered**.



Now we should recognize, that the macrostate (and its energy) can be changed by external influence even if no external parameter is changed and so no macroscopic work is done. We conclude some other form of work can be performed and we call it heat. In general we get the first law of thermodynamics

$$\delta Q = dE + \delta A$$

Going beyond the phenomenological thermodynamics the interpretation is easy, from the notion of microstate we now what are the external parameters and the corresponding conjugate variables and we know (from mechanics, say) what is the energy of the microstate, and therefore we also know what is the macroscopic energy E being the average over microstates

$$E = \langle E_i \rangle$$

Staying at the level of phenomenological thermodynamics the situation is more complicated. We can macroscopically well identify the macroscopic work, but we do not have direct access to neither macroscopic energy nor heat. We need more assumptions and usually we assume that we intuitively feel in which situations $\delta Q = 0$.

The problem is that an arbitrary final state cannot be reached from the reference state via the quasi-static adiabatic curve.

Energy conservation law holds for irreversible processes as well. So we can measure in principle the energy of any macrostate.

We first go from the reference state by adiabatic process to the state which has the same values of the external parameters as the desired final state, but is still a different state differing for example by the value of the conjugated parameters (practically it means by temperature). Now we keep the external parameters constant and perform heat to get to the desired final state. We perform (”produce”) heat by an irreversible process (Joule) performing mechanical work on some other (macroscopic but very small with respect to our system of interest) auxiliary system which is in contact with our system and exchanges heat with it, with the auxiliary system performing an irreversible process where external work is performed upon it. The auxiliary system is small, so even if its temperature is gradually increasing its energy increases by negligible amount.

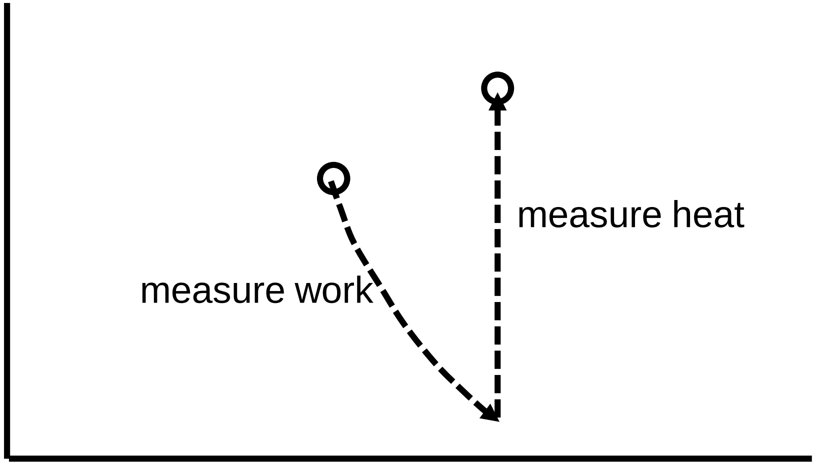
Historically the procedure was still different from the procedure using the Joule irreversible (work to heat transfer) process as described above.
The notion of heat was older then the fact that it can be measured using the units for work. The notion of heat developed in **calorimetric measurements**.

$$c_1(t_1 - t) = c_2(t - t_2)$$

Problem: heat capacity may depend on temperature, so the more realistic calorimetric equation is

$$\int_{t_1}^{t} c_1(t)dt = \int_{t}^{t_2} c_2(t)dt$$

Idea: parametrize heat capacity by a function with a few free parameters, then perform a fit.



How to measure heat. Fixed volume, no work. Contact with hot water until desired temperature reached. Measure heat exchanged by measuring the hot water the hot water final temperature.

Historical problem: work and heat measured by different units.

First law of thermodynamics:
there exists a universal unit converting constants that the sum lof work and converted calories to get from an initial state to a final state does not depend on trajectory

Therefore there exists a state function E, which can be determined as

$$\Delta E = W + c * Q$$

thus the energy conservation law was discovered: any energy change is balanced by the work (mechanical work + heat). To perform work there must be an agent. The work performed by an agent on the system is the same but with an opposite sign as the work performed by the system on the agent. **Therefore the energy change of the agent is just opposite to the energy change of the system and the overall energy of system + agent is conserved**.

Second law of thermodynamics:

$$\Delta S = \int \frac{\delta Q}{T}$$

does not depend on the trajectory. So entropy is a state function.

Cyclic reversible engine:

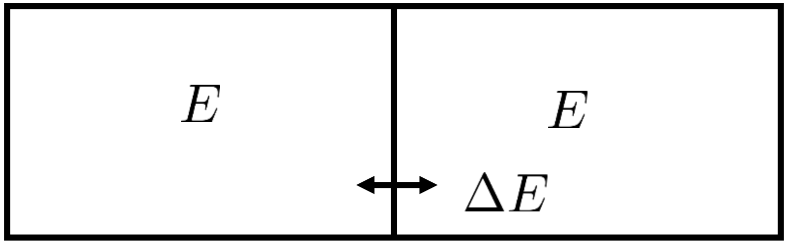
$$\Delta S = 0 \Rightarrow Q_1/T_1 + Q_2/T_2 = 0 \Rightarrow \eta = \frac{Q_1 + Q_2}{Q_1} = \frac{T_1 - T_2}{T_1}$$

Properties of thermodynamic potentials

We now investigate the situation when all the parameters are extensive quantities, we shall generically call them E and V. We are going to prove that the entropy

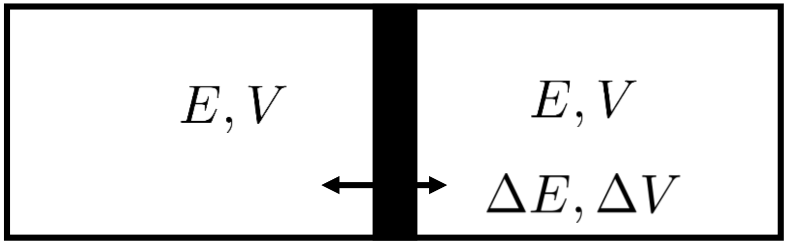
$$S(E, V)$$

is a concave function of its parameters.



$$S(E - \Delta E) + S(E + \Delta E) \leq 2S(E)$$

this is just the definition of the concave function of one variable.



$$S(E - \Delta E, V - \Delta V) + S(E + \Delta E, V + \Delta V) \leq 2S(E, V)$$

Expanding into Taylor serie we get

$$\frac{\partial^2 S}{\partial E^2}(\Delta E)^2 + \frac{\partial^2 S}{\partial V^2}(\Delta V)^2 + 2\frac{\partial^2 S}{\partial E \partial V}\Delta E \Delta V \leq 0$$

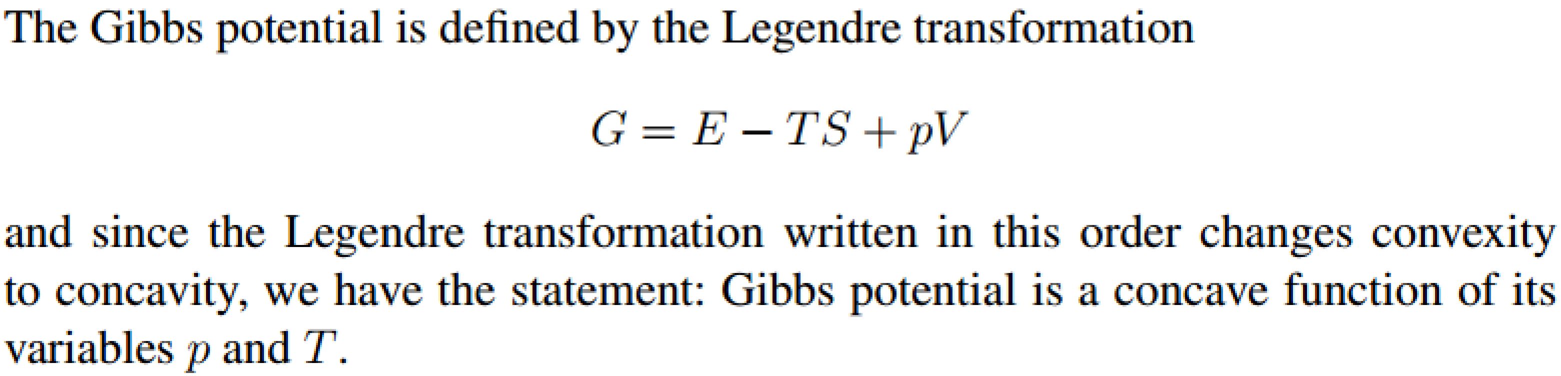
$$\begin{pmatrix} \frac{\partial^2 S}{\partial E^2} & \frac{\partial^2 S}{\partial E \partial V} \\ \frac{\partial^2 S}{\partial E \partial V} & \frac{\partial^2 S}{\partial V^2} \end{pmatrix}$$

must be negative definite, so we get the conditions for the concavity

$$\frac{\partial^2 S}{\partial E^2} \leq 0, \quad \frac{\partial^2 S}{\partial V^2} \leq 0, \quad \left(\frac{\partial^2 S}{\partial E \partial V}\right)^2 - \frac{\partial^2 S}{\partial E^2}\frac{\partial^2 S}{\partial V^2} \leq 0$$

These conditions are also called conditions of stability.

What was essential to derive the concavity property was that the parameters were extensive.
Other thermodynamic potentials do not have extensive natural parameters. However, they are derived by Legendre transformation. And the Legendre transformation preserves (or changes the sighn) of concavity. So we get concavity property for other thermodynamic potentials as well.



The Gibbs potential is defined by the Legendre transformation

$$G = E - TS + pV$$

and since the Legendre transformation written in this order changes convexity to concavity, we have the statement: Gibbs potential is a concave function of its variables p and T.

We remind the "statistical thermodynamics definitions"

$$\frac{1}{T} = \left(\frac{\partial S}{\partial E}\right)_{V,N}$$

$$\frac{p}{T} = \left(\frac{\partial S}{\partial V}\right)_{E,N}$$

$$\frac{\mu}{T} = -\left(\frac{\partial S}{\partial N}\right)_{E,V}$$

differentiating the first equation by E we get

$$-\frac{1}{T^2}\left(\frac{\partial T}{\partial E}\right)_{V,N} = \left(\frac{\partial^2 S}{\partial E^2}\right)_{V,N} \leq 0$$

where we used the condition of concavity so we get

$$\left(\frac{\partial T}{\partial E}\right)_{V,N} \geq 0$$

Let us consider two systems (S_1, E_1, T_1) and (S_2, E_2, T_2). Let put them into thermal contact, however with the wall between them that is only slightly diathermic, so the energy flow between the two system is small. Then, even the global process is irreversible, both subsystems individually are in equilibrium state. So the total entropy can be calculated as the sum of two equilibrium entropies. The system as a whole is isolated so the total energy is conserved. Therefore for the energy changes of the subsystems we get

$$dE_1 = -dE_2$$

The change of the total entropy will be

$$dS = \frac{\partial S_1}{\partial E_1} dE_1 + \frac{\partial S_2}{\partial E_2} dE_2 = \left(\frac{\partial S_1}{\partial E_1} - \frac{\partial S_2}{\partial E_2} \right) dE_1 = \left(\frac{1}{T_1} - \frac{1}{T_2} \right) dE_1 > 0$$

If $T_2 > T_1$ then $dE_1 > 0$ so the energy flows from the hotter subsystem to the colder one.

Now about the derivation of the equation of state.

We start with the equation

$$\frac{1}{T} = \left(\frac{\partial S}{\partial E}\right)_{V,N}$$

The condition of convexity gives

$$\left(\frac{\partial^2 S}{\partial E^2}\right)_{V,N} < 0$$

so the first equation can be inverted and the energy can be expressed in terms of temperature and volume. This expression then can be inserted into the equation

$$\frac{p}{T} = \left(\frac{\partial S}{\partial V}\right)_{E,N}$$

Energy is eliminated and we get p in terms of V, T, that is the equation of state.

We remind here the definition of thermodynamic potentials

$$F = E - TS$$

$$G = E - TS + pV$$

$$\Omega = E - TS - \mu N$$

In the last equation we have introduced so called grand potential, which is relevant for the grand-canonical distribution: the grand statistical sum can be expressed in terms of the grand potential

$$\mathcal{Z} = \sum \exp(\frac{\mu N_i - E_i}{kT}) = \exp(-\frac{\Omega}{kT})$$

The first and second law of thermodynamics give (for general, even irreversible)processes

$$TdS \geq dE + pdV - \mu dN$$

$$dE \leq TdS - pdV + \mu dN$$

$$dF \leq -SdT - pdV + \mu dN$$

$$dG \leq -SdT + Vdp + \mu dN$$

$$d\Omega \leq -SdT - pdV - Nd\mu$$

The Gibbs potential is a function of intensive variables p, T and one extensive variable N. So scaling the system by λ we get

$$\lambda G(p, T, N) = G(p, T, \lambda N)$$

Differentiating by λ and putting $\lambda = 1$ we get (for equilibrium potential)

$$G = N \frac{\partial G}{\partial N} = N\mu$$

and

$$dG = dN\mu + Nd\mu = -SdT + Vdp + \mu dN$$

and we get finally the Gibbs Duham relation

$$Nd\mu + SdT - Vdp = 0$$

$$dG \leq -SdT + Vdp + \mu dN$$

$$\frac{\partial^2 G}{\partial p \partial T} = \frac{\partial^2 G}{\partial T \partial p}$$

Since

$$S = -\left(\frac{\partial G}{T}\right)_{p,N}$$

$$V = \left(\frac{\partial G}{p}\right)_{T,N}$$

$$-\left(\frac{\partial S}{\partial p}\right)_{T,N} = \left(\frac{\partial V}{\partial T}\right)_{p,N}$$

This relation is called the Maxwell relation and similar Maxwell relations can be obtained doing the second derivatives of other potential.

The second derivatives of the thermodynamic potentials are often easily measurable physical quantities, so they often get a special name.

For example

$$-T\left(\frac{\partial^2 F}{\partial T^2}\right)_V = T\left(\frac{\partial S}{\partial T}\right)_V = \left(\frac{\delta Q}{\partial T}\right)_V = C_V$$

So we get for the specific heat at constant volume

$$C_V = -T\left(\frac{\partial^2 F}{\partial T^2}\right)_V$$

Expansion coefficient

$$\alpha = \frac{1}{V}\left(\frac{\partial V}{\partial T}\right)_p$$

Coefficient of isothermal compressibility

$$\kappa_T = -\frac{1}{V}\left(\frac{\partial V}{\partial p}\right)_T$$

Coefficient of adiabatic compressibility (for constant entropy)

$$\kappa_S = -\frac{1}{V}\left(\frac{\partial V}{\partial p}\right)_S$$

In addition to Maxwell relations we also often use the following properties of partial derivatives

$$\left(\frac{\partial x}{\partial y}\right)_z \left(\frac{\partial y}{\partial x}\right)_z = 1$$

which is simply the theorem about the derivative of inverse function.

$$\left(\frac{\partial x}{\partial y}\right)_z \left(\frac{\partial y}{\partial z}\right)_x \left(\frac{\partial z}{\partial x}\right)_y = -1$$

This property can be easily derived considering

$$dy = \left(\frac{\partial y}{\partial x}\right)_z dx + \left(\frac{\partial y}{\partial z}\right)_x dz$$

Now we put $dy = 0$ and we get

$$0 = \left(\frac{\partial y}{\partial x}\right)_z + \left(\frac{\partial y}{\partial z}\right)_x \left(\frac{\partial z}{\partial x}\right)_y$$

which is just the required relation if we invert the partial derivative $\left(\frac{\partial y}{\partial x}\right)_z$.

We start with the entropy expressed as a function of (non/natural) variables p, T.

$$dS = \left(\frac{\partial S}{\partial T}\right)_p dT + \left(\frac{\partial S}{\partial p}\right)_T dp$$

$$TdS = C_p dT + T\left(\frac{\partial S}{\partial p}\right)_T dp$$

Now we use the Maxwell relation

$$\left(\frac{\partial S}{\partial p}\right)_T = -\left(\frac{\partial V}{\partial T}\right)_p$$

and we get

$$TdS = C_p dT - T\left(\frac{\partial V}{\partial T}\right)_p dp$$

$$TdS = C_p dT - T\left(\frac{\partial V}{\partial T}\right)_p dp$$

Now we use the above general equation for a isochoric process at constant V. We get

$$T\left(\frac{\partial S}{\partial T}\right)_V = C_p + -T\left(\frac{\partial V}{\partial T}\right)_p \left(\frac{\partial p}{\partial T}\right)_V$$

So we get

$$C_p = C_V + T\left(\frac{\partial V}{\partial T}\right)_p \left(\frac{\partial p}{\partial T}\right)_V$$

Now we substitute from the general relation

$$\left(\frac{\partial p}{\partial T}\right)_V + \left(\frac{\partial V}{\partial T}\right)_p \left(\frac{\partial p}{\partial V}\right)_T = 0$$

and using the definition od expansion coefficients we get the final formula

$$C_p - C_V = \frac{TV\alpha^2}{\kappa_T}$$

We shall now investigate the consequences of the concavity of G, we get so called stability conditions. The concavity conditions are

$$\frac{\partial^2 G}{\partial T^2} \leq 0, \quad \frac{\partial^2 G}{\partial p^2} \leq 0, \quad \left(\frac{\partial^2 G}{\partial p \partial T}\right)^2 - \frac{\partial^2 G}{\partial T^2}\frac{\partial^2 G}{\partial p^2} \leq 0$$

From the first equation we get

$$\frac{\partial}{\partial T}\left(\frac{\partial G}{\partial T}\right)_p = -\left(\frac{\partial S}{\partial T}\right)_p = -\frac{1}{T}C_p \leq 0$$

and so we get the stability condition

$$C_p \geq 0$$

$$\frac{\partial^2 G}{\partial p^2} \leq 0$$

From the second concavity condition we get

$$\frac{\partial}{\partial p}\left(\frac{\partial G}{\partial p}\right)_T = V\frac{1}{V}\left(\frac{\partial V}{\partial p}\right)_T = -V\kappa_T \leq 0$$

And we get another stability condition

$$\kappa_T \geq 0$$

what means that increasing pressure must decrease the volume.

From the third concavity condition we get

$$\left(\frac{\partial^2 G}{\partial p \partial T}\right)^2 - \frac{\partial^2 G}{\partial T^2}\frac{\partial^2 G}{\partial p^2} \leq 0$$

$$\left(\frac{\partial^2 G}{\partial p \partial T}\right)^2 - (-\frac{C_p}{T})(-V\kappa_T) \leq 0$$

$$\frac{C_p V \kappa_T}{T} - \left(\left(\frac{\partial V}{\partial T}\right)_p\right)^2 \geq 0$$

$$C_p - \frac{\alpha^2 V T}{\kappa_T} \geq 0$$

So we get another stability condition

$$C_V \geq 0$$

Free expansion

We consider isoleted system of a gas performing free expansion into vacuum. It is an irreversible process during which the volume is increased $V \mapsto (V + dV)$. Since the system is isolated, its energy remains constant $dU = 0$. Both the initial and the final state are equilibrium states, for which we can write

$$dU = \left(\frac{\partial U}{\partial T}\right)_V dT + \left(\frac{\partial U}{\partial V}\right)_T dV$$

we put $dU = 0$ and get

$$\left(\frac{\partial T}{\partial V}\right)_U = -\frac{\left(\frac{\partial U}{\partial V}\right)_T}{\left(\frac{\partial U}{\partial T}\right)_V}$$

The l.h.s. is called the Joule coefficient and describes how much the temperature is changed during free expansion.

From molecular consideration it is clear, that for ideal gas

$$\left(\frac{\partial U}{\partial V}\right)_T = 0$$

$$TdS = dU + pdV$$

$$\left(\frac{\partial S}{\partial V}\right)_T = \frac{1}{T}\left(\frac{\partial U}{\partial V}\right)_T + \frac{p}{T}$$

Now we use the Maxwell relation (derived from the F potential)

$$\left(\frac{\partial S}{\partial V}\right)_T = \left(\frac{\partial p}{\partial T}\right)_V$$

and get

$$\left(\frac{\partial p}{\partial T}\right)_V = \frac{1}{T}\left(\frac{\partial U}{\partial V}\right)_T + \frac{p}{T}$$

$$\left(\frac{\partial U}{\partial V}\right)_T = T\left(\frac{\partial p}{\partial T}\right)_V - p$$

$$\left(\frac{\partial U}{\partial V}\right)_T = T\left(\frac{\partial p}{\partial T}\right)_V - p$$

The trick was that using the Maxwell relation we got rid of the calorical variables and everything on the r.h.s. can be evaluated just from the equation of state. If we use the equation of state for the ideal gas

$$pV = RT$$

we finally get

$$\left(\frac{\partial U}{\partial V}\right)_T = 0$$

IF we consider the Van der Waals gas

$$\left(p+\frac{a}{V^2}\right)(V-b)=RT$$

$$p=-\frac{a}{V^2}+\frac{RT}{V-b}$$

$$\left(\frac{\partial p}{\partial T}\right)_V=\frac{R}{V-b}$$

$$\left(\frac{\partial U}{\partial V}\right)_T=0=\frac{TR}{V-b}+\frac{a}{V^2}-\frac{RT}{V-b}=\frac{a}{V^2}>0$$

to evaluate the Joule coefficient for the Van der Waals gas we still need

$$\left(\frac{\partial U}{\partial T}\right)_V$$

what is the specific heat at constant volume. To get it we first prove, that the specific heat for the Van der Waals gas does not depend on volume at fixed temperature.

We start with the Maxwell relation

$$\left(\frac{\partial S}{\partial V}\right)_T = \left(\frac{\partial p}{\partial T}\right)_V$$

Differentiating by T

$$\left(\frac{\partial}{\partial V}\right)_T \left(\frac{\partial S}{\partial T}\right)_V = \left(\frac{\partial^2 p}{\partial T^2}\right)_V$$

$$\left(\frac{\partial}{\partial V}\right)_T T \left(\frac{\partial S}{\partial T}\right)_V = T \left(\frac{\partial^2 p}{\partial T^2}\right)_V$$

$$\left(\frac{\partial C_V}{\partial V}\right)_T = \left(\frac{\partial^2 p}{\partial T^2}\right)_V$$

$$p = -\frac{a}{V^2} + \frac{RT}{V - b}$$

we get

$$\left(\frac{\partial^2 p}{\partial T^2}\right)_V = 0$$

and

$$\left(\frac{\partial C_V}{\partial V}\right)_T = 0$$

$$\left(\frac{\partial C_V}{\partial V}\right)_T = 0$$

So the specific heat at constant temperature does not depend on volume, it can be evaluated for large volume where the gas is sufficiently dilute and behaves like ideal gas, so we get

$$C_V = \left(\frac{\partial U}{\partial T}\right)_V = \frac{3}{2}R$$

For the Joule constant we finally get

$$\left(\frac{\partial T}{\partial V}\right)_U = -\frac{\left(\frac{\partial U}{\partial V}\right)_T}{\left(\frac{\partial U}{\partial T}\right)_V} = -\frac{\frac{a}{V^2}}{\frac{3}{2}R} = -\frac{2a}{3RV^2}$$

Throttling

Suppose that initially we have a volume $V1$ of gas in the left compartment at pressure p_1 and no gas in the right compartment. At the end we shall have the volume V_2 in the right compartment at pressure p_2 and no gas in the left compartment. The total work performed by the gas during the process will be

$$A = p_2 V_2 - p_1 V_1$$

Since heat is zero the following energy conservation law must hold

and we get

$$E_1 + p_1 V_1 = E_2 + p_2 V_2$$

So the process is isenthalpic. For infinitesimal process we get

$$dH = TdS + Vdp$$

Now let us consider entropy being a function of T, p, so that

$$dS = \left(\frac{\partial S}{\partial T}\right)_p dT + \left(\frac{\partial S}{\partial p}\right)_T dp$$

$$dH = T\left(\frac{\partial S}{\partial T}\right)_p dT + \left(V + T\left(\frac{\partial S}{\partial p}\right)_T\right) dp$$

$$dH = T\left(\frac{\partial S}{\partial T}\right)_p dT + \left(V + T\left(\frac{\partial S}{\partial p}\right)_T\right) dp$$

from the Maxwell relation (for G)

$$\left(\frac{\partial S}{\partial p}\right)_T = -\left(\frac{\partial V}{\partial T}\right)_p$$

$$dH = T\left(\frac{\partial S}{\partial T}\right)_p dT + \left(V - T\left(\frac{\partial V}{\partial T}\right)_p\right) dp$$

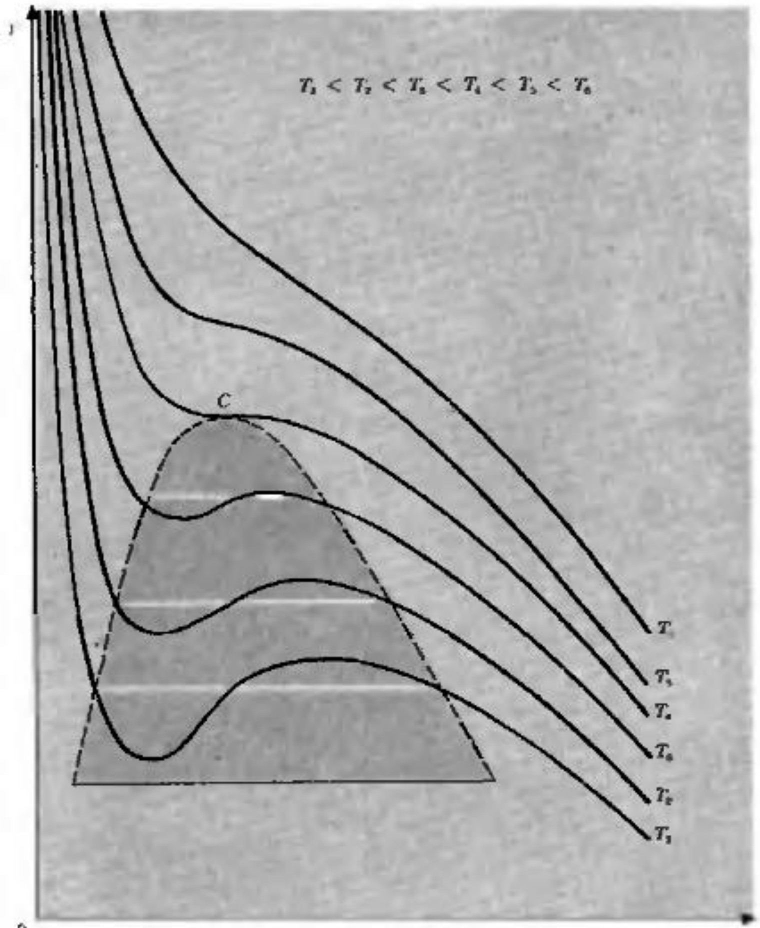
$$dH = C_p dT + \left(V - T\left(\frac{\partial V}{\partial T}\right)_p\right) dp$$

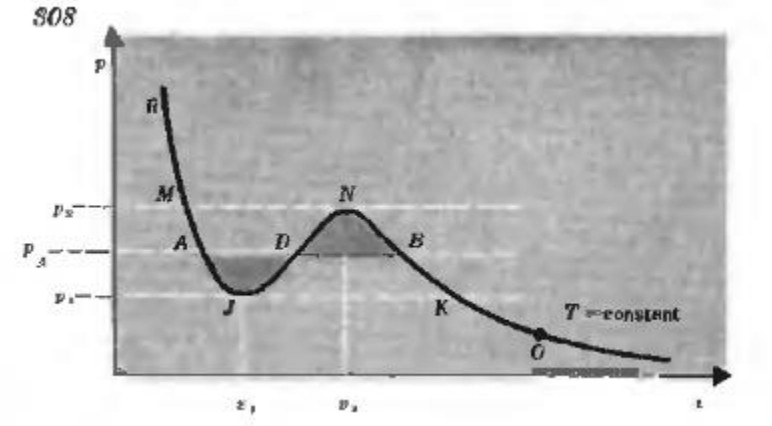
putting $dH = 0$ we get for the Joule Kelvin coefficient

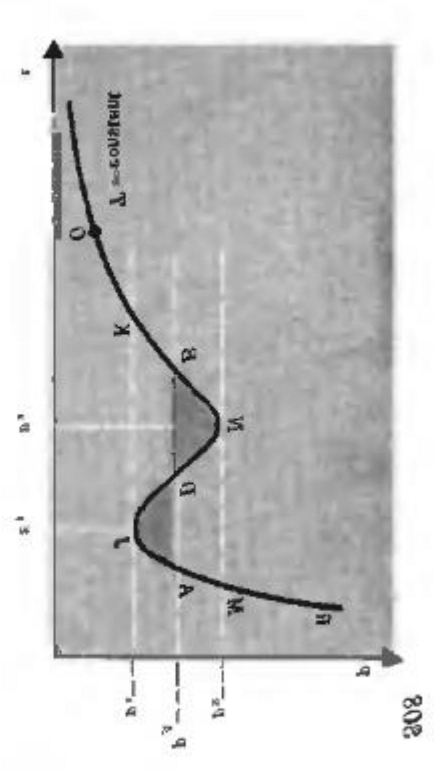
$$\left(\frac{\partial T}{\partial p}\right)_H = \frac{1}{C_p}\left(T\left(\frac{\partial V}{\partial T}\right)_p - V\right)$$

Maxwell construction

$$\left(p + \frac{a}{v^2}\right)(v - b) = RT$$



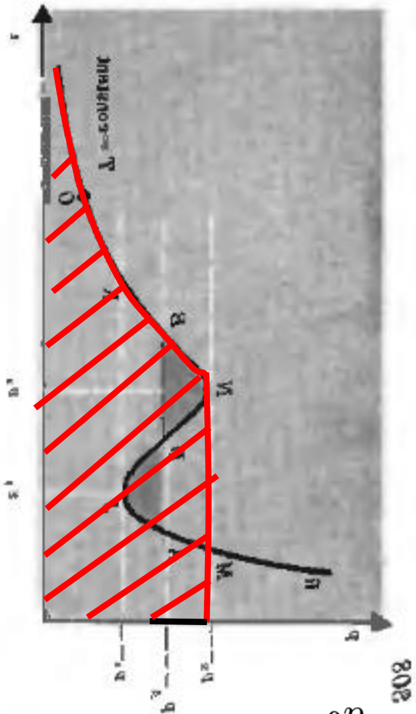


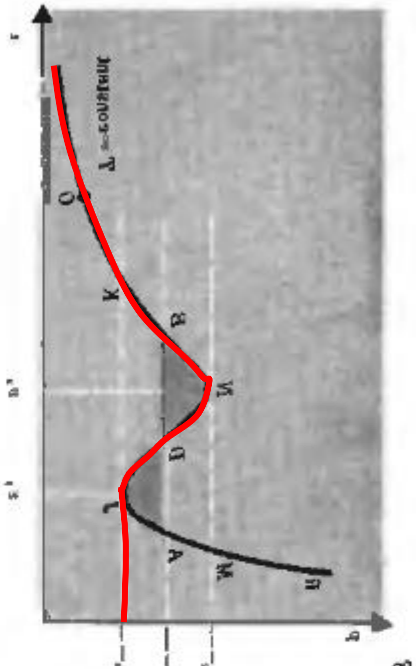


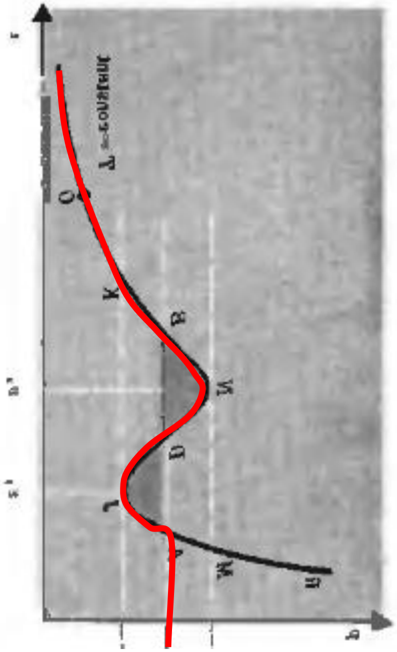
$$dG = -SdT + Vdp$$

na izoterme

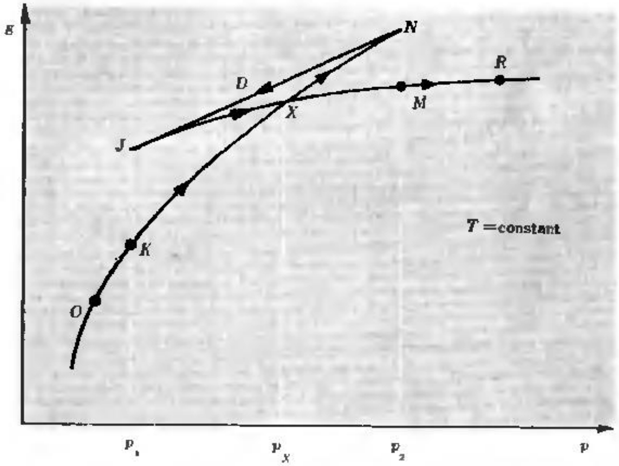
$$G(p) = G(p_0) + \int_{p_0}^{p} V dp$$



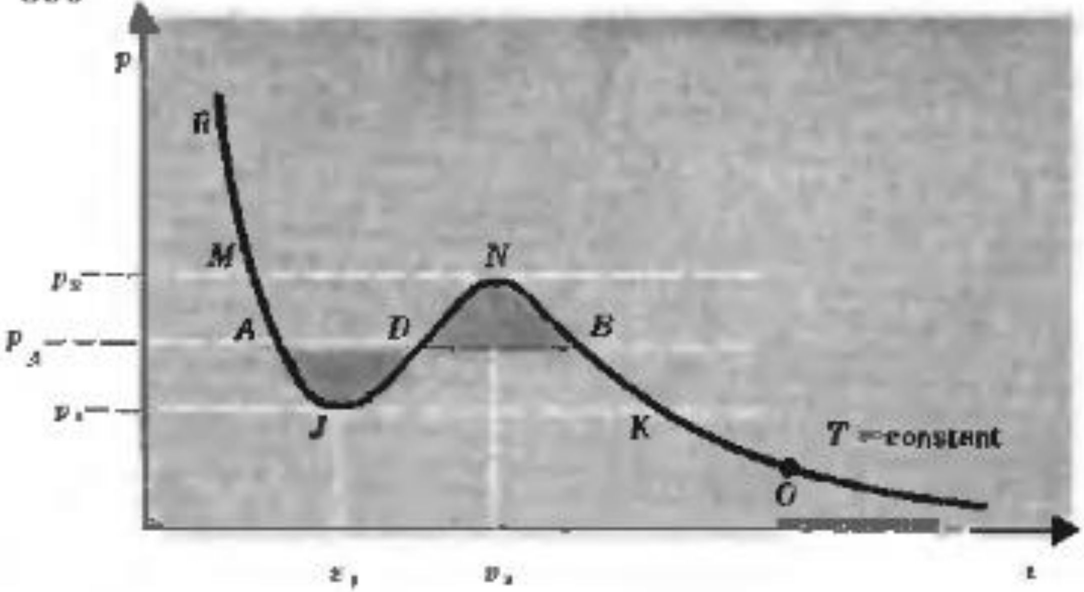


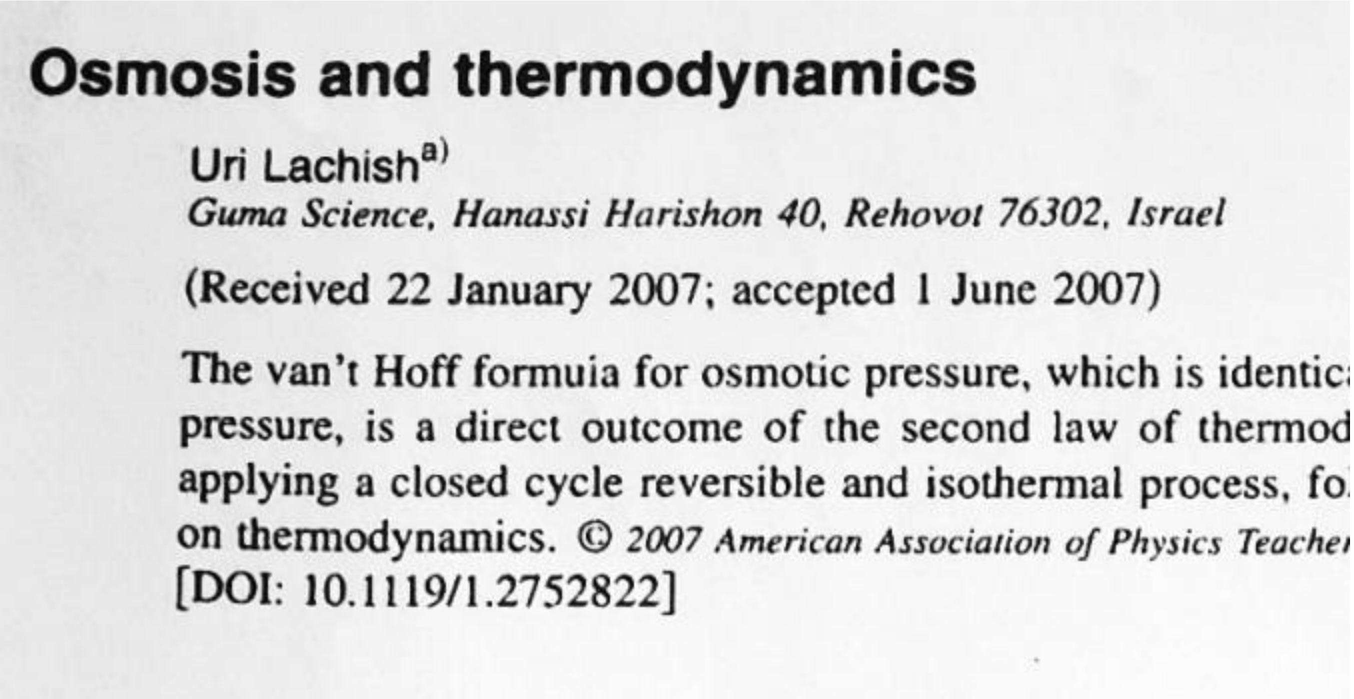


$$G(p) = G(p_0) + \int_{p_0}^{p} V\,dp$$



308





Osmosis and thermodynamics

Uri Lachish[a)]
Guma Science, Hanassi Harishon 40, Rehovot 76302, Israel

(Received 22 January 2007; accepted 1 June 2007)

The van't Hoff formula for osmotic pressure, which is identic
pressure, is a direct outcome of the second law of thermod
applying a closed cycle reversible and isothermal process, fo
on thermodynamics.
[DOI: 10.1119/1.2752822]

Am. J. Phys. **75** (11), November 2007

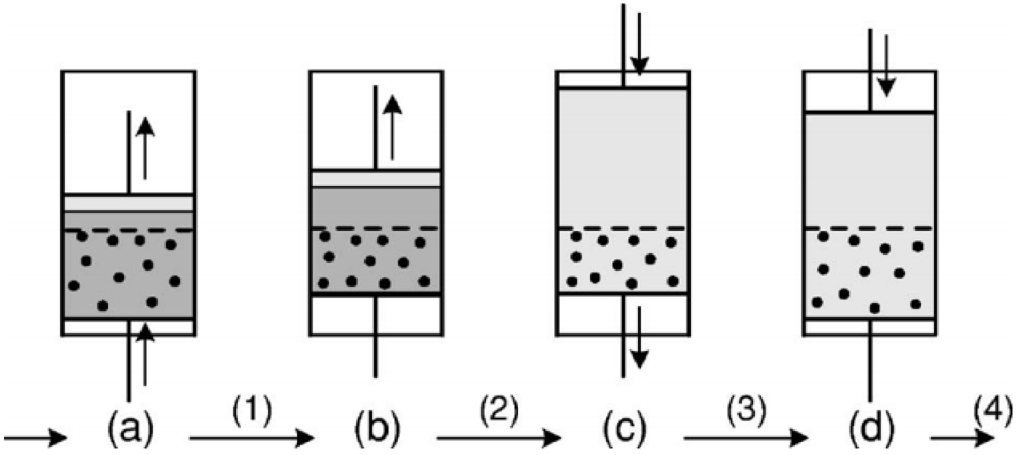
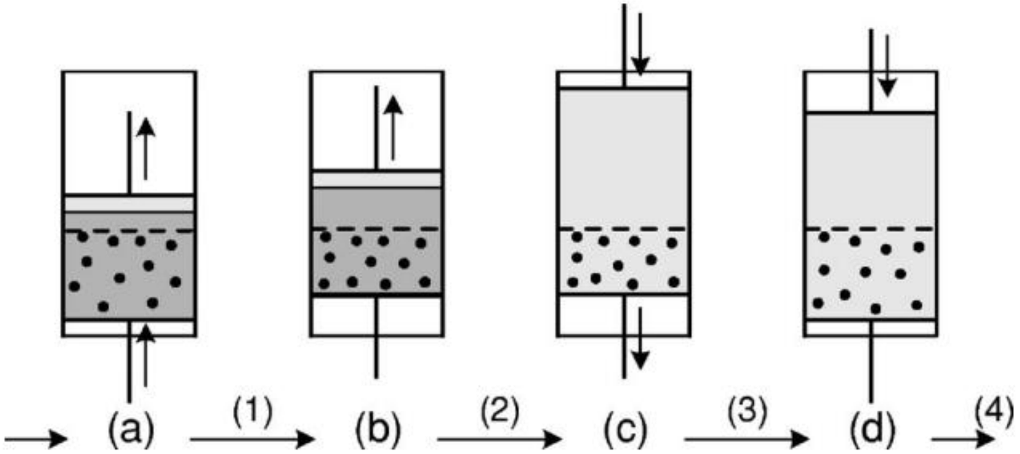
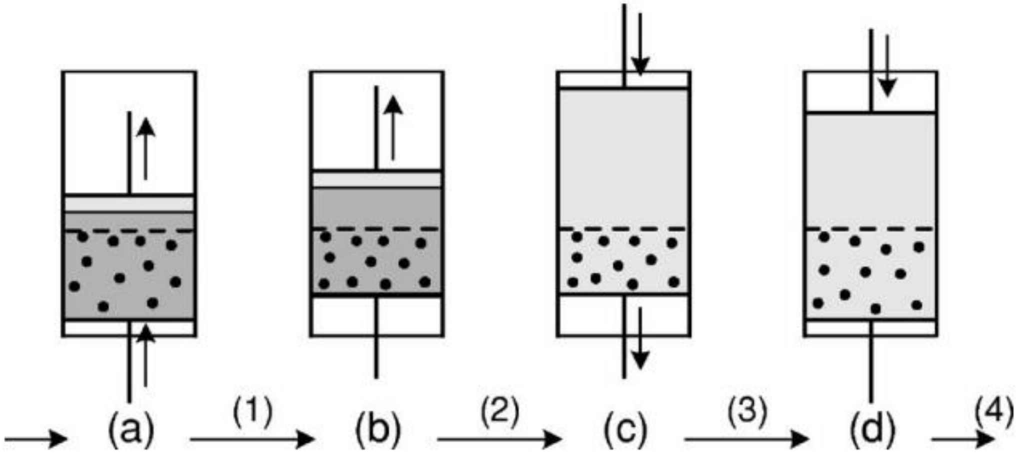


Fig. 1. Cyclic isothermal compression and expansion of particles within a fluid. The fluid is free to cross the semi-permeable membrane. Step 1: Compression of the particles within the liquid (dark gray). The fluid volume is constant. Step 2: Expansion of the fluid to evaporate all the liquid. Step 3: Expansion of the confined particles within a gas (light gray). The fluid volume is constant. Step 4: Compression of the gas to partially condense it into a liquid. The work of fluid expansion (step 2) is equal to the work of fluid compression (step 4). Therefore, by the second law of thermodynamics, the work of particle compression within the liquid must be equal to the work of particle expansion within the gas.



(1) The confined particles in the lower volume are dissolved in the liquid. The two pistons move up together without changing the fluid volume, and therefore there is no work done on it. The lower piston compresses the dissolved particles between it and the membrane, and does an amount of work $\pi \Delta V$, where π is the osmotic pressure, V is the volume of the dissolved particles, and ΔV is a small volume change.



(2) The upper piston moves up and induces evaporation of the liquid until it is transformed into a gas. It is assumed that the dissolved particles are now also in a gas phase.

(3) The two pistons move down together, again without changing the fluid volume. The confined particles push the lower piston, expand, and do a work $p\Delta V$ where p is the gas pressure and ΔV is the same volume change as in step 1.

(4) The upper piston moves down, compresses the fluid, and induces condensation of the gas into a liquid until it reaches the initial state of the completed cycle.

Because the process is reversible and isothermal, the net work done during the closed cycle must be zero. Otherwise it would violate the second law of thermodynamics as postulated by Kelvin.[7]

$$0 = \Delta S = \int \frac{\delta Q}{T} = \frac{1}{T} \int \delta Q$$

$$0 = \Delta E = Q - A \implies \boxed{A = 0}$$

Density of states

We started with one-particle states in a box, which are sinus waves with zeros at the box boundaries. The states are labeled by three positive integer numbers n_1, n_2, n_3 and their energy is

$$\varepsilon = \frac{\pi^2\hbar^2}{2mL^2}(n_1^2 + n_2^2 + n_3^2)$$

States correspond to points with integer-valued coordinates in positive octant of the n-space, so the total number of one-particle states with energy less then ε is

$$\varphi(\varepsilon) = \frac{1}{8}\frac{4}{3}\pi\left(\frac{2mL^2}{\pi^2\hbar^2}\varepsilon\right)^{3/2}$$

And the density of states in the ε-space is

$$\varphi'(\varepsilon) = \frac{2\pi V}{(2\pi\hbar)^3}(2m)^{3/2}\varepsilon^{1/2}$$

So the density of states in the k-space is

$$\frac{V}{(2\pi)^3}$$

and in p-space

$$\frac{V}{(2\pi\hbar)^3}$$

To switch to the ε-space one needs the dispersion relation, which for free particles reads

$$\varepsilon = \frac{p^2}{2m} = \frac{\hbar^2 k^2}{2m}$$

and the substitution in the integrals reads

$$\frac{V}{(2\pi)^3}\, d^3\overrightarrow{k} = \frac{V}{(2\pi)^3} 4\pi k^2\, dk = \frac{V}{(2\pi)^3} 4\pi k^2\, \frac{d\varepsilon}{\frac{d\varepsilon}{dk}}$$

Substituting from the dispersion relation we get the same density of states in the ε-space as before.

$$\frac{V}{(2\pi)^3}\, d^3\vec{k} = \frac{V}{(2\pi)^3} 4\pi k^2\, dk = \frac{V}{(2\pi)^3} 4\pi k^2\, \frac{d\varepsilon}{\frac{d\varepsilon}{dk}}$$

grupová rýchlosť

Ak rátam balistický prúd v nanofyzike, hustotu stavov treba vynásobiť grupovou rýchlosťou a ostane univerzálny vzorec nezávislý na tom, o akú tuhú látku ide.

Ideal Bose gas

We calculate grandcanonical distribution for one one-particle state. The ”state of the state j” is given by the occupation number n_j and the relevant grandcanonical sum is

$$\mathcal{Z}_j = \sum_{n_j} \exp(\frac{\mu}{kT} n_j - \frac{\varepsilon_j}{kT} n_j) = \frac{1}{1 - \exp(\frac{\mu - \varepsilon_j}{kT})}$$

We introduce fugacity z by

$$z = \exp(\frac{\mu}{kT})$$

So we get

$$\mathcal{Z}_j = \frac{1}{1 - z\exp(-\beta\varepsilon_j)}$$

The one-particle states are independent, so the total grandcanonical sum is given by the product

$$\mathcal{Z} = \prod_j \mathcal{Z}_j$$

$$\ln \mathcal{Z} = -\sum_j \ln(1 - z\exp(-\beta\varepsilon_j))$$

$$\ln \mathcal{Z} = -\sum_j \ln(1 - z\exp(-\beta\varepsilon_j))$$

We remind that

$$(Z) = \exp(\frac{pV}{kT})$$

and so

$$\frac{pV}{kT} = -\sum_j \ln(1 - z\exp(-\beta\varepsilon_j))$$

We pretend to know the value of μ, but experimentally we rather know the number of particles

$$N = \sum_j \overline{N}_j = \sum_j \frac{1}{\exp(\frac{\varepsilon_j - \mu}{kT}) - 1} = \sum_j \frac{1}{z^{-1}\exp(\beta\varepsilon_j) - 1}$$

This equation should be inverted: one should calculate μ in terms of N and T, substitute it to the previous equation and get the equation of state.

To perform this program, we need to replace the summations by integration in the energy space, using the known expression of the density of states in the ε-space. We get

$$\frac{p}{kT} = -\frac{2\pi}{(2\pi\hbar)^3}(2m)^{3/2}\int_0^\infty \varepsilon^{1/2}\ln(1 - z\exp(-\beta\varepsilon))d\varepsilon$$

$$\frac{N}{V} = \frac{2\pi}{(2\pi\hbar)^3}(2m)^{3/2}\int_0^\infty \frac{\varepsilon^{1/2}}{z^{-1}\exp(\beta\varepsilon) - 1}d\varepsilon$$

We switch to dimensionless integration variable

$$\frac{p}{kT} = -\frac{2}{\sqrt{\pi}}\frac{(2\pi mkT)^{3/2}}{(2\pi\hbar)^3}\int_0^\infty x^{1/2}\ln(1 - z\exp(-x))dx$$

$$\frac{N}{V} = \frac{2}{\sqrt{\pi}}\frac{(2\pi mkT)^{3/2}}{(2\pi\hbar)^3}\int_0^\infty \frac{x^{1/2}}{z^{-1}\exp(x) - 1}dx$$

We introduce the notation (de'Broglie thermal length)

$$\lambda = \frac{2\pi\hbar}{(2\pi mkT)^{1/2}}$$

and using the relation

$$\Gamma(\frac{3}{2}) = \frac{\sqrt{\pi}}{2}$$

$$\frac{p}{kT} = -\frac{2}{\sqrt{\pi}}\frac{(2\pi mkT)^{3/2}}{(2\pi\hbar)^3}\int_0^\infty x^{1/2}\ln(1 - z\exp(-x))dx$$

$$\frac{N}{V} = \frac{2}{\sqrt{\pi}}\frac{(2\pi mkT)^{3/2}}{(2\pi\hbar)^3}\int_0^\infty \frac{x^{1/2}}{z^{-1}\exp(x) - 1}dx$$

$$\frac{p}{kT} = -\frac{1}{\lambda^3}\frac{1}{\Gamma(3/2)}\int_0^\infty x^{1/2}\ln(1 - z\exp(-x))dx$$

$$\frac{N}{V} = \frac{1}{\lambda^3}\frac{1}{\Gamma(3/2)}\int_0^\infty \frac{x^{1/2}}{z^{-1}\exp(x) - 1}dx$$

Since similar integrals are often met in statistical physics, new special functions are introduced (called Einstein functions) as

$$g_\nu(z) = \frac{1}{\Gamma(\nu)}\int_0^\infty \frac{x^{\nu-1}}{z^{-1}\exp(x) - 1}dx$$

$$g_\nu(z) = \frac{1}{\Gamma(\nu)} \int_0^\infty \frac{x^{\nu-1}}{z^{-1}\exp(x) - 1} dx$$

which can be expanded into Taylor series:

$$\begin{aligned}
g_\nu(z) &= \frac{1}{\Gamma(\nu)} \frac{z}{e^x} \int_0^\infty \frac{x^{\nu-1}}{1 - z\exp(-x)} dx \\
&= \frac{1}{\Gamma(\nu)} \frac{z}{e^x} \int_0^\infty x^{\nu-1} \sum_{k=0}^\infty (ze^{-x})^k dx \\
&= \frac{1}{\Gamma(\nu)} \int_0^\infty x^{\nu-1} \sum_{k=1}^\infty (ze^{-x})^k dx \\
&= \frac{1}{\Gamma(\nu)} \sum_{k=1}^\infty \int_0^\infty x^{\nu-1} z^k e^{-kx} = \frac{1}{\Gamma(\nu)} \sum_{k=1}^\infty \frac{1}{k^\nu} z^k \Gamma(\nu)
\end{aligned}$$

$$g_\nu(z) = \sum_{k=1}^\infty \frac{z^k}{k^\nu}$$

$$\frac{N}{V} = \frac{1}{\lambda^3}\frac{1}{\Gamma(3/2)}\int_0^\infty \frac{x^{1/2}}{z^{-1}\exp(x)-1}dx$$

$$\frac{p}{kT} = -\frac{1}{\lambda^3}\frac{1}{\Gamma(3/2)}\int_0^\infty x^{1/2}\ln(1-z\exp(-x))dx$$

$$\frac{1}{\Gamma(3/2)}\int_0^\infty x^{1/2}\ln(1-z\exp(-x))dx = \frac{1}{\Gamma(3/2)}\int_0^\infty \frac{2}{3}x^{3/2}\frac{-z\exp(-x)}{1-z\exp(-x)}dx =$$

$$= -\frac{2}{3}\frac{1}{\Gamma(3/2)}\int_0^\infty \frac{x^{3/2}}{z^{-1}\exp(x)-1} = -\frac{1}{\Gamma(5/2)}\int_0^\infty \frac{x^{3/2}}{z^{-1}\exp(x)-1} = -g_{5/2}(z)$$

So we get

$$\boxed{\begin{aligned}\frac{p}{kT} &= \frac{1}{\lambda^3}g_{5/2}(z)\\ \frac{N}{V} &= \frac{1}{\lambda^3}g_{3/2}(z)\end{aligned}}$$

Now the idea is we should calculate z from the second equation and substitute it to the first one to get the equation of state, formally

$$\frac{pV}{NkT} = \frac{V}{N}\frac{1}{\lambda^3}g_{5/2}(g_{3/2}^{-1}(\lambda^3\frac{N}{V}))$$

method of ”series reversion”[1]

$$y = a_1 x + a_2 x^2 + a_3 x^3 + \dots$$

we can look for it inverse in the form

$$x = A_1 y + A_2 y^2 + A_3 y^3 + \dots$$

Plugging the second equation into the first one one gets

$$y = a_1 A_1 y + (a_2 A_1^2 + a_1 A_2) y^2 + \dots$$

Comparing coefficients on both sides one gets

$$A_1 = a_1^{-1}, \quad A_2 = -a_1^{-3} a_2, \quad \dots$$

$$\frac{N}{V} = \frac{1}{\lambda^3} g_{3/2}(z)$$

In the classical limit we have $z \ll 1$, so to get the first correction we expand to the second order in z

$$\lambda^3 \frac{N}{V} = z + \frac{z^2}{2^{3/2}}$$

We look for z in the form

$$z = \lambda^3 \frac{N}{V} + c\left(\lambda^3 \frac{N}{V}\right)^2$$

We find

$$c = -\frac{1}{2^{3/2}}$$

$$\frac{pV}{NkT} = \frac{V}{N}\frac{1}{\lambda^3} g_{5/2}(g_{3/2}^{-1}(\lambda^3 \frac{N}{V}))$$

$$\frac{pV}{NkT} = \frac{V/N}{\lambda^3}\left(z + \frac{1}{2^{5/2}} z^2\right) = \frac{V/N}{\lambda^3}\left(\frac{\lambda^3}{V/N} - \frac{1}{2^{3/2}}\left(\frac{\lambda^3}{V/N}\right)^2 + \frac{1}{2^{5/2}}\left(\frac{\lambda^3}{V/N}\right)^2\right)$$

$$\boxed{\frac{pV}{NkT} = 1 - \frac{1}{4\sqrt{2}}\frac{\lambda^3}{V/N}}$$

Van der Waals gas

we consider gas of molecules which interact via pairwise interaction:

$$U(\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_N) = \sum_{pairs} u(\boldsymbol{r}_i, \boldsymbol{r}_j)$$

$$Z = \frac{1}{N!} \int \frac{1}{(2\pi\hbar)^{3N}} \prod_i d^3\boldsymbol{p}_i d^3\boldsymbol{r}_i \exp\left(\frac{\sum_j \frac{\boldsymbol{p}_j^2}{2m} + U(\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_N)}{kT} \right)$$

$$Z = \frac{1}{N!} \frac{1}{(2\pi\hbar)^{3N}} \left(2\pi mkT\right)^{3N/2} Z_U$$

where

$$Z_U = \int \prod_i d^3\boldsymbol{r}_i \exp\left(\frac{-U}{kT}\right)$$

$$Z_U = \int \prod_i d^3\boldsymbol{r}_i \exp\left(\frac{-U}{kT}\right)$$

We calculate Z_U using the following trick with the mean potential energy

$$\bar{U} = -\frac{d}{d\beta}\ln Z_U(\beta)$$

and since $Z_U(0) = V^N$ we get

$$\ln Z_U(\beta) = N\ln V - \int_0^\beta \bar{U}(\beta')d\beta'$$

Now we assume that we can neglect correlations between pairs of particles and we get

$$\bar{U} = \frac{1}{2}N(N-1)\bar{u}$$

where

$$\bar{u} = \frac{\int d^3\boldsymbol{r}_1 d^3\boldsymbol{r}_2 u(\boldsymbol{r}_1,\boldsymbol{r}_2)\exp\left(-\frac{u(\boldsymbol{r}_1,\boldsymbol{r}_2)}{kT}\right)}{\int d^3\boldsymbol{r}_1 d^3\boldsymbol{r}_2 \exp\left(-\frac{u(\boldsymbol{r}_1,\boldsymbol{r}_2)}{kT}\right)}$$

where

$$\bar{u} = \frac{\displaystyle\int d^3\boldsymbol{r}_1 d^3\boldsymbol{r}_2 u(\boldsymbol{r}_1, \boldsymbol{r}_2) \exp\left(-\frac{u(\boldsymbol{r}_1, \boldsymbol{r}_2)}{kT}\right)}{\displaystyle\int d^3\boldsymbol{r}_1 d^3\boldsymbol{r}_2 \exp\left(-\frac{u(\boldsymbol{r}_1, \boldsymbol{r}_2)}{kT}\right)}$$

Using substitution

$$\boldsymbol{r}_1, \boldsymbol{r}_2 \quad \mapsto \quad \boldsymbol{R} = \frac{\boldsymbol{r}_1 + \boldsymbol{r}_2}{2}, \;\; \boldsymbol{r} = \boldsymbol{r}_1 - \boldsymbol{r}_2$$

we get

$$\bar{u} = \frac{\displaystyle\int d^3\boldsymbol{R} \int d^3\boldsymbol{r}\; u(\boldsymbol{r}) \exp\left(-\frac{u(\boldsymbol{r})}{kT}\right)}{\displaystyle\int d^3\boldsymbol{R} \int d^3\boldsymbol{r} \exp\left(-\frac{u(\boldsymbol{r})}{kT}\right)}$$

$$\bar{u} = -\frac{d}{\beta} \ln \int \exp(-\beta u) d^3\boldsymbol{r}$$

$$\int \exp(-\beta u) d^3\boldsymbol{r} = \int [1 + (\exp(-\beta u) - 1)] d^3\boldsymbol{r} = V + I = V(1 + \frac{I}{V})$$

$$I(\beta) = \int (\exp(-\beta u) - 1) d^3\boldsymbol{r} = \int_0^\infty (\exp(-\beta u) - 1)\, 4\pi r^2\, dr$$

$$\bar{u} = -\frac{d}{d\beta}[\ln V + \ln(1 + \frac{I(\beta)}{V})] \approx -\frac{d}{d\beta}\frac{I(\beta)}{V} = -\frac{1}{V}\frac{dI(\beta)}{d\beta}$$

$$\bar{U} = \frac{1}{2}N(N-1)\bar{u}$$

$$\bar{U}(\beta) = -\frac{1}{2}\frac{N^2}{V}\frac{dI(\beta)}{d\beta}$$

$$\ln Z_U(\beta) = N \ln V - \int_0^\beta \bar{U}(\beta')d\beta'$$

$I(0) = 0$

$$\ln Z_U(\beta) = N \ln V + \frac{1}{2}\frac{N^2}{V}\int \frac{dI(\beta)}{d\beta} = N \ln V + \frac{1}{2}\frac{N^2}{V}I(\beta)$$

$$\frac{p}{kT} = \frac{\partial \ln Z_U}{\partial V} = \frac{N}{V} - \frac{1}{2}\frac{N^2}{V^2}I(\beta)$$

$$I(\beta) = \int (\exp(-\beta u) - 1) d^3\boldsymbol{r} = \int_0^\infty (\exp(-\beta u) - 1)\, 4\pi r^2\, dr$$

$$u(r) = \infty \quad \text{for} \quad r < r_0$$

$$u(r) = f(r) \quad \text{for} \quad r > r_0$$

$$I(\beta) = -4\pi \int_0^{r_0} r^2 dr + 4\pi \int_{r_0}^\infty (\exp(-\beta f) - 1) r^2 dr = -\frac{4}{3}\pi r_0^3 - \frac{4\pi}{kT} \int_{r_0}^\infty f(r) r^2 dr$$

$$I(\beta) = -2b + \frac{2a}{kT}$$

where we have denoted

$$b = \frac{2\pi}{3} r_0^3$$

$$a = -2\pi \int_{r_0}^\infty f(r) r^2 dr$$

$$\frac{p}{kT} = \frac{\partial \ln Z_U}{\partial V} = \frac{N}{V} - \frac{1}{2}\frac{N^2}{V^2} I(\beta)$$

$$\frac{p}{kT} = \frac{N}{V} + \frac{N^2}{V^2}(b - \frac{a}{kT})$$

$$\frac{p}{kT} = \frac{N}{V} + \frac{N^2}{V^2}(b - \frac{a}{kT})$$

$$(p + a\frac{N^2}{V^2}) = \frac{N}{V}kT(1 + b\frac{N}{V}) \approx \frac{\frac{N}{V}kT}{1 - b\frac{N}{V}}$$

$$(p + a\frac{N^2}{V^2})(V - bN) = NkT$$

what is Van der Waals equation.

Critical phenomena

Spin-spin interaction

Classically the magnetic moment of a current loop is given as $m = IS$ (where I is the current, and S the surface of the loop). One can generalize this to a classical case of a point particle with charge q moving with a velocity v. We get $m = (q/2m)L$, where L is the particle (orbital) angular momentum $L = mrv$. If we expect that a similar relation holds for the internal angular momentum, spin, we get for the electron (having spin $\hbar/2$ the magnetic moment $e\hbar/4m$. Experiment (Einstein, de Haas) however shows that the electron (internal) magnetic moment is $\mu_B = e\hbar/2m$ (this value is called Bohr magneton). This value of the electron magnetic moment (the ”mysterious factor 2”) is explained by the Dirac equation.

Now back to the question of magnetic moment alignment in a ferromagnet. The first guess might be, that it is the interaction of magnetic moment with each other, which is responsible for the alignment. A quick order-of-magnitude estimate shows, that it is most probably not so.

$$\boldsymbol{B} = \frac{\mu_0}{4\pi}\left(\frac{3\boldsymbol{r}(\boldsymbol{m}\cdot\boldsymbol{r})}{r^5} - \frac{\boldsymbol{m}}{r^3}\right)$$

So $B \approx \mu_0 m/4\pi r^3$, therefore the interaction energy of a magnetic moment with its neighbours in a lattice is $W \approx z\mu_0 m^2/4\pi r^3$ (z is the number of nearest neighbours (In a cubic lattice $z = 6$). Numerically $\mu_B \doteq 9{,}27 \cdot 10^{-24}$ Am^2. So the interaction energy for $r = 2 \cdot 10^{-10}$ m is roughly 10^{-4} eV, what in temperature units means roughly 1 K. So we expect that at temperatures above 1 K the spin orientations would be randomized by thermal fluctuations and the ferromagnetism could not be observed at room temperatures. The solution of the puzzle is the effect of an electrostatic interaction combined with the Pauli exclusion principle.

Let us consider two non-interacting electrons in states with quantum numbers a and b, The state Ψ_a has energy E_a, the state Ψ_b energy E_b. The electron coordinates are denoted as $\boldsymbol{x}_1$ a $\boldsymbol{x}_2$, spins $\boldsymbol{S}_1$ a $\boldsymbol{S}_2$. If we consider two particle states, let us first forget about the fact that the two electrons are identical (indistinguishable). We can construct two states, the first will be $\Psi_{ab}(\boldsymbol{x}_1, \boldsymbol{x}_2) = \Psi_a(\boldsymbol{x}_1)\Psi_b(\boldsymbol{x}_2) \equiv \Psi_{ab}$ and the second one will be $\Psi_{ba}(\boldsymbol{x}_1, \boldsymbol{x}_2) = \Psi_b(\boldsymbol{x}_1)\Psi_a(\boldsymbol{x}_2) \equiv \Psi_{ba}$ The energies of these two states (as of any of their linear combinations) are the same $E(\Psi_{ab}) = E(\Psi_{ba}) = E_a + E_b$ For interacting electrons changes the situation considerably The interaction potential between two electrons must be symmetric $V(\boldsymbol{x}_1, \boldsymbol{x}_2) = V(\boldsymbol{x}_2, \boldsymbol{x}_1) \neq 0$, the Hamiltonian of the system is $\mathsf{H}_1 + \mathsf{H}_2 + V$. We can construct two different two-particle states, a symmetric and an antisymmetric one

$$\Psi_S = \frac{1}{\sqrt{2}}\big(\Psi_{ab} + \Psi_{ba}\big), \quad \Psi_A = \frac{1}{\sqrt{2}}\big(\Psi_{ab} - \Psi_{ba}\big).$$

Their mean energies we calculate as$\langle\Psi|\,\mathsf{H}_1 + \mathsf{H}_2 + V\,|\Psi\rangle$. As a result we get two different energies

$$E_S = E_a + E_b + I - J, \quad E_A = E_a + E_b + I + J,$$

where I and J denote the integrals (J is called the exchange integral)

$$I = \int \Psi_{ab}^* V \Psi_{ab} = \int \Psi_{ba}^* V \Psi_{ba}, \quad J = -\int \Psi_{ab}^* V \Psi_{ba} = -\int \Psi_{ba}^* V \Psi_{ab}.$$

The total antisymmetry of the state requires that we combine the symmetrical spin function with the antisymmetrical spatial function and the antisymmetrical spin function with the symmetrical space function. So legal two-particle states are $\Psi_S\chi^{(0)}$ and $\Psi_A\chi^{(1)}$. When we denote the total spin as s we get for the corresponding mean energy values

$$\begin{aligned} E_{s=0} &= E_S \\ E_{s=1} &= E_A \end{aligned}$$

We can combine these two expression formally into one

$$E = E_S + (E_A - E_S)\frac{s(s+1)}{2} \implies \mathsf{H} = E_S + (E_A - E_S)\frac{\mathsf{S}\cdot\mathsf{S}}{2}.$$

$$\mathsf{S}\cdot\mathsf{S} = (\mathsf{S}_1 + \mathsf{S}_2)\cdot(\mathsf{S}_1 + \mathsf{S}_2) = \underbrace{\mathsf{S}_1\cdot\mathsf{S}_1}_{\frac{1}{2}(\frac{1}{2}+1)} + \underbrace{\mathsf{S}_2\cdot\mathsf{S}_2}_{\frac{1}{2}(\frac{1}{2}+1)} + 2\mathsf{S}_1\cdot\mathsf{S}_2 = \frac{3}{2} + 2\mathsf{S}_1\cdot\mathsf{S}_2$$

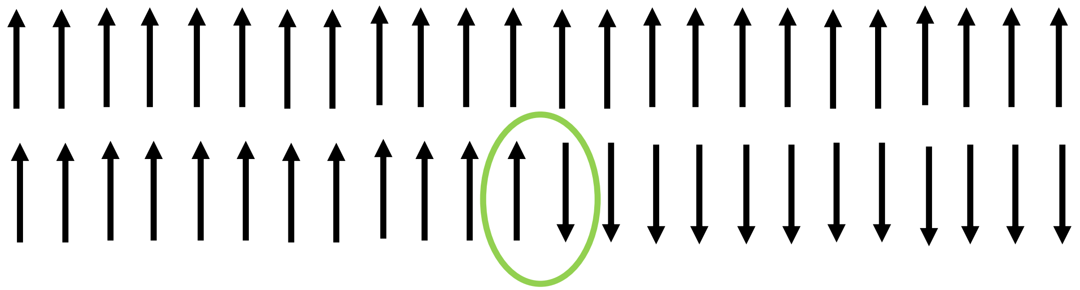
$$\mathsf{H} = E_S + (E_A - E_S)\Big(\frac{3}{4} + \mathsf{S}_1\cdot\mathsf{S}_2\Big) = E_a + E_b + I + \frac{J}{2} + \underbrace{(E_A - E_S)}_{2J}\mathsf{S}_1\cdot\mathsf{S}_2.$$

Ising model

A simple model with interaction between (classical) spins is the Ising model, where the interaction is limited to nearest neighbours. As already said the model is classical, what means, that in the hamiltonian there are no operators, just variables, which, however, can have only discrete values $-1, 1$. Energy of the system is given by the following formula

$$E(\{s\}) = -J \sum_{i \sim j} s_i s_j - \mu_B B \sum_i s_i,$$

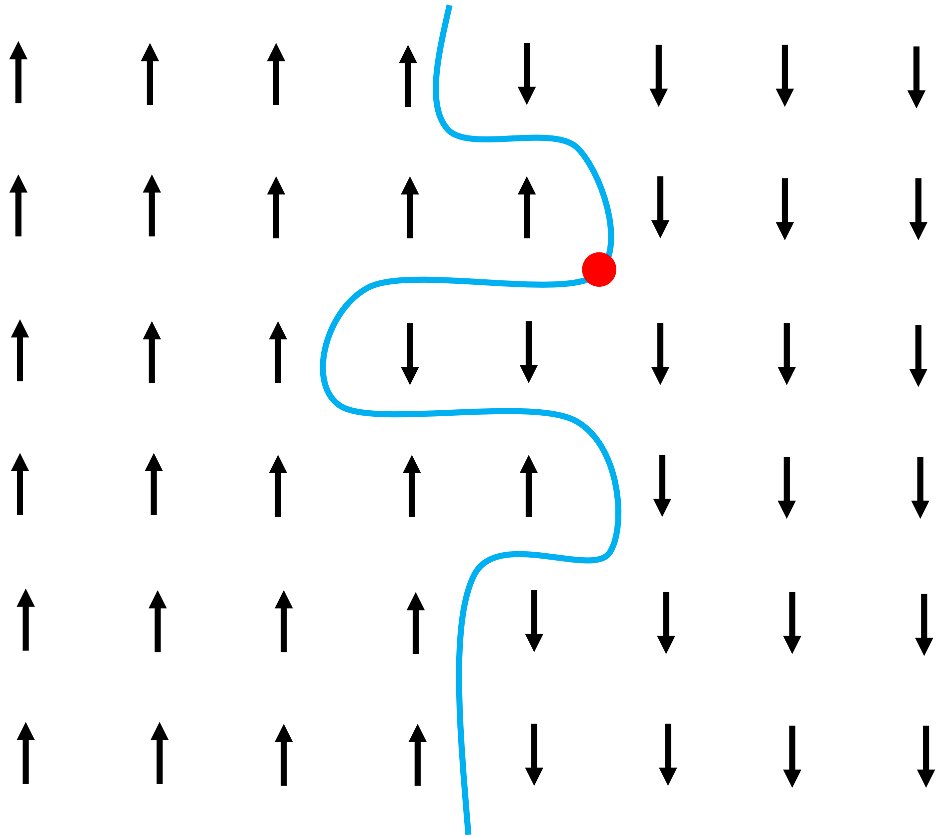
where $i \sim j$ means summation over the different pairs of neighbouring spins. The first author who formulated the model was Wilhelm Lenz (1920). Ernest Ising was his student who in his thesis described the first exact solution in the one-dimensional case (1924). The two-dimensional model (for zero external field) was solved by Lars Onsager (1944). For higher dimensions we do not know (up to now) an exact solution, we have just numerical simulations.



There is no phase transition in the one-dimensional Ising model. The lowest energy state is the state with no excited links. By excited links we mean here a link for which the spins on its two sites are oppositely oriented. In one-dimensional model the state closest to the lowest energy state is a state with exactly one link excited, the energy difference being $2J$. For such a state roughly one half of the spin is oriented ”up” and the other half is oriented ”down”. The magnetization for most of such states is equal to zero. A finite lattice with N sites has (roughly, for large N) N links. One excited link can be therefore chosen in N ways, so there are N states with energy larger from the lowest energy by $2J$. The entropy of the ”macrostate” with the energy $E_0 + 2J$ is therefore proportional to $\ln(N)$, so the free energy of such a state is

$$F = E_0 + 2J - T\ln(N)$$

For a non-zero temperature the state with higher energy has lower free energy and so for a non-zero temperature the mean magnetization is zero.



N^2 spins

Domain boundary is like a random walk from top to bottom.

At each step 3 possibilities

Each line typically N^2 steps

3^{N^2} dividing lines

Energy difference

$$2JN^2$$

Entropy

N^2ln3

$\Delta F \approx 2\mathrm{JN}^2 - \mathrm{TN\string^2ln3}$

Neither energy nor entropy is a clear winner

Phase transition possible

The Ising model in one dimension is explicitly calculable. Let us fix the boundary conditions so that the first spin from left has value $s_1 = 1$. Then the state of every other spin is given if we know about every link whether it is excited or not.

Let us introduce a dual lattice which correspond to the original lattice in such a way that to every link of the original lattice there corresponds a site of the new lattice with a spin having value q with value $q = 1$ if that link is not excited and $q = -1$ if that link is excited. The states of the dual lattices exactly correspond to each other If we require, that the energies of the corresponding states are equal then the energy of the new lattice should be given as

$$H = \sum_{\text{link}} -Jq_i$$

This, however, describes the model of independent spins, the statistical sum is easily calculated as

$$Z = (\exp(J/kT) + \exp(-J/kT))^{N-1}$$

and we see there is no singularity in the thermodynamic limit and so there is no phase transition.

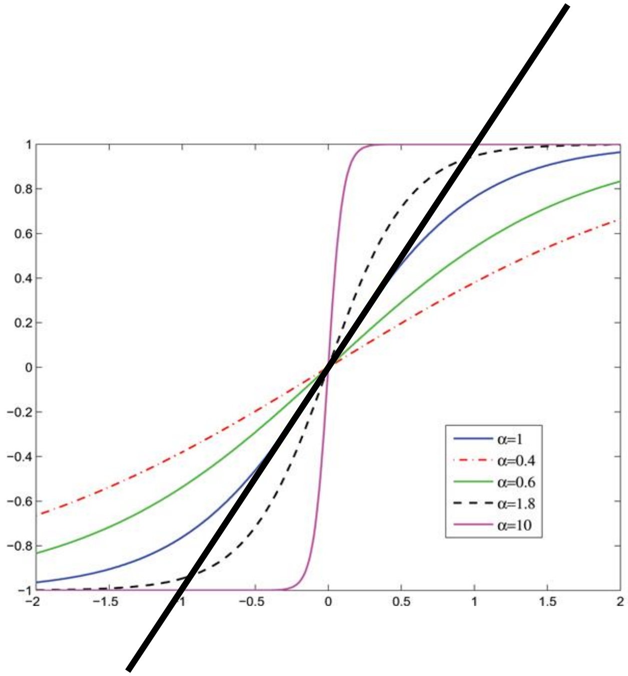
Mean field theory

$$H'_1(s) = \sum_{\text{neighbour}} -Js\langle s\rangle = -Jn\langle s\rangle s$$

where n si the number of nearest neighbours. Because of the symmetry reasons, the mean values for all the spins are equal. We have got the Hamiltonian of one spin in an external field $B = Jn\langle s\rangle$. This problem is easily solvable, and for the mean spin value we get a self-consistent equation[1]

$$\langle s\rangle = \tanh(B/kT) = \tanh(Jn\langle s\rangle/kT)$$

$$\langle s \rangle = \tanh(B/kT) = \tanh(Jn\langle s \rangle/kT)$$



The solution of this selfconsistent equation can be looked for graphically It is clear that for

$$\frac{Jn}{kT} < 1$$

there exists just one solution $\langle s \rangle = 0$, while for

$$\frac{Jn}{kT} > 1$$

there exist three solutions. The change happens in the critical point

$$\frac{Jn}{kT_c} = 1$$

[1] Let us discuss now an alternative argumentation, which, however, is wrong and leads to wrong answer. Let us try to solve the Ising model approximately, so that the real Hamiltonian is replaced by new (approximate) Hamiltonian

$$H' = \sum_{\text{link}} -J s_i \langle s_j \rangle$$

so one of the spins on the considered link is replaced by its mean value. Because of the symmetry reasons all the spins have the same mean value and we get

$$H' = \sum_{\text{link}} -J s_i \langle s \rangle$$

The sum over the links can be replaced by the sum over the spins if we realize that each link has two spins on its ends and so each link can be assigned to one of its spins as "being owned" by that spin. The lattice determines the number of nearest neighbours n. Each spin therefore "owns" $n/2$ links. We get

$$H' = \sum_{\text{i}} -J \frac{n}{2} s_i \langle s \rangle = \sum_{\text{i}} -B s_i$$

where $B = Jn\langle s \rangle / 2$ is the effective mean field felt by each spin. This is an easy problem of independent spins, with the (selfconsistent) solution

$$\langle s \rangle = \tanh(B/kT) = \tanh(Jn\langle s \rangle / (2kT))$$

So we again got a selfconsistent solution, however, different from the correct one, with twice smaller the selfconsistent field.

We begin by constructing a trial Hamiltonian

$$H' = \sum_{\text{i}} -Bs_i$$

where B is so far unknown parameter. The free energy for the model with the Hamiltonian H' is

$$F' = -kT\ln Z = -kT\ln((2\cosh(B/kT))^N) = -NkT\ln(2\cosh(B/kT))$$

The variation method still needs the mean value of the true Hamiltonian averaged over the canonical ensemble generated my the trial Hamiltonian (we denote this as $\langle H\rangle'$), and we also need the mean value of the trial Hamiltonian averaged over the same canonical ensemble generated by the same trial hamiltonian (we denote this as $\langle H'\rangle'$. We get

$$\langle H\rangle' = \langle\sum_{\text{link}} -Js_is_j\rangle' = \sum_{\text{link}} -J\langle s_is_j\rangle' = -\frac{n}{2}JN\langle s\rangle'^2$$

because the spins are independent in the trial ensemble and the lattice with N spins has $nN/2$ links.

$$\langle H'\rangle' = \langle\sum_{\text{i}} -Bs_i\rangle' = -BN\langle s\rangle'$$

Variation function is

$$\langle H\rangle' - \langle H'\rangle' + F' = -\frac{n}{2} JN\langle s\rangle'^2 + BN\langle s\rangle' - NkT\ln(2\cosh(B/kT))$$

The value of the variation parameter B is found by the minimization of the variation function. Differentiating with respect to B we get

$$-nJN\langle s\rangle' \frac{d\langle s\rangle'}{dB} + N\langle s\rangle' + BN\frac{d\langle s\rangle'}{dB} - N\tanh(B/kT) = 0$$

On the other hand the following relation holds

$$\langle s\rangle' = \tanh(B/kT)$$

and so we get

$$-nJN\langle s\rangle' \frac{d\langle s\rangle'}{dB} + BN\frac{d\langle s\rangle'}{dB} = 0$$

$$B = nJ\langle s\rangle'$$

We got the same expression es in the intuitive Boltzmann-like approach[2].

Landau theory of the phase transitions of the second kind

The essence of the Landau phenomenological approach is the idea where the non-analytic behaviour of the thermodynamic potentia can come from. Let us assume that we have an order parameter M like magnetization. In the equilibrium state this order parameter assumes some equilibrium value. One can calculate this value when one first deals with the non-equilibrium thermodynamic potential, which, in addition to its standard variables (p and T for the Gibbs potential) depends also on the (non-equilibrium) value of the order parameter M. Now we look for the minimum of the thermodynamic potential. The value of M for which the minimum is reached is just the equilibrium value of the order parameter.

$$G_{neq}(p, T, M)$$

$$M_{eq} = \text{argmin } G_{neq}(p, T, M)$$

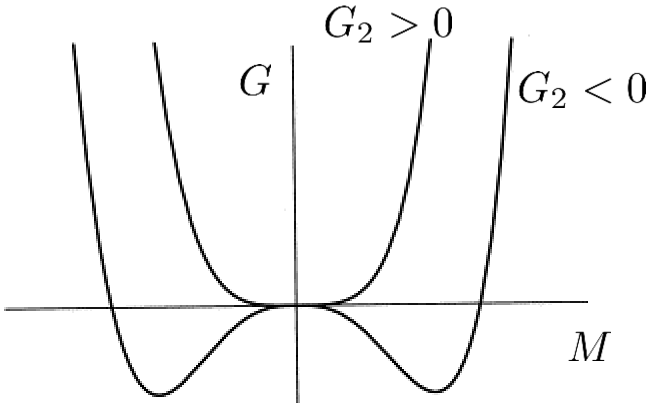
$$G_{eq} = G_{neq}(p, T, M_{neq})$$

Let us consider a case where the system is symmetric with respect to the change of sign of the order parameter

$$M \to -M$$

The pramater is chosen so that it assumes the value zero at the critical point. According to the assumption of analyticity we can expand the non-equilibrium potential (per one molecule) into the Taylor series in M in the form

$$G(p, T, M) = G_0(p, T) + G_2(p, T)M^2 + G_4(p, T)M^4$$



$$G_2(p, T) = \frac{a(p)}{T_c}(T - T_c)$$

$$t = \frac{T - T_c}{T_c}$$

$$M = (\frac{-at}{2G_4})^{1/2}$$

Critical exponent

Phase transitions occur at a certain temperature, called the critical temperature T_c. We want to describe the behaviour of a physical quantity f in terms of a power law around the critical temperature. So we introduce the reduced temperature $\tau := (T - T_c)/T_c$, which is zero at the phase transition, and define the critical exponent k:

$$k \overset{\text{def}}{=} \lim_{\tau \to 0} \frac{\log |f(\tau)|}{\log |\tau|}.$$

This results in the power law we were looking for:

$$f(\tau) \propto \tau^k, \quad \tau \approx 0.$$

It is important to remember that this represents the asymptotic behavior of the function $f(\tau)$ as $\tau \to 0$.

Critical exponents for τ > 0 (disordered phase)		**Critical exponents for τ < 0 (ordered phase)**	
Greek letter	relation	Greek letter	relation
α	$C \propto \tau^{-\alpha}$	α'	$C \propto (-\tau)^{-\alpha'}$
γ	$\chi \propto \tau^{-\gamma}$	β	$\Psi \propto (-\tau)^{\beta}$
ν	$\xi \propto \tau^{-\nu}$	γ'	$\chi \propto (-\tau)^{-\gamma'}$
		ν'	$\xi \propto (-\tau)^{-\nu'}$

Ψ	order parameter	χ	the susceptibility/compressibility/etc.; $\frac{\partial \Psi}{\partial J}$
C	specific heat; $-T\frac{\partial^2 f}{\partial T^2}$	ξ	correlation length

In the presence of an external (magnetic) field the (non-equilibrium) thermodynamic potential will be

$$G(p,T,M) = G_0(p,T) - HM + G_2(p,T)M^2 + G_4(p,T)M^4$$

The critical isotherm (a graph showing dependence of M on H for constant $T = T_c$) will be

$$M \sim H^{1/3}$$

The magnetic susceptibility

$$\left(\frac{\partial M}{\partial H}\right)_{T=T_c} \sim H^{-2/3}$$

diverges at the critical point.

Fázový prechod druhého druhu: nespojité (divergujúce)druhé derivácie termodynamického potenciálu

Transfer matrix

So let us investigate the one-dimensional Ising model with N spins in external magnetic field. The energy of the system is given as

$$E_N = -J \sum_{i=0}^{N-1} s_i s_{i+1} - \mu_B B \sum_{i=0}^{N-1} s_i,$$

with $s_i = \pm 1$. Let us assume periodic boundary conditions $s_N = s_0$ (topologically the spins are distributed along a circle) We expect that in the thermodynamic limit $N \to \infty$ this choice will not influence the result.

The trick consists in suitable rewriting of the statistical sum

$$\begin{aligned}
\mathrm{Z}_N &= \sum_{\{s\}} \exp\left[\beta J(s_0 s_1 + \ldots + s_{N-1}s_N) + \beta\mu_B B(s_0 + s_1 + \ldots + s_{N-1})\right] = \\
&= \sum_{\{s\}} \exp\left[\beta J s_0 s_1 + \beta\mu_B B\frac{s_0+s_1}{2}\right] \cdot \exp\left[\beta J s_1 s_2 + \beta\mu_B B\frac{s_1+s_2}{2}\right] \cdot \ldots \\
&\quad \ldots \cdot \exp\left[\beta J s_{N-1}s_0 + \beta\mu_B B(s_{N-1}+s_0)/2\right] \equiv \sum_{\{s\}} \mathsf{T}_{0,1}\mathsf{T}_{1,2}\ldots\mathsf{T}_{N-1,0}
\end{aligned}$$

where summing over $\{s\}$ is meant as summing over all possible spin configurations what means a multidimensional sum over $s_0, s_1, \ldots, s_{N-1}$. We have introduced a matrix T, with matrix elements $\mathsf{T}_{s_i,s_{i+1}} \equiv \mathsf{T}_{i,i+1}$ (so the matrix rows are numbered by the values of the spin s_i, the columns are numbered by values of the spin s_{i+1} In our specific case the matrix is

$$\mathsf{T} = \begin{pmatrix} \mathrm{e}^{\beta(J+\mu_B B)} & \mathrm{e}^{-\beta J} \\ \mathrm{e}^{-\beta J} & \mathrm{e}^{\beta(J-\mu_B B)} \end{pmatrix}.$$

The expression for Z_N can be simplified when we notice that summing over the spins effectively means matrix multiplication. What is left is only summation over s_0:

$$\mathrm{Z}_N = \sum_{s_0=\pm 1} \left(\mathsf{T}^N\right)_{0,0} = \mathrm{Sp}(\mathsf{T}^N) = \sum_{i=1}^{n} \lambda_i^N.$$

$$\mathrm{Z}_N = \sum_{s_0=\pm 1} \left(\mathsf{T}^N\right)_{0,0} = \mathrm{Sp}(\mathsf{T}^N) = \sum_{i=1}^{n} \lambda_i^N.$$

$$f = -kT \cdot \lim_{N\to\infty} \frac{1}{N} \ln \mathrm{Z}_N = -kT \cdot \lim_{N\to\infty} \frac{1}{N} \ln \left[\lambda_0^N \left(1 + \sum_{i=1}^{N-1} (\lambda_i/\lambda_0)^N\right) \right] = -kT \ln \lambda_0$$

Ising model on a square lattice

$$Z = \sum_{states} \prod_{links} \exp(J s_i s_j)$$

$$\exp(J), \exp(-J)$$

and these two values can be expressed using a suitable small variable, using the identity

$$\exp(J s_i s_j) = \frac{\exp(J) + \exp(-J)}{2} + s_i s_j \frac{\exp(J) - \exp(-J)}{2}$$

which holds for

$$s_i s_j \in \{1, -1\}$$

We have

$$\exp(J s_i s_j) = \cosh(J)(1 + s_i s_j t)$$

where

$$t = \tanh(J)$$

Small t means large T. High temperature expansion.

$$Z = \cosh^{2N}(J) \sum_{states} \prod_{links} (1 + s_i s_j t) = 2^N Z' \cosh^{2N}(J)$$

where

$$Z' = \frac{1}{2^N} \sum_{states} \prod_{links} (1 + s_i s_j t)$$

Expanding the product we get

$$Z' = \frac{1}{2^N} \sum_{s_1=\pm 1} \sum_{s_2=\pm 1} \cdots \sum_{s_N=\pm 1} \{1 + t \sum_{links} s_i s_j + t^2 \sum_{\text{link pairs}} (s_i s_j)(s_m s_n) + \ldots\}$$

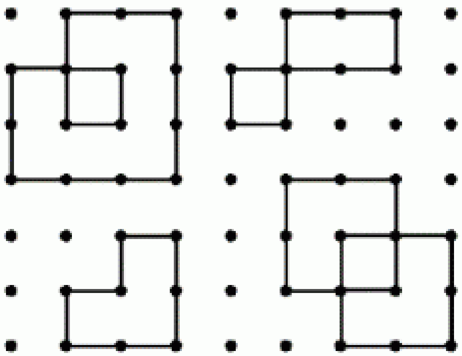
Now one has to realize that after summing over the spin states all the products where at least one spin is presented odd number of times (as a factor) give zero. If the spin appears in the product even number of times the product is equal to one and the sum over all the spin states gives 2^N.

So we get

$$Z' = \sum_L g(L) t^L$$

where

$$g(L)$$



So we get

$$Z' = \sum_{L} g(L) t^{L}$$

where

$$g(L)$$

denotes the number of possibilities how L links can be highlighted on a square lattice in such a way that every site is highlighted even number of times (zero count as even number here). Expressed differently, $g(L)$ is the number of graphs which can be drawn on a square lattice in such a way that each site on the graph is connected with the remaining sites by even number of edges.

Low temperature expansion

$$Z = \sum_{states} \prod_{links} \exp(j s_i s_j)$$

At low temperatures the spins prefer to be oriented in parallel if they sit on the same link. So we expect that at low temperatures those links contribute which have "their spins" oriented in parallel (both having value $+1$ or -1).

For a given spin state each link with parallelly oriented spins contributes to the statistical sum by the factor

$$\exp(j)$$

On the other hand each link with antiparallel spins contributes by the factor

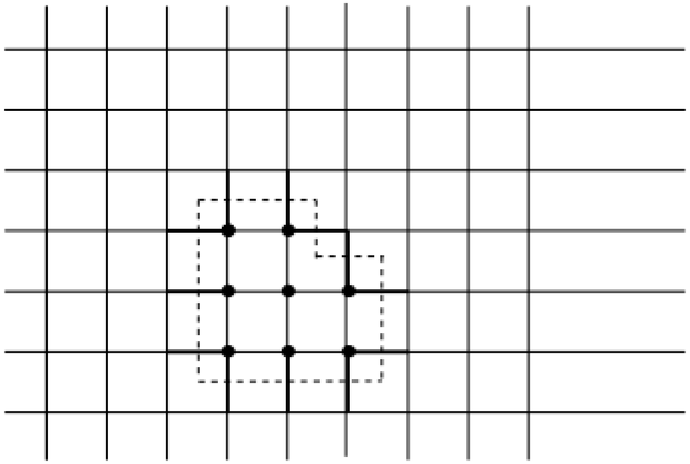
$$\exp(-j)$$

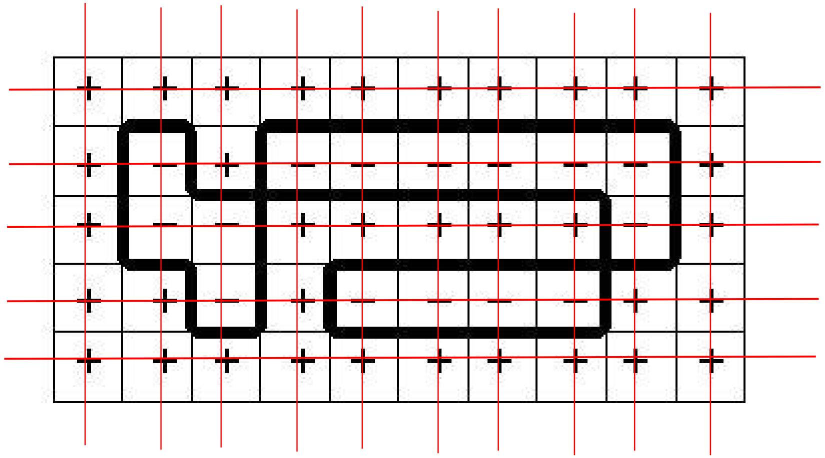
$$Z = \sum_{states} \exp((2N - L)j) \exp(-Lj)$$

$$Z = \sum_{states} \exp(2Nj)(\exp(-2j))^L$$

where $2N$ is the total number of links (for N sites) and the number L depends on the particular spin states and gives the number of links with antiparallel spins on their sites.

Dual latices





The statistical sum can be then rewritten as

$$Z = \exp(2Nj) \sum_L m(L)(\exp(-2j))^L$$

where $m(L)$ is the number of states to which (after highlighting) correspond a graph with exactly L links highlighted. Rephrased differently $m(L)$ je is the number of possibilities how a graph with L edges can be drawn on a lattice in such a

$$m(L) = g(L)$$

So we have got the high-temperature expansion

$$Z(J) = 2^N \cosh^{2N}(J) \sum_L g(L) t^L$$

where

$$t = \tanh(J)$$

and the low-temperature expansion

$$Z(j) = \exp(2Nj) \sum_L g(L) (\exp(-2j))^L$$

Up to now we have two arbitrary independent coupling constants J and j. We assume however that J is small and j is large. Let us now choose j related to J as

$$\exp(-2j) = \tanh(J)$$

Then we get a selfconsistent relation for the statistical sum

$$Z(J) = 2^N \cosh^{2N}(J) \tanh^N(J) Z(j)$$

where

$$j = -\frac{1}{2} \ln(\tanh(J)$$

Let us suppose now that in the thermodynamic limit $N \to \infty$ the free energy per site (only intensive variable can have thermodynamic limit) has a singularity at some temperature T_c. Let us also suppose that there is just one temperature value for which there is such a singularity. The selfconsistent formula for the statistical sum does not introduce any new singularity. So if there is only one singularity, it must be both at j and at J and the two values must be equal, that is $j = J$ and we get the relation

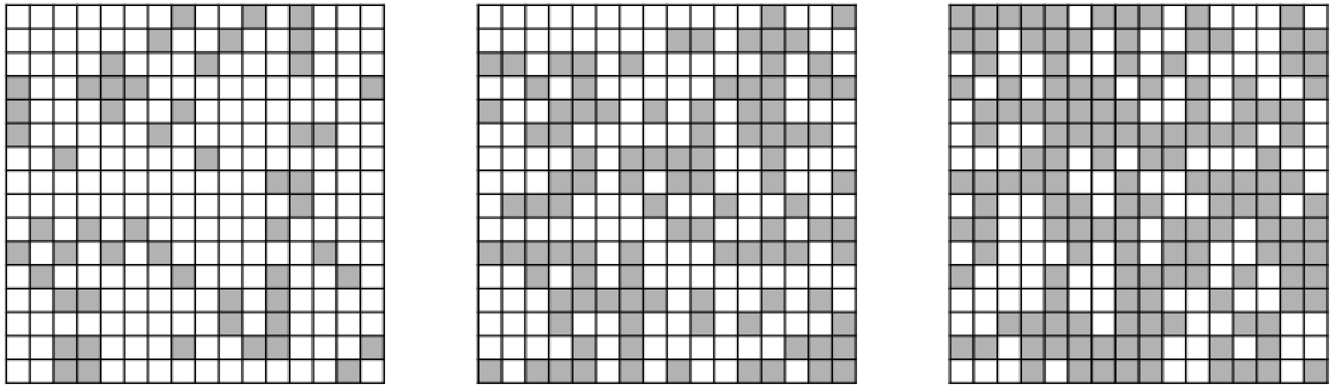
$$J_c = -\frac{1}{2}\ln(\tanh(J_c))$$

The solution gives for the critical value

$$J_c = \frac{1}{2}\ln(1+\sqrt{2})$$

Renormalization group

We start with a model which is more geometrical then physical (there is no temperature nor statistical sum involved). It is, however, very useful, since it is very illustrative. We shall investigate the percolation model. The lives on a lattice of square cells. Each cell can be ”empty” or ”occupied”. The probability p of a cell being occupied is independent on the occupation of any other cell. We introduce the notion of cluster: a set of occupied cells which are touching each other by at least one edge. Now we shall investigate a problem whether there exists a spanning cluster: a cluster which reaches from the top side of the lattice continuously to its bottom side. If yes, the lattice is considered to be percolative (imagine that the spanning cluster provides a path for water to percolate from top to bottom. The figure shows examples of possible lattice states for different values of p.



We shall investigate the limit for the size of the lattice going to infinity. We expect that there exists a critical probability p_c above which the lattice is percolative.

The main idea is that in the critical point the correlation length goes to infinity. The correlations between cells are present for any distances between them. We expect that the system looks qualitatively the same irrespective what zoom we ese to observe it. (So we expect a fractal like behaviour.) The key notion here is ”selfsimilarity”. The renormalization technique looks for the selfsimilar behaviour by averaging the image at some scale it hopes to find a selfsimilar image.

Let us combine the cells of the original lattice into larger square cells, containing $b \times b$ original cells. The new squares are cells of a new, renormalized lattice. Now we design a rule to specify which of the new larger cells are considered as occupied. There are many possible rules to be considered (those which have majority of the original cells occupied, those large cells which are percolative horizontally, or vertically, or in both directions . . .).

Let us take $b = 2$ and the rule: the renormalized cell is considered to be occupied if at least three of its original cells are occupied. Let the probability that the new renormalized cell is occupied is denoted as p'. Clearly

$$p' = p^4 + 4p^3(1 - p) \equiv R(p).$$

$$p' = p^4 + 4p^3(1-p) \equiv R(p).$$

If the original lattice had $p = 0{,}5$ we get $p' = 0.3125$. By next renormalization we would get $p' \doteq 0.0934$. The fix point of these iterative transformations is $p = 0$ what means an empty lattice.So in this specific case the selfsimilarity is not realized. The selfsimilarity requires to get $p' = p$. A numerical solution of this simple algebraic equation gives $p_* \doteq 0.7676$ (the true value is je $p_c \doteq 0.5927$).

To find the critical exponent for the correlation length ξ, one has to realize that all the corresponding lengths in the renormalized lattice are reduced b-times. Therefore $\xi' = \xi/b$. In the vicinity of the critical point we expect $\xi \sim |p - p_c|^{-\nu}$. Comparing the two expressions we get

$$|p' - p_*|^{-\nu} = \frac{1}{b}|p - p_*|^{-\nu}$$

We have replaced here the true value p_c with our fix point p_*, because for our choice of the renormalization rule the singular behaviour is expected for p_*. Now we have to find the relation between p and p' int the vicinity of the critical point

$$p' - p_* = R(p) - R(p_*) \approx (p - p_*) \left.\frac{\mathrm{d}R}{\mathrm{d}p}\right|_{p=p_*} \equiv \lambda(p - p_*).$$

$$|p' - p_*|^{-\nu} = \frac{1}{b}|p - p_*|^{-\nu}$$

$$p' - p_* = R(p) - R(p_*) \approx (p - p_*) \left.\frac{\mathrm{d}R}{\mathrm{d}p}\right|_{p=p_*} \equiv \lambda(p - p_*)$$

$$\lambda^{-\nu}|p - p_*|^{-\nu} = \frac{1}{b}|p - p_*|^{-\nu} \implies \lambda^{-\nu} = b^{-1} \implies \nu = \frac{\ln b}{\ln \lambda}$$

In our case $b = 2$ a $\lambda \doteq 1.6432$, so $\nu \doteq 1.40$ (the exact result is $\nu = 4/3$)

KANONICKÝ FORMALIZMUS PRE POHYB KLASICKEJ ČASTICE V ELEKTROMAGNETICKOM POLI

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \boldsymbol{v}} = \frac{\partial L}{\partial \boldsymbol{r}}$$

Pre pohyb častice s nábojom q v EM poli s intenzitami $\boldsymbol{E}(\boldsymbol{r}, t)$, $\boldsymbol{B}(\boldsymbol{r}, t)$ platí

$$m\frac{\mathrm{d}\boldsymbol{v}}{\mathrm{d}t} = q\boldsymbol{E} + q\boldsymbol{v} \times \boldsymbol{B} \tag{3}$$

ako sme už hovorili v predchádzajúcom článku. Intenzity $\boldsymbol{E}$, $\boldsymbol{B}$ môžeme vyjadriť pomocou potenciálov $\boldsymbol{A}$, φ:

$$\boldsymbol{E} = -\nabla\varphi - \frac{\partial \boldsymbol{A}}{\partial t}, \quad \boldsymbol{B} = \nabla \times \boldsymbol{A} \tag{4}$$

Ukážeme teraz, že Lagrangeovu funkciu nabitej častice s nábojom q v elektromagnetickom poli možno voliť v tvare

$$L = \frac{1}{2}m\boldsymbol{v}^2 + q\boldsymbol{A}.\boldsymbol{v} - q\varphi \tag{5}$$

$$L = \frac{1}{2} m\mathbf{v}^2 + q\mathbf{A}.\mathbf{v} - q\varphi$$

$$\frac{\partial L}{\partial \mathbf{v}} = m\mathbf{v} + q\mathbf{A}$$

$$\mathbf{p} = m\mathbf{v}$$

$$\mathbf{P} = \mathbf{p} + q\mathbf{A}$$

$$H(\mathbf{r},\mathbf{P}) = \frac{1}{2m}(\mathbf{P} - q\mathbf{A})^2 + q\varphi$$

$$\left(\frac{1}{c}\varphi, A^1, A^2, A^3\right) \qquad \frac{1}{c}\frac{\partial}{\partial t}, -\frac{\partial}{\partial x^1}, -\frac{\partial}{\partial x^2}, -\frac{\partial}{\partial x^3}$$

$$\left(\frac{i\hbar}{c}\frac{\partial}{\partial t}, \frac{\hbar}{i}\frac{\partial}{\partial x^1}, \frac{\hbar}{i}\frac{\partial}{\partial x^2}, \frac{\hbar}{i}\frac{\partial}{\partial x^3}\right) = \left(\frac{i\hbar}{c}\frac{\partial}{\partial t}, \mathrm{p}^1, \mathrm{p}^2, \mathrm{p}^3\right)$$

$$\boxed{\begin{aligned} & i\hbar\frac{\partial}{\partial t} \rightarrow i\hbar\frac{\partial}{\partial t} - \frac{1}{c}q\varphi \\ & -i\hbar\nabla \rightarrow -i\hbar\nabla - q\mathbf{A} \end{aligned}}$$

$$Z \propto \int \prod_i d^3 r_i d^3 p_i \, \exp\left(-\beta H(\mathbf{r}_1, \ldots; \mathbf{p}_1, \ldots)\right),$$

where H is the classical Hamilton function,

$$H = \frac{1}{2m} \sum_i \left(\mathbf{p}_i + e\mathbf{A}(\mathbf{r}_i)\right)^2 + V(\mathbf{r}_1, \ldots).$$

But now we can substitute $\mathbf{p}_i \to \tilde{\mathbf{p}}_i = \mathbf{p}_i + e\mathbf{A}(\mathbf{r}_i)$ in the integrals,

This is called the *Bohr-van Leeuwen theorem.*

$$Z \propto \int \prod_i d^3 r_i d^3 \tilde{p}_i \, \exp\left[-\beta\left(\frac{1}{2m} \sum_i \tilde{\mathbf{p}}_i^2 + V \right)\right]$$

Thus we have eliminated the vector potential $\mathbf{A}$ from the partition function.

This is called the *Bohr-van Leeuwen theorem.*

Fluctuation – Response Theorem

$$\bar{E} = \frac{1}{Z}\sum_i E_i \exp(-\beta E_i) = -\frac{\partial}{\partial\beta}\ln Z$$

$$\bar{E} = \frac{1}{Z}\sum_i E_i \exp(-\beta E_i) = -\frac{1}{Z}\frac{\partial Z}{\partial\beta} - \frac{\partial}{\partial\beta}\ln Z$$

$$C_V = \left(\frac{\partial\bar{E}}{\partial\beta}\right)_V \frac{\partial\beta}{\partial T} = -\frac{1}{kT^2}\left[-\frac{1}{Z}\frac{\partial^2 Z}{\partial\beta^2} + \frac{1}{Z^2}\left(\frac{\partial Z}{\partial\beta}\right)^2\right] = \frac{1}{kT^2}\left[\overline{E^2} - \bar{E}^2\right]$$

$$\boxed{C_V = \frac{1}{kT^2}\left[\overline{E^2} - \bar{E}^2\right]}$$

This is how one calculate C_V numerically, since calculating it via $\bar{E}(T+dT) - \bar{E}(T)$ is numerically intractable

Note: cumulant

$$\boxed{\left[\overline{E^2} - \bar{E}^2\right] = \left(-\frac{\partial}{\partial\beta}\right)^2 \ln Z}$$

Fluctuation – Response Theorem for arbitrary quantity

$$E_i \to E_i - A_i J$$

$$\bar{A} = \frac{1}{\beta}\frac{\partial}{\partial J}\ln Z$$

$$\chi_A = \frac{\partial \bar{A}}{\partial J} == \beta\frac{\partial}{\beta\partial J}\bar{A} = \beta\left(\frac{\partial}{\beta\partial J}\right)^2 \ln Z$$

$$\boxed{\chi_A = \beta\overline{\Delta A^2}}$$

Note: if A is extensive, then χ_A is extensive, so

$$\overline{\Delta A^2} \propto N$$

$$\boxed{\frac{\sqrt{\overline{\Delta A^2}}}{\bar{A}} \propto \frac{1}{\sqrt{N}}}$$

$Z(J)$ is a generating function of moments of A

$\ln(\mathrm{Z}(\mathrm{J}))$ is a generating function of cumulants of A.

Fluctuation – Response Theorem for arbitrary quantity

$$E_i \to E_i - A_i J$$

$$\bar{A} = \frac{1}{\beta}\frac{\partial}{\partial J}\ln Z$$

$$\chi_A = \frac{\partial \bar{A}}{\partial J} = \frac{1}{\beta}\frac{\partial}{\partial J}\ln Z = \beta\frac{\partial}{\beta\partial J}\bar{A} = \beta\left(\frac{\partial}{\beta\partial J}\right)^2 \ln Z$$

$$\boxed{\chi_A = \beta\overline{\Delta A^2}}$$

Note: if A is extensive, then χ_A is extensive, so

$$\overline{\Delta A^2} \propto N$$

$$\boxed{\frac{\sqrt{\overline{\Delta A^2}}}{\bar{A}} \propto \frac{1}{\sqrt{N}}}$$

$Z(J)$ is a generating function of moments of A

$\ln(\mathrm{Z}(\mathrm{J}))$ is a generating function of cumulants of A.

Probability of fluctuations

$$p(i) = \frac{1}{Z}\exp(-\beta E_i)$$

$$\bar{M} = \sum_i M_i p(i) = \frac{1}{Z}\sum_i M_i \exp(-\beta E_i)$$

$$p(M) = \frac{1}{Z}\sum_{i, M_i = M} \exp(-\beta E_i)$$

$$\bar{M} = \sum_M M p(M) = \frac{1}{Z}\sum_M M \sum_{i, M_i = M} \exp(-\beta E_i) \frac{1}{Z}\sum_M \sum_{i, M_i = M} M_i \exp(-\beta E_i) =$$

$$= \frac{1}{Z}\sum_i M_i \exp(-\beta E_i)$$

$$p(M) = \frac{1}{Z} \sum_{i, M_i = M} \exp(-\beta E_i) = \frac{1}{Z} Z_{neq}(T, V, M) = \frac{1}{Z} \exp(-\beta F_{neq}(T, V, M))$$

$$p(M) = \exp(-\beta \, \Delta F(M))$$

$$\Delta F(M) = F_{neq}(T, V, M) - F_{eq}(T, V)$$

$$Z = \exp(-\beta F_{eq}(T, V))$$

Be careful:

$$F_{eq}(T, V) \approx F_{neq}(T, V, M_{eq})$$

but

$$Z \neq \exp(-\beta F_{neq}(T, V, M_{eq}))$$

$$10^{23} + 100 \approx 10^{23}$$

$$\exp(10^{23}) \times \exp(100) \neq \exp(10^{23})$$

Sometimes we can reasonably guess the dependence of

$$F_{neq}(T, V, M)$$

on M. Like in the Landau model of phase transitions. Then we can have a phenomenological effective theory like

$$F_{neq}(T, V, M) = F_0(T, V) + F_2(T, V)M^2 + F_4(T, V)M^4$$

$$p(M) = \frac{1}{Z}\exp(-\beta(F_0(T, V) + F_2(T, V)M^2 + F_4(T, V)M^4))$$

Then Z need not be calculated from the original microscopic Hamiltonian, but as a normalization of an effective theory

$$Z = \sum_M \exp(-\beta(F_0(T, V) + F_2(T, V)M^2 + F_4(T, V)M^4))$$

$$p(M) = \frac{1}{Z}\exp(-\beta(F_0(T,V) + F_2(T,V)M^2 + F_4(T,V)M^4))$$

$$Z = \sum_M \exp(-\beta(F_0(T,V) + F_2(T,V)M^2 + F_4(T,V)M^4))$$

It looks like a canonical physics of a simple system with one degree of freedom, M, with a Hamiltonian

$$H(M) = F_0(T,V) + F_2(T,V)M^2 + F_4(T,V)M^4$$

The (nonequilibrium) free energy of the microscopic theory effectively becomes a Hamiltonian of the effective theory. This is why people in textbooks suddenly switch the language and start to speak about “Hamiltonian”.
This happens even more in effective field theories, when one introduces a nonequilibrium field

$$M(\vec{r})$$

The nonequilibrium free energy of the microscopic theory is then a functional

$$F_{neq}[M] = F_{neq}(T,V,M(.)])$$

Ginzburg-Landau Theory

Phenomenological hamiltonian

$$\beta H = \int d\mathbf{x} \left[\frac{t}{2}\mathbf{m}^2 + u\mathbf{m}^4 + \cdots \right.$$
$$\left. + \frac{K}{2}(\nabla \mathbf{m})^2 + \frac{L}{2}(\nabla^2 \mathbf{m})^2 + \frac{N}{2}\mathbf{m}^2(\nabla \mathbf{m})^2 + \cdots - \mathbf{h}\cdot\mathbf{m} \right]$$

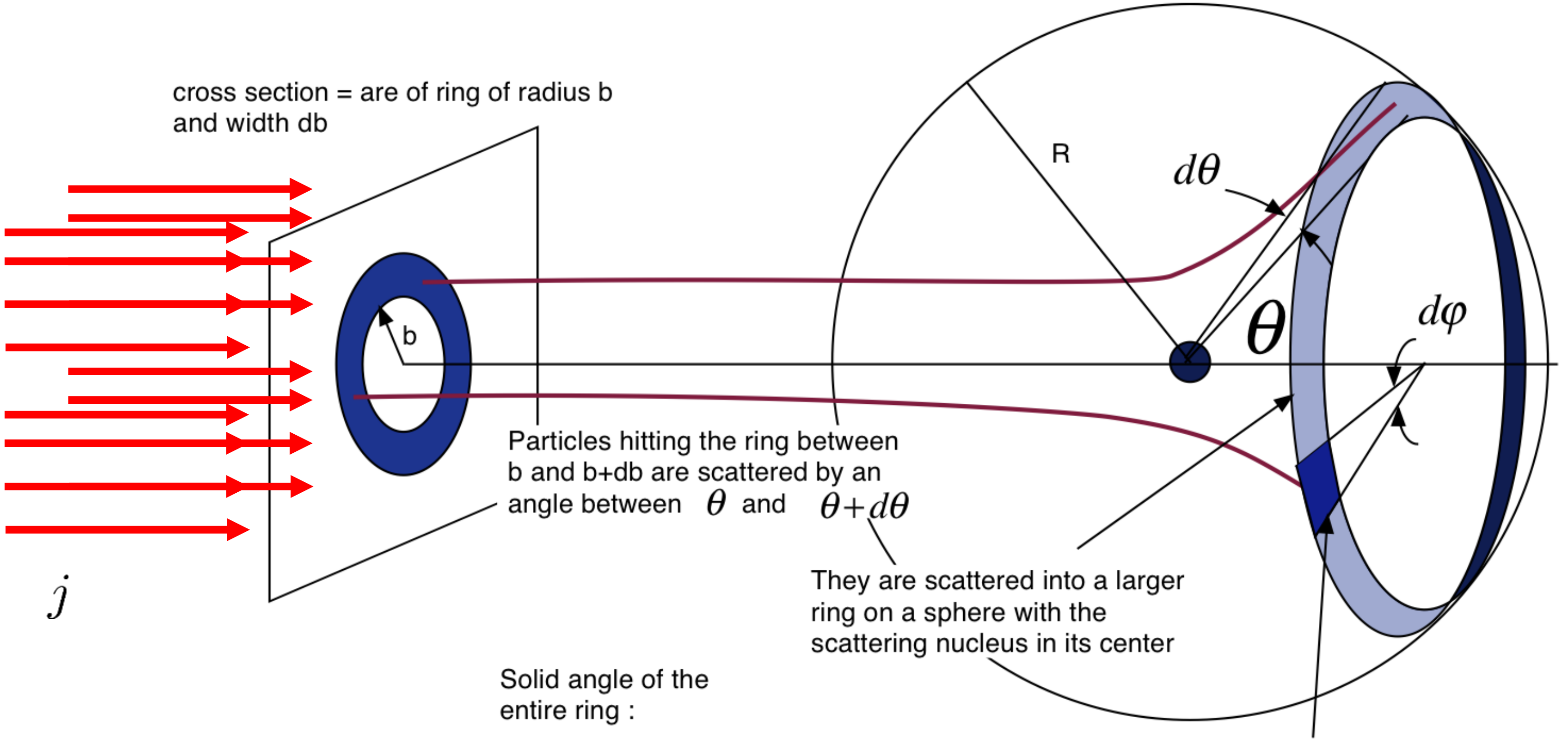
External field $h(\vec{r})$

Z is then a functional of h and is a generating functional of moments of $m(\vec{r})$.

$$\overline{m(\vec{r})} = \frac{\delta}{\delta h(\vec{r})} \ln Z[h]$$

All moments and correlations can be calculated as certain variational derivatives of the generating functional.

Scattering cross section



cross section = are of ring of radius b and width db

Particles hitting the ring between b and b+db are scattered by an angle between θ and $\theta + d\theta$

They are scattered into a larger ring on a sphere with the scattering nucleus in its center

Solid angle of the entire ring :

$$d\Omega = \frac{2\pi R \sin(\theta) R d\theta}{R^2} = 2\pi \sin(\theta) d\theta$$

solid angle of small area:

$$d\Omega = \frac{d\varphi R \sin(\theta) R d\theta}{R^2} = \sin(\theta) d\theta d\varphi$$

j particles/cm^{-2}/s

n detector_clicks/s

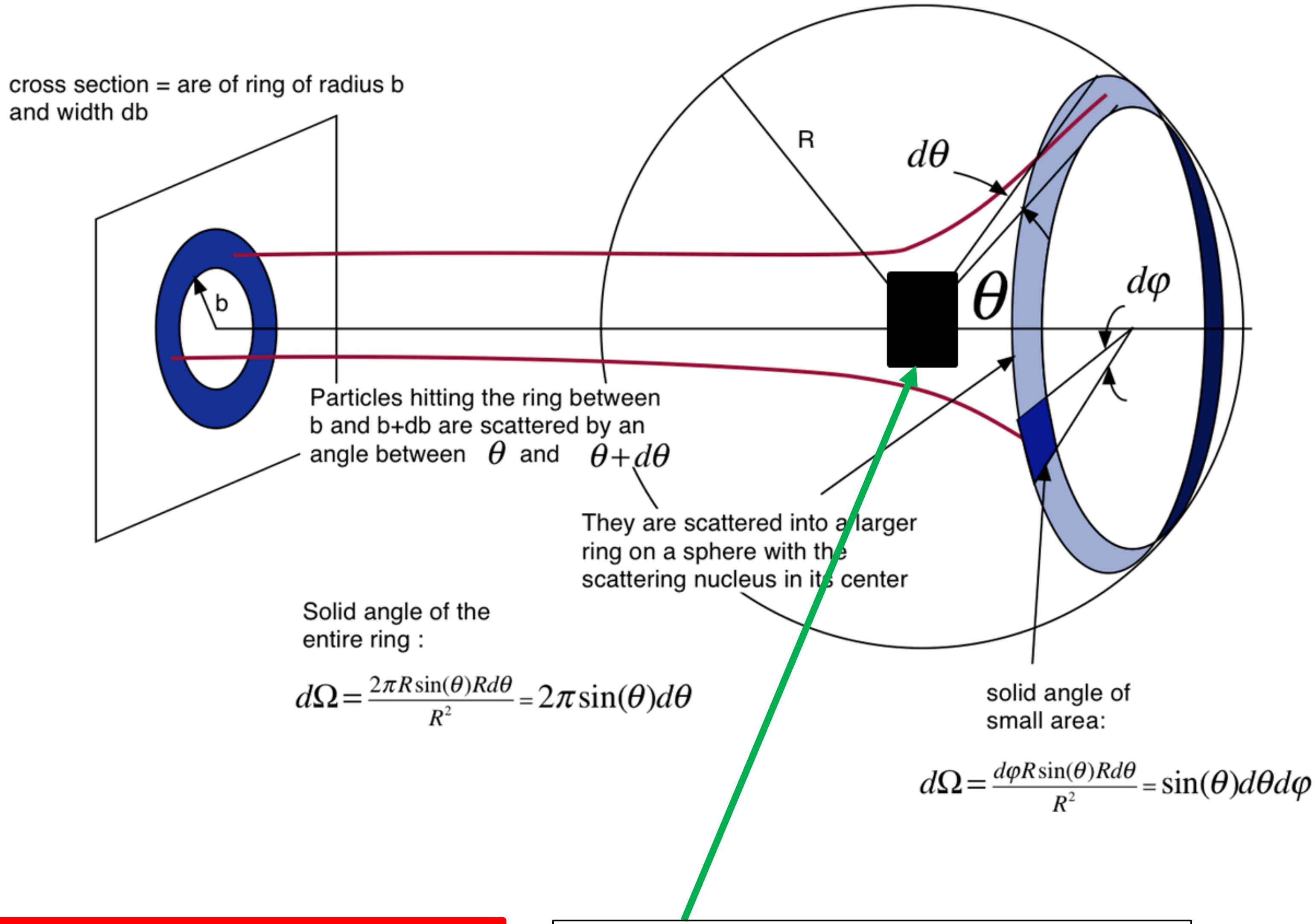
$$\boxed{n = j\sigma}$$

$$\boxed{n = j \frac{d\sigma}{d\Omega} d\Omega}$$

$$n = j\sigma$$

Assumption: all collision events are independent of each other

This may not be true. If a collision destroys the target particle, the target particle is no more available for colliding with the next beam particles, so there would be only one click of the detectors eventually.



cross section = are of ring of radius b and width db

Particles hitting the ring between b and b+db are scattered by an angle between θ and $\theta + d\theta$

They are scattered into a larger ring on a sphere with the scattering nucleus in its center

Solid angle of the entire ring :

$$d\Omega = \frac{2\pi R \sin(\theta) R d\theta}{R^2} = 2\pi \sin(\theta) d\theta$$

solid angle of small area:

$$d\Omega = \frac{d\varphi R \sin(\theta) R d\theta}{R^2} = \sin(\theta) d\theta d\varphi$$

$$\boxed{n = \int_{target} j(\vec{r}) \sigma \rho(\vec{r}) dV}$$

massive target ρ volume density of scattering centers

$$n = \int\limits_{target} j(t,\vec{r})\sigma\rho(\vec{r})dV$$

$$n(t) = L(t)\sigma$$

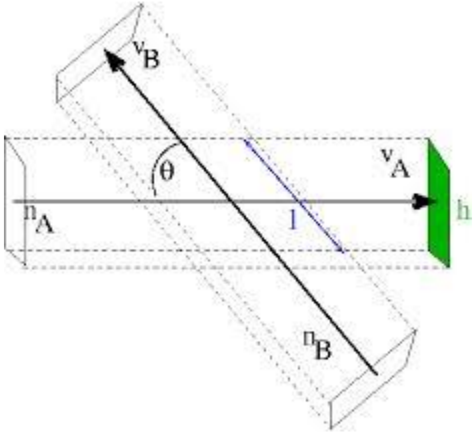
$$N = \int dtL(t)\sigma = L_{int}\sigma$$

L luminosity s^{-1}cm^{-2}
L_{int} integrated luminosity cm^{-2} (inverse milibarns)

Luminosity is difficult to measure directly, beam profile not uniform, fluctuates in time

Solution: calibrate by a well known process with well known cross section

Colliders



$$n = \int_{target} j(\vec{r})\sigma\rho(\vec{r})dV$$

for head-on collisions $$n = \int \rho_a \rho_B |v_A + v_B| \sigma dV$$

$$n = \int \rho_a \rho_B |v_A + v_B| \sigma dV$$

σ may be differential cross section or even may contain δ-functions like

$$\sigma(\vec{v}_1, \vec{v}_2; \vec{v}_1', \vec{v}_2')$$

If detectors measure $\vec{v}_1', \vec{v}_2'$ they measure with certain precision (discrete histograms, not points). So the number of detector counts is the integral over the sensitivity volume, the delta functions for energy momentum conservation get integrated-out.

The first obvious property is the invariance with respect to time inversion

$$\sigma'(\boldsymbol{v}, \boldsymbol{v}_1 \to \boldsymbol{v}', \boldsymbol{v}_1') = \sigma'(-\boldsymbol{v}', -\boldsymbol{v}_1' \to -\boldsymbol{v}, -\boldsymbol{v}_1).$$

Next obvious property is the invariance with respect to reflections.

$$\sigma'(\boldsymbol{v}, \boldsymbol{v}_1 \to \boldsymbol{v}', \boldsymbol{v}_1') = \sigma'(-\boldsymbol{v}, -\boldsymbol{v}_1 \to -\boldsymbol{v}', -\boldsymbol{v}_1'),$$

Combining these two invariances we get the invariance with respect to the inverese scattering which we get from the original proces by exchanging the initial and final state.

$$\sigma'(\boldsymbol{v}, \boldsymbol{v}_1 \to \boldsymbol{v}', \boldsymbol{v}_1') = \sigma'(\boldsymbol{v}', \boldsymbol{v}_1' \to \boldsymbol{v}, \boldsymbol{v}_1).$$

Boltzmann equation

Up to now we have investigated only equilibrium statistical systems. Here we start investigating non-equilibrium system. We shall not present any general theory, we limit ourselves to investigating dilute gas not far from equilibrium in classical (non-quantum) approximation.

We consider a classical dilute gas consisting of just one type of molecules. Classically, we know the state of the gas if wee know the position and the velocity of each of the molecules. If we consider an ensemble of microstates representing some (in general non-equilibrium) macrostate we should describe it by a probability density which is the function of 6N variables. We shall limit ourselves to a drastic approximation: a one-particle description.

So let us look to an infinitesimal region of space around the vector $\boldsymbol{r}$ and ask what is the mean density of molecules at $\boldsymbol{r}$ at time t: let us denote it as $n(\boldsymbol{r}, t)$. Then we ask what is the probability distribution of the velocities of individual molecules in this region. We arrive at the function

$$f(\boldsymbol{r}, \boldsymbol{v}, t)$$

$$f(\boldsymbol{r}, \boldsymbol{v}, t)$$

whose meaning is the following. $f(\boldsymbol{r}, \boldsymbol{v}, t)\, d^3\boldsymbol{r}\, d^3\boldsymbol{v}$ is the mean number of molecules present at time t in the infinitesimal region around $\boldsymbol{r}$ and having velocities in the infinitesimal neighbourhood of vector $\boldsymbol{v}$. The following (normalization condition) holds

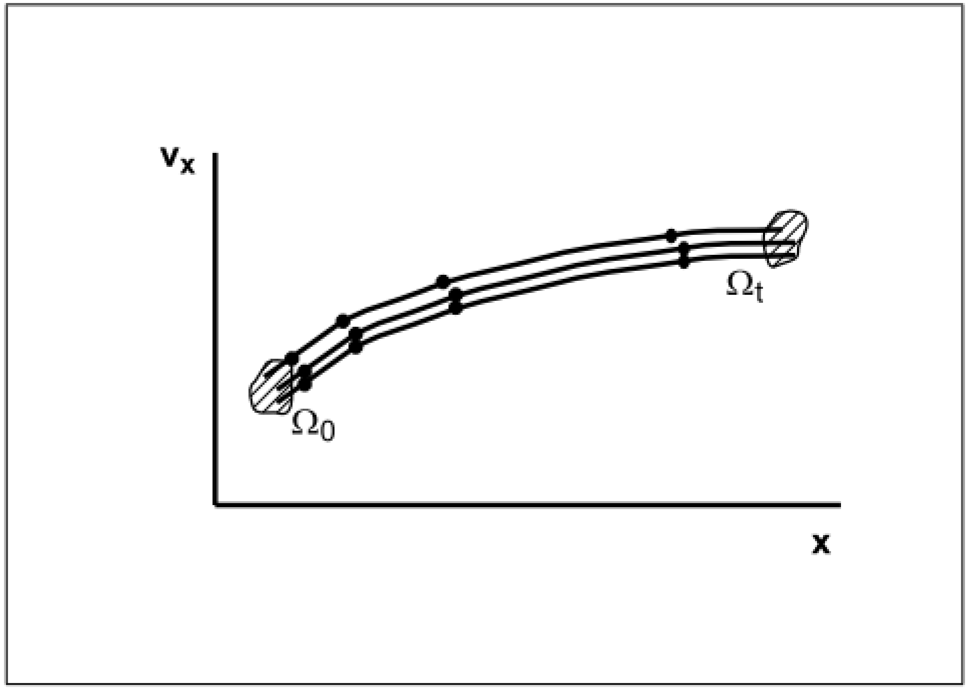
$$n(\boldsymbol{r}) = \int f(\boldsymbol{r}, \boldsymbol{v}, t)\, d^3\boldsymbol{v}$$

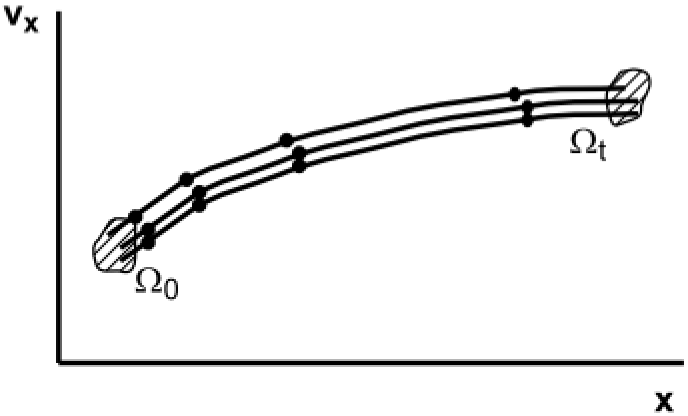
It becomes to be the **oneparticle approximation** when we say that it contains **whole statistical information**. By that we mean, that all statistical results (like mean values of all the physical variables) can be obtained just from function
***f*(*r*, *v*, *t*).**

$$f_2(\boldsymbol{r_1}, \boldsymbol{v_1}, \boldsymbol{r_2}, \boldsymbol{v_2}, t) = f(\boldsymbol{r_1}, \boldsymbol{v_1}, t) f(\boldsymbol{r_2}, \boldsymbol{v_2}, t)$$

Ideal gas, no collisions

$$
\begin{aligned}
\boldsymbol{v}(t) &= \boldsymbol{v}_0 + \frac{\boldsymbol{F}}{m} t \\
\boldsymbol{r}(t) &= \boldsymbol{r}_0 + \boldsymbol{v}_0 t + \frac{1}{2}\frac{\boldsymbol{F}}{m} t^2
\end{aligned}
$$





$$\int_{\Omega_0} f(\boldsymbol{r},\boldsymbol{v},t=0)d^3\boldsymbol{r}d^3\boldsymbol{v} = \int_{\Omega_t} f(\boldsymbol{r},\boldsymbol{v},t)d^3\boldsymbol{r}d^3\boldsymbol{v}$$

The integration variables at the right-hand side can be denoted arbitrarily, so let us denote them as $\boldsymbol{r}_t, \boldsymbol{v}_t$ and we write

$$\int_{\Omega_0} f(\boldsymbol{r},\boldsymbol{v},t=0)d^3\boldsymbol{r}d^3\boldsymbol{v} = \int_{\Omega_t} f(\boldsymbol{r}_t,\boldsymbol{v}_t,t)d^3\boldsymbol{r}_t d^3\boldsymbol{v}_t$$

$$\int_{\Omega_0} f(\boldsymbol{r}, \boldsymbol{v}, t=0) d^3\boldsymbol{r} d^3\boldsymbol{v} = \int_{\Omega_t} f(\boldsymbol{r}_t, \boldsymbol{v}_t, t) d^3\boldsymbol{r}_t d^3\boldsymbol{v}_t$$

In the integral at the right-hand side we perform a substitution. The variables $\boldsymbol{r}_t$, $\boldsymbol{v}_t$ will be substituted by variables $\boldsymbol{r}_0, \boldsymbol{v}_0$. The substitution relations will be those of (8.1),

$$\begin{aligned} \boldsymbol{v}_t &= \boldsymbol{v}_0 + \frac{\boldsymbol{F}}{m} t \\ \boldsymbol{r}_t &= \boldsymbol{r}_0 + \boldsymbol{v}_0 t + \frac{1}{2} \frac{\boldsymbol{F}}{m} t^2 \end{aligned}$$

When the "old" variables $\boldsymbol{r}_t, \boldsymbol{v}_t$ run through the region Ω_t, the "new" variables $\boldsymbol{r}_0, \boldsymbol{v}_0$ run through the regionΩ_0 and we get

$$\int_{\Omega_0} f(\boldsymbol{r}, \boldsymbol{v}, t=0) d^3\boldsymbol{r} d^3\boldsymbol{v} = \int_{\Omega_0} f(\boldsymbol{r}_t, \boldsymbol{v}_t, t) J \, d^3\boldsymbol{r}_0 d^3\boldsymbol{v}_0$$

It is understood that the variables $\boldsymbol{r}_t, \boldsymbol{v}_t$ are expressed through the integration variables $\boldsymbol{r}_0, \boldsymbol{v}_0$ and J is the Jacobian of the substitution transformation

According to the Liouville theorem

$$J = 1$$

a so we get

$$\int_{\Omega_0} f(\boldsymbol{r}, \boldsymbol{v}, t = 0) d^3\boldsymbol{r} d^3\boldsymbol{v} = \int_{\Omega_0} f(\boldsymbol{r}_t, \boldsymbol{v}_t, t)\, d^3\boldsymbol{r}_0 d^3\boldsymbol{v}_0$$

Now we change the notation of the integration variables on the left-hand side for $\boldsymbol{r}_0, \boldsymbol{v}_0$ and we get

$$\int_{\Omega_0} f(\boldsymbol{r}_0, \boldsymbol{v}_0, t = 0) d^3\boldsymbol{r}_0 d^3\boldsymbol{v}_0 = \int_{\Omega_0} f(\boldsymbol{r}_t, \boldsymbol{v}_t, t)\, d^3\boldsymbol{r}_0 d^3\boldsymbol{v}_0$$

This relation holds for an arbitrary region of integration Ω_0 thus the integrands must be equal

$$f(\boldsymbol{r}_0, \boldsymbol{v}_0, t = 0) = f(\boldsymbol{r}_t, \boldsymbol{v}_t, t)$$

This relation holds for arbitrary t, therefore

$$\frac{d}{dt} f(\boldsymbol{r}_t, \boldsymbol{v}_t, t) = 0$$

From there we get the ”Boltzmann equation without the right-hand side”

$$\frac{\partial}{\partial t} f + \boldsymbol{v}.\frac{\partial}{\partial \boldsymbol{r}} f + \frac{\boldsymbol{F}}{m} \cdot \frac{\partial}{\partial \boldsymbol{v}} f = 0$$

This equation holds for ideal gas when we neglect collisions between molecules. In a general case the following equation holds

$$\frac{\partial}{\partial t} f + \boldsymbol{v}.\frac{\partial}{\partial \boldsymbol{r}} f + \frac{\boldsymbol{F}}{m} \cdot \frac{\partial}{\partial \boldsymbol{v}} f = D_c f$$

where at the right-hand side one writes so called collision term, which we shall discuss later. The differential operator at the left-hand side is usually denoted as D

$$D = \frac{\partial}{\partial t} + \boldsymbol{v}.\frac{\partial}{\partial \boldsymbol{r}} + \frac{\boldsymbol{F}}{m} \cdot \frac{\partial}{\partial \boldsymbol{v}}$$

And the Boltzmann equation is then written in a compact form

$$Df = D_C f$$

Mean time between collisions

Molecules in gas collide with each other. Let us suppose that the collisions are independent and express the probability that a given particle collides with some other particle in the next infinitesimal time interval dt as

$$\frac{1}{\tau}dt$$

We denote as $\mathcal{P}(t)$ the probability that the given particle does not collide in the time interval t from the present moment. Then

$$\mathcal{P}(t+dt) = \mathcal{P}(t)(1 - \frac{1}{\tau}dt)$$

This differential equation has the solution

$$\mathcal{P}(t) = \exp(-t/\tau)$$

Now let us denote as $P(t)dt$ the probability that the given particle collides for the first time (from now) at the time interval$(t, t+dt)$. Then

$$P(t)dt = \frac{1}{\tau}\exp(-t/\tau)dt$$

$$P(t)dt = \frac{1}{\tau}\exp(-t/\tau)dt$$

This probability is (of course) correctly normalized as

$$\int dt P(t) = 1$$

The mean time up to the next collision is given as

$$\langle t\rangle = \int dt\, t P(t) = \tau$$

This time is also called ”mean time between two collisions” and one can easily check that the notion is correct. Indeed, due to independence of collisions the ”mean time between two collisions” and the ”mean time from now to the next collision” is the same. If the collisions are independent, then the two notions ”from now” and ”from now, when the particle just collided” have the same meaning what concerns future.

Thought experiments

- catch a molecule. Ask her to send you the time of the next collision from now
- catch a molecule. Ask her to tell you the time of the last collision before now
- catch a molecule. Ask her to tell you the time of the last collision and then the time of the next collision. Subtract the two numbers
- catch a molecule. Ask her to give you the recordings of the times of collisions. Subtract always the two consecutive times

Now we shall describe a simple approximate way how to include collisions into the Boltzmann equation. To simplify the explanation, let us assume the gas considered lives in a one-dimensional world without external field. In such a world the molecules between collisions move uniformly (and linearly, but everything is linear in a one-dimensional world).

The expression

$$f(x, v, t)dxdv$$

gives the number of molecules which in the time moment t are in the position interval $(x, x + dx)$ and have velocities from the interval $(v, v + dv)$. Let us ask where these molecules were in past in the time moment $t - t'$, $t' > 0$. If there are no collisions then during the time t' they moved uniformly and so at the time moment $t - t'$ they had velocity $v' = v$ and were positioned in the neighbourhood of the point $x' = x - vt'$. Using the Liouville theorem we get

$$dx\ dv = dx'\ dv'$$

and thus the identity

$$f(x, v, t) = f(x - vt', v, t - t')$$

If particles collide, then every particle which at the time t is at the phase-space point (x, v) have collided somewhere sometimes the last time before the arrival to the point (x, v). After that last collision the particle moved uniformly. So in that last collision the particle has to obtain just the velocity v and if the last collision happened at the time $t - t'$, then the collision must have happened at the position $x' = x - vt'$. Each particle comes from the point of its last collision, so we can write

$$f(x, v, t) = \int dt' \tilde{f}(x - vt', v, t - t') \exp(-\frac{t'}{\tau})$$

Where $\tilde{f}(x, v, t)dx\ dv\ dt$ gives the number of particles which at time t are in the interval $(x, x + dx)$, have velocity from $(v, v + dv)$ and in the infinitesimal time interval $(t - dt, t)$ they just collided, so the velocity v was obtained just in that collision. Other molecules which at the time t are present in the interval $(x, x+dx)$ with velocity from $(v, v+dv)$ are those which did not just collide and their number is

$$f(x, v, t)dxdv$$

The normalization is the following

$$dt\ dx \int dv \tilde{f}(x, v, t) = \frac{dt}{\tau} dx \int dv f(x, v, t)$$

Now the question is how $\tilde{f}$ depends on velocity. Here we make a strong assumption that the distribution $\tilde{f}$, describing those particles which have just collided is Maxwellian By that we assume that the particles which just collided are thermalized, so they are a sample from some equilibrium distribution, which is the Maxwell distribution.

We get the expression for the Boltzmann distribution in the form of integral through the trajectory

$$f(x, v, t) = \int dt' \frac{1}{\tau} f_0(x - vt', v, t - t') \exp(-\frac{t'}{\tau})$$

where the index 0 denotes, that the function is Maxwellian what means

$$f_0(x, v) = n(x) \left(\frac{m\beta(x)}{2\pi} \right)^{3/2} \exp(-\frac{1}{2}\beta(x)v^2)$$

If the functions $\beta(x), n(x)$ were known, then the above integral would explicitly define the Boltzmann distribution function. These function, however, are not known a priori. That means that the integral represents a selfconsistent equation,

$$f_0(x,v) = n(x)\left(\frac{m\beta(x)}{2\pi}\right)^{3/2}\exp(-\frac{1}{2}\beta(x)v^2)$$

If the functions $\beta(x), n(x)$ were known, then the above integral would explicitly define the Boltzmann distribution function. These function, however, are not known a priori. That means that the integral represents a selfconsistent equation, which determines the function f my through $\beta(x)$ a $n(x)$, but these functions are determined through f via the consistency equations

$$\begin{aligned} n(x) &= \int dv\ f(x,v) \\ \frac{1}{2}\frac{1}{\beta(x)} &= \int dv \frac{1}{2}mv^2 f(x,v) \end{aligned}$$

Now we derive a differential equation satisfied by the Boltzmann function as defined by the integral through the trajectory. We write the expression defining f for two close points in the phase space which are connected by a particle trajectory. We shall assume the external field is zero, to get simple solutions for the trajectory.

$$f(x,v,t) = \int dt' \frac{1}{\tau} f_0(x - vt', v, t - t') \exp(-\frac{t'}{\tau})$$

$$f(x + v\, dt, v, t + dt) = \int dt' \frac{1}{\tau} f_0(x - vt' + v\, dt, v, t + dt - t') \exp(-\frac{t'}{\tau})$$

In the second expression we make the substitution $t'' = t' - dt$ and we get

$$f(x + v\, dt, v, t + dt) = \int_{-dt}^{\infty} dt'' \frac{1}{\tau} f_0(x - vt'', v, t - t'') \exp(-\frac{t'' + dt}{\tau})$$

We expand the exponential up to the first order and write t' instead of t''. We get

$$\begin{aligned} f(x + v\, dt, v, t + dt) = &\int_{-dt}^{\infty} dt' \frac{1}{\tau} f_0(x - vt', v, t - t') \exp(-\frac{t'}{\tau}) + \\ &+ \int_{-dt}^{\infty} dt' \frac{1}{\tau} f_0(x - vt', v, t - t') \exp(-\frac{t'}{\tau}) \frac{dt}{\tau} \end{aligned}$$

In the second integral the integrand is of the first order in dt so we can shift the lower limit infinitesimally and get

$$\begin{aligned} f(x+v\,dt,v,t+dt) &= \int_{-dt}^{\infty} dt'\frac{1}{\tau} f_0(x-vt',v,t-t')\exp(-\frac{t'}{\tau}) + \\ &\quad -\frac{dt}{\tau}\int_0^{\infty} dt'\frac{1}{\tau} f_0(x-vt',v,t-t')\exp(-\frac{t'}{\tau}) \end{aligned}$$

$$f(x,v,t) = \int dt'\frac{1}{\tau} f_0(x-vt',v,t-t')\exp(-\frac{t'}{\tau})$$

$$\begin{aligned} f(x+v\,dt,v,t+dt) - f(x,v,t) &= \int_{-dt}^{0} dt'\frac{1}{\tau} f_0(x-vt',v,t-t')\exp(-\frac{t'}{\tau}) + \\ &\quad -\frac{dt}{\tau}\int_0^{\infty} dt'\frac{1}{\tau} f_0(x-vt',v,t-t')\exp(-\frac{t'}{\tau}) \\ &= \frac{dt}{\tau} f_0(x,v,t) - \frac{dt}{\tau} f(x,v,t) \end{aligned}$$

We have got the equation

$$Df = -\frac{f-f_0}{\tau}$$

We have got the equation

$$Df = -\frac{f - f_0}{\tau}$$

Boltzmann equation with this choice of the collision term at the right-hand side is called Baltzmann equation in the approximation of relaxation time and is usually just postulated after presenting a few more-or-less plausible arguments. Our approach presented here is taken from Reif.

$$\boldsymbol{u}(\boldsymbol{r},t) \equiv \langle \boldsymbol{v}(\boldsymbol{r},t) \rangle = \frac{1}{n(\boldsymbol{r},t)} \int d^3\boldsymbol{v}\, f(\boldsymbol{r},\boldsymbol{v},t)\boldsymbol{v} \tag{13·1·5}$$

This velocity $\boldsymbol{u}(\boldsymbol{r},t)$ describes the mean velocity of flow of the gas at a given point. (This is just the "hydrodynamic velocity" of the fluid described by macroscopic hydrodynamics.) It is useful to measure the velocity $\boldsymbol{v}$ of a molecule with respect to this mean velocity. We shall therefore define the "peculiar velocity" $\boldsymbol{U}$ of a molecule by the relation

$$\boldsymbol{U} \equiv \boldsymbol{v} - \boldsymbol{u} \tag{13·1·6}$$

Thus it follows, by the definition (13·1·5), that

$$\langle \boldsymbol{U} \rangle = \langle \boldsymbol{v} \rangle - \boldsymbol{u} = 0 \tag{13·1·7}$$

In considering transport phenomena, one is interested in calculating the fluxes of various quantities. Consider at time t and at the point $\boldsymbol{r}$ an infinitesimal element of area dA whose normal is denoted by the unit vector $\hat{\boldsymbol{n}}$.

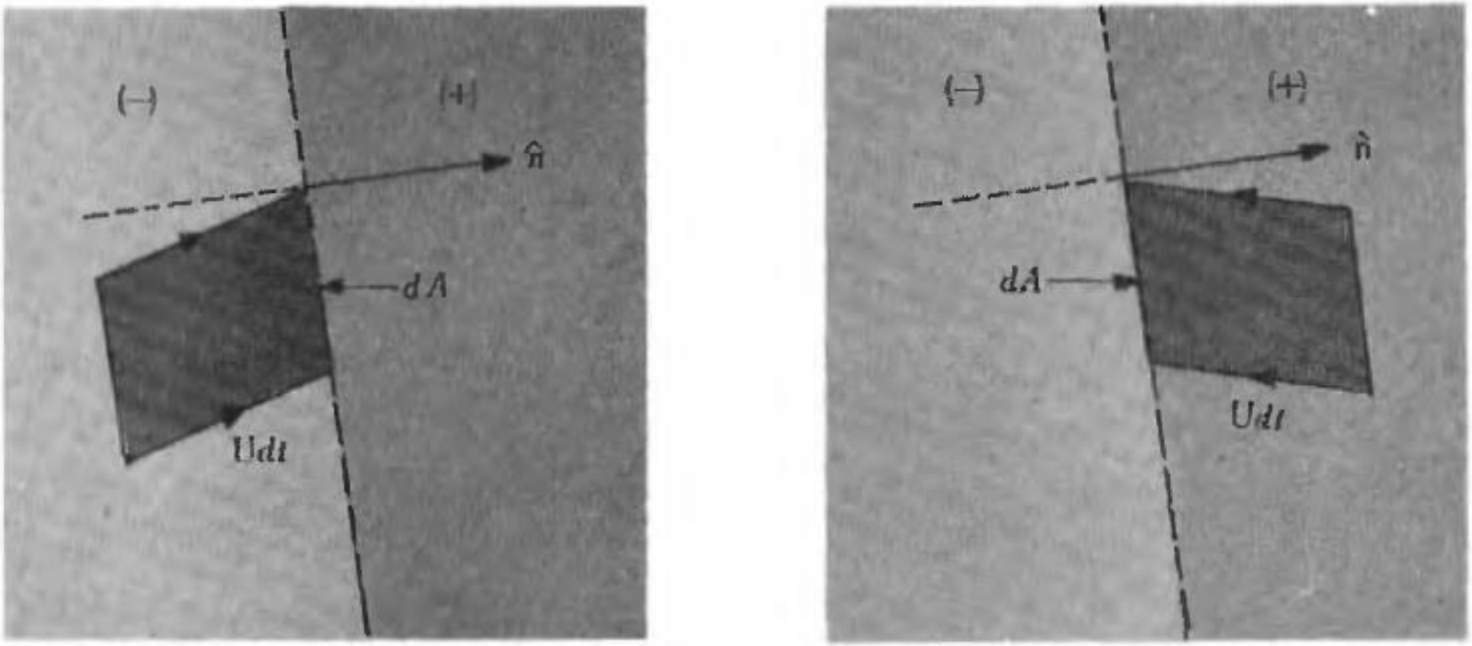


***Fig. 13·1·1** The element of area dA with normal $\hat{n}$ divides the gas into a (+) and (−) region and moves with velocity u. The figure illustrates molecules crossing the element of area in time dt from the (−) to the (+) side (left diagram), and from the (+) to the (−) side (right diagram).*

$\mathfrak{F}_n(\mathbf{r},t) \equiv$ the net amount of χ which is transported per unit time per unit area of an element of area (oriented with its normal along $\hat{\mathbf{n}}$) from its (−) to its (+) side.

▶
$$\mathcal{F}_n(\mathbf{r},t) = \int d^3\mathbf{v}\; f\hat{\mathbf{n}} \cdot \mathbf{U}\chi \tag{13·1·12}$$

In terms of the definition (13·1·4) this can be written

▶
$$\mathcal{F}_n(\mathbf{r},t) = n\,\langle \hat{\mathbf{n}} \cdot \mathbf{U}\chi \rangle \tag{13·1·13}$$

Thus $\mathcal{F}_n$ can be regarded as the $\hat{\mathbf{n}}$ component of a flux *vector* $\boldsymbol{\mathcal{F}}$ such that

$$\mathcal{F}_n = \hat{\mathbf{n}} \cdot \boldsymbol{\mathcal{F}} \tag{13·1·14}$$

where
$$\boldsymbol{\mathcal{F}} = n\langle \mathbf{U}\chi \rangle \tag{13·1·15}$$

Examples In calculating the viscosity of a gas (as we did in Sec. 12·3), one is interested in finding $P_{z\alpha}$, the α component of the mean stress exerted, on a unit area of surface with normal oriented along the z axis, by the fluid below this surface on the fluid above this surface. The corresponding rate of change of momentum is given by the net flux of α component of molecular momentum transported from below to above the surface. The quantity transported is thus $\chi = mv_\alpha$, while $\hat{\mathbf{n}} \cdot \mathbf{U} = U_z$. Hence this stress, or momentum flux, is by (13·1·13)

$$P_{z\alpha} = nm\langle U_z v_\alpha \rangle \tag{13·1·16}$$

$u_x \neq 0$ and $u_z = 0$. Then $U_z = v_z$, and (13·1·16) becomes simply

$$P_{z\alpha} = nm\langle v_z v_\alpha \rangle = m\int d^3\mathbf{v}\, f v_z v_\alpha$$

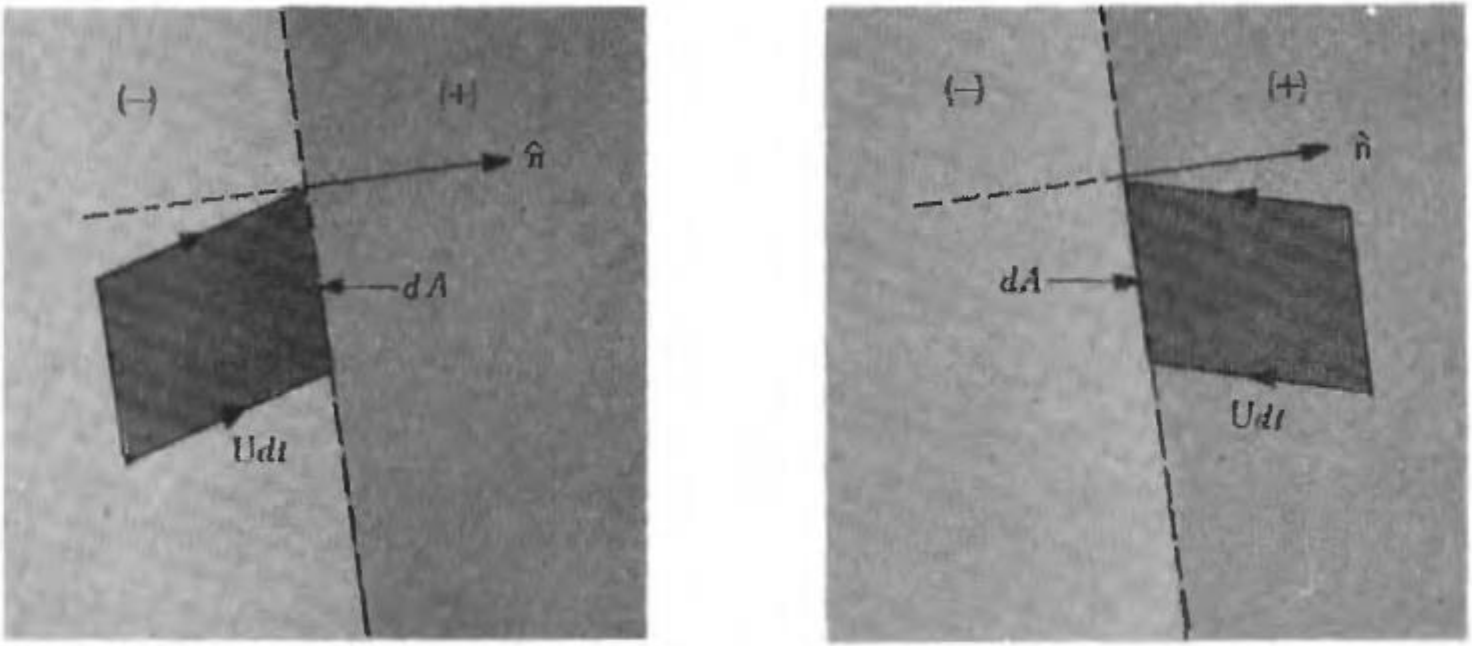


Fig. 13·1·1 ***The element of area dA with normal $\hat{n}$ divides the gas into a (+) and (−) region and moves with velocity u. The figure illustrates molecules crossing the element of area in time dt from the (−) to the (+) side (left diagram), and from the (+) to the (−) side (right diagram).***

$\mathfrak{F}_n(\mathbf{r},t) \equiv$ the net amount of χ which is transported per unit time per unit area of an element of area (oriented with its normal along $\hat{n}$) from its (−) to its (+) side.

▶
$$\mathcal{F}_n(\mathbf{r},t) = \int d^3\mathbf{v}\; f\hat{\mathbf{n}} \cdot \mathbf{U}\chi \qquad (13\cdot 1\cdot 12)$$

In terms of the definition (13·1·4) this can be written

▶
$$\mathcal{F}_n(\mathbf{r},t) = n\,\langle \hat{\mathbf{n}} \cdot \mathbf{U}\chi \rangle \qquad (13\cdot 1\cdot 13)$$

Thus $\mathcal{F}_n$ can be regarded as the $\hat{\mathbf{n}}$ component of a flux *vector* $\boldsymbol{\mathcal{F}}$ such that

$$\mathcal{F}_n = \hat{\mathbf{n}} \cdot \boldsymbol{\mathcal{F}} \qquad (13\cdot 1\cdot 14)$$

where

$$\boldsymbol{\mathcal{F}} = n\langle \mathbf{U}\chi \rangle \qquad (13\cdot 1\cdot 15)$$

Examples In calculating the viscosity of a gas (as we did in Sec. 12·3), one is interested in finding $P_{z\alpha}$, the α component of the mean stress exerted, on a unit area of surface with normal oriented along the z axis, by the fluid below this surface on the fluid above this surface. The corresponding rate of change of momentum is given by the net flux of α component of molecular momentum transported from below to above the surface. The quantity transported is thus $\chi = mv_\alpha$, while $\hat{\mathbf{n}} \cdot \mathbf{U} = U_z$. Hence this stress, or momentum flux, is by (13·1·13)

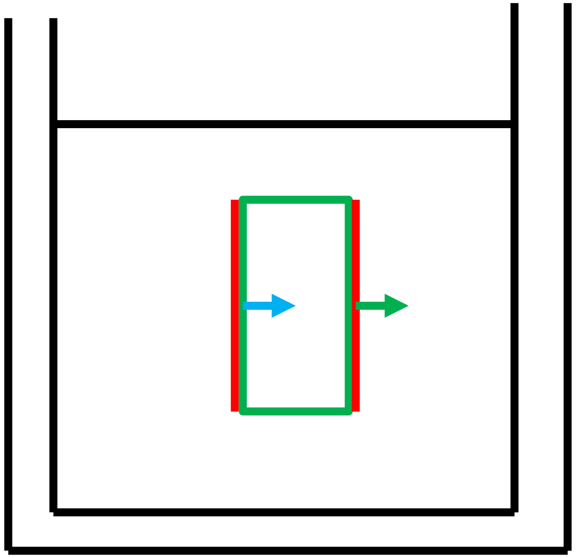
$$P_{z\alpha} = nm\langle U_z v_\alpha \rangle \qquad (13\cdot 1\cdot 16)$$

$u_x \neq 0$ and $u_z = 0$. Then $U_z = v_z$, and (13·1·16) becomes simply

$$P_{z\alpha} = nm\langle v_z v_\alpha \rangle = m\int d^3\mathbf{v}\; f v_z v_\alpha$$

Tlak: $P_{xx} = nm < v_x^2 >$

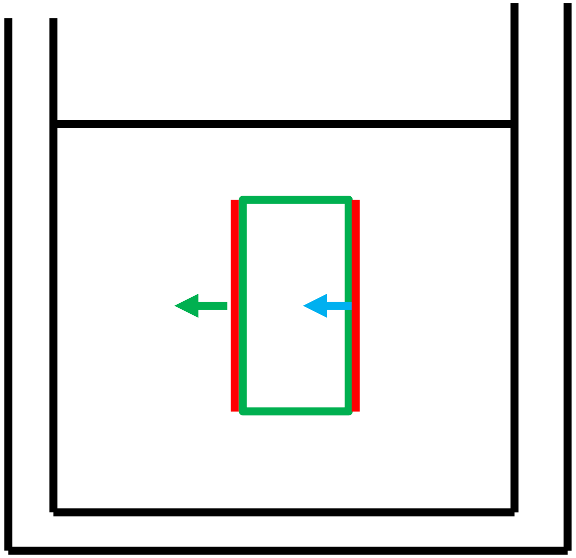
Note: even power of v! Nonzero also for Maxwell (equilibrium) distribution. How it is possible to have non-zero flux of some quantity in equilibrium. Symmetry prohibits equilibrium flux for a scalar quantity, but for a vector quantity like momentum it is possible. Look to the glass of water in equilibrium



$$P_{xx} = nm < v_x^2 > \; > 0$$

Tlak: $P_{zz} = nm < v_z^2 >$

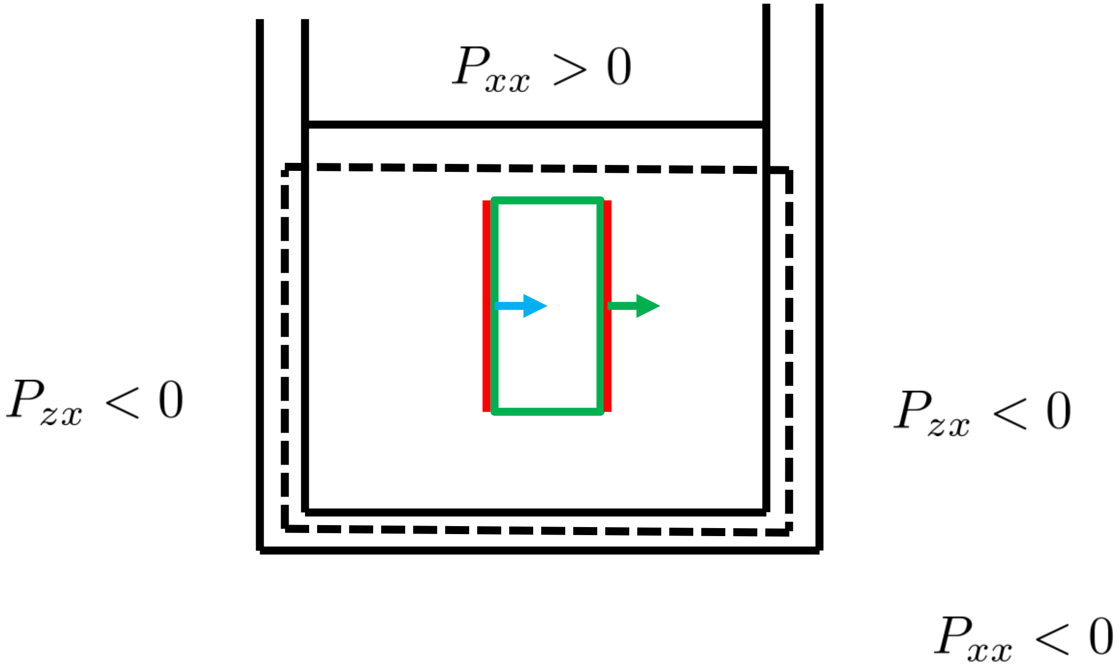
Note: even power of v! Nonzero also for Maxwell (equilibrium) distribution. How it is possible to have non-zero flux of some quantity in equilibrium. Symmetry prohibits equilibrium flux for a scalar quantity, but for a vector quantity like momentum it is possible. Look to the glass of water in equilibrium

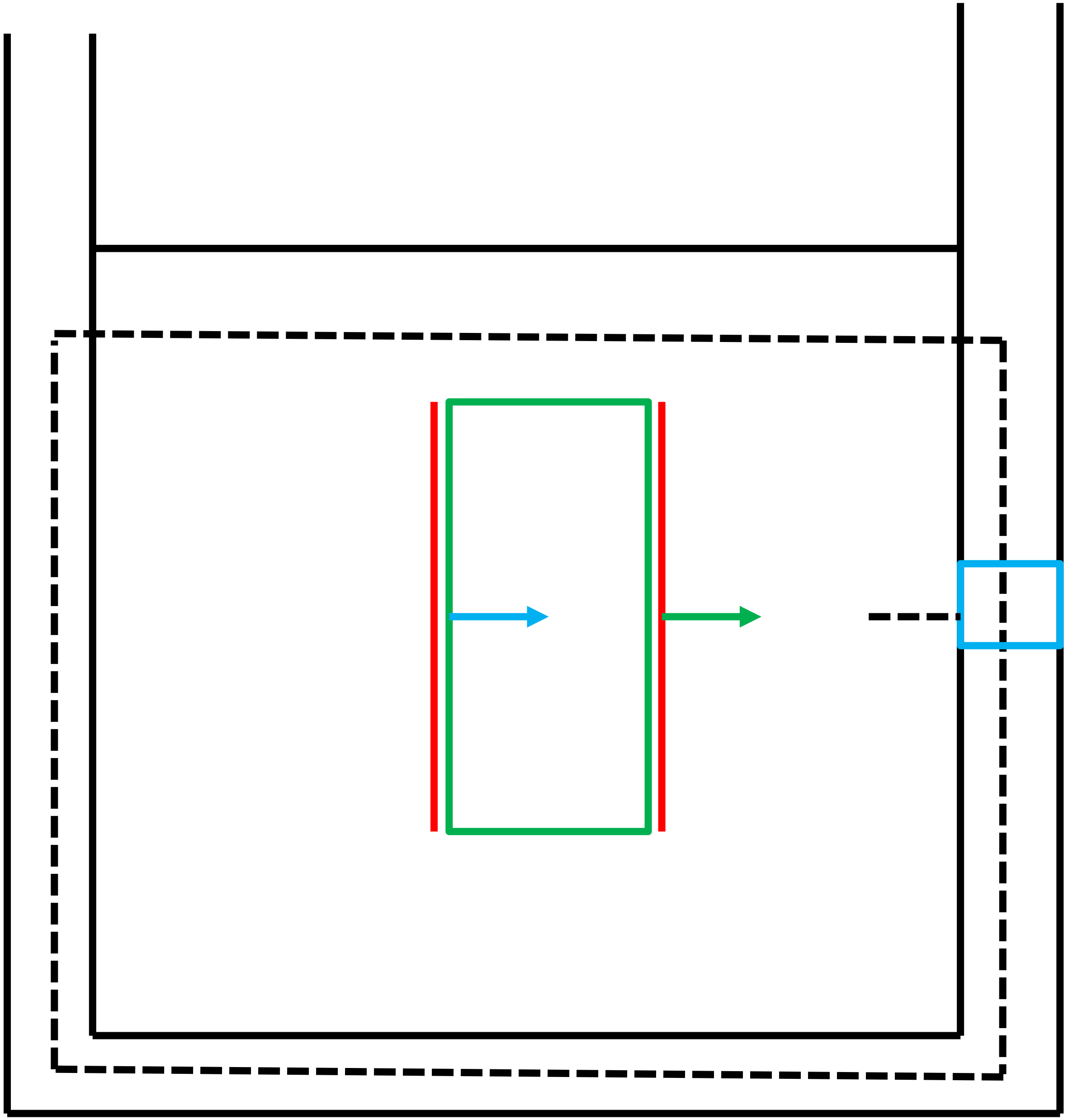


$$P_{xx} = nm < v_x^2 > \; > 0$$

Tlak:
$$P_{zz} = nm < v_z^2 >$$

Note: even power of v! Nonzero also for Maxwell (equilibrium) distribution. How it is possible to have non-zero flux of some quantity in equilibrium. Symmetry prohibits equilibrium flux for a scalar quantity, but for a vector quantity like momentum it is possible. Look to the glass of water in equilibrium





Total outflow of p_x from the blue box is zero, box does not move in the x-direction

Heat conduction in the approximation of relaxation time

As an illustration we present here the problem of energy ("heat") conduction in a pipe filled with gas, which is inserted between two thermal reservoirs, one at temperature T_1, the other at temperature T_2. After certain time a stationary regime is established: a constant energy current will flow from the hotter end towards the colder end. Along the pipe a time-constant temperature distribution will be formed, with temperature linearly decreasing from the hotter end towards the colder one.

We shall be looking for the Boltzmann distribution function $f(x, \vec{v})$, which is stationary, that is it does not depend on time explicitly. Non-trivial spatial distribution develops only along the pipe axis, in the direction of the energy flow. We select the x-axis to be in this direction.

We look for the solution of the equation

$$f(x, \overrightarrow{v}) = \int \frac{dt'}{\tau} \exp(-\frac{t'}{\tau}) f_0(x - v_x t', \overrightarrow{v})$$

where f_0 has the form of the Maxwell distribution

$$f_0(x, \overrightarrow{v}) = n(x) \left(\frac{m\beta(x)}{2\pi} \right)^{3/2} \exp(-\frac{1}{2}\beta(x) m \overrightarrow{v}^2)$$

$\beta(x)$ and $n(x)$ are unknown functions which should be found while solving for f. They correspond to the distribution function $f(x, \overrightarrow{v})$ as follows

$$\begin{aligned} n(x) &= \int d^3\overrightarrow{v} f(x, \overrightarrow{v}) \\ \frac{3}{2}\frac{1}{\beta(x)} &= \int d^3\overrightarrow{v} \frac{1}{2} m \overrightarrow{v}^2 f(x, \overrightarrow{v}) \end{aligned}$$

We look for the solution of the equation

$$f(x,\overrightarrow{v}) = \int \frac{dt'}{\tau} \exp(-\frac{t'}{\tau}) f_0(x - v_x t', \overrightarrow{v})$$

The trick leading to the solution is to do the per-partes transformation

$$f(x,\overrightarrow{v}) = f_0(x,\overrightarrow{v}) + \int \frac{df_0}{dt'} \exp(-\frac{t'}{\tau}) dt'$$

where f_0 has the form of the Maxwell distribution

$$f_0(x,\overrightarrow{v}) = n(x) \left(\frac{m\beta(x)}{2\pi}\right)^{3/2} \exp(-\frac{1}{2}\beta(x) m \overrightarrow{v}^2)$$

Next one has to realize that formally we are integrating through the region $(0,\infty)$, but practically only through the interval of the order $(0,\tau)$. Outside this interval the integrand is exponentially small Performing the differentiation with respect to t' we get

$$
\begin{aligned}
f(x,\overrightarrow{v}) &= f_0(x,\overrightarrow{v}) \\
&\quad + \int \left(-\frac{dn}{ndx} v_x - \frac{3}{2}\frac{d\beta}{\beta dx} v_x + \frac{1}{2} m \overrightarrow{v}^2 \frac{d\beta}{dx} v_x \right) f_0(x,\overrightarrow{v}) \exp(-\frac{t'}{\tau}) dt'
\end{aligned}
$$

All the terms in the bracket are proportional to the mean free path $v_x\tau$, therefore the variable x in the integrand in functions n, β, f_0 is no more shifted by $-v_x t'$.Such a shift would lead to corrections of higher order in τ.

$$f(x,\overrightarrow{v}) = f_0(x,\overrightarrow{v}) + \int\left(-\frac{dn}{ndx}v_x - \frac{3}{2}\frac{d\beta}{\beta dx}v_x + \frac{1}{2}m\overrightarrow{v}^2\frac{d\beta}{dx}v_x\right) f_0(x,\overrightarrow{v})\exp(-\frac{t'}{\tau})dt'$$

The integration is therefore trivial and we get

$$f(x,\overrightarrow{v}) = f_0(x,\overrightarrow{v}) + \tau\left(-\frac{dn}{ndx}v_x - \frac{3}{2}\frac{d\beta}{\beta dx}v_x + \frac{1}{2}m\overrightarrow{v}^2\frac{d\beta}{dx}v_x\right) f_0(x,\overrightarrow{v})$$

The distribution function is expressed through a Maxwell-like distributions, so all the momenta will be expressed through the Maxwellian momenta which are

$$\begin{aligned}
\langle v_x\rangle_0 &= 0\\
\langle v_x^2\rangle_0 &= \frac{1}{m\beta}\\
\langle v_x^4\rangle_0 &= \frac{3}{(m\beta)^2}\\
\langle v_x^6\rangle_0 &= \frac{3\times 5}{(m\beta)^3}
\end{aligned}$$

$$f(x,\overrightarrow{v}) = f_0(x,\overrightarrow{v}) + \tau\left(-\frac{dn}{ndx}v_x - \frac{3}{2}\frac{d\beta}{\beta dx}v_x + \frac{1}{2}m\overrightarrow{v}^2\frac{d\beta}{dx}v_x\right)f_0(x,\overrightarrow{v})$$

We are interested only in the stationary situation, when there is no flow of particles along the pipe. So we want a distribution function satisfying the relation

$$\langle v_x\rangle = \frac{1}{n}\int d^3\overrightarrow{v}\, v_x f(x,\overrightarrow{v}) = 0$$

$$\begin{aligned} 0 = \langle v_x\rangle &= \tau\left(-\frac{dn}{ndx}\langle v_x^2\rangle_0 - \frac{3}{2}\frac{d\beta}{\beta dx}\langle v_x^2\rangle_0 + \frac{1}{2}m\frac{d\beta}{dx}\langle v_x^4 + v_x^2v_y^2 + v_x^2v_z^2\rangle_0\right) \\ &= \frac{\tau}{m\beta}\left(-\frac{dn}{ndx} + \frac{d\beta}{\beta dx}\right) \end{aligned}$$

From there

$$\begin{aligned} \frac{dn}{ndx} &= \frac{d\beta}{\beta dx} \\ nkT &= \text{const} \end{aligned}$$

Let us now calculate the density of energy flow, which is given by the relation

$$j_E = \frac{1}{2} nm \langle \overrightarrow{v}^2 v_x \rangle$$

$$f(x, \overrightarrow{v}) = f_0(x, \overrightarrow{v}) + \tau \left(-\frac{dn}{ndx} v_x - \frac{3}{2} \frac{d\beta}{\beta dx} v_x + \frac{1}{2} m \overrightarrow{v}^2 \frac{d\beta}{dx} v_x \right) f_0(x, \overrightarrow{v})$$

$$\begin{aligned}
j_E &= \frac{1}{2} nm\tau \left(-\frac{d\beta}{\beta dx} \langle \overrightarrow{v}^2 v_x^2 \rangle_0 - \frac{3}{2} \frac{d\beta}{\beta dx} \langle \overrightarrow{v}^2 v_x^2 \rangle_0 + \frac{1}{2} m \frac{d\beta}{dx} \langle \overrightarrow{v}^2 \overrightarrow{v}^2 v_x^2 \rangle_0 \right) \\
&= \frac{1}{2} nm\tau \left(-\frac{25}{2} \frac{1}{(m\beta)^2} \frac{d\beta}{\beta dx} + \frac{1}{2} m \frac{d\beta}{dx} \langle v_x^6 + 2v_x^4 v_y^2 + 2v_x^4 v_z^2 + v_x^2 v_y^4 + v_x^2 v_z^4 + 2v_x^2 v_y^2 v_z^2 \rangle_0 \right) \\
&= \frac{5\tau}{(m\beta)^2} \frac{d\beta}{dx} \\
&= \frac{5}{2} \frac{\tau k^2 T n}{m} \left(-\frac{dT}{dx} \right)
\end{aligned}$$

So the coefficient of the heat conduction is

$$\varkappa = \frac{5}{2} \frac{\tau k^2 T n}{m}$$

Electrical conductivity The physical situation is the one described in Sec. 13·4. In the absence of an external electric field $\mathcal{E}$, the distribution function is given by

$$f^{(0)} = g(\epsilon), \qquad \epsilon = \tfrac{1}{2}mv^2 \tag{13·8·1}$$

where $g(\epsilon)$ is the MB distribution (13·4·2) in the case of ions or the FD distribution in the case of electrons in a metal. If a spatially uniform time-independent electric field $\mathcal{E}$ is applied in the z direction, one expects that the new distribution function $f(\mathbf{r},\mathbf{v},t)$ will still be independent of $\mathbf{r}$ and t. Then the Boltzmann equation (13·6·3) becomes simply, since $\boldsymbol{F} = e\boldsymbol{\mathcal{E}}$ has only a z component,

$$\frac{e\mathcal{E}}{m}\frac{\partial f}{\partial v_z} = -\frac{f - f^{(0)}}{\tau} \tag{13·8·2}$$

Let us assume that $\mathcal{E}$ is quite small. Then one expects that f differs only slightly from $f^{(0)} = g$. Thus we put

$$f = g + f^{(1)}, \qquad \text{where } f^{(1)} \ll g \tag{13·8·3}$$

$$f^{(1)} = f - g = -\frac{e\mathcal{E}\tau}{m}\frac{\partial g}{\partial v_z} = -e\mathcal{E}\tau v_z\frac{dg}{d\epsilon}$$

$$f^{(1)} = f - g = -\frac{e\mathcal{E}\tau}{m}\frac{\partial g}{\partial v_z} = -e\mathcal{E}\tau v_z \frac{dg}{d\epsilon}$$

▶

$$f(\mathbf{r},\mathbf{v},t) = g(\epsilon) - e\mathcal{E}\tau v_z \frac{dg}{d\epsilon} \tag{13·4·5}$$

The current density j_n in the $\hat{\mathbf{n}}$ direction is the flux of charge through an element of area directed along $\hat{\mathbf{n}}$. Thus

$$j_n = e \int d^3\mathbf{v}\, f v_n \tag{13·4·6}$$

▶

$$\sigma_{\text{el}} \equiv \frac{j_z}{\mathcal{E}} = -e^2 \int d^3\mathbf{v}\, \frac{dg}{d\epsilon}\tau {v_z}^2 \tag{13·4·7}$$

The ratio $j_z/\mathcal{E} = \sigma_{\text{el}}$ is, by definition, the electrical conductivity of the particles. As one would expect for the case of sufficiently low electric fields, (13·4·7) shows that $j_z \propto \mathcal{E}$.

When g is the Maxwell-Boltzmann distribution of (13·4·2), as it would be for ions or for (sufficiently dilute) electrons in a gas,

$$\frac{dg}{d\epsilon} = -\beta g \tag{13·4·8}$$

Then (13·4·7) becomes

$$\sigma_{\text{el}} = \beta e^2 \int d^3\mathbf{v}\, g\tau {v_z}^2 \tag{13·4·9}$$

$$\sigma_{\mathrm{el}} \approx \beta e^2 \bar{\tau} \int d^3\mathbf{v}\, g v_z^{\,2} = \beta e^2 \bar{\tau} (n \overline{v_z^{\,2}})$$

Here the average is calculated with the equilibrium function g so that the equipartition result $\frac{1}{2} m \overline{v_z^{\,2}} = \frac{1}{2} kT$ applies. Thus

$$\sigma_{\mathrm{el}} = \frac{ne^2}{m} \bar{\tau} \tag{13·4·10}$$

Collision term

Without collisions the particles move on Newtonian trajectories and Boltzmann equation without the right-hand side holds.

$$Df = 0$$

$$D = \frac{\partial}{\partial t} + \boldsymbol{v}.\frac{\partial}{\partial \boldsymbol{r}} + \frac{\boldsymbol{F}}{m} \cdot \frac{\partial}{\partial \boldsymbol{v}}$$

With collisions on the right-hand side there is the collision term

$$Df = D_C f$$

Let us imagine two infinitesimally close regions of the phase space such that they are connected by Newtonian trajectories. If the regions have equal volumes, then each trajectory which begins in the first region ends in the second region and vice versa. When collision take place, the particles do not follow the Newtonian trajectories. Some particles which start in the first region fall out of their trajectory before arriving to the second region and never arrive at the second region. On the other hand, some particles arrive to the second region which never started in the first region. They change the velocity in a collision and get onto a trajectory which leads to the second region somewhere between the two regions. The two regions we are speaking about are infinitesimal regions around the phase space point $\boldsymbol{r}$,$\boldsymbol{v}$.

Let us firs consider a particle which collides in the point $\boldsymbol{r}$ which before the collision had velocity $\boldsymbol{v}$. It collided with a particle which was present i the same space point $\boldsymbol{r}$ and had before the collision velocity $\boldsymbol{v}_1$.

The definition of the collision cross section says that the number of such collisions within a time interval dt in the volume element $d^3\boldsymbol{r}$ is

$$D_C^{(-)} f(\boldsymbol{r}, \boldsymbol{v}, t) d^3\boldsymbol{r} d^3\boldsymbol{v} dt = d^3\boldsymbol{r} d^3\boldsymbol{v} dt \int\int\int d^3\boldsymbol{v}_1 d^3\boldsymbol{v}' d^3\boldsymbol{v}_1' \\ |\boldsymbol{v} - \boldsymbol{v}_1| f(\boldsymbol{r}, \boldsymbol{v}, t) f(\boldsymbol{r}, \boldsymbol{v}_1, t) \sigma(\boldsymbol{v}, \boldsymbol{v}_1; \boldsymbol{v}', \boldsymbol{v}_1')$$

We have integrated over the velocities after the collisions and also over the velocity $\boldsymbol{v}_1$, since we are just interested in the fact, that there was a particle with the velocity $\boldsymbol{v}$ which escaped from the phase space point irrespective what happened to it and its collision partner.

On the other hand, the number of collisions in the considered space volume element such that one of the particles gained the velocity $\boldsymbol{v}$ after the collision is

$$D_C^{(+)} f(\boldsymbol{r}, \boldsymbol{v}, t) d^3\boldsymbol{r} d^3\boldsymbol{v} dt = d^3\boldsymbol{r} d^3\boldsymbol{v} dt \int\int\int d^3\boldsymbol{v}_1 d^3\boldsymbol{v}' d^3\boldsymbol{v}_1' \\ |\boldsymbol{v}' - \boldsymbol{v}_1'| f(\boldsymbol{r}, \boldsymbol{v}', t) f(\boldsymbol{r}, \boldsymbol{v}_1', t) \sigma(\boldsymbol{v}', \boldsymbol{v}_1'; \boldsymbol{v}, \boldsymbol{v}_1)$$

The energy momentum conservation gives

$$|\boldsymbol{v} - \boldsymbol{v}_1| = |\boldsymbol{v}' - \boldsymbol{v}_1'|$$

and P,T invariance of the collision process gives

$$\sigma(\boldsymbol{v}, \boldsymbol{v}_1; \boldsymbol{v}', \boldsymbol{v}_1') = \sigma(\boldsymbol{v}', \boldsymbol{v}_1'; \boldsymbol{v}, \boldsymbol{v}_1)$$

The total collision term will be

$$D_C = D_C^{(+)} - D_C^{(-)}$$

and so

$$D_C f(\boldsymbol{r}, \boldsymbol{v}, t) = \int\int\int d^3\boldsymbol{v}_1 d^3\boldsymbol{v}' d^3\boldsymbol{v}_1' \\ |\boldsymbol{v}' - \boldsymbol{v}_1'| (f(\boldsymbol{r}, \boldsymbol{v}', t) f(\boldsymbol{r}, \boldsymbol{v}_1', t) - f(\boldsymbol{r}, \boldsymbol{v}, t) f(\boldsymbol{r}, \boldsymbol{v}_1, t)) \sigma(\boldsymbol{v}', \boldsymbol{v}_1'; \boldsymbol{v}, \boldsymbol{v}_1)$$

$$D_C f(\boldsymbol{r},\boldsymbol{v},t) = \int\int\int d^3\boldsymbol{v}_1 d^3\boldsymbol{v}' d^3\boldsymbol{v}_1'$$

$$|\boldsymbol{v}' - \boldsymbol{v}_1'|(f(\boldsymbol{r},\boldsymbol{v}',t)f(\boldsymbol{r},\boldsymbol{v}_1',t) - f(\boldsymbol{r},\boldsymbol{v},t)f(\boldsymbol{r},\boldsymbol{v}_1,t))\sigma(\boldsymbol{v}',\boldsymbol{v}_1';\boldsymbol{v},\boldsymbol{v}_1)$$

The notation is usually make shorter as

$$D_C f(\boldsymbol{r},\boldsymbol{v},t) = \int\int\int d^3\boldsymbol{v}_1 d^3\boldsymbol{v}' d^3\boldsymbol{v}_1' |\boldsymbol{v}' - \boldsymbol{v}_1'|(f' f_1' - f f_1)\sigma$$

when we added to the symbol f the index and/or prime sign according to what velocity is to be inserted into the function f as the velocity variable.

Equilibrium distribution

$$D_C f(\boldsymbol{r},\boldsymbol{v},t) = \int\int\int d^3\boldsymbol{v}_1 d^3\boldsymbol{v}' d^3\boldsymbol{v}_1' |\boldsymbol{v}' - \boldsymbol{v}_1'|(f'f_1' - ff_1)\sigma$$

$$D_C f = 0$$

$$f(v^2)f(v_1^2) = f(v'^2)f(v_1'^2)$$

$$v^2 + v_1^2 = v'^2 + v_1'^2$$

$$f(x)f(E-x) = \mathrm{const}$$

$$f'(x)f(E-x) - f(x)f'(E-x) = 0$$

$$f'(x)/f(x) = f'(E-x)/f(E-x)$$

$$f'(x)/f(x) = \mathrm{const}$$

$$f(v^2) = C\exp(-\alpha v^2)$$

Maxwell

Conservation laws

Let us consider som one-particle physical quantity $\chi(\boldsymbol{r}, \boldsymbol{v}, t)$. By that we mean a quantity which has the value χ when a particle is found at the time t in the phase-space point $\boldsymbol{r}, \boldsymbol{v}$. Then the mean value of the quantity χ in the space point $\boldsymbol{r}$ atr time t will be

$$\langle \chi(\boldsymbol{r}, t) \rangle = \frac{1}{n(\boldsymbol{r}, t)} \int d^3\boldsymbol{v} f(\boldsymbol{r}, \boldsymbol{v}, t) \chi(\boldsymbol{r}, \boldsymbol{v}, t)$$

where

$$n(\boldsymbol{r}, t) = \int d^3\boldsymbol{v} f(\boldsymbol{r}, \boldsymbol{v}, t)$$

Let us start from the Boltzmann equation

$$Df = D_C f$$

We get

$$\int d^3\boldsymbol{v} \chi D f = \int d^3\boldsymbol{v} \chi D_C f$$

$$\int d^3\boldsymbol{v}\chi Df = \int d^3\boldsymbol{v}\chi D_C f$$

At the left-had side we get several terms which we rewrite as follows

$$\int d^3\boldsymbol{v}\chi\frac{\partial f}{\partial t} = \int d^3\boldsymbol{v}[\frac{\partial}{\partial t}(f\chi) - f\frac{\partial\chi}{\partial t}] = \frac{\partial}{\partial t}(n\langle\chi\rangle) - n\langle\frac{\partial\chi}{\partial t}\rangle$$

$$\int d^3\boldsymbol{v}\ \ \boldsymbol{v}\cdot\frac{\partial f}{\partial\boldsymbol{r}}\chi = \int d^3\boldsymbol{v}\ v_\alpha\frac{\partial f}{\partial x_\alpha}\chi = \frac{\partial}{\partial x_\alpha}(n\langle v_\alpha\chi\rangle) - n\langle v_\alpha\frac{\partial\chi}{\partial x_\alpha}\rangle$$

The third term we rewrite assuming that external forces are independent of velocities

$$\int d^3\boldsymbol{v}\ \ \frac{\boldsymbol{F}}{m}\cdot\frac{\partial f}{\partial\boldsymbol{v}}\chi = \int d^3\boldsymbol{v}\ [\frac{\partial}{\partial v_\alpha}(\frac{F_\alpha}{m}f\chi) - \frac{F_\alpha}{m}f\frac{\partial\chi}{\partial v_\alpha}]$$

Here the first term after the integration gives zero on the integration boundaries and so we get

$$\int d^3\boldsymbol{v}\ \ \frac{\boldsymbol{F}}{m}\cdot\frac{\partial f}{\partial\boldsymbol{v}}\chi = -\frac{F_\alpha}{m}n\langle\frac{\partial\chi}{\partial v_\alpha}\rangle$$

$$\int d^3\boldsymbol{v}\chi Df = \int d^3\boldsymbol{v}\chi D_C f$$

Now we look at the right-hand side of the Boltzmann equation.

$$\int d^3\boldsymbol{v}\chi D_C f = \iiiint d^3\boldsymbol{v}d^3\boldsymbol{v}_1 d^3\boldsymbol{v}' d^3\boldsymbol{v}_1'(f'f_1' - ff_1)|\boldsymbol{v}-\boldsymbol{v}_1|\sigma(\boldsymbol{v},\boldsymbol{v}_1;\boldsymbol{v}',\boldsymbol{v}_1')\chi(\boldsymbol{r},\boldsymbol{v},t)$$

We change the notation for the integration variables $\boldsymbol{v} \leftrightarrow \boldsymbol{v}_1$ and $\boldsymbol{v}' \leftrightarrow \boldsymbol{v}_1'$. We get

$$\int d^3\boldsymbol{v}\chi D_C f = \iiiint d^3\boldsymbol{v}d^3\boldsymbol{v}_1 d^3\boldsymbol{v}' d^3\boldsymbol{v}_1'(f'f_1' - ff_1)|\boldsymbol{v}_1-\boldsymbol{v}|\sigma(\boldsymbol{v}_1,\boldsymbol{v};\boldsymbol{v}_1',\boldsymbol{v}')\chi(\boldsymbol{r},\boldsymbol{v}_1,t)$$

This change of notation, however, does not change the values neither of the relative velocity, nor of the cross section, so we get

$$\int d^3\boldsymbol{v}\chi D_C f = \frac{1}{2}\iiiint d^3\boldsymbol{v}d^3\boldsymbol{v}_1 d^3\boldsymbol{v}' d^3\boldsymbol{v}_1'(f'f_1' - ff_1)|\boldsymbol{v}_1-\boldsymbol{v}|\sigma(\chi+\chi_1)$$

where for the quantity χ we used the same short notation as for f, that is we used indices to denote what velocity should be used as the variable in the function evaluation.

$$\int d^3\boldsymbol{v}\chi D_C f = \frac{1}{2}\iiiint d^3\boldsymbol{v}d^3\boldsymbol{v}_1 d^3\boldsymbol{v}' d^3\boldsymbol{v}_1'(f'f_1' - ff_1)|\boldsymbol{v}_1 - \boldsymbol{v}|\sigma(\chi + \chi_1)$$

Let us use further symmetry of the above relation: we change variables $\boldsymbol{v} \leftrightarrow \boldsymbol{v}'$ and $\boldsymbol{v}_1 \leftrightarrow \boldsymbol{v}_1'$. This gives the inverse collision which has the same cross section and we finally get

$$\int d^3\boldsymbol{v}\chi D_C f = \frac{1}{4}\iiiint d^3\boldsymbol{v}d^3\boldsymbol{v}_1 d^3\boldsymbol{v}' d^3\boldsymbol{v}_1'(f'f_1' - ff_1)|\boldsymbol{v}_1 - \boldsymbol{v}|\sigma(\chi' + \chi_1' - \chi - \chi_1)$$

The result is particularly interesting when the quantity χ is conserved in collisions. Then

$$\chi' + \chi_1' - \chi - \chi_1 = 0$$

and so

$$\int d^3\boldsymbol{v}\chi D_C f = 0$$

$$\int d^3\boldsymbol{v}\chi Df = \int d^3\boldsymbol{v}\chi D_C f$$

At the left-had side we get several terms which we rewrite as follows

$$\int d^3\boldsymbol{v}\chi\frac{\partial f}{\partial t} = \int d^3\boldsymbol{v}[\frac{\partial}{\partial t}(f\chi) - f\frac{\partial\chi}{\partial t}] = \frac{\partial}{\partial t}(n\langle\chi\rangle) - n\langle\frac{\partial\chi}{\partial t}\rangle$$

$$\int d^3\boldsymbol{v}\ \ \boldsymbol{v}\cdot\frac{\partial f}{\partial\boldsymbol{r}}\chi = \int d^3\boldsymbol{v}\ v_\alpha\frac{\partial f}{\partial x_\alpha}\chi = \frac{\partial}{\partial x_\alpha}(n\langle v_\alpha\chi\rangle) - n\langle v_\alpha\frac{\partial\chi}{\partial x_\alpha}\rangle$$

The third term we rewrite assuming that external forces are independent of velocities

$$\int d^3\boldsymbol{v}\ \ \frac{\boldsymbol{F}}{m}\cdot\frac{\partial f}{\partial\boldsymbol{v}}\chi = \int d^3\boldsymbol{v}\ [\frac{\partial}{\partial v_\alpha}(\frac{F_\alpha}{m}f\chi) - \frac{F_\alpha}{m}f\frac{\partial\chi}{\partial v_\alpha}]$$

Here the first term after the integration gives zero on the integration boundaries and so we get

$$\int d^3\boldsymbol{v}\ \ \frac{\boldsymbol{F}}{m}\cdot\frac{\partial f}{\partial\boldsymbol{v}}\chi = -\frac{F_\alpha}{m}n\langle\frac{\partial\chi}{\partial v_\alpha}\rangle$$

$$\boxed{\frac{\partial}{\partial t}\langle n\chi\rangle + \frac{\partial}{\partial x_\alpha}\langle nv_\alpha\chi\rangle = n\langle D\chi\rangle}$$

conservation of mass $\chi = m$

The mass conservation gives the equation

$$\frac{\partial}{\partial t}\langle nm \rangle + \frac{\partial}{\partial x_\alpha}\langle nmv_\alpha \rangle = 0$$

The quantity n is independent of velocities, so it can be taken in front of the averaging angle brackets and we get

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial x_\alpha}\rho\langle v_\alpha \rangle = 0$$

If the mean velocity (drift) is denoted as

$$\boldsymbol{u} = \langle \boldsymbol{v} \rangle$$

we get the equation of continuity

$$\frac{\partial \rho}{\partial t} + \nabla.(\rho \boldsymbol{u}) = 0$$

$$\frac{\partial}{\partial t}\langle nmv_\gamma\rangle + \frac{\partial}{\partial x_\alpha}\langle nmv_\alpha v_\gamma\rangle = \langle nm\ Dv_\gamma\rangle = nm\langle\frac{F_\alpha}{m}\frac{\partial v_\gamma}{\partial v_\alpha}\rangle$$

If the external force does not depend on velocity we get

$$\frac{\partial}{\partial t}(\rho u_\gamma) + \frac{\partial}{\partial x_\alpha}(\rho\langle v_\alpha v_\gamma\rangle) = \frac{F_\gamma}{m}$$

We separate now the chaotic and the drift velocity

$$\boldsymbol{v} = \boldsymbol{u} + \boldsymbol{U}$$

then

$$((\langle v_\alpha v_\gamma\rangle = u_\alpha u_\gamma + \langle U_\alpha U_\gamma\rangle$$

We introduce the tensor of tensions as

$$P_{\alpha\gamma} = \rho\langle U_\alpha U_\gamma\rangle$$

and we get the Euler hydrodynamic equation

$$\frac{\partial}{\partial t}(\rho u_\gamma) + \frac{\partial}{\partial x_\alpha}(\rho u_\alpha u_\gamma) = -\frac{\partial P_{\alpha\gamma}}{\partial x_\alpha} + \frac{F_\gamma}{m}$$

H-theorem

Let us consider the Boltzmann distribution function satisfying the Boltzmann equation with the collision term on the right-hand side.

$$Df = D_C f$$

where

$$Df = \frac{\partial}{\partial t} f + \boldsymbol{v}.\frac{\partial}{\partial \boldsymbol{r}} f + \frac{\boldsymbol{F}}{m} \cdot \frac{\partial}{\partial \boldsymbol{v}} f$$

and

$$D_C f = \int d^3\boldsymbol{v}_1 d^3\boldsymbol{v}' d^3\boldsymbol{v}_1' |\boldsymbol{v} - \boldsymbol{v}_1| (f' f_1' - f f_1) \sigma(\boldsymbol{v}, \boldsymbol{v}_1; \boldsymbol{v}', \boldsymbol{v}_1')$$

Let us now investigate the expression

$$H(t) = \int d^3\boldsymbol{r} d^3\boldsymbol{v} f(\boldsymbol{r}, \boldsymbol{v}, t) \ln(f(\boldsymbol{r}, \boldsymbol{v}, t))$$

$$H(t) = \int d^3\boldsymbol{r}_t d^3\boldsymbol{v}_t f(\boldsymbol{r}_t, \boldsymbol{v}_t, t) \ln(f(\boldsymbol{r}_t, \boldsymbol{v}_t, t))$$

$$H(t) = \int d^3\boldsymbol{r}_t d^3\boldsymbol{v}_t f(\boldsymbol{r}_t, \boldsymbol{v}_t, t) \ln(f(\boldsymbol{r}_t, \boldsymbol{v}_t, t))$$

where we introduced a formal index t. Now we make a substitution in the integral so that instead of variables $\boldsymbol{r}_t, \boldsymbol{v}_t$ we introduce new variables $\boldsymbol{r}_0, \boldsymbol{v}_0$ via the relations

$$\boldsymbol{r}_t = \boldsymbol{R}(\boldsymbol{r}_0, \boldsymbol{v}_0, t)$$

$$\boldsymbol{v}_t = \boldsymbol{V}(\boldsymbol{r}_0, \boldsymbol{v}_0, t)$$

where the functions $\boldsymbol{R}(\boldsymbol{r}_0, \boldsymbol{v}_0, t)$ and $\boldsymbol{V}(\boldsymbol{r}_0, \boldsymbol{v}_0, t)$ denote solutions of the Newton equations of motion with initial conditions $\boldsymbol{r}_0, \boldsymbol{v}_0$. The Jacobian of this transformation is equal to unity, the integration region is again the whole phase space, so we get

$$H(t) = \int d^3\boldsymbol{r}_0 d^3\boldsymbol{v}_0 f(\boldsymbol{R}(\boldsymbol{r}_0, \boldsymbol{V}(\boldsymbol{r}_0, \boldsymbol{v}_0, t), t), \boldsymbol{v}_t, t) \ln(f(\boldsymbol{R}(\boldsymbol{r}_0, \boldsymbol{V}(\boldsymbol{r}_0, \boldsymbol{v}_0, t), t), \boldsymbol{v}_t, t))$$

$$\frac{d}{dt} H(t) = \int d^3\boldsymbol{r}_0 d^3\boldsymbol{v}_0 (Df) \ln(f) + \int d^3\boldsymbol{r}_0 d^3\boldsymbol{v}_0 Df$$

$$\frac{d}{dt}H(t) = \int d^3\boldsymbol{r}_0 d^3\boldsymbol{v}_0 \ln(f) D_C f$$

Now we make the "inverse trick" and we go over to integration variable $\boldsymbol{r}, \boldsymbol{v}$ and we get

$$\frac{d}{dt}H(t) = \int d^3\boldsymbol{r} d^3\boldsymbol{v} d^3\boldsymbol{v}_1 d^3\boldsymbol{v}' d^3\boldsymbol{v}_1' \ln(f)|\boldsymbol{v} - \boldsymbol{v}_1|(f'f_1' - ff_1)\sigma(\boldsymbol{v}, \boldsymbol{v}_1; \boldsymbol{v}', \boldsymbol{v}_1')$$

Now we proceed the same way as we did while deriving the continuity equation we get

$$\frac{d}{dt}H(t) = \frac{1}{4}\int d^3\boldsymbol{r} d^3\boldsymbol{v} d^3\boldsymbol{v}_1 d^3\boldsymbol{v}' d^3\boldsymbol{v}_1' \ln(\frac{ff_1}{f'f_1'})|\boldsymbol{v} - \boldsymbol{v}_1|(f'f_1' - ff_1)\sigma(\boldsymbol{v}, \boldsymbol{v}_1; \boldsymbol{v}', \boldsymbol{v}_1')$$

$$\frac{d}{dt}H(t) = -\frac{1}{4}\int d^3\boldsymbol{r} d^3\boldsymbol{v} d^3\boldsymbol{v}_1 d^3\boldsymbol{v}' d^3\boldsymbol{v}_1' |\boldsymbol{v} - \boldsymbol{v}_1| ff_1 \ln(\frac{f'f_1'}{ff_1})(\frac{f'f_1'}{ff_1} - 1)\sigma(\boldsymbol{v}, \boldsymbol{v}_1; \boldsymbol{v}', \boldsymbol{v}_1')$$

The logarithmic function satisfies the inequality (draw the graph of the function)

$$\ln(x)(x-1) \geq 0$$

and so we get

$$\frac{d}{dt}H(t) \leq 0$$

Langevin equation

The Langevin equation describes movement of a classical particle in the field of a random force.

$$m\frac{d\boldsymbol{v}}{dt} = \boldsymbol{F} + \boldsymbol{f}(t)$$

where $\boldsymbol{F}$ is deterministic external force and $\boldsymbol{f}(t)$ is a random force describing for example influence of the collisions of the observed particle with surrounding molecules. The typical correlation time for that force is of the order 10^{-13}s.

We investigate the problem by introducing statistical ensembles. So let as imagine an ensemble where $\bar{\boldsymbol{v}}$ is the same for all the ensemble members and the ensemble averages of the fluctuating force is $\bar{\boldsymbol{f}}$ So we have

$$\boldsymbol{f} = \bar{\boldsymbol{f}} + \boldsymbol{f}'$$

If $\boldsymbol{v} = 0$ then $\bar{\boldsymbol{f}} = 0$ because there is no preferred direction. Therefore in the lowest approximation tho following must hold

$$\bar{\boldsymbol{f}} = -\alpha\boldsymbol{v}$$

and we get the Langevin equation

$$m\frac{d\boldsymbol{v}}{dt} = \boldsymbol{F} - \alpha\boldsymbol{v} + \boldsymbol{f}'$$

To simplify the calculations let us consider one-dimensional case without external deterministic force. We get

$$m\frac{d\dot{x}}{dt} = -\alpha\dot{x} + f'$$

$$mx\frac{d\dot{x}}{dt} = -\alpha x\dot{x} + xf'$$

$$m[\frac{d}{dt}(x\dot{x}) - \dot{x}^2] = -\alpha x\dot{x} + xf'$$

Averaging we get (since $\langle f' \rangle = 0$ independently of x and $\dot{x}$

$$m\langle\frac{d}{dt}(x\dot{x})\rangle = \langle m\dot{x}^2\rangle - \alpha\langle x\dot{x}\rangle$$

$$\frac{d}{dt}\langle x\dot{x}\rangle = \frac{kT}{m} - \frac{\alpha}{m}\langle x\dot{x}\rangle$$

Denoting

$$\varphi(t) = \langle x\dot{x}\rangle - \frac{kT}{\alpha}$$

we get the equation

$$\dot{\varphi}(t) = -\frac{\alpha}{m}\varphi(t)$$

$$\varphi(t) = \langle x\dot{x}\rangle - \frac{kT}{\alpha}$$

we get the equation

$$\dot{\varphi}(t) = -\frac{\alpha}{m}\varphi(t)$$

$$\langle x\dot{x}\rangle = C\exp(-\gamma t) + \frac{kT}{\alpha}$$

where

$$\gamma = \frac{\alpha}{m}$$

Let us suppose, that all the particles in the ensemble start at $x = 0$, $\dot{x} = 0$. Then

$$0 = C + \frac{kT}{\alpha}$$

$$\langle x\dot{x}\rangle = \frac{kT}{\alpha}(1 - \exp(-\gamma t))$$

$$\frac{d}{dt}\langle x^2\rangle = \frac{kT}{\alpha}(1 - \exp(-\gamma t))$$

$$\langle x^2\rangle = \frac{2kT}{\alpha}(t - \frac{1}{\gamma}(1 - \exp(-\gamma t)))$$

$$\langle x^2 \rangle = \frac{2kT}{\alpha}(t - \frac{1}{\gamma}(1 - \exp(-\gamma t)))$$

This means that for $t < \frac{1}{\gamma}$

$$\langle x^2 \rangle = \frac{kT}{m} t^2$$

and for $t \gg \frac{1}{\gamma}$

$$\langle x^2 \rangle = \frac{2kT}{\alpha} t$$

$$m\frac{d\dot{x}}{dt} = -\alpha\dot{x} + f'$$

For strong dumping we can neglect in the Langevin equation the inertial term and we get a Aristotle type of equation

$$0 = -\alpha\dot{x} + f'$$
$$\dot{x} = \frac{1}{\alpha}f'$$
$$x(t) = x(0) + \int_0^t d\tau\frac{f'(\tau)}{\alpha}$$

For the initial condition $x(0) = 0$ we get

$$\langle x^2(t)\rangle = \frac{1}{\alpha^2}\int_0^t d\tau_1 d\tau_2\langle f'(\tau_1)f'(\tau_2)\rangle$$

Now the correlation depends only on the time difference and we get

$$\langle x^2(t)\rangle = \frac{1}{\alpha^2}\int_0^t d\frac{\tau_1+\tau_2}{2}d(\tau_1-\tau_2)\langle f'(0)f'(\tau_1-\tau_2)\rangle$$

$$\langle x^2(t)\rangle = \frac{1}{\alpha^2}\int_0^t d\frac{\tau_1+\tau_2}{2}d(\tau_1-\tau_2)\langle f'(0)f'(\tau_1-\tau_2)\rangle$$

$$\langle x^2(t)\rangle = \frac{t}{\alpha^2}\int_{-\infty}^{\infty} d\tau\langle f'(0)f'(\tau)\rangle$$

where we extended the region of integration formally to $(-\infty,\infty)$ since the correlation function is different from zero only on a very small interval and so the integration region does not play any role.

We, however, know from previous calculation that

$$\langle x^2\rangle = \frac{2kT}{\alpha}t$$

and comparing the two results we get

$$\alpha = \frac{1}{2kT}\int_{-\infty}^{\infty} d\tau\langle f'(0)f'(\tau)\rangle$$

So the dissipation coefficient α is given by the correlations of the fluctuation force. This is so called fluctuation-dissipation theorem.

Now we return back to the expression

$$x(t) = \int_0^t d\tau \frac{f'(\tau)}{\alpha}$$

The integral is the sum of many random variables, so $x(t)$ must be normally distributed with zero mean and variance

$$\langle x^2 \rangle = \frac{2kT}{\alpha} t$$

so the distribution is

$$\rho(x,t) = \frac{1}{\sqrt{4\pi Dt}} \exp(-\frac{x^2}{4Dt})$$

where

$$D = \frac{kT}{\alpha}$$

This is called Einstein dissipation theorem. The distribution density satisfies the diffusion equation

$$\frac{\partial \rho}{\partial t} = D \triangle \rho$$

TRANSITION PROBABILITIES AND MASTER EQUATION

Consider an isolated system A. Let its Hamiltonian (or energy) be

$$\mathcal{H}_0 = \mathcal{H} + \mathcal{H}_i \qquad (15\cdot 1\cdot 1)$$

where $\mathcal{H}$ is the main part of the Hamiltonian and $\mathcal{H}_i \ll \mathcal{H}$ is a small additional part describing some weak interactions not included in $\mathcal{H}$.

Let the quantum states of $\mathcal{H}$ be denoted by r and their corresponding energy levels by E_r. If $\mathcal{H}_i = 0$, these states would be quantum states of the total Hamiltonian so that the system A would remain in any such state indefinitely. The presence of the additional interaction $\mathcal{H}_i$ makes this no longer true, since $\mathcal{H}_i$ is capable of inducing transitions between the various unperturbed states r. If $\mathcal{H}_i$ is small, if there is a nearly continuous distribution of accessible energy levels,* and if one considers time intervals which are not too small, then there exists a well-defined transition probability W_{rs} per unit time from the unperturbed state r to the unperturbed state s of system A. By conservation of energy W_{rs} is such that

$$\text{if } E_r \neq E_s, \qquad W_{rs} = 0 \qquad (15\cdot 1\cdot 2)$$

Furthermore, there is a symmetry property relating this transition to the inverse transition* from state s to state r,

▶ $$W_{sr} = W_{rs} \tag{15·1·3}$$

NESTACIONÁRNA PORUCHOVÁ METÓDA

$$\mathrm{H}(t) = \mathrm{H}_0 + \mathrm{H}'(t)$$

$$i\hbar \frac{\partial \psi(\boldsymbol{r},t)}{\partial t} = \mathrm{H}(t)\psi(\boldsymbol{r},t)$$

$$\psi(\boldsymbol{r},t) = \sum_n a_n(t)\mathrm{e}^{-\mathrm{i}E_n t/\hbar}\Phi_n(\boldsymbol{r})$$

$$a_n^{(0)}(t) = \delta_{nm} \qquad \dot{a}_k(t) = \frac{1}{\mathrm{i}\hbar}\sum_n H'_{kn}(t)a_n(t)\mathrm{e}^{\mathrm{i}\omega_{kn}t}, \quad k = 1, 2, \ldots$$

kde bodka označuje deriváciu podľa času, pričom

$$\omega_{kn} = (E_k - E_m)/\hbar$$

$$H'_{kn}(t) = \int \Phi_k^*(\boldsymbol{r})\mathrm{H}(t)\Phi_n(\boldsymbol{r})\,\mathrm{d}^3\boldsymbol{r}$$

$$\mathrm{i}\hbar\, a_k^{(1)}(t) = \int_{-\infty}^{t} \mathrm{d}t'\, H'_{km}(t')\, \mathrm{e}^{\mathrm{i}\omega_{km}t'} \tag{6}$$

Next, it is assumed that the perturbing force described by H' "turns on" at $t = 0$ and is constant over the interval $0 \le t' \le t$. Equation (6) can then be integrated to give:

$$\mathrm{i}\hbar\, a_k^{(1)}(t) \approx 2H'_{km}\, \mathrm{e}^{\mathrm{i}\omega_{km}t/2} \left(\frac{\sin \omega_{km}t/2}{\omega_{km}} \right)$$

The probability $P_k(t)$ that the system undergoes a transition from state m to state k is:

$$P_k(t) = |a_k(t)|^2 \approx \frac{4\,|H'_{km}|^2}{\hbar^2} \frac{\sin^2 \omega_{km}t/2}{\omega_{km}^2} \tag{7}$$

The mean *rate* for the transition is given by $w_k = P_k(t)/t$.

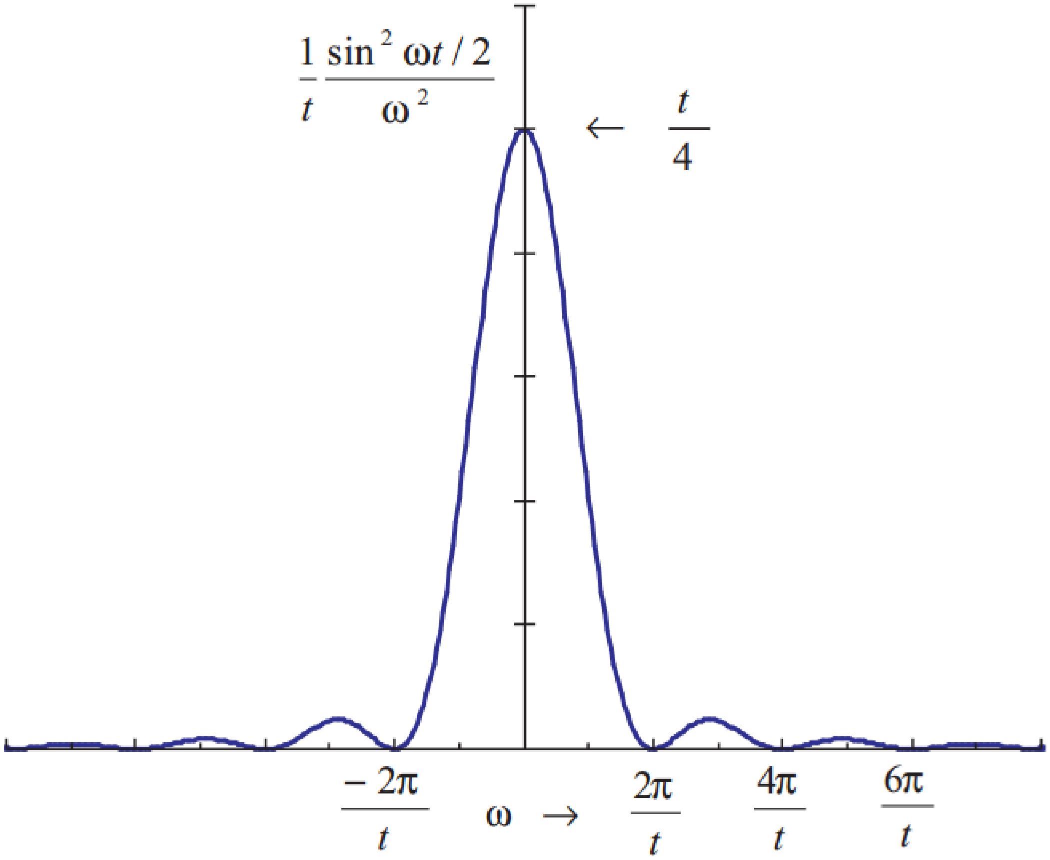


Figure 1: Behavior of the function $g(\omega, t) = \frac{1}{t}\frac{\sin^2 \omega t/2}{\omega^2}$ versus ω. g has the effect of enforcing energy-conservation because in the limit $t \rightarrow \infty$, $g \rightarrow \frac{\pi}{2}\delta(\omega)$; it explicitly demonstrates the Heisenberg uncertainty relation between energy and time through, for example, the half-width of the peak $\Delta\omega$ and the "lifetime" t of the perturbation: $\Delta\omega\, t \sim \pi$.

that states to which transitions can occur must have $\omega_{km} \approx 0$, forcing energy conservation.

In general, there will be some number of states $\mathrm{d}n$ within an interval $\mathrm{d}\omega_{km}$. The number of possible transition states can be written:

$$\mathrm{d}n = \rho(k)\,\mathrm{d}E_k$$

where $\rho(k) = \mathrm{d}n/\mathrm{d}E_k$ is the "density of states" per unit energy interval near E_k; $\mathrm{d}\omega_{km}$ and $\mathrm{d}E_k$ are related by $\mathrm{d}\omega_{km} = \mathrm{d}E_k/\hbar$. It is expected that $\rho(k)$ and H'_{km} are smoothly varying functions of momentum or energy near the state k.

The physically meaningful quantity is the *total* transition rate to states *near* the state k:

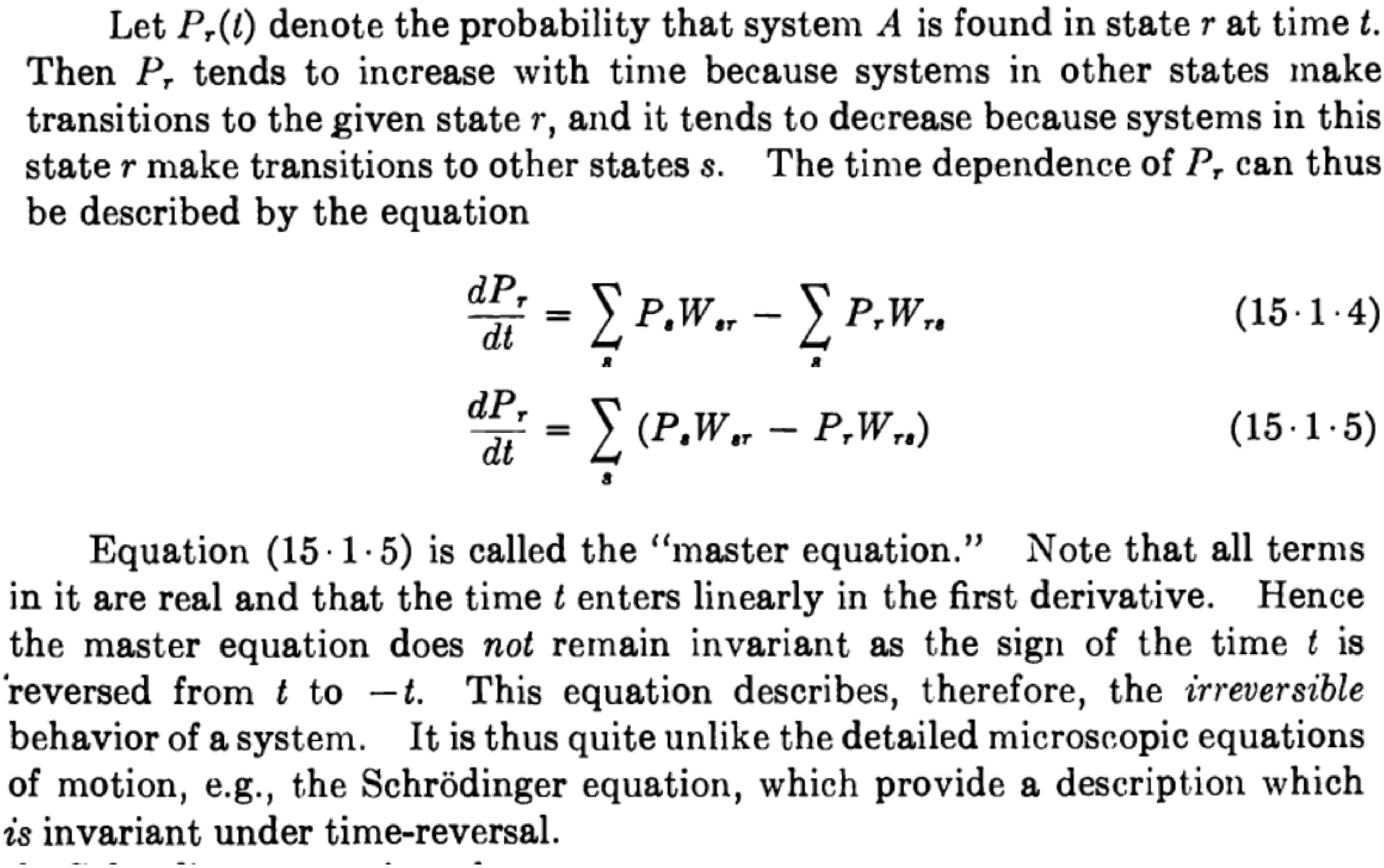
$$W_k = \frac{1}{t} \sum_{k' \text{ near } k} P_{k'}(t)$$

This summation can be replaced by an integral over $\mathrm{d}E_k$:

$$\begin{aligned} W_k &= \frac{1}{t}\int P_{k'}(t)\rho(k')\,\mathrm{d}E_{k'} \\ &= \int \mathrm{d}E_k\,\rho(k)\,\frac{4\left|H'_{km}\right|^2}{\hbar^2}\,\frac{1}{t}\left(\frac{\sin^2\omega_{km}t/2}{\omega_{km}^2}\right) \\ &= \frac{4}{\hbar}\left|H'_{km}\right|^2\rho(k)\int\limits_{-\infty}^{\infty}\mathrm{d}\omega\frac{1}{t}\frac{\sin^2\omega t/2}{\omega^2} \end{aligned}$$

As can be anticipated from Figure 1, the last integral has the value $\pi/2$ and we arrive at Fermi's "Second Golden Rule":

$$W_k = \frac{2\pi}{\hbar}\left|H'_{km}\right|^2\rho(k) \tag{8}$$



Let $P_r(t)$ denote the probability that system A is found in state r at time t. Then P_r tends to increase with time because systems in other states make transitions to the given state r, and it tends to decrease because systems in this state r make transitions to other states s. The time dependence of P_r can thus be described by the equation

$$\frac{dP_r}{dt} = \sum_s P_s W_{sr} - \sum_s P_r W_{rs} \tag{15·1·4}$$

$$\frac{dP_r}{dt} = \sum_s (P_s W_{sr} - P_r W_{rs}) \tag{15·1·5}$$

Equation (15 · 1 · 5) is called the "master equation." Note that all terms in it are real and that the time t enters linearly in the first derivative. Hence the master equation does *not* remain invariant as the sign of the time t is reversed from t to $-t$. This equation describes, therefore, the *irreversible* behavior of a system. It is thus quite unlike the detailed microscopic equations of motion, e.g., the Schrödinger equation, which provide a description which *is* invariant under time-reversal.

$$\frac{dP_r}{dt} = \sum_s (P_s W_{sr} - P_r W_{rs}) \qquad (15\cdot 1\cdot 5)$$

detailed balance

$$P_r W_{rs} = P_s W_{sr} \qquad W_{sr} = W_{rs}$$

$$P_r = P_s$$

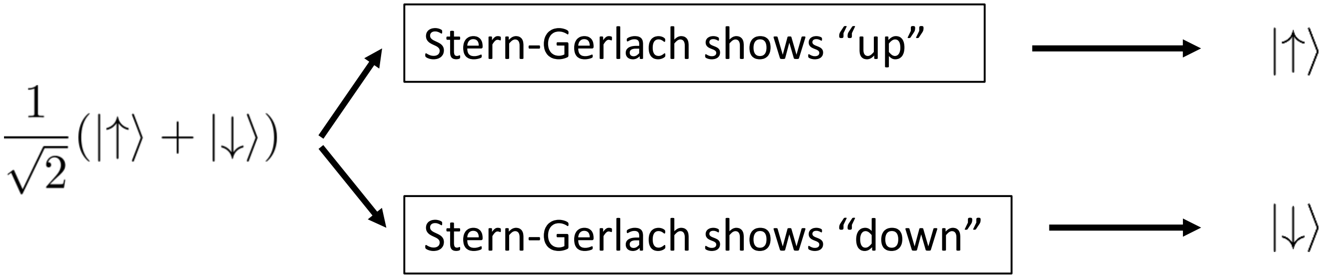
Decoherence,
quantum measurement,
master equation, exponential decay
and all that

Quantum mechanics and measurement

- states (pure states) are vectors in Hilbert space $|\psi\rangle$
- measuring devices correspond to Hermitian operators $\hat{A}$
- superposition states

$$\frac{1}{\sqrt{2}}(|\uparrow\rangle + |\downarrow\rangle)$$

- measuring devices are classical devices, described by classical, not quantum mechanics
- quantum system state collapses to an eigenstate of the measuring device's operator



$$\frac{1}{\sqrt{2}}(|\uparrow\rangle + |\downarrow\rangle)$$

Stern-Gerlach shows "up" $\longrightarrow$ $|\uparrow\rangle$

Stern-Gerlach shows "down" $\longrightarrow$ $|\downarrow\rangle$

Problem of measurement in QM:
We would like to describe the measuring device also by quantum mechanics (by Schrodinger equation), but this is incompatible with state collapse as shown by von Neuman because **state collapse is a non-unitary operation**

Density matric, statistical

$$\overline{A} = \sum_x p(x) \langle x|\, \hat{A} |x\rangle$$

$$\hat{\rho} = \sum_x |x\rangle p(x) \langle x|$$

$$\overline{A} = Tr(\hat{\rho}\hat{A})$$

$$\begin{aligned}
Tr(\rho\hat{A}) &= \sum_n \langle n|\, \hat{\rho}\hat{A} |n\rangle = \sum_n \langle n|\, \sum_x |x\rangle p(x) \langle x|\, \hat{A}|n\rangle = \\
&= \sum_n \sum_x \langle n \mid x\rangle\, p(x) \langle x|\, \hat{A}|n\rangle = \\
&= \sum_n \sum_x p(x) \langle x|\, \hat{A}|n\rangle\, \langle n \mid x\rangle = \\
&= \sum_x p(x) \langle x|\, \hat{A}|x\rangle = \overline{A}
\end{aligned}$$

The density matrix is evidently a hermitian operator, it can be expressed in any base (in general it would not be diagonal) as

$$\hat{\rho} = \sum_{m,n} |m\rangle \rho_{mn} \langle n|$$

The time evolution of a macrostate will be described by a time dependent density matrix

$$\begin{aligned}\hat{\rho}(t) &= \sum_{x} |x(t)\rangle p(x) \langle x(t)| = \\ &= \sum_{x} \exp(-\frac{i}{\hbar}\hat{H}t)|x\rangle p(x)\langle x| \exp(\frac{i}{\hbar}\hat{H}t)\end{aligned}$$

Differentiating with respect to time we get the equation of motion

$$i\hbar\frac{\partial}{\partial t}\hat{\rho}(t) = [\hat{H}, \hat{\rho}]$$

Density matric, subsystem

Compound system (S,E)

$$|\psi\rangle_{SE} = \sum_{ik} c_{ik} \, |i\rangle_S \, |k\rangle_E$$

$$\hat{H} = \hat{H}_S + \hat{H}_E + \hat{H}_{\text{int}}$$

$$|\psi(t)\rangle_{SE} = \sum_{ij} c_{ij}(t) \, |i\rangle_S \, |j\rangle_E$$

Suppose we have a measuring device which measures some quantity of the system S. It means that its operator $\hat{A}$ operates just in the subspace of the compound Hilbert space. This subspace is spanned on vectors $|i\rangle_S$. It means it is completely determined by the matrix elements

$$A_{ij} = \langle i|_S \, \hat{A} \, |j\rangle_S$$

the mean A value is then calculated as

$$\langle\psi|_{SE} \, \hat{A} \, |\psi\rangle_{SE} = \sum_{i'k'} c^*_{i'k'} \, \langle i'|_S \, \langle k'|_E \, \hat{A} \sum_{ik} c_{ik} \, |i\rangle_S \, |k\rangle_E$$

Density matric, subsystem

$$\langle\psi|_{SE}\hat{A}|\psi\rangle_{SE}=\sum_{i'k'}c^*_{i'k'}\langle i'|_S\langle k'|_E\hat{A}\sum_{ik}c_{ik}|i\rangle_S|k\rangle_E=\sum_{ii'}A_{ii'}\sum_k c^*_{i'k}c_{ik}$$

So we see that all information about the state is hidden in the expression

$$\varrho_{i'i}=\sum_k c^*_{i'k}c_{ik}$$

We can introduce a density matrix operator

$$\hat{\varrho}=\sum_{i'i}\varrho_{i'i}|i'\rangle\langle i|$$

and the mean A value can be written as

$$\boxed{\langle\psi|_{SE}\hat{A}|\psi\rangle_{SE}=Tr(\hat{A}\hat{\varrho})}$$

This is similar to what we had for the statistical density matrix. A big difference between the two matrices is that the time evolution of the subsystem density matrix is not given just by the system Hamiltonian $\widehat{H}_S$. To calculate the time evolution of a subsystem density matrix is not simple.

Density matric, subsystem

The time evolution of a subsystem density matrix can clearly be calculated by keeping the pure state of the overall system SE:

$$|\psi\rangle_{SE} = \sum_{ik} c_{ik} \, |i\rangle_S \, |k\rangle_E$$

Using the total Hamiltonian $\hat{H} = \hat{H}_S + \hat{H}_E + \hat{H}_{\text{int}}$

we calculate its time development (using the Schrödinger equation)

$$|\psi(t)\rangle_{SE} = \sum_{ij} c_{ij}(t) \, |i\rangle_S \, |j\rangle_E$$

and then calculate the subsystem density matrix in each time instant anew:

$$\varrho_{i'i}(t) = \sum_{k} c^*_{i'k}(t) c_{ik}(t)$$

$$\hat{\varrho}(t) = \sum_{i'i} \varrho_{i'i}(t) \, |i'\rangle \, \langle i|$$

$$\bar{A}(t) = Tr(\hat{A}\hat{\varrho}(t))$$

The time development of the subsystem density matrix is clearly **non-unitary.**

If at some time instant t_0 system S becomes decoupled from E, then its state is given by the subsystem density matrix $\hat{\varrho}(t_0)$ which becomes the statistical density matrix at time t_0, and its further time development is **unitary** corresponding to the Hamiltonian $\widehat{H}_S$.

$$i\hbar \frac{\partial}{\partial t} \hat{\varrho}(t) = [\hat{H}_S, \hat{\varrho}(t)]$$

Measurement, wave function collapse, statistical density matrix

Suppose we have a pure quantum state of some system S given as a superposition of eigenstates $|\varphi_i\rangle$ of a Hermitian operator $\hat{A}$

$$\hat{A}\,|\varphi_i\rangle = a_i\,|\varphi_i\rangle$$

$$|\psi\rangle = \sum_i c_i\,|\varphi_i\rangle$$

We perform the measurement of the quantity A and the measuring device shows the value a_k (it always shows one of the eigenvalues a_i). The wave function collapse happens immediately after the measurement the system wave function becomes $|\varphi_k\rangle$.
The probability of this to happen is $|c_k|^2$.
If we repeat the identical measurement experiment many times, we get various vales a_i with probabilities $|c_i|^2$ leading the wave function collapse into the state $|\varphi_i\rangle$.
So repeating the measurement experiment we get a statistical ensemble of states described by the statistical density matrix

$$\hat{\varrho} = \sum_i |\varphi_i\rangle\,|c_i|^2\,\langle\varphi_i|$$

Measurement, wave function collapse, statistical density matrix

So a measurement of a quantity A, leads to the non-unitary wave function collapse, the state of the system after the measurement can be described by a statistical density matrix

$$\boxed{|\psi\rangle = \sum_i c_i \, |\varphi_i\rangle \longrightarrow \hat{\varrho} = \sum_i |\varphi_i\rangle \, |c_i|^2 \, \langle \varphi_i|}$$

Problem of measurement in QM:
We would like to describe the measuring device also by quantum mechanics (by Schrodinger equation), but this is incompatible with state collapse as shown by von Neuman because **state collapse is a non-unitary operation**

Our aim is to show that a process of quantum decoherence can lead to “effective” wave function collapse and thus can “mimic” a measurement process.

Decoherence due to interaction with environment

$$\psi(0) = \sum_{ij} c_{ij}(0) \left|s_i\right\rangle \left|\varphi_j\right\rangle_E \qquad \hat{H} = \hat{H}_S + \hat{H}_E + \hat{H}_{\text{int}}$$

System and environment. $\left|s_i\right\rangle$ is an arbitrary base in system subspace, $\left|\varphi_j\right\rangle_E$ is a base in environment subspace. Time development is given by the total Hamiltonian (via Schrödinger equation)

$$\psi(t) = \sum_{ij} c_{ij}(t) \left|s_i\right\rangle \left|\varphi_j\right\rangle_E$$

$$\psi(t) = \sum_{i} \left|s_i\right\rangle \sum_{j} c_{ij}(t) \left|\varphi_j\right\rangle_E$$

$\sum_{j} c_{ij}(t) \left|\varphi_j\right\rangle_E$ is a set of some vectors in the subspace of the environment, indexed by i.

These vectors **do not form a base** in the environment subspace: the environment has much more degrees of freedom. We can introduce normalized vectors in the environment subspace as $\left|e_i(t)\right\rangle = \lambda_i(t) \sum_{j} c_{ij}(t) \left|\varphi_j\right\rangle_E$ We get

$$\psi(t) = \sum_{i} \lambda_i(t) \left|s_i\right\rangle \left|e_i(t)\right\rangle$$

Suppose we have a physical quantity which from the system subspace (not sensitive to the state of environment. Then

$$\bar{A}(t) = \sum_{ii'} \lambda_i^*(t) \lambda_{i'}(t) \left\langle e_i(t)\right| \left\langle s_i\right| \hat{A} \left|s_i'\right\rangle \left|e_i'(t)\right\rangle$$

Decoherence: collapse to diagonal density matrix

Now we can introduce a subsystem density matrix:

$$\bar{A}(t) = \sum_{ii'} \lambda_i^*(t)\lambda_{i'}(t) \langle e_i(t)| \langle s_i| \hat{A} |s_i'\rangle |e_i'(t)\rangle = \sum_{ii'} \varrho_{i'i} A_{ii'} = Tr(\hat{\varrho}\hat{A})$$

$$\varrho_{i'i} = \lambda_i^*(t)\lambda_{i'}(t) \langle e_i(t)| e_i'(t)\rangle$$

What if

$$\langle e_i(t)| e_i'(t)\rangle \rightarrow \delta_{ii'}$$

$$\hat{\varrho} = \sum_i |s_i\rangle |\lambda_i|^2 \langle s_i|$$

This matrix is diagonal !!!

Compaer this with the wave function collapse after measurement

$$|\psi\rangle = \sum_i c_i |\varphi_i\rangle \;\; \longrightarrow \;\; \hat{\varrho} = \sum_i |\varphi_i\rangle |c_i|^2 \langle \varphi_i|$$

Decoherence: collapse to diagonal density matrix

$$\psi(t) = \sum_{ij} c_{ij}(t) \left|s_i\right\rangle \left|\varphi_j\right\rangle_E$$

$$\psi(t) = \sum_{i} \lambda_i(t) \left|s_i\right\rangle \left|e_i(t)\right\rangle$$

$$\varrho_{i'i} = \lambda_i^*(t)\lambda_{i'}(t) \left\langle e_i(t)\right| \left. e_{i'}(t)\right\rangle$$

$$\left\langle e_i(t)\right| \left. e_i'(t)\right\rangle \to \delta_{ii'}$$

For a different choice of the basis $\left|ss_i\right\rangle$ we would get different accompanying environment vectors $\left|ee_i\right\rangle$. So even if for some choice of $\left|s_i\right\rangle$ we get $\left\langle e_i(t)\right| \left. e_i'(t)\right\rangle \to \delta_{ii'}$, for a different choice $\left|ss_i\right\rangle$, the accompanying environment vectors will not converge to diagonal matrix

$$\left\langle ee_i(t)\right| \left. ee_i'(t)\right\rangle \nrightarrow \delta_{ii'}$$

It means that we can have a decoherent collapse for some choice of basis states of the system and not have a decoherent collapse for some other choice. Or perhaps a bit weaker: **for some choice we could observe a decoherence collapse after just a small time of interaction with environment and for other choice of basis states the decoherence time would be much larger**

Decoherence: collapse to diagonal density matrix

$$\bar{A}(t) = \sum_{ii'} \lambda_i^*(t)\lambda_{i'}(t) \langle e_i(t)| \langle s_i| \hat{A} |s_i'\rangle |e_i'(t)\rangle = \sum_{ii'} \varrho_{i'i} A_{ii'} = Tr(\hat{\varrho}\hat{A})$$

$$\varrho_{i'i} = \lambda_i^*(t)\lambda_{i'}(t) \langle e_i(t)|\, e_i'(t)\rangle$$

Do notice that the definition of the density matrix does not depend at all on what operator $\hat{A}$ we have chosen for the derivation of the density matrix.

If the density matrix is diagonal, it means that all measurements on the system S can discover superpositions of the basis states.
Be careful: we cannot **deduce** from the results of any measurement, that the system S was in some superposition of states $|s_i\rangle$ **before the measurement was performed**. Of course **after the measurement** the system would be in an eigenstate of the measured quantity, which **can** be a superposition of the states $|s_i\rangle$.

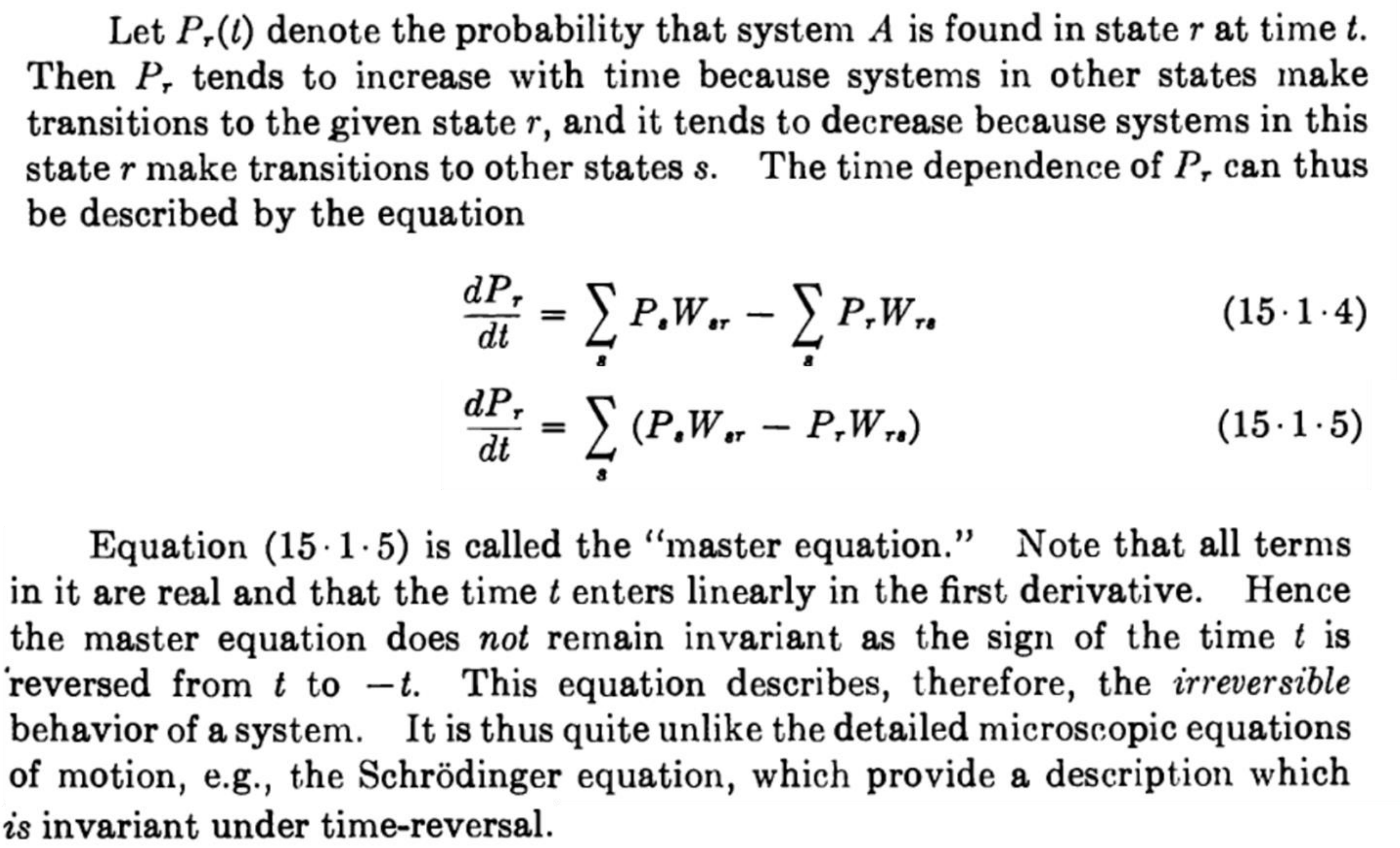
Be careful: we **did not prove that decoherence exists**. We did not prove that

$$\langle e_i(t)|\, e_i'(t)\rangle \to \delta_{ii'}$$

We just said that **"perhaps it may happen"**.

Decoherence and quantum kinetics: master equation

The following is a snippet from Reif: Fundamentals of statistical and thermal physics



Let $P_r(t)$ denote the probability that system A is found in state r at time t. Then P_r tends to increase with time because systems in other states make transitions to the given state r, and it tends to decrease because systems in this state r make transitions to other states s. The time dependence of P_r can thus be described by the equation

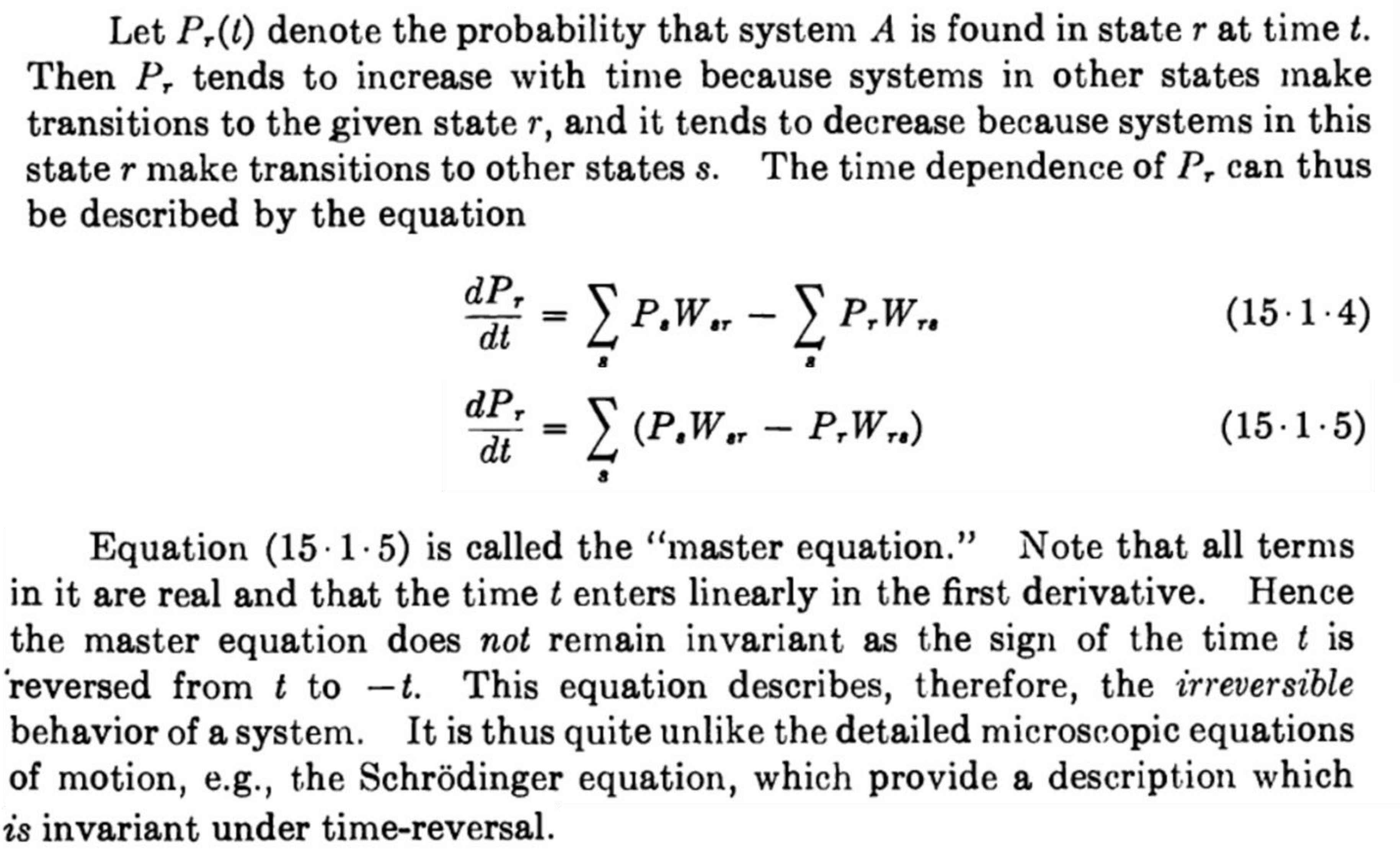
$$\frac{dP_r}{dt} = \sum_s P_s W_{sr} - \sum_s P_r W_{rs} \qquad (15\cdot1\cdot4)$$

$$\frac{dP_r}{dt} = \sum_s (P_s W_{sr} - P_r W_{rs}) \qquad (15\cdot1\cdot5)$$

Equation $(15\cdot1\cdot5)$ is called the "master equation." Note that all terms in it are real and that the time t enters linearly in the first derivative. Hence the master equation does *not* remain invariant as the sign of the time t is reversed from t to $-t$. This equation describes, therefore, the *irreversible* behavior of a system. It is thus quite unlike the detailed microscopic equations of motion, e.g., the Schrödinger equation, which provide a description which *is* invariant under time-reversal.

Decoherence and quantum kinetics: master equation

The following is a snippet from Reif: Fundamentals of statistical and thermal physics



Let $P_r(t)$ denote the probability that system A is found in state r at time t. Then P_r tends to increase with time because systems in other states make transitions to the given state r, and it tends to decrease because systems in this state r make transitions to other states s. The time dependence of P_r can thus be described by the equation

$$\frac{dP_r}{dt} = \sum_s P_s W_{sr} - \sum_s P_r W_{rs} \qquad (15\cdot 1\cdot 4)$$

$$\frac{dP_r}{dt} = \sum_s (P_s W_{sr} - P_r W_{rs}) \qquad (15\cdot 1\cdot 5)$$

Equation (15 · 1 · 5) is called the "master equation." Note that all terms in it are real and that the time t enters linearly in the first derivative. Hence the master equation does *not* remain invariant as the sign of the time t is reversed from t to $-t$. This equation describes, therefore, the *irreversible* behavior of a system. It is thus quite unlike the detailed microscopic equations of motion, e.g., the Schrödinger equation, which provide a description which *is* invariant under time-reversal.

Decoherence and quantum kinetics: master equation

Reif stresses the fact of time irreversibility of the master equation used to describe the kinetics of a quantum object.
For me, reading the book, a more important question is this: **where are the superposition states?**

$$\frac{dP_r}{dt} = \sum_s (P_s W_{sr} - P_r W_{rs})$$

The master equation deals only with probabilities, not with amplitudes! How this can be justified for a quantum object?

The answer may be this:

Consider the time steps in the differential master equation as not infinitesimally small. So net derivatives in a strict mathematical sense. Just small time denominators (Newton had no ε , δ , just fractions of small quantities). If the small time increments are larger than decoherence time due to continuous interaction with environment, we get it!
In this way an exponential decay is acceptable, which otherwise can hardly be justified in "orthodox" quantum mechanics.

System in contact with a heat reservoir

Consider a system A in thermal contact with a much larger system A'. The total Hamiltonian of the combined system $A^{(0)} = A + A'$ is

$$\mathcal{H}^{(0)} = \mathcal{H} + \mathcal{H}' + \mathcal{H}_i$$

where $\mathcal{H}$ is the Hamiltonian of A, $\mathcal{H}'$ is that of the heat reservoir A', and $\mathcal{H}_i$ is very small and describes the weak interaction between A and A'. In the absence of interaction when $\mathcal{H}_i = 0$, denote the energy of A in state r by E_r, and the energy of A' in state r' by $E'_{r'}$. The presence of the interaction $\mathcal{H}_i$ induces transitions between these states and is responsible for bringing about equilibrium between A and A'.

Let P_r be the probability that A is in state r, and $P'_{r'}$ the corresponding probability that A' is in state r'. The state of the combined system $A^{(0)}$ is described by the pair of numbers r and r'; the probability of $A^{(0)}$ being in this state is $P^{(0)}_{rr'} = P_r P'_{r'}$. In this combined system the interaction $\mathcal{H}_i$ causes transitions between states. Under assumptions similar to those of the last subsection, there exists a well-defined transition probability $W^{(0)}(rr' \to ss')$ per unit time from state rr' to state ss'. By conservation of energy,

$$\text{if } E_r + E'_{r'} \neq E_s + E'_{s'}, \qquad W^{(0)}(rr' \to ss') = 0 \qquad (15 \cdot 2 \cdot 1)$$

$$W^{(0)}(ss' \rightarrow rr') = W^{(0)}(rr' \rightarrow ss')$$

detailed balance

$$P_r P'_{r'} W^{(0)}(rr' \rightarrow ss') = P_s P'_{s'} W^{(0)}(ss' \rightarrow rr')$$

$$\frac{P_r}{P_s} = \frac{P'_{s'}}{P'_{r'}}$$

$$\frac{f(E_r)}{f(E_s)} = \frac{f(E_{s'})}{f(E_{r'})}$$

$$f(E_r)f(E_{r'}) = f(E_s)f(E_{s'})$$

$$f(E_r)f(E - E_r) = f(E_s)f(E - E_s)$$

$$f(E_r)f(E - E_r) = \text{const}$$

$$f'(E_r)f(E - E_r) - f(E_r)f'(E - E_r) = 0$$

$$\frac{f'(E_r)}{f(E_r)} = \text{comst}$$

$$f(E) = C\exp(-\beta E)$$

$$\frac{P_r}{P_s} = e^{-\beta(E_r - E_s)} = \frac{e^{-\beta E_r}}{e^{-\beta E_s}}$$

The transition probability W_{rs} can be obtained by multiplying the transition probability $W^{(0)}(rr' \to ss')$ for the combined system $A^{(0)}$ by the probability $P'_{r'}$ that A' is in the particular state r' and then summing over all possible initial states r' in which A' can be found and all the possible final states s' in which it can end up. Thus

$$W_{rs} = \sum_{r's'} P_{r'} W^{(0)}(rr' \to ss') = C \sum_{r's'} e^{-\beta E'_{r'}} W^{(0)}(rr' \to ss') \qquad (15\cdot 2\cdot 8)$$

$$W_{sr} = C \sum_{r's'} e^{-\beta E'_{s'}} W^{(0)}(ss' \to rr')$$

But the energy conservation (15·2·1) implies that $E'_{s'} = E'_{r'} + E_r - E_s$. Using the symmetry relation (15·2·2), the expression (15·2·9) becomes then

$$W_{sr} = C \sum_{r's'} e^{-\beta E_{r'}}\, e^{-\beta(E_r - E_s)} W^{(0)}(rr' \to ss') = e^{-\beta(E_r - E_s)} W_{rs}$$

Thus

▶

$$\frac{W_{sr}}{W_{rs}} = e^{-\beta(E_r - E_s)} = \frac{e^{-\beta E_r}}{e^{-\beta E_s}} \qquad (15\cdot 2\cdot 10)$$

$$\frac{W_{sr}}{W_{rs}} = e^{-\beta(E_r - E_s)} = \frac{e^{-\beta E_r}}{e^{-\beta E_s}}$$

detailed balance

$$P_r W_{rs} = P_s W_{sr}$$

$$\frac{P_r}{P_s} = e^{-\beta(E_r - E_s)} = \frac{e^{-\beta E_r}}{e^{-\beta E_s}}$$

It is convenient to make the relationship $(15\cdot 2\cdot 10)$ apparent by introducing a quantity $\lambda_{rs} = \lambda_{sr}$ defined by

$$e^{-\beta E_s} W_{sr} = e^{-\beta E_r} W_{rs} \equiv \lambda_{rs} \equiv \lambda_{sr}$$

Then one can write

$$W_{sr} = e^{\beta E_s} \lambda_{sr}, \qquad W_{rs} = e^{\beta E_r} \lambda_{rs} \tag{15$\cdot$2$\cdot$12}$$

and $(15\cdot 2\cdot 10)$ follows automatically since $\lambda_{sr} = \lambda_{rs}$.

The probabilities for system A satisfy again the master equation $(15\cdot 1\cdot 5)$

$$\frac{dP_r}{dt} = \sum_s (P_s W_{sr} - P_r W_{rs}) = \sum_s \lambda_{sr} (P_s \, e^{\beta E_s} - P_r \, e^{\beta E_r}) \tag{15$\cdot$2$\cdot$13}$$

Magnetic resonance

An instructive and important example of the ideas developed in the preceding sections is that of magnetic resonance. Consider a substance containing N noninteracting nuclei (or electrons) of spin $\frac{1}{2}$ and magnetic moment μ. If the substance is placed in an applied magnetic field H, each spin can point either "up" (i.e., parallel to H) or "down." We denote the corresponding states by $+$ and $-$, respectively. The two possible energies of each nucleus are then

$$\epsilon_{\pm} = \mp \mu H \qquad (15 \cdot 3 \cdot 1)$$

Let n_+ be the mean number of spins pointing up and n_- be the mean number of spins pointing down. Clearly, $n_+ + n_- = N$.

The total Hamiltonian of the system can be written as

$$\mathcal{H}^{(0)} = \mathcal{H}_n + \mathcal{H}_L + \mathcal{H}_i$$

Here $\mathcal{H}_n$ is the Hamiltonian expressing the interaction between the nuclear moments and the external field H; $\mathcal{H}_L$ is the Hamiltonian describing the "lattice," i.e., all non–spin degrees of freedom of the nuclei and all other atoms in the substance. The Hamiltonian $\mathcal{H}_i$ describes the interaction between the spins of the nuclei and the lattice, and causes transitions between the possible spin states of the nuclei. (For example, the magnetic moment of a *moving* nucleus produces a fluctuating magnetic field at the positions of other nuclei and this field causes transitions.) Let W_{+-} be the transition probability per unit time that a nucleus flips its spin from "up" to "down" as a result of interaction with the lattice. The lattice itself is a large system which can be regarded as being always very close to internal equilibrium at the absolute temperature $T = (k\beta)^{-1}$. Thus (15·2·10) allows one to write the relationship

$$\frac{W_{-+}}{W_{+-}} \doteq \frac{e^{-\beta\epsilon_+}}{e^{-\beta\epsilon_-}} = e^{\beta(\epsilon_- - \epsilon_+)} \qquad (15\cdot3\cdot2)$$

$$\frac{W_{-+}}{W_{+-}} = \frac{e^{-\beta\epsilon_+}}{e^{-\beta\epsilon_-}} = e^{\beta(\epsilon_- - \epsilon_+)} \tag{15·3·2}$$

Now (15 · 3 · 1) gives $\beta(\epsilon_- - \epsilon_+) = 2\beta\mu H$. For nuclei the magnetic moment $\mu \approx 5 \cdot 10^{-24}$ ergs/gauss so that in laboratory fields of the order of $H = 10^4$ gauss

$$\beta\mu H = \frac{\mu H}{kT} \approx \frac{5 \cdot 10^{-4}}{T} \ll 1$$

for all but *exceedingly* low temperatures. Even for electronic moments, which are about 1000 times larger, this inequality is almost always well satisfied. By expanding the exponential, the relation (15 · 3 · 2) can then be written in the form

$$\begin{aligned} &W_{+-} \equiv W \\ \text{and} \qquad &W_{-+} = W(1 + 2\beta\mu H), \qquad \text{where } \beta\mu H \ll 1 \end{aligned} \tag{15·3·3}$$

Finally, there may also be present an externally applied alternating magnetic field of angular frequency ω. If $\hbar\omega \approx \epsilon_- - \epsilon_+ = 2\mu H$, this field will induce transitions between the spin states of a nucleus. (If $H \approx 10^4$ gauss, ω is typically a radio frequency (rf) of the order of 10^8 sec^{-1}.) Let w_{+-} be the transition probability per unit time for the "up" to "down" transition induced by this rf field. Then one again has the symmetry property (15·1·3)

Here $w = w(\omega)$ is only appreciable if ω satisfies the resonance condition $\hbar\omega \approx 2\mu H$.

The master equations for $n_+(t)$ and $n_-(t)$ then become

$$\left.\begin{aligned} \frac{dn_+}{dt} &= n_-(W_{-+} + w) - n_+(W_{+-} + w) \\ \frac{dn_-}{dt} &= n_+(W_{+-} + w) - n_-(W_{-+} + w) \end{aligned}\right\} \qquad (15\cdot3\cdot5)$$

$$\left.\begin{aligned}\frac{dn_+}{dt} &= n_-(W_{-+} + w) - n_+(W_{+-} + w)\\ \frac{dn_-}{dt} &= n_+(W_{+-} + w) - n_-(W_{-+} + w)\end{aligned}\right\} \qquad (15\cdot3\cdot5)$$

By subtracting the second equation from the first, one obtains

$$\frac{d}{dt}(n_+ - n_-) = -2n_+(W_{+-} + w) + 2n_-(W_{-+} + w) \qquad (15\cdot3\cdot6)$$

Introducing the population difference

$$n \equiv n_+ - n_- \qquad (15\cdot3\cdot7)$$

and using (15 · 3 · 3), the relation (15 · 3 · 6) becomes

$$\frac{dn}{dt} = -2(W + w)n + 2\beta\mu HWN \qquad (15\cdot3\cdot8)$$

Here we have put $4\beta\mu HWn_- = 4\beta\mu HW\ (\frac{1}{2}\ N - n) \approx 2\beta\mu HWN$, since one always has $n \ll N$ in the temperature range of interest.

$$\frac{dn}{dt} = -2(W + w)n + 2\beta\mu HWN \tag{15·3·8}$$

Here we have put $4\beta\mu HWn_- = 4\beta\mu HW\ (\frac{1}{2}\ N - n) \approx 2\beta\mu HWN$, since one always has $n \ll N$ in the temperature range of interest.

Let us now investigate various cases of interest. Consider first the equilibrium situation in the absence of an applied rf field, i.e., with $w = 0$. Then $dn/dt = 0$, and (15·3·8) yields for the *equilibrium* excess number of spins

$$n_0 = N\beta\mu H \tag{15·3·9}$$

This is, of course, the result which follows from the canonical distribution according to which one has *in equilibrium*

$$n_\pm = N\frac{e^{\pm\beta\mu H}}{e^{\beta\mu H} + e^{-\beta\mu H}} \approx N\frac{1 \pm \beta\mu H}{2} = \frac{1}{2}N(1 \pm \beta\mu H)$$

so that $n_0 = n_+ - n_-$ assumes the value (15·3·9)

Hence (15·3·8) can be written in the form

$$\frac{dn}{dt} = -2W(n - n_0) - 2wn \tag{15·3·10}$$

▶
$$\frac{dn}{dt} = -2W(n - n_0) - 2wn \tag{15·3·10}$$

In the absence of a rf field when $w = 0$, this yields upon integration

$$n(t) = n_0 + [n(0) - n_0]\, e^{-2Wt} \tag{15·3·11}$$

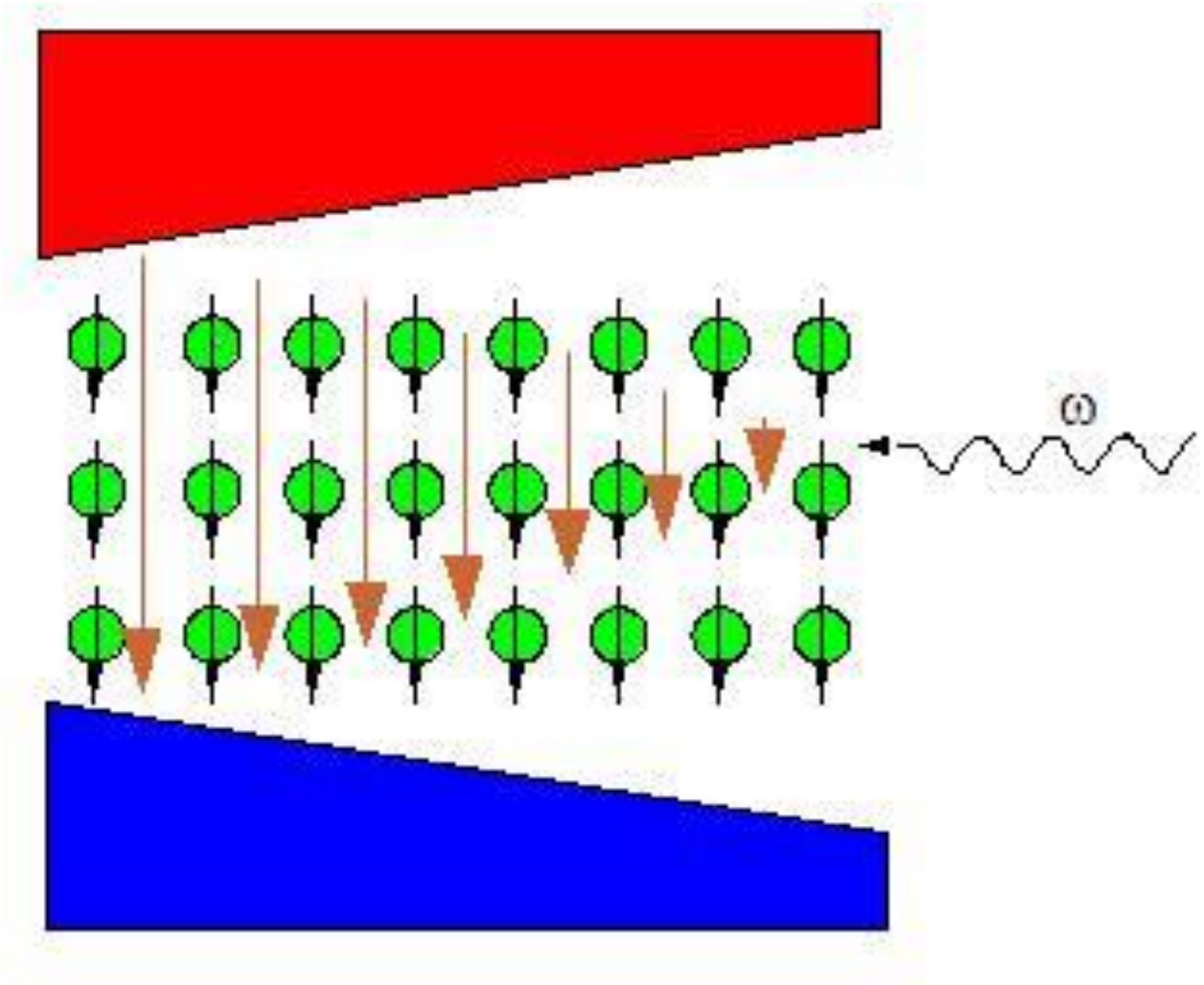
where $n(0)$ is the population difference at the initial time $t = 0$. Thus $n(t)$ approaches its equilibrium value n_0 exponentially with a characteristic "relaxation time" $(2W)^{-1}$. Obviously, the larger the interaction W of the spins with the lattice heat reservoir, the shorter the relaxation time.

Suppose now that the interaction of the spins with the lattice is very weak, so that $W \approx 0$, and that a rf magnetic field is applied. Then (15·3·10) becomes

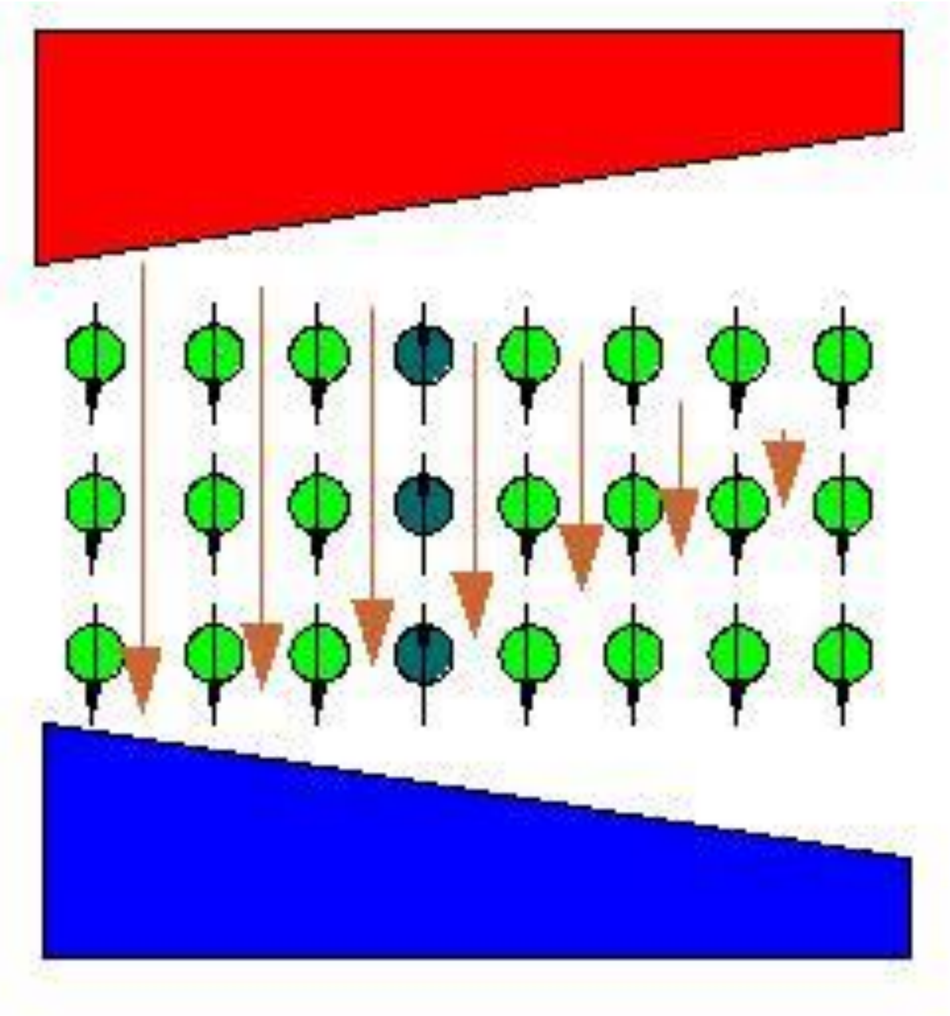
$$\frac{dn}{dt} = -2wn$$

so that

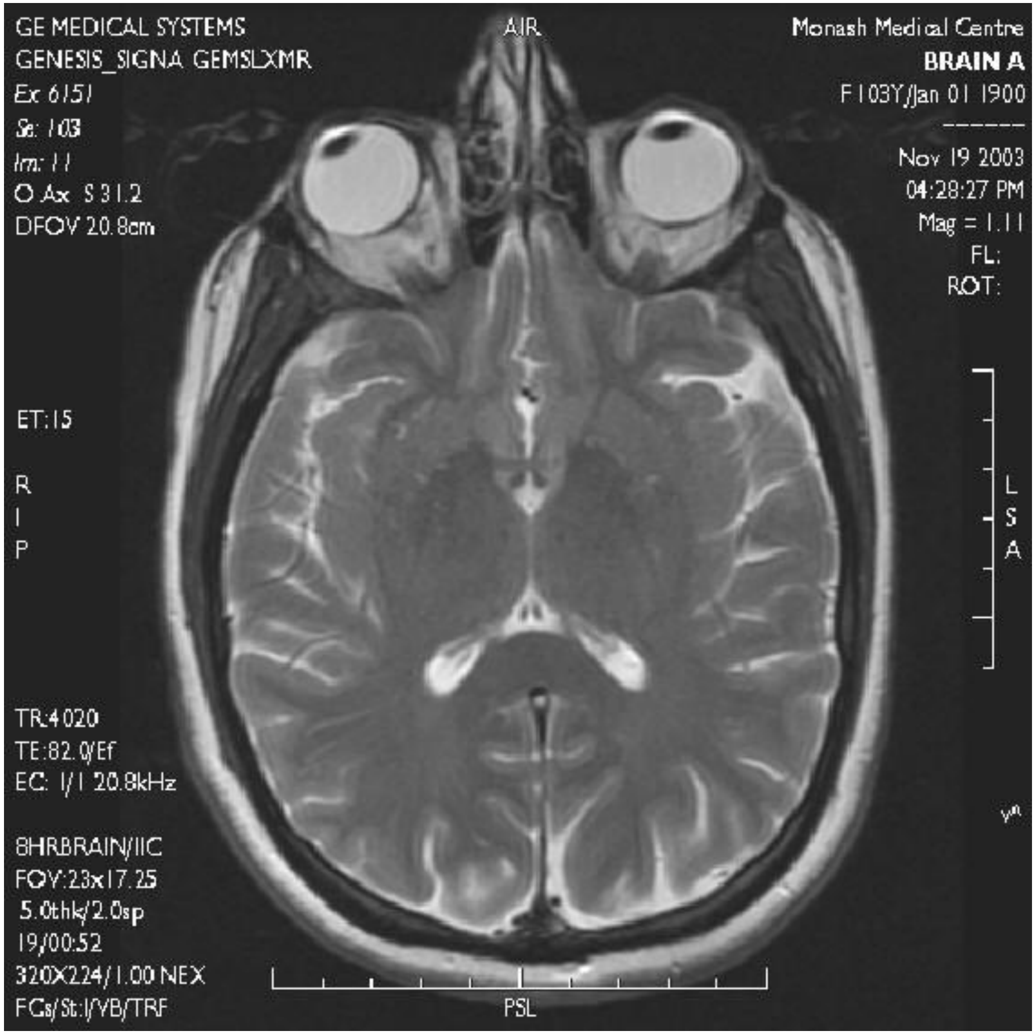
$$n(t) = n(0)\, e^{-2wt} \tag{15·3·12}$$



$\omega=\gamma B$



$\omega = \gamma B$


GE MEDICAL SYSTEMS
GENESIS_SIGNA GEMSLXMR
Ex 6151
Se: 103
Im: 11
O Ax S 31.2
DFOV 20.8cm
AIR
Monash Medical Centre
BRAIN A
F 103Y/Jan 01 1900

Nov 19 2003
04:28:27 PM
Mag = 1.11
FL:
ROT:
ET:15
R
I
P
L
S
A
TR:4020
TE:82.0/Ef
EC: 1/1 20.8kHz
8HRBRAIN/IIC
FOV:23x17.25
5.0thk/2.0sp
19/00:52
320X224/1.00 NEX
FCs/St:I/VB/TRF
PSL

Dynamic nuclear polarization; Overhauser effect

Nonequilibrium methods for achieving nuclear polarization (the so-called "Overhauser effect" being one such method) provide a particularly illuminating illustration of the ideas presented in the last few sections. Consider a substance containing both nuclei of spin $\frac{1}{2}$ and magnetic moment μ_n, and also unpaired electrons of spin $\frac{1}{2}$ and magnetic moment μ_e ($\mu_e < 0$). The substance is placed in an external magnetic field H pointing in the z direction. Suppose that the principal interaction of a nucleus is with an electron through "hyperfine interaction," i.e., through the magnetic field produced by the electron at the position of the nucleus. (This interaction is described by a Hamiltonian of the form $\mathcal{H}_{ne} = a\mathbf{I}\cdot\mathbf{S}$ where $\mathbf{I}$ is the nuclear and $\mathbf{S}$ the electronic spin angular momentum.) Since no z component of external torque acts on the system consisting of nucleus and electron (i.e., since the total Hamiltonian of this system is invariant under rotations about the z direction), the *total* z component of angular momentum ($I_z + S_z$) of this system is a constant of the motion. Thus the transitions induced by the interaction between a nucleus and an electron must always be such that whenever the nucleus flips its spin from "up" to "down," the electron must flip its spin from "down" to "up," and vice versa. We shall denote the transition probability per unit time due to this interaction by $W_{ne}(+\,\boldsymbol{-} \to -\,\boldsymbol{+})$ where $+$ and $-$ indicate up and down orientations of the nucleus n, and $\boldsymbol{+}$ and $\boldsymbol{-}$ up and down orientations of the electron e.

The nuclei interact then predominantly with the electron spins, which in turn interact appreciably with the lattice heat reservoir. It is through this chain of interactions that the nuclear spins attain the thermal equilibrium situation corresponding to the lattice temperature $T = (k\beta)^{-1}$. Let n_+ and n_- denote the mean number of nuclear "up" and "down" spins; let N_+ and N_- denote the mean number of electron "up" and "down" spins. In thermal equilibrium one then obtains, for the nuclei of energy $\epsilon_\pm = \mp \mu_n H$,

$$\frac{n_+}{n_-} = \frac{e^{\beta\mu_n H}}{e^{-\beta\mu_n H}} = e^{2\beta\mu_n H} \tag{15·4·1}$$

Also for the electrons of energy $E_\pm = \mp \mu_e H$,

$$\frac{N_+}{N_-} = \frac{e^{\beta\mu_e H}}{e^{-\beta\mu_e H}} = e^{2\beta\mu_e H} \tag{15·4·2}$$

The degree of polarization of the nuclei and electrons can be measured by the respective ratios defined by

$$\xi_n \equiv \frac{n_+ - n_-}{n_+ + n_-} \qquad \text{and} \qquad \xi_e \equiv \frac{N_+ - N_-}{N_+ + N_-} \tag{15·4·3}$$

The degree of polarization of the nuclei and electrons can be measured by the respective ratios defined by

$$\xi_n \equiv \frac{n_+ - n_-}{n_+ + n_-} \qquad \text{and} \qquad \xi_e \equiv \frac{N_+ - N_-}{N_+ + N_-} \tag{15·4·3}$$

Each ξ lies in the range $-1 \le \xi \le 1$. Since $|\mu_e| \approx 1000|\mu_n|$, it is clear by (15·4·1) and (15·4·2) that $\xi_n \ll \xi_e$. Even if one goes to such high applied fields H and low temperatures T that the electrons are appreciably polarized, the degree of nuclear polarization is thus still very small, in particular much too small to do nuclear physics experiments on polarized nuclei.

From the point of view of the transition probabilities one can regard the combined system $(n + e)$ of nucleus and electron as being in thermal contact with the lattice heat reservoir. It then follows by (15·2·10) that the transition probabilities must satisfy the relation

$$\frac{W_{ne}(+\,-\to-\,+)}{W_{ne}(-\,+\to+\,-)} = e^{-\beta(\epsilon_- + E_+ - \epsilon_+ - E_-)} = e^{-2\beta(\mu_n - \mu_e)H} \tag{15·4·4}$$

detailed balance

$$n_+ N_- W_{ne}(+\,\boldsymbol{-} \to -\,\boldsymbol{+}) = n_- N_+ W_{ne}(-\,\boldsymbol{+} \to +\,\boldsymbol{-})$$

$$\frac{n_+}{n_-}\frac{N_-}{N_+} = e^{2\beta(\mu_n - \mu_e)H}$$

Imagine, however, that an rf field is applied at the electron-spin resonance frequency and suppose that it is strong enough to saturate the electron-spin system so that $N_+ = N_-$. If one assumes that the detailed balance condition (15·4·5) remains valid in this steady-state situation, (15·4·6) becomes

▶ $$\boxed{\frac{n_+}{n_-} = e^{2\beta(\mu_n - \mu_e)H} \approx e^{-2\beta\mu_e H}} \qquad (15\cdot 4\cdot 7)$$

Thermodynamics of Irreversible Processes and the Onsager Reciprocal Relations

In many reversible processes, such as the dielectric polarization of a crystal by an electric field, it is well known that the response matrix is symmetric. Thus if in an anisotropic crystal the polarization P is related to the electric field E by

$$\begin{aligned} P_x &= \chi_{11}E_x + \chi_{12}E_y + \chi_{13}E_z; \\ P_y &= \chi_{21}E_x + \chi_{22}E_y + \chi_{23}E_z; \\ P_z &= \chi_{31}E_x + \chi_{32}E_y + \chi_{33}E_z; \end{aligned} \tag{33.1}$$

there exist among the matrix elements of the susceptibility tensor χ the relations

$$\chi_{ij} = \chi_{ji}, \tag{33.2}$$

$$dU_A = \tau\, d\sigma + \mathbf{P} \cdot d\mathbf{E},$$

$$F_A = U_A - \tau\sigma;$$

$$dF_A = \mathbf{P} \cdot d\mathbf{E} - \sigma\, d\tau,$$

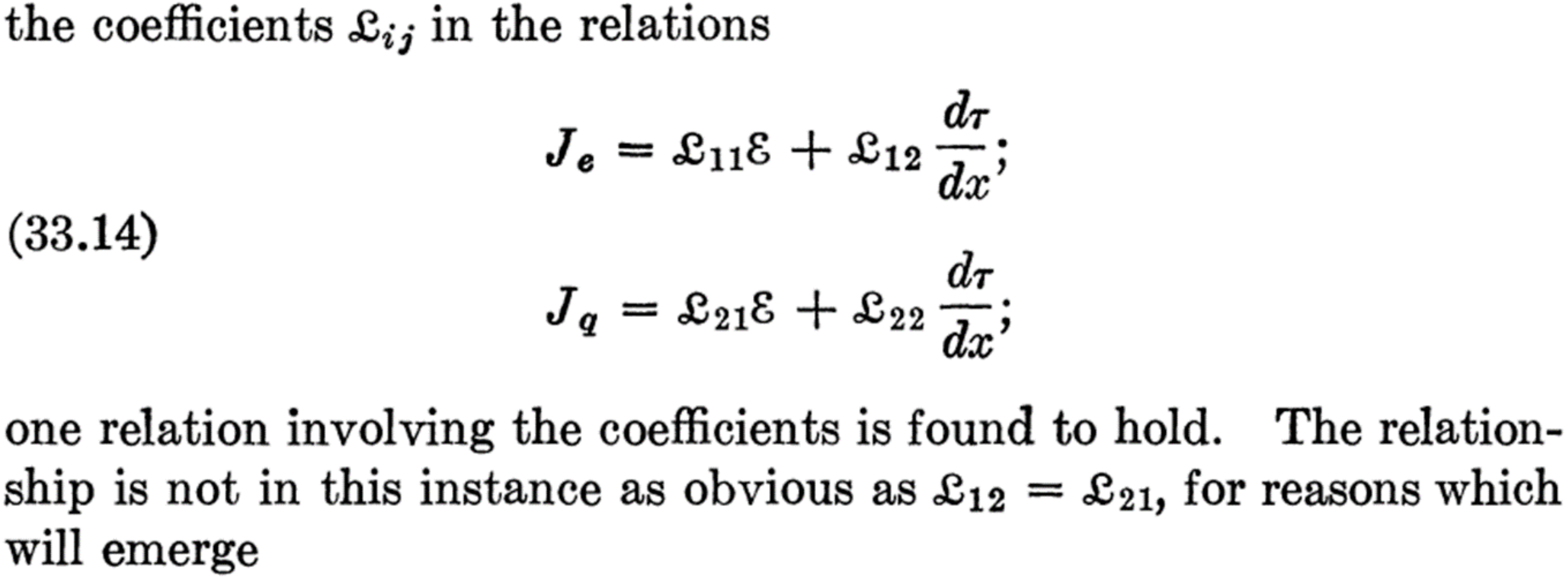
$$P_i = \left(\frac{\partial F_A}{\partial E_i}\right)_\tau \qquad (i = x, y, z).$$

Consequently

$$\chi_{ij} = \left(\frac{\partial P_i}{\partial E_j}\right)_\tau = \left(\frac{\partial^2 F_A}{\partial E_j\, \partial E_i}\right)_\tau, \tag{33.7}$$

which must be equal to

$$\chi_{ji} = \left(\frac{\partial^2 F_A}{\partial E_i\, \partial E_j}\right)_\tau, \tag{33.8}$$



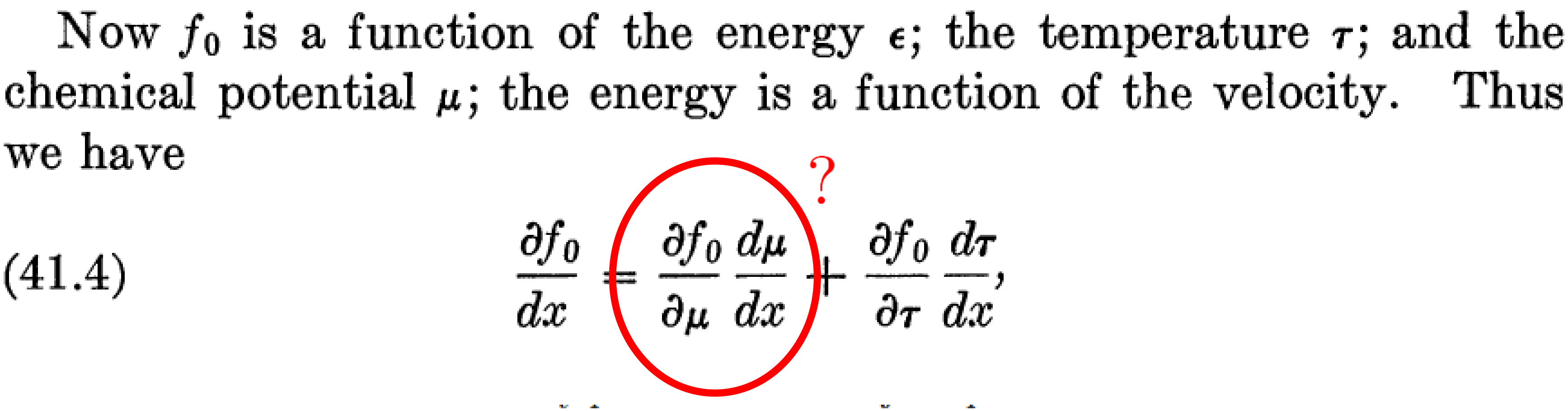
the coefficients $\mathcal{L}_{ij}$ in the relations

$$J_e = \mathcal{L}_{11}\mathcal{E} + \mathcal{L}_{12}\frac{d\tau}{dx};$$

$$J_q = \mathcal{L}_{21}\mathcal{E} + \mathcal{L}_{22}\frac{d\tau}{dx}; \qquad (33.14)$$

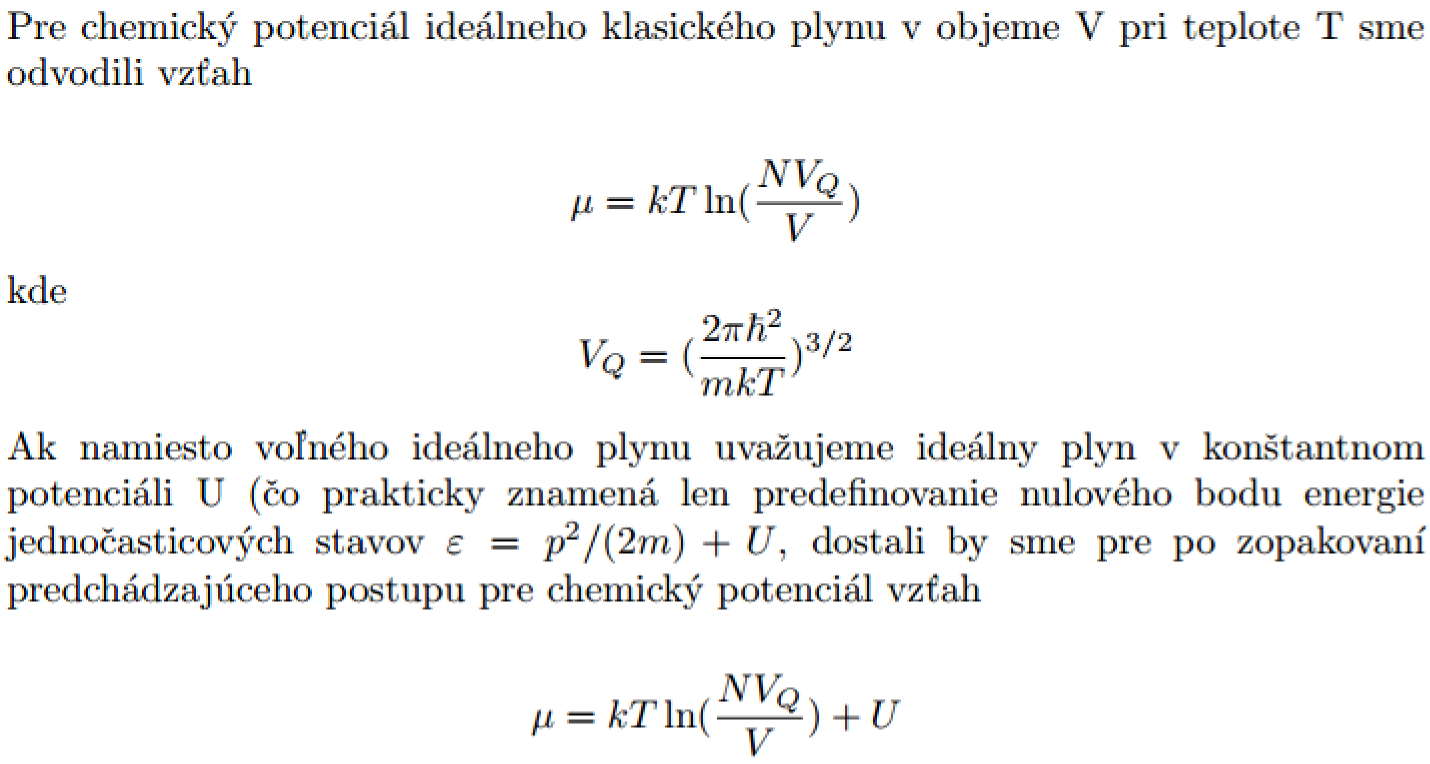
one relation involving the coefficients is found to hold. The relationship is not in this instance as obvious as $\mathcal{L}_{12} = \mathcal{L}_{21}$, for reasons which will emerge

Vsuvka



Now f_0 is a function of the energy ϵ; the temperature τ; and the chemical potential μ; the energy is a function of the velocity. Thus we have

$$\frac{\partial f_0}{dx} = \boxed{\frac{\partial f_0}{\partial \mu}\frac{d\mu}{dx}}^{\,?} + \frac{\partial f_0}{\partial \tau}\frac{d\tau}{dx}, \tag{41.4}$$



Pre chemický potenciál ideálneho klasického plynu v objeme V pri teplote T sme odvodili vzťah

$$\mu = kT\ln(\frac{NV_Q}{V})$$

kde

$$V_Q = (\frac{2\pi\hbar^2}{mkT})^{3/2}$$

Ak namiesto voľného ideálneho plynu uvažujeme ideálny plyn v konštantnom potenciáli U (čo prakticky znamená len predefinovanie nulového bodu energie jednočasticových stavov $\varepsilon = p^2/(2m) + U$, dostali by sme pre po zopakovaní predchádzajúceho postupu pre chemický potenciál vzťah

$$\mu = kT\ln(\frac{NV_Q}{V}) + U$$

Klasická limita

Všimnime si teraz klasický prípad, t.j. situáciu keď stredná obsadenosť všetkých jednočasticových hladín je oveľa menšia ako jedna, teda keď

$$\overline{n} = \frac{1}{\exp(\frac{\varepsilon-\mu}{kT}) \pm 1} \ll 1$$

Takáto situácia nastane nezávisle od toho, či uvažujeme fermióny alebo bozóny vtedy, keď platí

$$\overline{n}(\varepsilon) \approx \exp(\frac{\mu - \varepsilon}{kT}) \ll 1$$

Predpokladajme teda takúto situáciu, uvažujeme klasický ideálny plyn. Našim cieľom je nájsť vyjadrenie pre chemický potenciál ideálneho klasického plynu.

Stredný počet častíc v celom uvažovanom objeme plynu dostaneme sumovaním strednej obsadenosti cez všetky jednočasticové stavy. Bude teda

$$\overline{N} = \sum_i \exp(\frac{\mu - \varepsilon_i}{kT})$$

Sumovanie cez jednočasticové stavy môžeme nahradiť sumovaním cez jednočasticové energie

$$\overline{N} = \sum_{\varepsilon} \Omega(\varepsilon) \exp(\frac{\mu - \varepsilon_i}{kT})$$

Sumovanie teraz možno približne nahradiť integrálom cez spojité hodnoty energie

$$\overline{N} = \int \mathrm{d}\varepsilon \varphi'(\varepsilon) \exp(\frac{\mu - \varepsilon_i}{kT})$$

kde $\varphi(\varepsilon)$ udáva počet jednočasticových stavov s energiou menšou ako ε.

Podľa modelu "častica v jame" dostaneme pre častice so spinom 0

$$\varphi(\varepsilon) = \frac{1}{6}\pi \left(\frac{2mL^2}{\pi^2\hbar^2}\right)^{3/2} \varepsilon^{3/2}$$

Poznamenajme, že hoci uvažujeme klasickú limitu, pri ktorej sa stiera rozdiel medzi fermiónmi a bozónmi, fakt degenerácie hladín vzhľadom na spinové stavy prežije i klasickú limitu. Pre častice s rôznym spinom sa preto i v klasickej limite dostanú rôzne výsledky líšiace sa však len faktorom degenerácie vo výraze pre počet stavov. Preto budeme uvažovať najjednoduchší prípad bez spinovej degenerácie. Po dosadení a preintegrovaní dostaneme

$$\overline{N} = \frac{V}{V_Q} \exp(\frac{\mu}{kT})$$

kde sme označili

$$V_Q = \left(\frac{2\pi\hbar^2}{mkT}\right)^{3/2}$$

$$\overline{N} = \frac{V}{V_Q}\exp(\frac{\mu}{kT})$$

$$V_Q = \left(\frac{2\pi\hbar^2}{mkT}\right)^{3/2}$$

$$\exp(\frac{\mu}{kT}) = N\frac{V_Q}{V}$$

$$\mu = kT\ln N\frac{V_Q}{V}$$

Chemický potenciál vo vonkajšom poli

Pre chemický potenciál ideálneho klasického plynu v objeme V pri teplote T sme odvodili vzťah

$$\mu = kT\ln(\frac{NV_Q}{V})$$

kde

$$V_Q = (\frac{2\pi\hbar^2}{mkT})^{3/2}$$

Ak namiesto voľného ideálneho plynu uvažujeme ideálny plyn v konštantnom potenciáli U (čo prakticky znamená len predefinovanie nulového bodu energie jednočasticových stavov $\varepsilon = p^2/(2m) + U$, dostali by sme pre po zopakovaní predchádzajúceho postupu pre chemický potenciál vzťah

$$\mu = kT\ln(\frac{NV_Q}{V}) + U$$

Zmena energie o konštantnú hodnotu nemá fyzikálne dôsledky, no predchádzajúci vzťah môžeme veľmi užitočne využiť, ak uvažujeme rôzne systémy v kontakte s výmenou častíc. Predstavme si dva ideálne klasické plyny zložené z rovnakých častíc, z ktorých jeden je vo vonkajšom poli U_1 a druhý vo vonkajšom poli U_2. Ak sú tie dva plyny v rovnováhe, pričom si môžu vymieňať teplo i častice, vieme, že rovnováha sa dosiahne, ak v obidvoch podsystémoch bude rovnaký chemický potenciál. Musí teda platiť

$$\mu_1 = \mu_2$$

$$kT\ln(\frac{N_1 V_Q}{V_1}) + U_1 = kT\ln(\frac{N_2 V_Q}{V_2}) + U_2$$

keď sme už využili fakt, že systémy v tepelnom kontakte v rovnováhe musia mať rovnakú teplotu. Pre objemové hustoty častíc $n_i = N_i/V_i$ teda dostaneme

$$\frac{n_1}{n_2} = \exp(-\frac{U_1 - U_2}{kT})$$

Toto je vzťah, ktorý už poznáme ako barometrickú formulu a odvodili sme ho z kánonického rozdelenia. Tento alternatívny postup cez chemický potenciál v niektorých situáciách je transparentnejší.

Electrical and Thermal Conductivity in an Electron Gas

We consider a specimen with an electric field $\mathcal{E}$ in the x direction and a temperature gradient $d\tau/dx$. Our program is to solve approximately the Boltzmann equation for the distribution function and then to find the flux of electric charge and of energy. We restrict ourselves to the steady state (dc conditions), so that $\partial f/\partial t = 0$. Then the transport equation (40.10) becomes

$$\frac{e\mathcal{E}}{m}\frac{\partial f}{\partial u} + u\,\frac{\partial f}{\partial x} = -\,\frac{f - f_0}{\tau_c}, \tag{41.1}$$

where u is the x component of the velocity, and e is the charge on the electron. Rewriting (41.1) we have

$$f = f_0 - \tau_c\left(\frac{e\mathcal{E}}{m}\frac{\partial f}{\partial u} + u\,\frac{\partial f}{\partial x}\right). \tag{41.2}$$

$$f = f_0 - \tau_c\left(\frac{e\mathcal{E}}{m}\frac{\partial f_0}{\partial u} + u\,\frac{\partial f_0}{\partial x}\right). \tag{41.3}$$

$$f = f_0 - \tau_c \left(\frac{e\mathcal{E}}{m} \frac{\partial f_0}{\partial u} + u \frac{\partial f_0}{\partial x} \right). \tag{41.3}$$

Now f_0 is a function of the energy ϵ; the temperature τ; and the chemical potential μ; the energy is a function of the velocity. Thus we have

$$\frac{\partial f_0}{dx} = \frac{\partial f_0}{\partial \mu} \frac{d\mu}{dx} + \frac{\partial f_0}{\partial \tau} \frac{d\tau}{dx}, \tag{41.4}$$

$$\frac{\partial f}{\partial u} = \frac{\partial f_0}{\partial \epsilon} \frac{d\epsilon}{du} = mu \frac{\partial f_0}{\partial \epsilon}. \tag{41.5}$$

The electrical conductivity is usually defined under the conditions $d\tau/dx = 0$ and $dN/dx = 0$, where N is the carrier concentration. Then $\partial f_0/\partial x = 0$, and (41.3) reduces to

$$f = f_0 - \tau_c e\mathcal{E}u \, \partial f_0/\partial \epsilon. \tag{41.6}$$

The electric current density is given by

$$J_e = \int euf \, d\mathbf{v} = -\tau_c e^2 \mathcal{E} \int u^2 (\partial f_0/\partial \epsilon) \, d\mathbf{v}, \tag{41.7}$$

The electric current density is given by

$$J_e = \int euf\,d\mathbf{v} = -\tau_c e^2 \mathcal{E} \int u^2(\partial f_0/\partial \epsilon)\,d\mathbf{v}, \tag{41.7}$$

For the Maxwellian distribution

$$f_0 = N\left(\frac{m}{2\pi\tau}\right)^{3/2} e^{-mv^2/2\tau}, \tag{41.8}$$

where v is the magnitude of the velocity:

$$v^2 = v_x{}^2 + v_y{}^2 + v_z{}^2. \tag{41.9}$$

We note that for the Maxwellian distribution

$$\frac{\partial f_0}{\partial \epsilon} = -\frac{1}{\tau} f_0, \tag{41.10}$$

so that, from (41.7),

$$J_e = \frac{\tau_c e^2 \mathcal{E}}{\tau} \int u^2 f_0\,d\mathbf{v}.$$

$$J_e = \frac{\tau_c e^2 \mathcal{E}}{\tau} \int u^2 f_0 \, d\mathbf{v}.$$

But

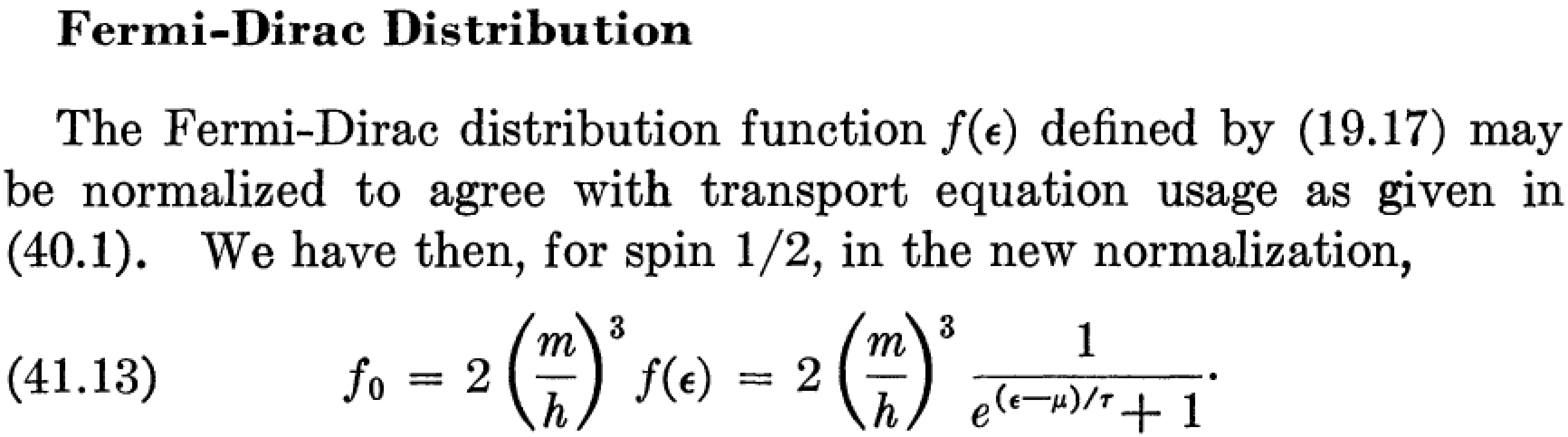
$$\tfrac{1}{2} m \int u^2 f_0 d\mathbf{v} = \tfrac{1}{2} N \tau,$$

and thus

(41.11) $$J_e = \frac{N e^2 \tau_c}{m} \mathcal{E}$$

and

(41.12) $$\sigma = N e^2 \tau_c / m.$$



Fermi-Dirac Distribution

The Fermi-Dirac distribution function $f(\epsilon)$ defined by (19.17) may be normalized to agree with transport equation usage as given in (40.1). We have then, for spin 1/2, in the new normalization,

$$f_0 = 2\left(\frac{m}{h}\right)^3 f(\epsilon) = 2\left(\frac{m}{h}\right)^3 \frac{1}{e^{(\epsilon-\mu)/\tau}+1}. \qquad (41.13)$$

$$\varepsilon = \frac{\pi^2\hbar^2}{2mL^2}(n_1^2+n_2^2+n_3^2)$$

$$\varphi(\varepsilon) = 2\frac{1}{8}\frac{4}{3}\pi r_n^3 = 2\frac{1}{8}\frac{4}{3}\pi(\frac{2m}{\pi^2\hbar^2})^{3/2}L^3\varepsilon^{3/2}$$

$$n(\varepsilon,\varepsilon+d\varepsilon)d\varepsilon = \frac{1}{L^3}f_{FD}(\varepsilon)\varphi'(\varepsilon)d\varepsilon = f_{FD}(\varepsilon)\frac{1}{3}\pi(\frac{2m}{\pi^2\hbar^2})^{3/2}\frac{3}{2}\varepsilon^{1/2}d\varepsilon$$

$$n(\varepsilon,\varepsilon+d\varepsilon)d\varepsilon = f_{FD}(\varepsilon)\frac{1}{2}\pi(\frac{2m}{\pi^2\hbar^2})^{3/2}(\frac{1}{2}mv^2)^{1/2}mvdv$$

$$n(\varepsilon,\varepsilon+d\varepsilon)d\varepsilon = f_{FD}(\varepsilon)\frac{m^3}{\pi^3\hbar^3}\frac{1}{4}4\pi v^2dv$$

$$\boxed{f_0(\varepsilon) = 2\frac{m^3}{h^3}f_{FD}(\varepsilon)}$$

The electric current density is given by

(41.7) $$J_e = \int euf\, d\mathbf{v} = -\tau_c e^2 \mathcal{E} \int u^2 (\partial f_0/\partial\epsilon)\, d\mathbf{v},$$

Now for $\tau \ll \mu$ the function

(41.14) $$\frac{\partial}{\partial\epsilon} \frac{1}{e^{(\epsilon-\mu)\tau/} + 1} \cong -\delta(\epsilon - \mu)$$

(41.15) $$J_e = -2e^2\mathcal{E} \left(\frac{m}{h}\right)^3 \int \tau_c u^2 \frac{\partial}{\partial\epsilon} \frac{1}{e^{(\epsilon-\mu)/\tau} + 1}\, d\mathbf{v}.$$

Now

(41.16) $$u^2\, d\mathbf{v} = \frac{4\pi}{3} v^4\, dv = \frac{4\pi}{3} \left(\frac{2\epsilon}{m}\right)^{3/2} \frac{1}{m}\, d\epsilon$$

$$= \frac{8\pi}{3} \sqrt{2}\, m^{-5/2} \epsilon^{3/2}\, d\epsilon,$$

$$J_e = e^2\, \mathcal{E}\, \tau_c(\mu) \mu^{3/2} (16\pi/3) \sqrt{2}\, m^{-5/2} (m/h)^3, \tag{41.17}$$

$$\mu_0 = (h^2/8\pi^2 m)(3\pi^2 N)^{2/3}$$

so that (at $\tau \ll \mu$)

$$\sigma = J_e/\mathcal{E} = Ne^2\tau_c/m, \tag{41.18}$$

which is identical with the result (41.12) for the Maxwellian distribution, except that now we do not have to assume τ_c independent of velocity; it is only the value of τ_c at the Fermi surface that enters the conductivity.

Thermal Conductivity, Maxwellian Distribution

We now carry through in the Maxwellian limit the problem of determining the coefficients $\mathfrak{L}_{ij}$ in the general relations (34.1):

$$J_e = \mathfrak{L}_{11}\mathcal{E} + \mathfrak{L}_{12}\frac{d\tau}{dx};$$

(41.19)

$$J_Q = \mathfrak{L}_{21}\mathcal{E} + \mathfrak{L}_{22}\frac{d\tau}{dx}.$$

We bear in mind that $\mathcal{E} = -\partial\varphi/\partial x$. We assume for simplicity that the relaxation time τ_c is independent of velocity. We have already evaluated $\mathfrak{L}_{11} = \sigma = Ne^2\tau_c/m$. Next consider the term $\mathfrak{L}_{12}$. From (41.8),

$$f_0 = N\left(\frac{m}{2\pi\tau}\right)^{3/2} e^{-mv^2/2\tau},$$

(41.20)
$$\frac{\partial f_0}{\partial x} = \left(\frac{\epsilon}{\tau} - \frac{3}{2}\right) f_0 \frac{1}{\tau}\frac{d\tau}{dx},$$

$$J_e = \int euf \, d\mathbf{v}$$

$$J_Q = \int u\epsilon f \, d\mathbf{v}$$

$$f = f_0 - \tau_c \left(\frac{e\mathcal{E}}{m} \frac{\partial f_0}{\partial u} + u \frac{\partial f_0}{\partial x} \right).$$

$$\frac{\partial f}{\partial u} = \frac{\partial f_0}{\partial \epsilon} \frac{d\epsilon}{du} = mu \frac{\partial f_0}{\partial \epsilon}.$$

$$\frac{\partial f_0}{\partial \epsilon} = -\frac{1}{\tau} f_0,$$

$$\frac{\partial f_0}{\partial x} = \left(\frac{\epsilon}{\tau} - \frac{3}{2} \right) f_0 \frac{1}{\tau} \frac{d\tau}{dx},$$

$$J_Q = \int u\epsilon f \, d\mathbf{v} = \frac{5Ne\tau_c\tau\mathcal{E}}{2m} - \frac{5N\tau_c\tau}{m} \frac{d\tau}{dx}.$$

the coefficients $\mathfrak{L}_{ij}$ in the relations

$$(33.14)\qquad \begin{aligned} J_e &= \mathfrak{L}_{11}\mathcal{E} + \mathfrak{L}_{12}\frac{d\tau}{dx}; \\ J_q &= \mathfrak{L}_{21}\mathcal{E} + \mathfrak{L}_{22}\frac{d\tau}{dx}; \end{aligned}$$

$$\mathfrak{L}_{11} = \sigma = \frac{Ne^2\tau_c}{m};$$

$$\mathfrak{L}_{12} = -\frac{Ne\tau_c}{m};$$

$$\mathfrak{L}_{21} = \frac{5Ne\tau_c\tau}{2m};$$

$$\mathfrak{L}_{22} = -\frac{5N\tau_c\tau}{m}.$$

A basic theory which explains why reciprocal relations are found in irreversible processes has been given by Onsager. The technique is to look at the decay of a fluctuation. We consider an adiabatically insulated system, represented by a microcanonical ensemble. We look at it at an instant of time at which, by malice or by chance, the ensemble is in the throes of a fluctuation from statistical average behavior. Let $\alpha_1, \alpha_2, \cdots, \alpha_n$ denote the deviation of appropriate physical parameters from their equilibrium values. Then

$$\Delta\sigma = -\tfrac{1}{2} \Sigma\, g_{ik}\alpha_i\alpha_k, \tag{33.15}$$

where g_{ik} is a positive definite form so that $\Delta\sigma$ in a fluctuation from equilibrium is never positive. Terms linear in the α's are excluded because the entropy is a maximum in the equilibrium state.

Recalling the connection $\sigma = \log \Delta\Gamma$, the probability of finding the system with parameters in $d\boldsymbol{\alpha}$ at $\boldsymbol{\alpha}$ is, according to Sec. 26,

$$P(\boldsymbol{\alpha})\, d\boldsymbol{\alpha} = \frac{e^\sigma\, d\boldsymbol{\alpha}}{\int e^\sigma\, d\boldsymbol{\alpha}}. \tag{33.16}$$

(33.15) $$\Delta\sigma = -\tfrac{1}{2}\sum g_{ik}\alpha_i\alpha_k,$$

Now introduce the quantity

(33.17) $$X_i = \frac{\partial\sigma}{\partial\alpha_i} = -\sum g_{ik}\alpha_k,$$

which acts as a driving force for an irreversible or transport process. We shall call X_i the *generalized force*. If g^{ik} is the reciprocal matrix to g_{ik}, we have

(33.18) $$\alpha_i = -\sum g^{ik}X_k.$$

$$\boxed{P(\boldsymbol{\alpha})\,d\boldsymbol{\alpha} = \frac{e^{\sigma}\,d\boldsymbol{\alpha}}{\int e^{\sigma}\,d\boldsymbol{\alpha}}}$$

Now consider the ensemble average

(33.19) $$\overline{\alpha_i X_j} = \int \alpha_i X_j P\,d\boldsymbol{\alpha} = \int \alpha_i \frac{\partial P}{\partial\alpha_j}\,d\boldsymbol{\alpha},$$

using (33.16) and (33.17). On integrating by parts,

(33.20) $$\overline{\alpha_i X_j} = -\int P\delta_{ij}\,d\boldsymbol{\alpha} = -\delta_{ij}.$$

Further, using this result,

(33.21) $$\overline{X_i X_j} = -\sum_k g_{ik}\overline{\alpha_k X_j} = g_{ij},$$

$$\overline{\alpha_i\alpha_j} = -\sum_k g^{ik}\overline{X_k\alpha_j} = g^{ij}. \tag{33.22}$$

Because of the invariance of the equations of motion under time reversal, the correlation function must have the property

$$\overline{\alpha_i(t)\,\alpha_j(t+\tau)} = \overline{\alpha_i(t)\,\alpha_j(t-\tau)}. \tag{33.23}$$

Here and below in this section the averages are to be taken over a subensemble of all systems with the given values of the α's at time t. If magnetic fields or Coriolis forces are present, it is understood that the sign of such fields or forces is to be reversed between the two sides of this equation. This preserves the invariance of the equations of motion. An alternate statement of (33.23) is that

$$\overline{\alpha_i(t)\,\alpha_j(t+\tau)} = \overline{\alpha_i(t+\tau)\alpha_j(t)}, \tag{33.24}$$

so that

$$\overline{\alpha_i(t)[\alpha_j(t+\tau) - \alpha_j(t)]} = \overline{\alpha_j(t)[\alpha_i(t+\tau) - \alpha_i(t)]}. \tag{33.25}$$

(33.26) $$\overline{d\alpha_i/dt} = \Sigma\, C_{ij}\alpha_j,$$

or

(33.27) $$J_i = \overline{d\alpha_i/dt} = \Sigma\, L_{ij}X_j;$$

here the time derivative is defined as

(33.28) $$\overline{d\alpha_i/dt} = \tau^{-1}\overline{[\alpha_i(t+\tau) - \alpha_i(t)]}.$$

Here the time τ is to be taken as longer than the collision time for an individual process but shorter than the time for the decay of a macroscopic fluctuation. Combining (33.17) and (33.27), we have

(33.29) $$d\sigma/dt = \Sigma\, X_iJ_i.$$

(33.17) $$X_i = \frac{\partial\sigma}{\partial\alpha_i}$$

$$\overline{\alpha_i(t)[\alpha_j(t+\tau)-\alpha_j(t)]} = \overline{\alpha_j(t)[\alpha_i(t+\tau)-\alpha_i(t)]} \tag{33.25}$$

$$J_i = \overline{d\alpha_i/dt} = \Sigma\, L_{ij}X_j; \tag{33.27}$$

here the time derivative is defined as

$$\overline{d\alpha_i/dt} = \tau^{-1}\overline{[\alpha_i(t+\tau)-\alpha_i(t)]}. \tag{33.28}$$

Here the time τ is to be taken as longer than the collision time for an individual process but shorter than the time for the decay of a macroscopic fluctuation. Combining (33.17) and (33.27), we have

$$d\sigma/dt = \Sigma\, X_iJ_i. \tag{33.29}$$

Now from (33.25) and (33.27),

$$\overline{\alpha_i(t)\sum_k L_{jk}X_k} = \overline{\alpha_j(t)\sum_k L_{ik}X_k}. \tag{33.30}$$

$$\overline{\alpha_i X_j} = -\int P\delta_{ij}\, d\boldsymbol{\alpha} = -\delta_{ij}. \tag{33.20}$$

$$\overline{\alpha_i(t) \sum_k L_{jk} X_k} = \overline{\alpha_j(t) \sum_k L_{ik} X_k}. \tag{33.30}$$

Using (33.20),

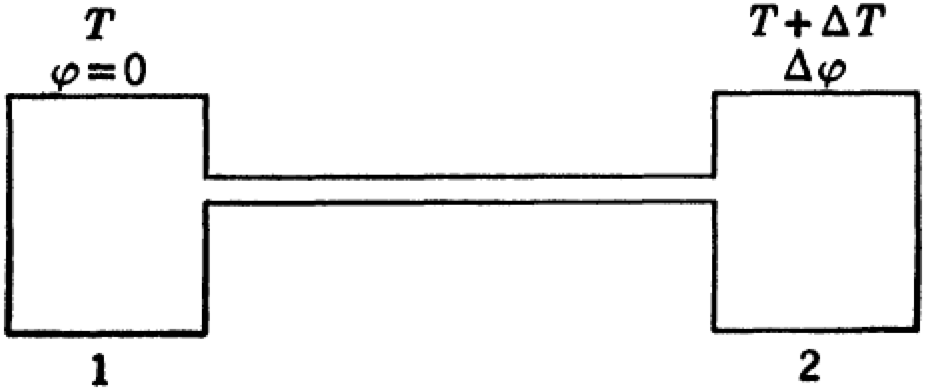
$$-\sum_k L_{jk}\delta_{ki} = -\sum_k L_{ik}\delta_{kj},$$

or

$$L_{ji} = L_{ij}. \tag{33.31}$$

This is the Onsager relation. We note that it obtains *only* when the forces X_i are defined by (33.17) and the fluxes J_j by (33.26) and (33.27). It is rarely a trivial problem to find the correct choice of forces and fluxes applicable to the Onsager relation. In the next section we discuss the forces and fluxes which describe charge and energy transport in a homogeneous conductor.

Application of the Onsager Relations to Charge and Energy Transport in a Homogeneous Conductor

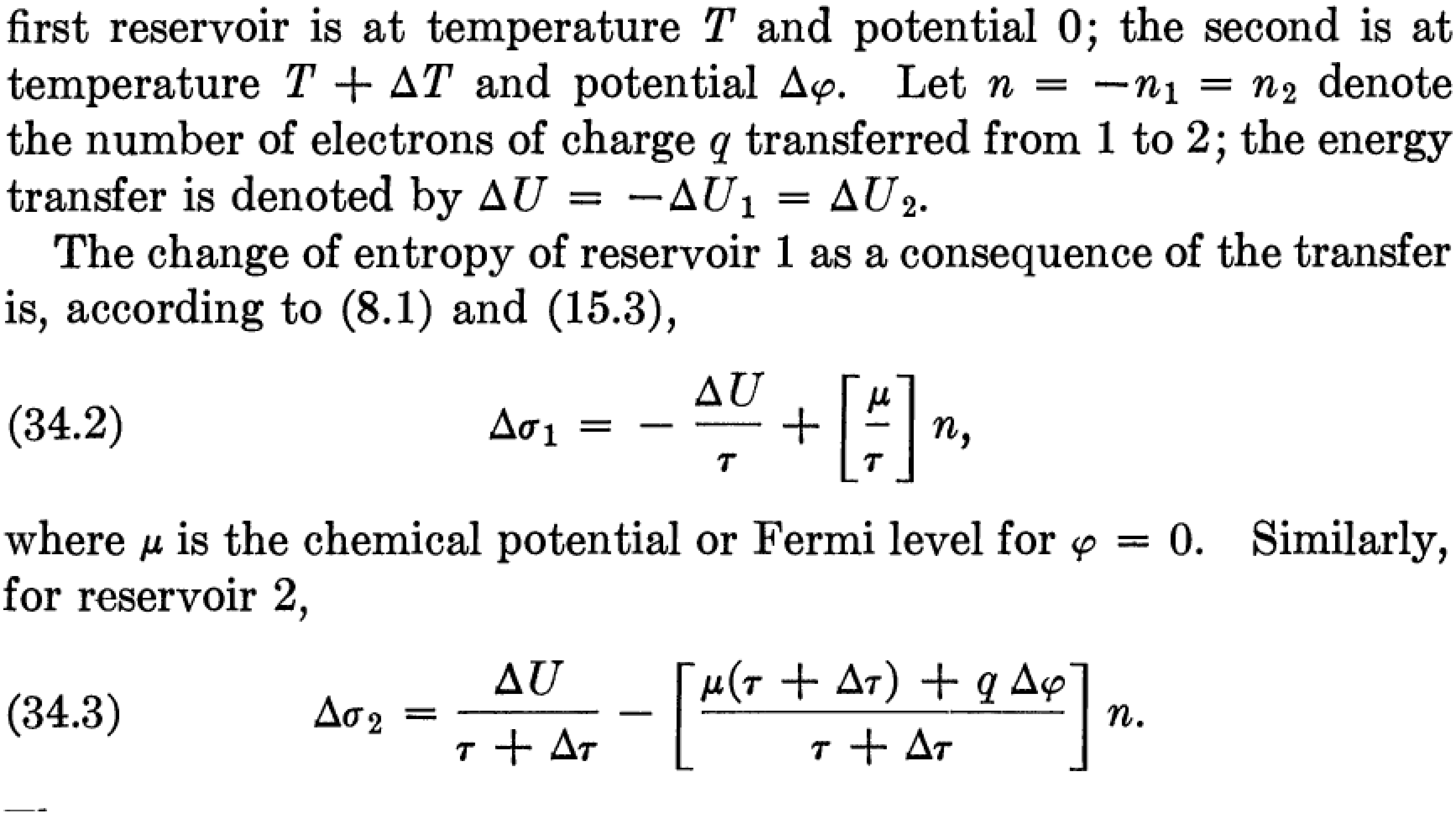


$$I = l_{11}\,\Delta\varphi + l_{12}\,\Delta T$$

$$W = l_{21}\,\Delta\varphi + l_{22}\,\Delta T$$

$$J_e = \mathfrak{L}_{11}\mathcal{E} + \mathfrak{L}_{12}\,\frac{d\tau}{dx}$$

$$J_q = \mathfrak{L}_{21}\mathcal{E} + \mathfrak{L}_{22}\,\frac{d\tau}{dx}$$



first reservoir is at temperature T and potential 0; the second is at temperature $T + \Delta T$ and potential $\Delta\varphi$. Let $n = -n_1 = n_2$ denote the number of electrons of charge q transferred from 1 to 2; the energy transfer is denoted by $\Delta U = -\Delta U_1 = \Delta U_2$.

The change of entropy of reservoir 1 as a consequence of the transfer is, according to (8.1) and (15.3),

$$\Delta\sigma_1 = -\frac{\Delta U}{\tau} + \left[\frac{\mu}{\tau}\right] n, \tag{34.2}$$

where μ is the chemical potential or Fermi level for $\varphi = 0$. Similarly, for reservoir 2,

$$\Delta\sigma_2 = \frac{\Delta U}{\tau + \Delta\tau} - \left[\frac{\mu(\tau + \Delta\tau) + q\,\Delta\varphi}{\tau + \Delta\tau}\right] n. \tag{34.3}$$

$$\Delta\sigma = \Delta\sigma_1 + \Delta\sigma_2 = \Delta U \left[\frac{1}{\tau + \Delta\tau} - \frac{1}{\tau} \right] - n \left[\frac{\mu(\tau + \Delta\tau)}{\tau + \Delta\tau} - \frac{\mu}{\tau} + \frac{q\,\Delta\varphi}{\tau + \Delta\tau} \right] \cong \Delta U \left[-\frac{\Delta\tau}{\tau^2} \right] - nq \left[\frac{\Delta\tau}{q} \frac{\partial}{\partial\tau} \left(\frac{\mu}{\tau} \right) + \frac{\Delta\varphi}{\tau} \right]. \tag{34.4}$$

$$\frac{d}{dt}\Delta\sigma = \frac{d(\Delta U)}{dt} \left(-\frac{\Delta\tau}{\tau^2} \right) - q\,\frac{dn}{dt} \left[\frac{\Delta\varphi}{\tau} + \frac{\Delta\tau}{q} \frac{\partial}{\partial\tau} \left(\frac{\mu}{\tau} \right) \right] = \sum_i J_i X_i, \tag{34.5}$$

where

$$\text{(34.5)} \qquad \frac{d}{dt}\Delta\sigma = \frac{d(\Delta U)}{dt}\left(-\frac{\Delta\tau}{\tau^2}\right) - q\frac{dn}{dt}\left[\frac{\Delta\varphi}{\tau} + \frac{\Delta\tau}{q}\frac{\partial}{\partial\tau}\left(\frac{\mu}{\tau}\right)\right]$$

$$= \sum_i J_i X_i,$$

where

$$\text{(34.6)} \quad J_1 = \frac{d(\Delta U)}{dt} = \text{Energy (thermal) current} = W;$$

$$\text{(34.7)} \quad X_1 = \left(-\frac{\Delta\tau}{\tau^2}\right) = \text{Generalized force for energy current;}$$

$$\text{(34.8)} \quad J_2 = q\frac{dn}{dt} = \text{Electric current} = I;$$

$$\text{(34.9)} \quad X_2 = -\frac{\Delta\varphi}{\tau} - \frac{\Delta\tau}{q}\frac{\partial}{\partial\tau}\left(\frac{\mu}{\tau}\right) = \text{Generalized force for electric current.}$$

Using this definition of currents and forces we can write the currents J_i as $J_i = \sum_j L_{ij} X_j$, where the L_{ij} satisfy the Onsager relation $L_{12} = L_{21}$:

$$I = L_{11}\left[-\frac{\Delta\varphi}{\tau} - \frac{\Delta\tau}{q}\frac{\partial}{\partial\tau}\left(\frac{\mu}{\tau}\right)\right] + L_{12}\left(-\frac{\Delta\tau}{\tau^2}\right); \tag{34.10}$$

$$W = L_{21}\left[-\frac{\Delta\varphi}{\tau} - \frac{\Delta\tau}{q}\frac{\partial}{\partial\tau}\left(\frac{\mu}{\tau}\right)\right] + L_{22}\left(-\frac{\Delta\tau}{\tau^2}\right). \tag{34.11}$$

$$J_e = \mathfrak{L}_{11}\mathcal{E} + \mathfrak{L}_{12}\frac{d\tau}{dx};$$

$$J_q = \mathfrak{L}_{21}\mathcal{E} + \mathfrak{L}_{22}\frac{d\tau}{dx}$$

$$\frac{L_{21}}{\tau} = \mathfrak{L}_{21}$$

$$-\frac{L_{21}}{e}\frac{\partial}{\partial\tau}\left(\frac{\mu}{\tau}\right) - \frac{L_{22}}{\tau^2} = \mathfrak{L}_{22}$$

$$\frac{L_{11}}{\tau} = \mathfrak{L}_{11};$$

$$-\frac{L_{11}}{e}\frac{\partial}{\partial\tau}\left(\frac{\mu}{\tau}\right) - \frac{L_{12}}{\tau^2} = \mathfrak{L}_{12}$$

$$\frac{L_{21}}{\tau} = \mathfrak{L}_{21}$$

$$-\frac{L_{21}}{e}\frac{\partial}{\partial\tau}\left(\frac{\mu}{\tau}\right) - \frac{L_{22}}{\tau^2} = \mathfrak{L}_{22}$$

$$\frac{L_{11}}{\tau} = \mathfrak{L}_{11}$$

$$-\frac{L_{11}}{e}\frac{\partial}{\partial\tau}\left(\frac{\mu}{\tau}\right) - \frac{L_{12}}{\tau^2} = \mathfrak{L}_{12}$$

Now, if $L_{21} = L_{12}$, we must have from these results

$$-\frac{\tau\mathfrak{L}_{11}}{e}\frac{\partial}{\partial\tau}\left(\frac{\mu}{\tau}\right) - \frac{\mathfrak{L}_{21}}{\tau} = \mathfrak{L}_{12}. \tag{41.30}$$

Now, if $L_{21} = L_{12}$, we must have from these results

$$-\frac{\tau\mathcal{L}_{11}}{e}\frac{\partial}{\partial\tau}\left(\frac{\mu}{\tau}\right) - \frac{\mathcal{L}_{21}}{\tau} = \mathcal{L}_{12}. \tag{41.30}$$

$$\mu = kT\ln\left(\frac{NV_Q}{V}\right)$$

$$V_Q = \left(\frac{2\pi\hbar^2}{mkT}\right)^{3/2}$$

$$\frac{\partial}{\partial\tau}\left(\frac{\mu}{\tau}\right) = -\frac{3}{2\tau}$$

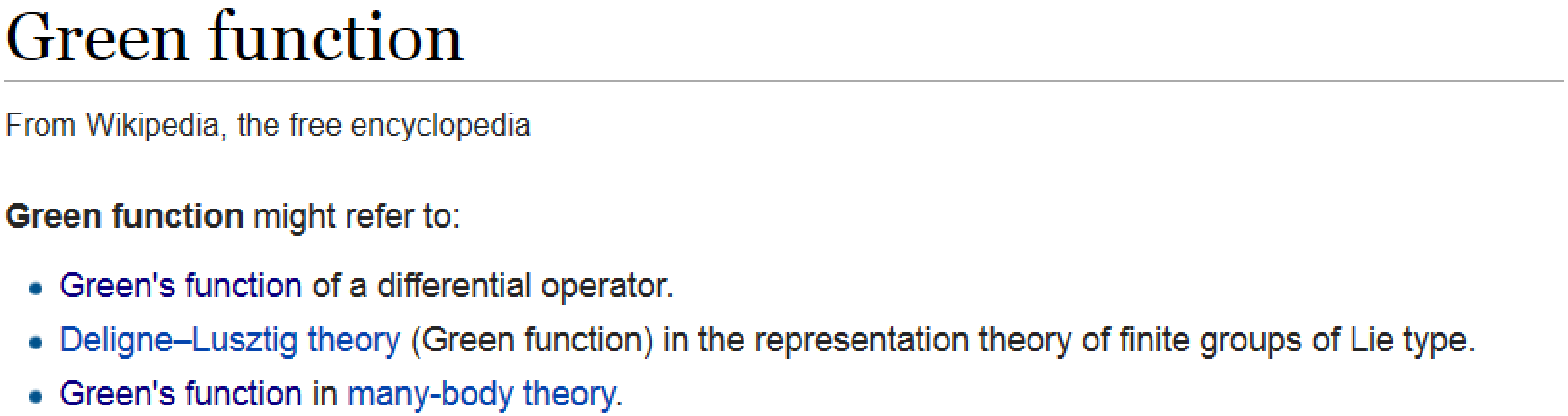
$$\mathcal{L}_{11} = \sigma = \frac{Ne^2\tau_c}{m};$$

$$\mathcal{L}_{12} = -\frac{Ne\tau_c}{m};$$

$$\mathcal{L}_{21} = \frac{5Ne\tau_c\tau}{2m};$$

$$\mathcal{L}_{22} = -\frac{5N\tau_c\tau}{m}.$$

$$\boxed{\frac{3Ne\tau_c}{2m} - \frac{5Ne\tau_c}{2m} = -\frac{Ne\tau_c}{m},}$$

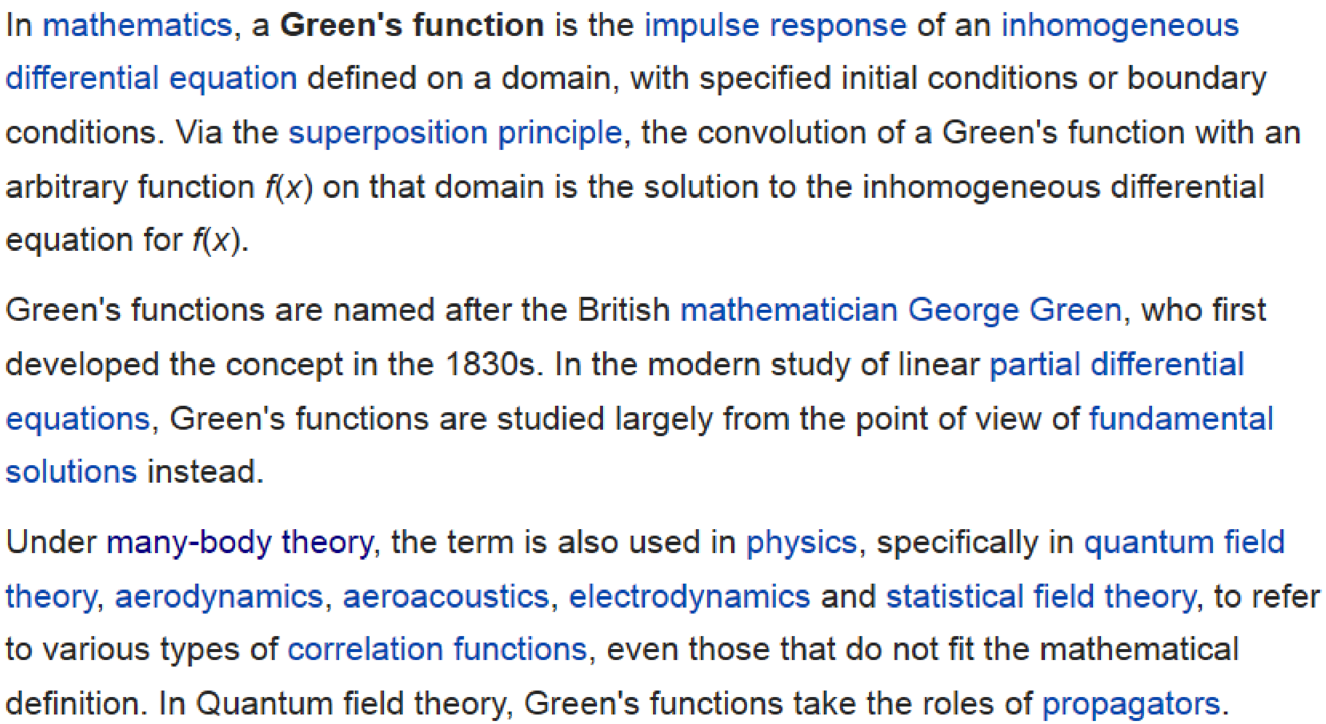


Green function

From Wikipedia, the free encyclopedia

Green function might refer to:

- Green's function of a differential operator.
- Deligne–Lusztig theory (Green function) in the representation theory of finite groups of Lie type.
- Green's function in many-body theory.

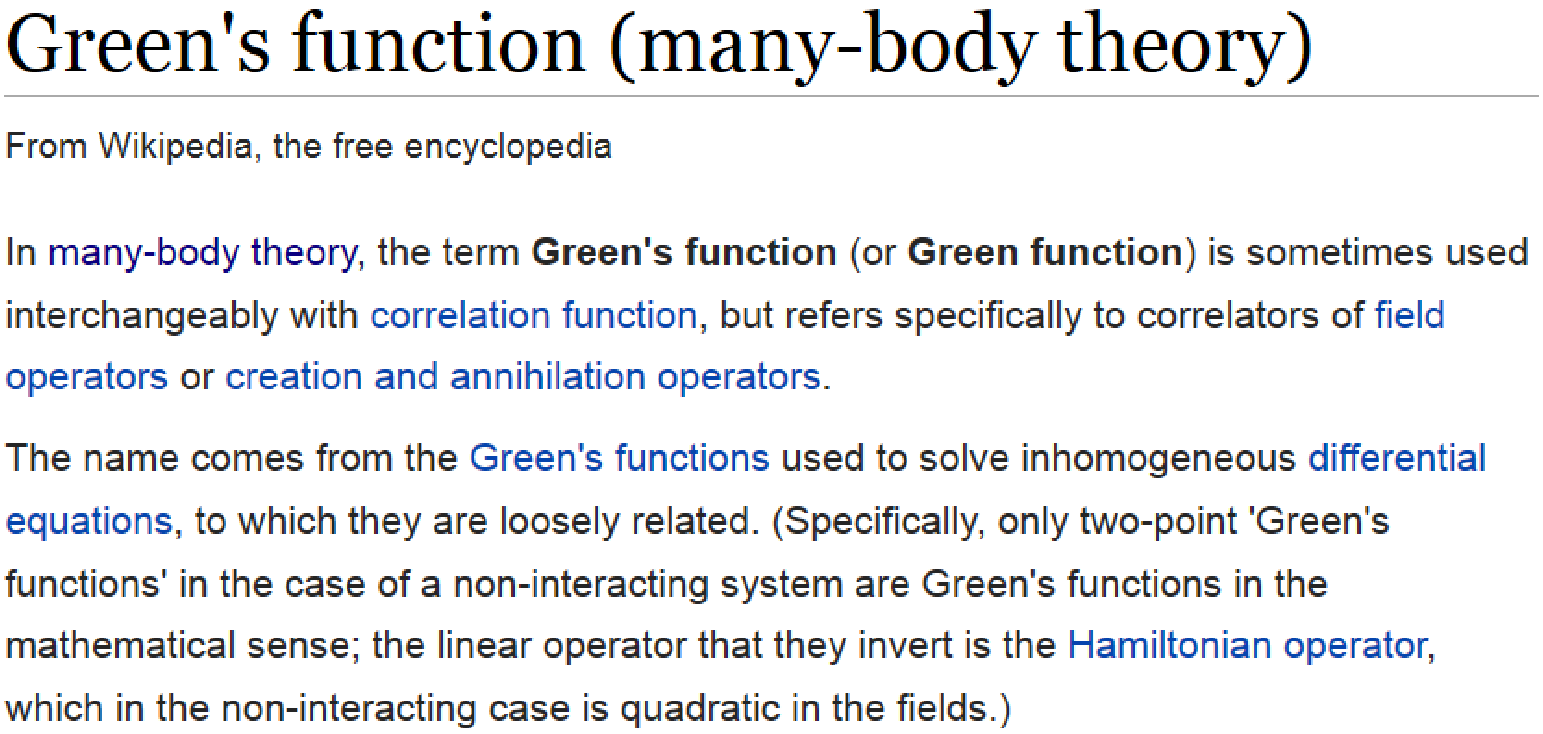


In mathematics, a **Green's function** is the impulse response of an inhomogeneous differential equation defined on a domain, with specified initial conditions or boundary conditions. Via the superposition principle, the convolution of a Green's function with an arbitrary function *f*(*x*) on that domain is the solution to the inhomogeneous differential equation for *f*(*x*).

Green's functions are named after the British mathematician George Green, who first developed the concept in the 1830s. In the modern study of linear partial differential equations, Green's functions are studied largely from the point of view of fundamental solutions instead.

Under many-body theory, the term is also used in physics, specifically in quantum field theory, aerodynamics, aeroacoustics, electrodynamics and statistical field theory, to refer to various types of correlation functions, even those that do not fit the mathematical definition. In Quantum field theory, Green's functions take the roles of propagators.

$$LG(x,s) = \delta(x-s)$$



Green's function (many-body theory)

From Wikipedia, the free encyclopedia

In many-body theory, the term **Green's function** (or **Green function**) is sometimes used interchangeably with correlation function, but refers specifically to correlators of field operators or creation and annihilation operators.

The name comes from the Green's functions used to solve inhomogeneous differential equations, to which they are loosely related. (Specifically, only two-point 'Green's functions' in the case of a non-interacting system are Green's functions in the mathematical sense; the linear operator that they invert is the Hamiltonian operator, which in the non-interacting case is quadratic in the fields.)

The retarded Green's function of the Schrödinger equation $G^R(x_f, x_i; t_f, t_i)$ is defined (up to a factor of $-i$) as a probability amplitude of a particle to hop from the point x_i at a time t_i to the point x_f at a time t_f, where $t_f > t_i$. In other words, the Green's function coincides with a wave function $\psi(x, t)$ of a particle evaluated at a point $x = x_f$ and time $t = t_f$ whose wave function at time t_i was $\psi(x, t_i) = -i\delta(x - x_i)$ (The coefficient $-i$ is introduced because of a standard convention). At the same time, if $t_f < t_i$, the retarded Green's function is identically zero. The regarded Green's function, by construction, must satisfy the Schrödinger equation

$$i\frac{\partial G^R(x, x_i; t, t_i)}{\partial t} = \hat{H}(x, t)\, G(x, x_i; t, t_i), \;\; t > t_i. \tag{1.1}$$

It must also satisfy the boundary condition

$$\lim_{t_f \to t_i + 0} G(x_f, x_i; t_f, t_i) = -i\delta(x_f - x_i). \tag{1.2}$$

It is easy to check that the retarded Green's function has the following expression in terms of the eigenfunctions ψ_n and the eigenvalues E_n of the Hamiltonian, as long as $t_f \geq t_i$ (the assumption here is that the Hamiltonian does not change in time):

$$G^R(x_f, x_i; t_f, t_i) = -i \sum_n \exp\left(-iE_n\left(t_f - t_i\right)\right) \psi_n(x_f)\psi_n^*(x_i). \tag{1.3}$$

Indeed, if $t_f > t_i$, this expression explicitly satisfies the Schrodinger equation. At the same time, if $t_f \to t_i$, then using a well known relation

$$\sum_n \psi_n(x_f)\psi_n^*(x_i) = \delta(x_f - x_i) \tag{1.4}$$

gives us (1.2).

It follows from this definition that, if a particle's wave function is $\psi_i(x)$ at time $t = t_i$, its wave function at $t = t_f > t_i$ is

$$\psi_f(x) = i \int dy\, G^R(x, y; t_f, t_i)\psi_i(y). \tag{1.5}$$

The Fourier transform of the Green's function is very useful. In case of the retarded Green's function, this is defined by

$$G^R(x_f, x_i, E) = \int_0^\infty dt\, e^{iEt} G^R(x_f, x_i, t). \tag{1.6}$$

Substituting (1.3) we observe that we need to compute integrals of the form $\int_0^\infty dt\ e^{iEt}$, and these are not well defined if E is real. However, if E is complex, and if its imaginary part is positive, then this integral is well defined. It is customary to denote the real part of E simply by E, and the imaginary part by ϵ. This gives

$$G^R(x_f, x_i, E) = \sum_n \frac{\psi_n(x_f)\psi_n^*(x_i)}{E + i\epsilon - E_n}. \tag{1.7}$$

It is easy to see that

$$\left(E + i\epsilon - \hat{H}(x)\right) G^R(x, x_i, \omega) = \delta\left(x - x_i\right). \tag{1.8}$$

Formally this means that the Green's function can be understood as an operator

$$\hat{G}^R = \frac{1}{E + i\epsilon - \hat{H}}. \tag{1.9}$$

E.S.Abers, B.W.Lee, Phys.Rep. 9,1 (1973). Je tam urobená základná filozofia integrálov po trajektóriách v QM a prechod do euklidovskej formulácie ako analyt. predĺženie. Odtiaľ to odpísal a trochu okomentoval Amit

D.J.Amit, Field theory, renormalization group and critical phenomena, McGrawHill 1978

Trochu matematickejší je Vasiliev: A.N.Vasiliev, Funkcion. metody v kvantovoj teorii polja, Izd. Leningradskogo Universiteta 1976

B. de Wit, Functional methods in quantum field theory, NIKHEF-H/81-27

Logika je taká: Od Feynmana vieme, že (pre jednoduchosť v jednorozmernom prípade) platí

$$\langle q', t' | q, t \rangle = \int \mathcal{D}q \, \exp\left\{ i \int_t^{t'} \mathcal{L}(q, \dot{q}) \, d\tau \right\} \qquad (*)$$

Teda, že nejaký maticový element prechodu sa ráta cez funkcionálny integrál. Táto vec potrebuje komentár. $|q, t\rangle$ je Heisenbergovský stav, ktorý označuje stav s touto vlastnosťou: keď urobíme v čase t meranie súradnice na tomto stave, dostaneme ostrú hodnotu q.

Treba si uvedomiť, že stav $|q, t\rangle$ ako vektor v Hilbertovom priestore nezávisí na čase. Symbol t nie je časová závislosť vektora ale časť mena toho stavu (časť hieroglyphu).

Časová závislosť je potom schovaná v časovej závislosti operátora súradnice $\hat{q}(\tau)$. (t je fixná hodnota, τ je premenná, označujúca čas, v ktorom sa systém vyvíja.). Platí teda

$$\hat{q}(\tau) |q, t\rangle \Big|_{\tau = t} = q \, |q, t\rangle$$

Všimnime si ešte maticové elementy typu

$$\langle q',t'|T[\hat{q}(t_1)\cdot\hat{q}(t_2)]|q,t\rangle \qquad \text{pre} \quad t<t_1, t_2<t'$$

Pre $\boxed{\underline{\underline{t_1>t_2}}}$ vsunieme na vhodné miesta úplné systémy stavov a dostaneme.

$$\langle q',t'|T[\hat{q}(t_1)\cdot\hat{q}(t_2)]|q,t\rangle = \langle q',t'|\hat{q}(t_1)\,\hat{q}(t_2)|q,t\rangle =$$

$$= \int dq_1\,dq_2\,dq_3\,dq_4\;\langle q',t'|q_1,t_1\rangle\langle q_1,t_1|\hat{q}(t_1)|q_2,t_1\rangle\langle q_2,t_1|q_3,t_2\rangle\cdot$$

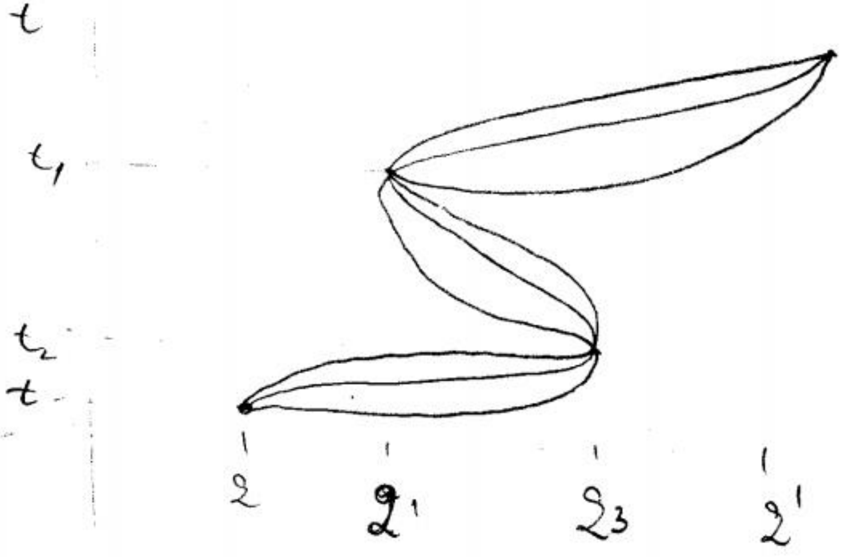
$$\cdot\langle q_3,t_2|\hat{q}(t_2)|q_4,t_2\rangle\langle q_4,t_2|q,t\rangle =$$

$$= \int dq_1\,dq_2\,dq_3\,dq_4\;\langle q',t'|q_1,t_1\rangle\, q_2\cdot\delta(q_1-q_2)\langle q_2,t_1|q_3,t_2\rangle\cdot$$

$$\cdot\, q_4\,\delta(q_3-q_4)\langle q_4,t_2|q,t\rangle =$$

$$= \int dq_1 \, dq_3 \, \langle q', t' | q_1, t_1 \rangle \, q_1 \, \langle q_1, t_1 | q_3, t_2 \rangle \, q_3 \, \langle q_3, t_2 | q, t \rangle =$$

Teraz by bolo treba dosadiť vyjadrenia tých maticových elementov cez kontinuálne integrály. Ukázalo by sa, že súčin tých integrálov je to isté ako jeden kontinuálny integrál cez trajektórie z q do q', ale len cez také trajektórie, ktoré v čase t_2 idú cez bod q_3 a v čase t_1 cez bod q_1 .



$$= \int dq_1\, dq_3 \, \langle q', t' | q_1, t_1 \rangle q_1 \langle q_1, t_1 | q_3, t_2 \rangle q_3 \langle q_3, t_2 | q, t \rangle =$$

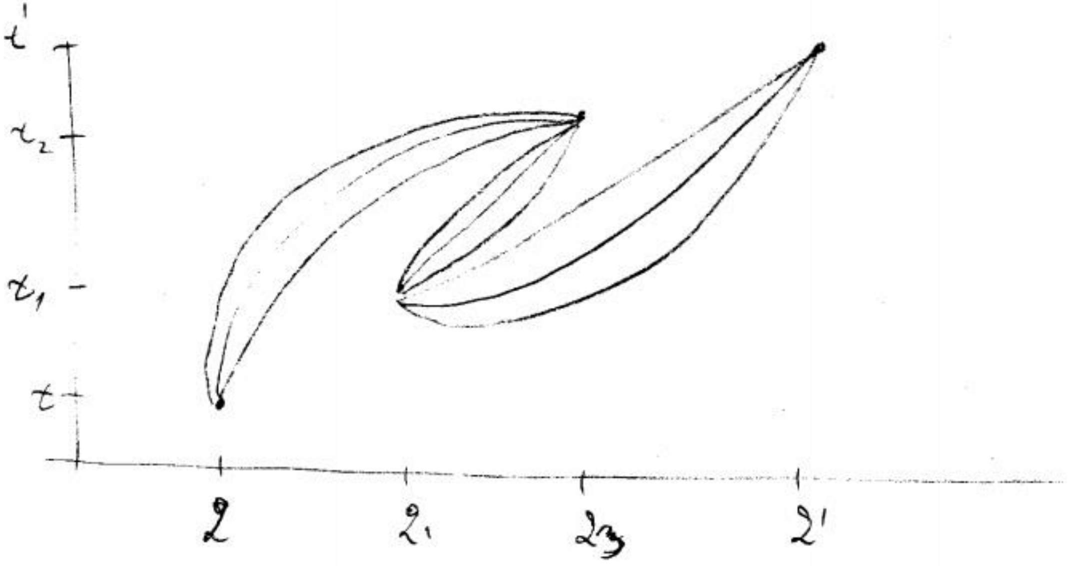
a po preintegrovaní cez q_1, q_3 sa dostanú ľubovoľné trajektórie spájajúce q a q'.

$$\langle q', t' | \, \hat{q}(t_1) \hat{q}(t_2) | q, t \rangle \underset{\text{pre } t_1 > t_2}{=} \int \mathcal{D}q \; q(t_1) q(t_2) \exp (i \int_t^{t'} \mathcal{L}(q(\tau), \dot{q}(\tau)) d\tau$$

keby sme rátali

$$\langle q', t' | \hat{q}(t_1)\, \hat{q}(t_2) | q, t \rangle \quad \text{pre} \quad t_2 > t_1$$

teda to, čo T produkt nepustí, dostali by sme v obdobnom medzikroku



toto sa ale nedá spojiť do jednej trajektórie, lebo tá by v časti išla v čase dozadu

Preto kontinuálny integrál

$$\int \mathcal{D}q \; q(t_1)\, q(t_2)\, \exp\left(+i \int_t^{t'} \mathcal{L}(q(\tau), \dot{q}(\tau))\, d\tau \right) \qquad (*)$$

ráta obloženú hodnotu T-produktu, hoci to z tohto výrazu vôbec nevidno. Je to schované v tom, že uvažujeme len trajektórie idúce v čas stále len dopredu.

Pomocou kontinuálnych integrálov teda vieme rátať amplitúdy (a s prípadne vloženými T-produktmi) typu

$$\langle q', t' | q, t \rangle \qquad (**)$$

Často nás ale viac zaujíma amplitúda (prípadne s vloženými T-produktmi)typu

$$\langle 0, t' | 0, t \rangle \qquad (***)$$

kde $|0\rangle$ označuje základný stav. Všetci vieme, že (**) aj (***) sa rátajú pomocou výrazov typu (*). Ibaže (**) a (***) sú rôzne veci, tak sa musia rátať dajako odlišne

V málo starostlivých textoch sa v kontinuálnych integráloch nepíšu integračné hranice. Ale v rôznych hraniciach práve spočíva trik rozlišujúci (**) a (***).

Povieme najprv výsledok, potom bude komentár

$$\langle q', +\infty | q, -\infty \rangle = \int \mathcal{D}q \; \exp\left(+i \int_{-\infty}^{\infty} \mathcal{L}(q(\tau), \dot{q}(\tau))\, d\tau \right) \qquad (**)$$

$$\langle \phi, +\infty | \phi, -\infty \rangle = \int \mathcal{D}q \; \exp\left(+i \int_{+i\infty}^{-i\infty} \mathcal{L}(q(\tau), \dot{q}(\tau))\, d\tau \right) \qquad (***)$$

Let $|q, t\rangle_H$ be the Heisenberg picture state vector describing a state which at time t is an eigenstate of the coordinate Q_H with eigenvalue q:

$$Q_H(t)|q, t\rangle_H = q|q, t\rangle_H ,$$

$$Q_H(t) = e^{iHt} Q_S e^{-iHt} , \tag{11.1}$$

where Q_S is the time-independent position operator in the Schroedinger picture, and H in the exponent is the Hamiltonian. The state

$$|q\rangle = e^{-iHt}|q, t\rangle_H$$

is an eigenstate of Q_S with eigenvalue q:

$$Q_S|q\rangle = q|q\rangle$$

and

$$|q, t\rangle_H = e^{+iHt}|q\rangle . \tag{11.2}$$

The transformation matrix element

$$F(q', t'; q, t) = {}_{\mathrm{H}}\langle q', t' | q, t\rangle_{\mathrm{H}} = \langle q' | \exp\{-\mathrm{i}H(t'-t)\} | q\rangle \tag{11.3}$$

plays a fundamental role in quantum mechanics. We are going to express $F(q', t'; q, t)$ as a path integral. We shall subdivide the time interval into $n+1$ equal segments, and define

$$t_l = l\,\epsilon + t, \qquad t' = (n+1)\epsilon + t. \tag{11.4}$$

We make use of the completeness of the state vectors $|q_l, t_l\rangle$ to write

$$F(q', t'; q, t) = \int \mathrm{d}q_1(t_1) \int \mathrm{d}q_2(t_2) \int \ldots \int \mathrm{d}q_n(t_n) \langle q', t' | q_n, t_n\rangle\langle q_n, t_n | q_{n-1}, t_{n-1}\rangle \ldots \langle q_1, t_1 | q, t\rangle. \tag{11.5}$$

$$\langle q', \epsilon | q, 0\rangle = \langle q' | \mathrm{e}^{-\mathrm{i}\epsilon H} | q\rangle = \delta(q-q') - \mathrm{i}\,\epsilon\,\langle q' | H | q\rangle + \mathrm{O}(\epsilon^2) \tag{11.6}$$

where the first equality follows from (11.3).

The Hamiltonian $H = H(P, Q)$ is a function of the operators P and Q. Consider the case when H is of the form

$$H = \tfrac{1}{2}P^2 + V(Q). \tag{11.7}$$

In this case

$$\langle q' | H(P, Q) | q\rangle = \int \frac{\mathrm{d}p}{2\pi} \exp\{\mathrm{i}p(q'-q)\}\,[\tfrac{1}{2}p^2 + V(q)] = \int \frac{\mathrm{d}p}{2\pi} \exp\{\mathrm{i}p(q'-q)\} H(p, \tfrac{1}{2}(q+q')) \tag{11.8}$$

where $H(p, q)$ is the classical Hamiltonian. We can write eq. (11.6) correct up to first order in ϵ, as

$$\langle q_l, t_l | q_{l-1}, t_{l-1}\rangle \approx \int \frac{\mathrm{d}p}{2\pi} \exp[\mathrm{i}\{p(q_l - q_{l-1}) - \epsilon\, H(p, \tfrac{1}{2}(q_l + q_{l-1}))\}]\,. \tag{11.9}$$

$$F(q',t';q,t) = \lim_{n\to\infty}\int \prod_{i=1}^{n} \mathrm{d}q_i \prod_{i=1}^{n+1} \frac{\mathrm{d}p_i}{2\pi} \exp\left\{ \mathrm{i} \sum_{j=1}^{n+1} [p_j(q_j - q_{j-1}) - H(p_j, \tfrac{1}{2}(q_j + q_{j-1}))\,(t_j - t_{j-1})] \right\} \tag{11.10}$$

$$F(q', t'; q, t) = \int \left[\frac{\mathrm{d}q\mathrm{d}p}{2\pi\hbar}\right] \exp\left\{\frac{\mathrm{i}}{\hbar}\int_{t}^{t'} (p\dot{q} - H(p, q))\mathrm{d}\tau\right\} \tag{11.11}$$

which is a suggestive shorthand notation for the operation implied by the right-hand side of eq. (11.10). In eq. (11.11)

When the Hamiltonian has the form of eq. (11.7), the p-integration on the right-hand-side of eq. (11.10) can be performed explicitly by making use of the formula

$$\int_{-\infty}^{\infty} \frac{\mathrm{d}p}{2\pi} \exp\{\mathrm{i}\epsilon(p\dot{q} - \tfrac{1}{2}p^2)\} = [2\pi\mathrm{i}\epsilon]^{-1/2} \exp(\tfrac{1}{2}\mathrm{i}\epsilon\dot{q}^2). \tag{11.13}$$

$$F(q', t'; q, t) = \lim_{n \to \infty} \int \prod_{i=1}^{n} \frac{\mathrm{d}q_i}{[2\pi \mathrm{i}\epsilon]^{1/2}} \exp\left\{\mathrm{i} \sum_{j=1}^{n+1} \epsilon \left[\tfrac{1}{2}\left(\frac{q_j - q_{j-1}}{\epsilon}\right)^2 - V\left(\frac{q_j + q_{j+1}}{2}\right)\right]\right\}$$

$$= \int \left[\frac{\mathrm{d}q}{[2\pi \mathrm{i}\epsilon]^{1/2}}\right] \exp\left\{\mathrm{i} \int_t^{t'} L(q, \dot{q})\mathrm{d}t\right\} \tag{11.14}$$

where L is the Lagrangian,

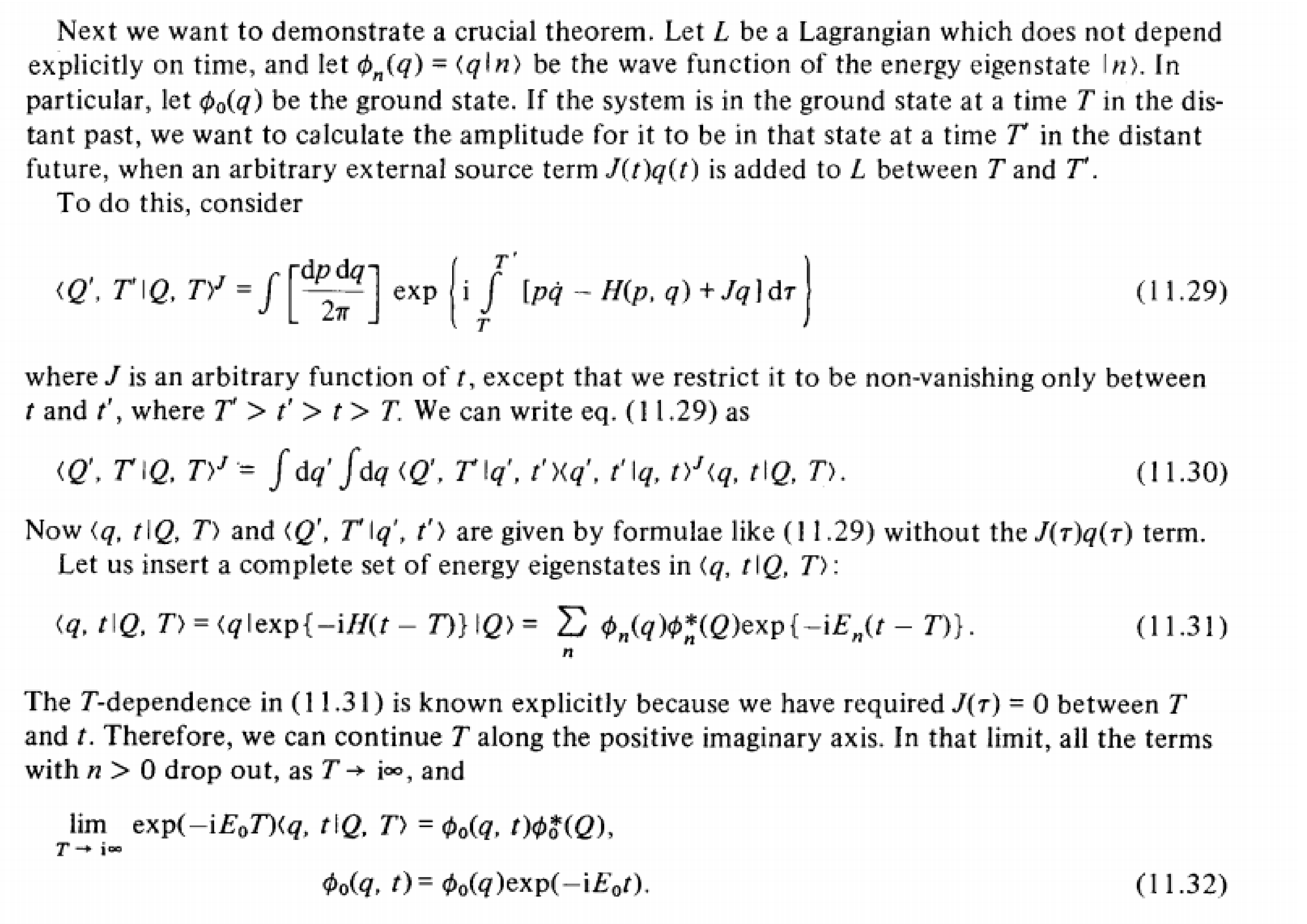
$$L = \tfrac{1}{2}\dot{q}^2 - V(q) \tag{11.15}$$

and $q_o = q(t)$ and $q_{n+1} = q'(t'_{n+1})$.

The quantity

$$S = \int L(q, \dot{q})\mathrm{d}t \tag{11.16}$$

is the action which generates the temporal development of the quantum mechanical system described by the Lagrangian (11.15).



Next we want to demonstrate a crucial theorem. Let L be a Lagrangian which does not depend explicitly on time, and let $\phi_n(q) = \langle q|n\rangle$ be the wave function of the energy eigenstate $|n\rangle$. In particular, let $\phi_0(q)$ be the ground state. If the system is in the ground state at a time T in the distant past, we want to calculate the amplitude for it to be in that state at a time T' in the distant future, when an arbitrary external source term $J(t)q(t)$ is added to L between T and T'.

To do this, consider

$$\langle Q', T'|Q, T\rangle^J = \int \left[\frac{\mathrm{d}p\,\mathrm{d}q}{2\pi}\right] \exp\left\{ \mathrm{i}\int_T^{T'} [p\dot{q} - H(p, q) + Jq]\,\mathrm{d}\tau \right\} \tag{11.29}$$

where J is an arbitrary function of t, except that we restrict it to be non-vanishing only between t and t', where $T' > t' > t > T$. We can write eq. (11.29) as

$$\langle Q', T'|Q, T\rangle^J = \int \mathrm{d}q' \int \mathrm{d}q\, \langle Q', T'|q', t'\rangle\langle q', t'|q, t\rangle^J \langle q, t|Q, T\rangle. \tag{11.30}$$

Now $\langle q, t|Q, T\rangle$ and $\langle Q', T'|q', t'\rangle$ are given by formulae like (11.29) without the $J(\tau)q(\tau)$ term.

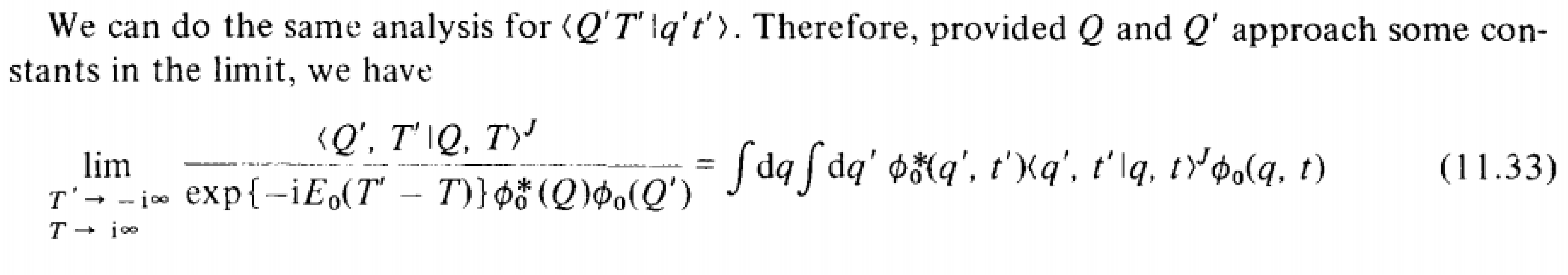
Let us insert a complete set of energy eigenstates in $\langle q, t|Q, T\rangle$:

$$\langle q, t|Q, T\rangle = \langle q|\exp\{-\mathrm{i}H(t-T)\}|Q\rangle = \sum_n \phi_n(q)\phi_n^*(Q)\exp\{-\mathrm{i}E_n(t-T)\}. \tag{11.31}$$

The T-dependence in (11.31) is known explicitly because we have required $J(\tau) = 0$ between T and t. Therefore, we can continue T along the positive imaginary axis. In that limit, all the terms with $n > 0$ drop out, as $T \to \mathrm{i}\infty$, and

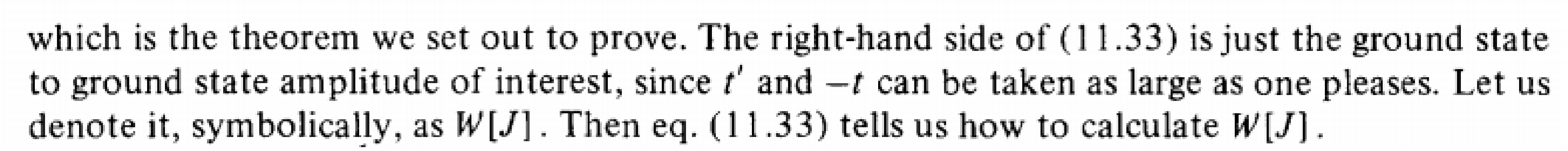
$$\lim_{T\to \mathrm{i}\infty} \exp(-\mathrm{i}E_0T)\langle q, t|Q, T\rangle = \phi_0(q, t)\phi_0^*(Q),$$

$$\phi_0(q, t) = \phi_0(q)\exp(-\mathrm{i}E_0t). \tag{11.32}$$



We can do the same analysis for $\langle Q'T'|q't'\rangle$. Therefore, provided Q and Q' approach some constants in the limit, we have

$$\lim_{\substack{T'\to -i\infty \\ T\to i\infty}} \frac{\langle Q', T'|Q, T\rangle^J}{\exp\{-iE_0(T'-T)\}\phi_0^*(Q)\phi_0(Q')} = \int dq \int dq'\, \phi_0^*(q', t')\langle q', t'|q, t\rangle^J \phi_0(q, t) \qquad (11.33)$$



which is the theorem we set out to prove. The right-hand side of (11.33) is just the ground state to ground state amplitude of interest, since t' and $-t$ can be taken as large as one pleases. Let us denote it, symbolically, as $W[J]$. Then eq. (11.33) tells us how to calculate $W[J]$.

$$\langle q', +\infty | q, -\infty\rangle = \int \mathcal{D}q \exp\left(+i\int_{-\infty}^{\infty} \mathcal{L}(q(\tau), \dot{q}(\tau))\, d\tau\right)$$

$$\langle \phi, +\infty | \phi, -\infty\rangle = \int \mathcal{D}q \exp\left(+i\int_{+i\infty}^{-i\infty} \mathcal{L}(q(\tau), \dot{q}(\tau))\, d\tau\right)$$

(A)

(B)

Integrál (B) sa myslí tak, že najprv sa urobí v reálnych časoch

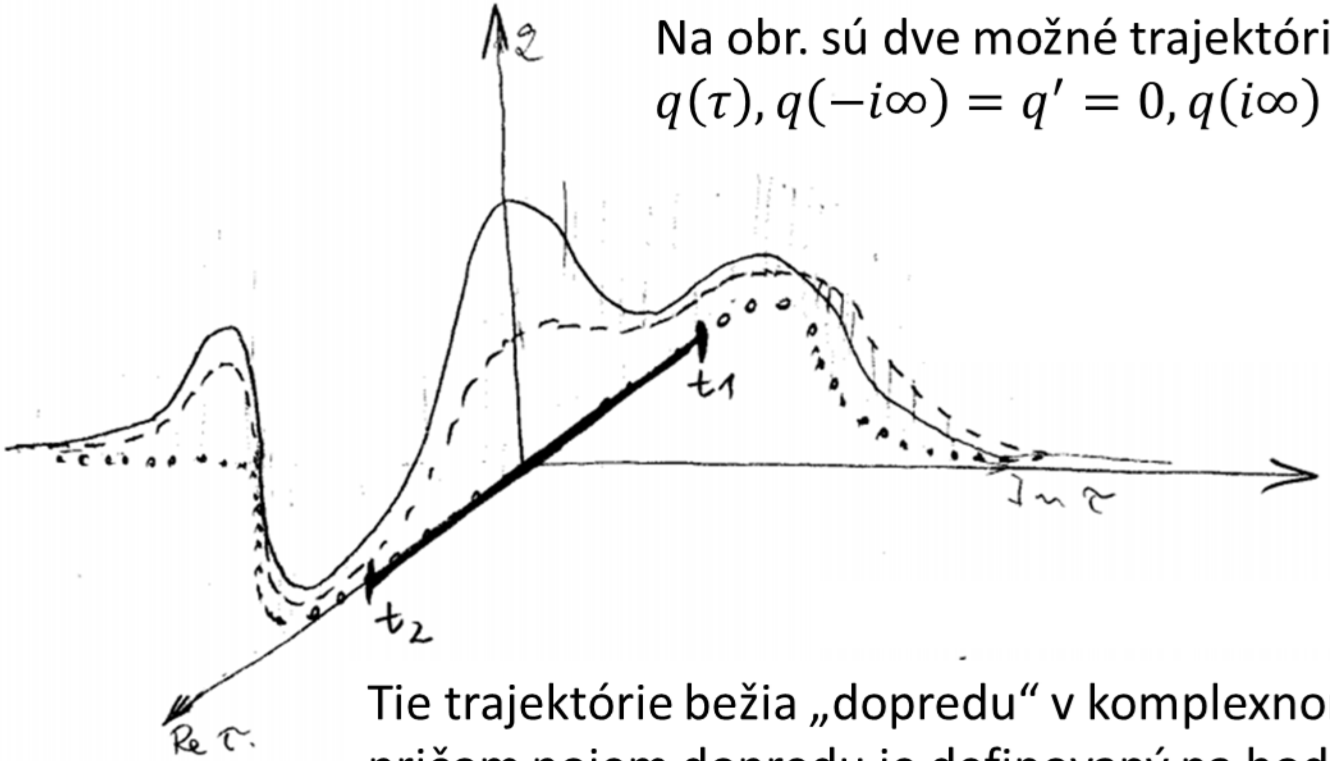
$$\langle q', t' \mid q, t \rangle = F(t, t')\Big|_{q, q' \text{ fix}}$$

a potom sa nájde analytické predĺženie funkcie $F(t, t')$ do bodu $F(+i\infty, -i\infty)$ v komplexnej časovej rovine. To čo sa takto dostane je maticový element

$$\langle \phi, +\infty \mid \phi, -\infty \rangle$$

až na nejaký triviálny faktor, o ktorom bude reč neskôr.

Na obr. sú dve možné trajektórie $q(\tau), q(-i\infty) = q' = 0, q(i\infty) = q = 0$



Tie trajektórie bežia „dopredu“ v komplexnom čase, pričom pojem dopredu je definovaný na bodkovanej čiare v komplexnej rovine τ.

Samozrejme, vec ako $\langle \phi, t' | \phi, t \rangle$
by nebola zaujímavá, keby $|\phi\rangle$ bol základný stav celého Hamiltoniánu, lebo potom by bola tá časová závislosť v maticovom elemente triviálna. To, čo máme naozaj na mysli je

$$\mathcal{L} = \mathcal{L}_0 + J(\tau) \cdot q(\tau) \quad ; \quad J(\tau) = 0 \ \text{pre} \ \begin{matrix} \tau > t_2 \\ \tau < t_1 \end{matrix}$$

Teda vonkajší zdroj zapnutý v (reálnom) čase $t_1 < \tau < t_2$ a $|\phi\rangle$ je základný stav príslušný k hamiltoniánu bez vonkajšieho zdroja. Zaujímame sa teda o pravdepodobnosť toho, že základný stav prejde na základný stav, ak zapneme a potom zas vypneme vonkajší zdroj J.

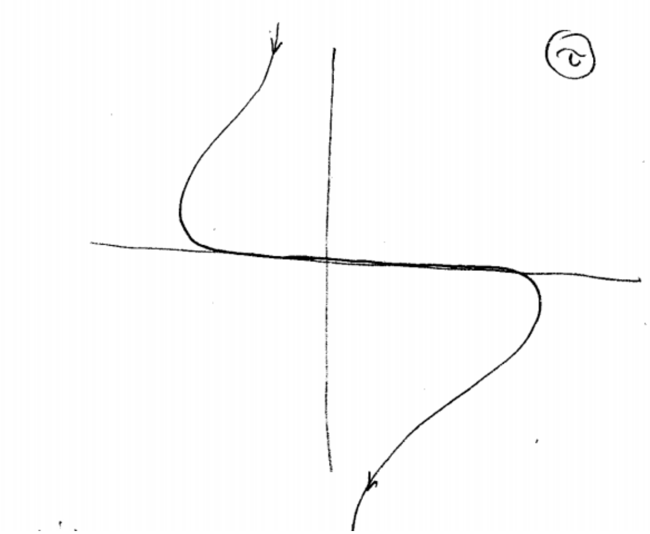
Dobré je to na to, že si tak vyrábame generujúci funkcionál greenových funkcií, napr.

$$i^2 \langle \phi | T(\hat{q}(t_1)\hat{q}(t_2)|\phi\rangle = \frac{\delta^2}{\delta J(t_1)\delta J(t_2)} \langle \phi, +\infty | \phi, -\infty \rangle \Big|_{J=0}$$

Nakoniec sa zaujímame len o connected greenove funkcie, ktoré sa dostanú predelením faktorom

$$\langle \phi, +\infty | \phi, -\infty \rangle$$

Neznámy faktor, o ktorom sme sa zmieňovali takto z hry vypadne.



Chceme mať reálne časové argumenty pri $\hat{q}(t)$ v T-produktoch, preto J musí byť funkcia reálneho argumentu, preto časová krivka na obrázku musí ísť po reálnej osi v tých časoch, keď je zapnuté J.

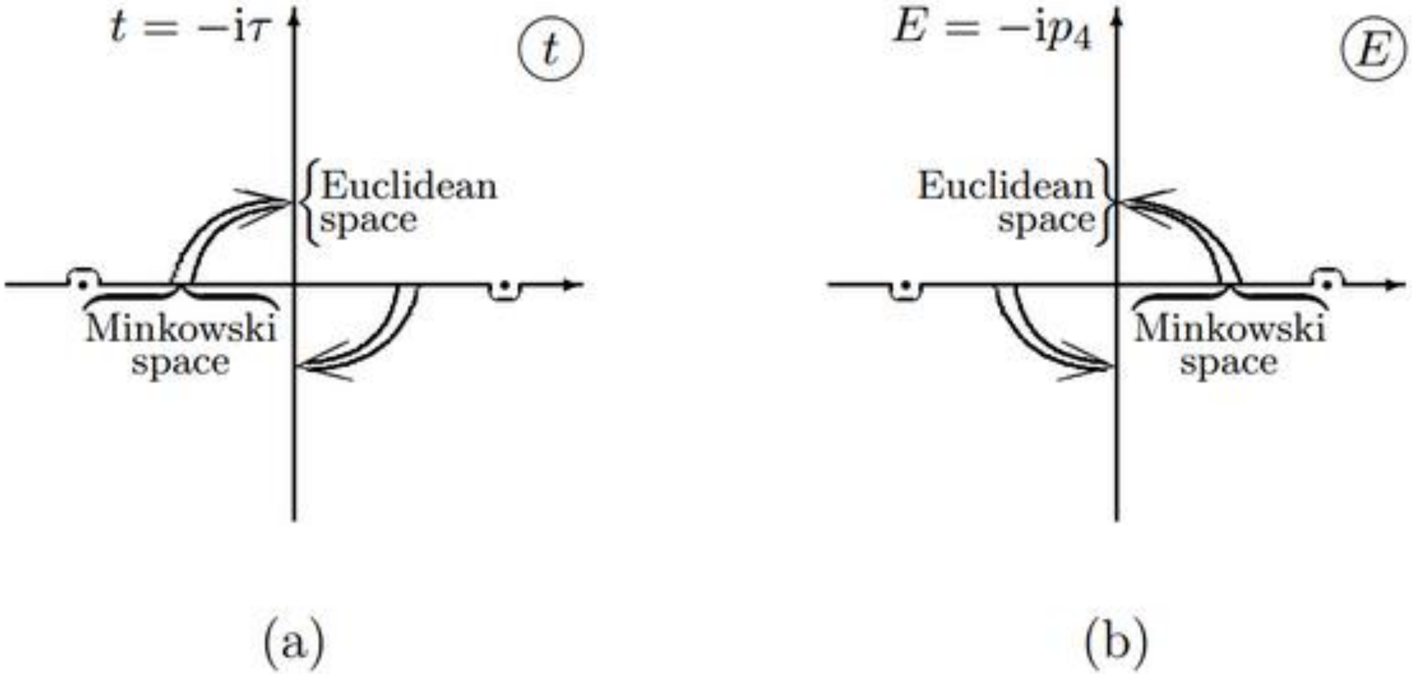


Fig. 1.1. Direction of Wick's rotation from Minkowski to Euclidean space (indicated by the arrows) for (a) time and (b) energy. The dots represent singularities of a free propagator in (a) coordinate and (b) momentum spaces. The contours of integration in Minkowski space are associated with causal Green functions. They can obviously be deformed in the directions of the arrows.

Je výpočtovo nepraktické prehupnúť sa ku koncom z reálnej časovej osi na imaginárnu. Jednoduchšie je mať celú trajektóriu na imaginárnej časovej osi, tam zrátať veci ako

$$\langle 0 | T(\hat{q}(-i\tau_1)\,\hat{q}(-i\tau_2))|0\rangle = f(\tau_1,\tau_2)$$

a potom urobiť analytické predĺženie v τ_1, τ_2 naspäť k reálnym časom. To je idea Wickovej rotácie a prechodu k Euklidovskej teórii poľa.

Zistili sme teda, že pre počítanie v kvantovej teórii môže byť užitočné rátať nejaké integrály v Euklidovskej oblasti. Má to byť len matematický trik pre uľahčenie výpočtov s tým, že nakoniec urobíme analytické predĺženie k fyzike v normálnej Minkovského teórii.

Prirodzená je však otázka, či tie Euklidovské kontinuálne integrály majú aj nejaký fyzikálny význame „samé o sebe, tak ako sú“.

Odpoveď je taká: majú a dokonca **dva úplne rôzne významy.**

Pozrime sa, čo urobí Wickova rotácia v podintegrálnej funkcii

$$e^{iS} \to ?$$

Napr. pre harmonický oscilátor

$$i\int \left\{ (\dot{q}^2 - q^2) + Jq \right\} dt$$

tohto označíme Wick. rotovanú funkciu

Po Wickovej rotácii:

$$q(t) \to q(-i\tau) \equiv \bar{q}(\tau)$$

$$\dot{q}(t) \to \dot{q}(-i\tau) = \left(\frac{dq(t)}{dt} \right)_{t=-i\tau} = \frac{dq(-i\tau)}{d(-i\tau)} = \frac{d\bar{q}(\tau)}{-i\,d\tau} =$$

$$= i\,\dot{q}_{\text{Eukl}}(\tau)$$

$$i\int\left\{\dot{q}^2-q^2+Jq\right\}dt \longrightarrow i\int\left\{\left(-\dot{q}_{eukleid}^2-q_{eukl}^2\right)+J_{eukl}\cdot q_{eukl}\right\}(-i\,d\tau)=$$

$$\doteq -\int\left\{\left(\dot{q}_{eukl}^2+q_{eu}^2\right)-J_{eukl}\cdot q_{eukl}\right\}d\tau \equiv -\int\left\{\mathcal{L}_E(q,\dot{q})-\bar{J}\,\underline{\bar{q}}\right\}d\tau$$

Tu sme definovali, čo nazývame Euklidovským lagranžiánom. Fór je v tom, že Euklidovký lagranžián je vlastne energia pôvodného kvantového systému.

Nešpecifikovali sme hranice integrálov, lebo zatiaľ sa len tak hrajkáme, robíme heuristickú investigatívu.

To že sme dostali v exponente energiu so záporným znamienkom znamená že začíname tušiť súvis so štatistickou fyzikou.

Vráťme sa k otázke hraníc

$$G(t',t)=\langle q',t' | q,t\rangle = \int \mathcal{D}q \, e^{i\int_t^{t'} \mathcal{L} + Jq \, d\tau}$$

je greenova funkcia rovnice

$$i\frac{\partial\psi(q)}{\partial t} = \hat{H}\psi \qquad G = \sum_n \phi_n(q')\phi^*_n(q)\exp\left(-iE_n(t'-t)\right)$$

Po Wickovej rotácii

$$G(q,q',t-t') \rightarrow d(q,q',\tau-\tau')$$

bude to zrejme greenova funkcia zrotovanej rovnice

$$-\frac{\partial\psi}{\partial\tau} = \hat{H}\psi \qquad (*)$$

$$d(q,q',\tau'-\tau) = \int \mathcal{D}q \exp\left(-\int_\tau^{\tau'} \mathcal{L}_E \, d\tau\right)$$

$$-\frac{\partial\psi}{\partial\tau} = \hat{H}\psi \qquad (*)$$

$$d(q,q',\tau'-\tau) = \int Dq \, \exp\left(-\int_{\tau}^{\tau'} \mathcal{L}_E \, d\tau\right)$$

a bude to

$$d(q,q') = \sum_n \phi_n(q)\,\phi_n^*(q)\,e^{-(\tau'-\tau)E_n}$$

keby tam v exponente namiesto $\tau' - \tau$ bolo β, tak by išlo o maticové elementy nenormalizovanej (bez Z) „matice hustoty“

$$\hat{\varrho} = \exp(-\beta\hat{H})$$

$$d(q',q) = \langle q'|\exp(-\beta\hat{H})\,|q\rangle$$

„Matica hustoty“ je zjavne operátor časového posunutia v rovnici (*), preto sa netreba diviť, že to tak vychádza.

Teda: ak urobíme Wickovu rotáciu v generujúcom funkcionáli a zoberieme časový interval

$$\Delta\tau = \beta = \frac{1}{kT}$$

dostaneme element matice hustoty

$$d(q', q) = \langle q'|\exp(-\beta\hat{H})\,|q\rangle$$

kde q, q' sú fixované koncové body trajektórií v kontinuálnom integráli.
Ak chceme dostať štatistickú sumu

$$Z = \int dq\, \rho_{qq} = \int dq\, \langle q|\hat{\rho}|q\rangle$$

To sa dá urobiť rovno v kontinuálnom integráli tak, že zoberieme trajektórie začínajúce a končiace v tom istom bode q (s periódou β) a potom ešte preintegrujeme aj cez q.

Záver: Ak chcem spočítať štatistickú sumu nejakého kvantoveho fyzikálneho systému pri inverznej teplote β, potom

- tvárim sa ako keby som chcel ten klasický systém kvantovať, teda robiť kvantové maticové elementy prechodu
- ale namiesto toho urobím v kontinuálnom integráli Wickovu rotáciu
- zoberiem všetky periodické trajektórie s periódou β
- preintegrujem cez všetky body trajektórií, teda nič nedržím fixné

Existuje ale ešte úplne iná štatistická fyzika, ktorá súvisí s tým kvantovým systémom a získamu je z Wickovsky zrotovaných hieroglyfov.

Na integrál

$$d(q,q',\tau,\tau') = \int \mathcal{D}q \, \exp\left(-\int_{\tau}^{\tau'} \mathcal{L}_E(q,\dot{q}) - Jq\right) \qquad (*)$$

Sa dá dívať ako na štatistickú sumu nejakého fiktívneho nekvantového systému, ktorý nijako nesúvisí s tým pôvodným nekvantovým systémom, ktorý sme chceli kvantovať pomocou kontinuálneho integrálu.

Ten fiktívny klasický objekt je čosi ako „struna“. Mikrostav tej struny je daný funkciou $q(\tau)$, pričom krajné body tej struny sú vo všetkých mikrostavov zafixované v bodoch q, q'. Energia struny v každom mikrostave je daná integrálom Euklidovského lagranžiánu pôvodného klasického systému.

Teraz niečo o tom na čo je to dobré:

V podstate kvôli dvom veciam: to prvé v štatistickej fyzike máme nejaký cit - ak intuitívne uhádneme výsledok, môžno ho aj preložíme do jazyka QM (urobíme opačnú Wick. rotáciu) a máme informáciu o tzv. chaoní systému.

Druhý dôvod je technicko-počtový. Tak QM podobne ako štat. mech. vedie na vyčíslovanie kontimnohých integrálov. V oboch prípadoch je to technicky ťažko-prevediteľný výpočet. Môže sa ale stať, že v štat. fyzike nás vlastne nezaujíma štatistická suma ale zaujíma nás niektoré charakteristiky základného stavu. (energia, magnetizácia, fluktuácie a.p.). Môžeme potom skúsiť "analógový výpočet" pokúsiť sa na počítači simulovať ~~fyzikálnu~~ fyzikálny systém a "experimentálne nájsť" charakteristiky rovnovážneho stavu, ktoré nás zaujímajú.

Pretože počítač je diskrétna vec, musíme spomínané "šrumu" modelovať ako diskrétnu sadu bodov, teda častice dostaneme ~~diskrét~~ mriežkový výpočet.

Keď takýmto spôsobom získame "experimentálne" hodnoty štat. problému, môžeme ich (analyt. predlžením) preložiť do jazyka kvantovej teórie.

Rovnovážne žiarenie v dutine a termodynamický dôkaz Weylovej vety

Abstrakt: Weylova veta je (pre mňa) ťažká veta z matematickej fyziky o spektre Laplaciánu. Bola ako hypotéza vyslovená okolo roku 1910 a veľký Hilbert sa vtedy vyjadril, že neverí, že by ju niekto ešte počas jeho života dokázal. Weyl mu za štyri mesiace dôkaz priniesol. Môj seminár nebude žiadna veda, len známe veci. Pokúšam sa vyrábať akýsi učebný text zo štatistickej fyziky a pri tom si musím na staré kolená všeličo ujasňovať. Mám dojem že sa mi aj podarilo niečo si ujasniť a o niečo starostlivejšie než je obvykle zvykom sformulovať nejaký termodynamický dôkaz a jeho súvis s Weylovou vetou. To pochopenie mi prinieslo istú radosť, ktorú by som vám rád komunikoval. Ak počas seminára usúdite, že miera mojej radosti je len mierou mojej už senility, vopred sa ospravedlňujem za nedostatok kritickej sebareflexie.

Statistical physics of electromagnetic field

Physics manifesto

We have a new physical system: electromagnetic field. It is a new animal in the physics ZOO. So we have to show how the physics manifesto works for this new animal.

- **Find how to describe "current" state of the animal**
- **Find the equation of motion**
- **Predict future states**
- **If you cannot write down current state in sufficient details do statistical physics instead**

State of electromagnetic field: two vector fields $\vec{E}(\vec{r})\,\vec{B}(\vec{r})$

What are the equations of motion, how to predict future?

$$\mathrm{rot}\vec{E}(\vec{r}) = -\frac{\partial \vec{B}(\vec{r})}{\partial t} \qquad \mathrm{rot}\,\vec{B}(\vec{r}) = \mu_0\vec{j}(\vec{r}) + \mu_0\varepsilon_0\frac{\partial \vec{E}(\vec{r})}{\partial t}$$

$$\frac{\partial \vec{E}}{\partial t} = \frac{1}{\mu_0\varepsilon_0}\mathrm{rot}\,\vec{B} - \frac{1}{\varepsilon_0}\vec{j} \qquad \frac{\partial \vec{B}}{\partial t} = -\mathrm{rot}\vec{E}$$

Suppose we know the state $\vec{E}(t,\vec{r}), \vec{B}(t,\vec{r})$. The state after time dt should be

$$\boxed{\begin{aligned}\vec{E}(t+dt,\vec{r}) &= \vec{E}(t,\vec{r}) + \frac{1}{\mu_0\varepsilon_0}\mathrm{rot}\,\vec{B}(t,\vec{r})\;dt - \frac{1}{\varepsilon_0}\vec{j}(t,\vec{r})\;dt\\ \vec{B}(t+dt,\vec{r}) &= \vec{B}(t,\vec{r}) - \mathrm{rot}\vec{E}(t,\vec{r})\;dt\end{aligned}}$$

Note: boundary conditions

$$\vec{E}(t+dt,\vec{r}) = \vec{E}(t,\vec{r}) + \frac{1}{\mu_0\varepsilon_0}\mathrm{rot}\,\vec{B}(t,\vec{r})\;dt - \frac{1}{\varepsilon_0}\vec{j}(t,\vec{r})\;dt$$
$$\vec{B}(t+dt,\vec{r}) = \vec{B}(t,\vec{r}) - \mathrm{rot}\vec{E}(t,\vec{r})\;dt$$

These equations can in principle be used to predict the state after a small time interval dt knowing the current state at the time t. And then iterate to further future, next dt, and next dt an d so on. We can try to do it at least numerically on a computer. However, we meet two problems:

- numerical instabilities. We either have be very careful with the choice of steps in time and space or we have to choose more sophisticated numerical methods. This is not a sign of some essential problem
- we will not be able to numerically approximate the rotations in points at the boundaries. We have to use boundary conditions for the boundary points instead. **This is essential and it helps us to understand why the partial differential equations force us to specify the boundary conditions. To know the initial state is not enough to predict the future.**

If you cannot write down current state in sufficient details do statistical physics instead

Heating stove: even if the fuel is burned out, no fire there, if you open the door you see red light. **There is electromagnetic field inside heated to temperature T**. This is all we know about the field inside. We do not know the detailed state (microstate). We just have very reduced information:

There is **radiation in the box of volume V, having temperature T**. This sentence describes the current macrostate. We assume that some relaxation time has already passed, so it is an **equilibrium macrostate**. A priori we do not know if V, T specify the equilibrium macrostate sufficiently, so that we can start statistical physics machinery. Just assume that yes and we shall see later that V, T is really enough.

Statistical physics machinery

One macrostate can be realized by tremendously high number of detailed microstates. Those microstates we denote just symbolically as i. Each microstate has well defined energy E_i. Statistical physics machinery looks like this

- Calculate
$$Z(V,T) = \sum_i \exp\left(-\frac{E_i}{kT}\right)$$
- Rewrite it as
$$Z(V,T) = \exp\left(-\frac{F(V,T)}{kT}\right)$$
- You can calculate anything of interest from $F(V,T)$ by suitable derivatives. For example radiation pressure

$$p = -\frac{\partial F}{\partial V}$$

Statistical physics machinery

$$Z(V,T) = \sum_i \exp\left(-\frac{E_i}{kT}\right)$$

Looks easy, if you know hot to calculate sums over all possible microstates

$$\sum_i$$

for the new physics animal: the electromagnetic field.
The states are given by fields. It is a field theory. The sum is a sum over all possible fields!

Wiener functional integrals !? ☹

A way out: write down the states differently in a simpler but equivalent way. At least for a **free electromagnetic field in a box (possibly of arbitrary shape).**
The first thing is to find what the states of the field look like.

Free field in a box

$$\mathrm{rot}\vec{E} = -\frac{\partial \vec{B}}{\partial t}$$

make rot rot

$$\frac{\partial^2 \vec{E}}{\partial t^2} = c^2 \Delta \vec{E}$$

We have all learned in quantum mechanics how to handle such equation. First we find Laplacian eigenfunctions

$$\Delta \vec{\Psi}(\vec{r}) = -k^2 \vec{\Psi}(\vec{r})$$

We need solutions inside the box satisfying the boundary conditions and an additional condition

$$\nabla . \vec{\Psi}(\vec{r}) = 0$$

We find that good solutions exist only for a discrete set of k values. Let denote them k_j. ***j* might be a multiindex**. We assume, that the eigenfunctions $\vec{\Psi}_j(\vec{r})$ form a complete set of functions, therefore the solutions of the wave equation satisfying the boundary conditions and the zero divergence condition can be looked for as superpositions

$$\vec{E}(t, \vec{r}) = \sum_j \Phi_j(t) \vec{\Psi}_j(\vec{r})$$

Free field in a box, states, discrete formulation

$$\vec{E}(t,\vec{r}) = \sum_i \Phi_j(t)\vec{\Psi}_j(\vec{r})$$

Inserting this into the wave equation (equation of motion) we get

$$\boxed{\frac{\partial^2 \Phi_j(t)}{\partial t^2} = -c^2 k_j^2 \Phi_j(t)}$$

These are equation of motion for independent harmonic oscillators whose “names” are j.

We see that **the state of the free field in a box can be given by writing down the states of a (discrete) set of independent harmonic oscillators as**

$$\{\Phi_j(0), \dot{\Phi}_j(0)\}$$

The frequency of the oscillator j is $\omega_j = ck_j$. The frequencies are directly given by the spectrum of the Laplacian corresponding to the shape of the box and the boundary conditions.

Free field in a box: how to do statistical physics

We have retold the whole story about our physics animal: the electromagnetic field in a box.

- its state is given as a set of states of independent harmonic oscillators with frequencies ω_j
- its future can be predicted by predicting the time development of the set of independent harmonic oscillators starting with their initial states. This is trivial.
- **statistical physics can be done by creating a canonical ensemble for the set of harmonic oscillators. This is trivial once the spectrum ω_j is found. However, we have a choice:**

The oscillators can be considered as classical or quantum objects

There is no general formula for the **complete** spectrum of the Laplacian. It has to be determined anew for each box shape and each boundary conditions. However, there is the **Weyl theorem describing the asymptotic distribution of ω_j for large frequencies**.

First let us investigate what we can say without knowing the spectrum in details.

Calculating mean energy

We shall not do complete statistical physics, we shall just calculate the mean energy at temperature T. Since wee have a system of independent harmonic oscillators, the key point is to calculate the mean energy of a harmonic oscillator having frequency ω at the temperature T. Let us denote this as $\bar{\varepsilon}(\omega, T)$. Then the mean total energy of the radiation in the cavity will be given by a sum over spectrum

$$\overline{E}(T) = \sum_j \bar{\varepsilon}(\omega_j, T)$$

It is more convenient to rewrite the sum as an integral **introducing Laplacian spectral density $\boldsymbol{\rho(\omega)}$** with the meaning that $\rho(\omega)d\omega$ is the number of oscillators $(j's)$ having frequencies ω_j in the interval $d\omega$ around ω. The integral is certainly a good approximation to the sum for large enough ω. If one wants to be more precise one can consider $\rho(\omega)$ to be a distribution consisting of delta functions for small ω plus an ordinary smooth function for large ω. Thus we get

$$\overline{E}(T) = \int \varepsilon(\omega, T)\rho(\omega)d\omega$$

From there we get for the **spectral energy density (energy per unit volume per unit frequency interval)**

$$u(\omega) = \frac{1}{V}\bar{\varepsilon}(\omega, T)\rho(\omega)$$

Energy spectral density

$$u(\omega) = \frac{1}{V}\bar{\varepsilon}(\omega, T)\rho(\omega)$$

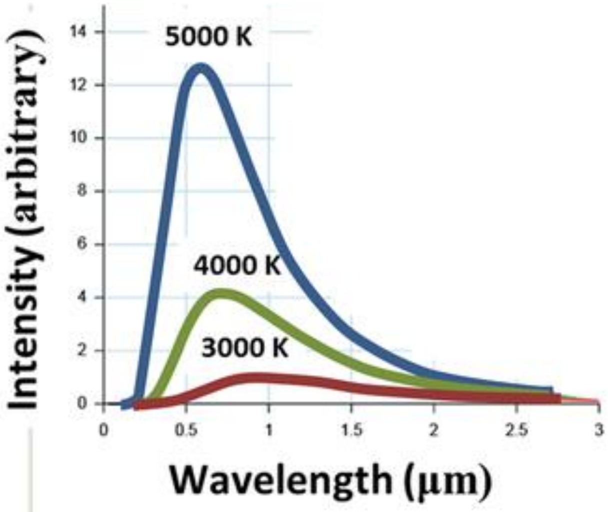
- **For classical oscillator we have** $$\bar{\varepsilon}(\omega, T) = kT$$

This **is disaster**: there is no dependence on ω. The portion of energies carried by very high frequencies relative to small frequencies does not depend on T!!! If you open the door on your stove, **you get irradiated by deadly ultraviolet, roentgen and gamma radiation.**

WRONG !

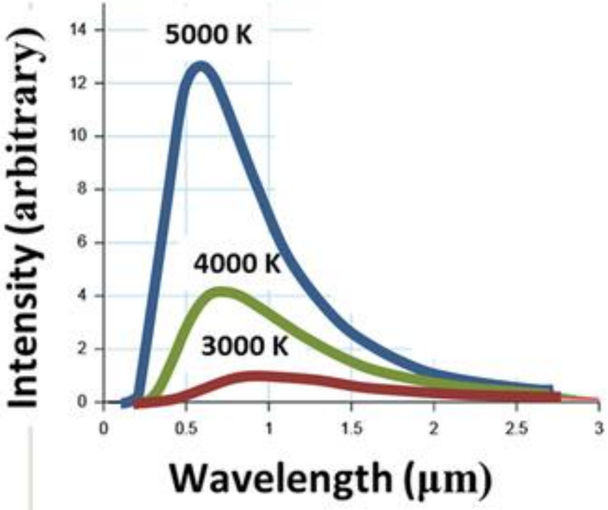
- **For quantum oscillator we have** $$\bar{\varepsilon}(\omega, T) = \frac{\hbar\omega}{\exp\left(\frac{\hbar\omega}{kT}\right) - 1} + \xcancel{\frac{1}{2}\hbar\omega}$$

If the Laplacian spectral density is not exponentially rising, the large ω asymptotic may be correct



Energy spectral density

$$u(\omega) = \frac{1}{V} \frac{\hbar\omega}{\exp\left(\frac{\hbar\omega}{kT}\right) - 1} \rho(\omega)$$



You can experimentally determine the Laplacian spectrum for your cavity:

- Measure the energy spectral density
- Divide by

$$\frac{\hbar\omega}{\exp\left(\frac{\hbar\omega}{kT}\right) - 1}$$

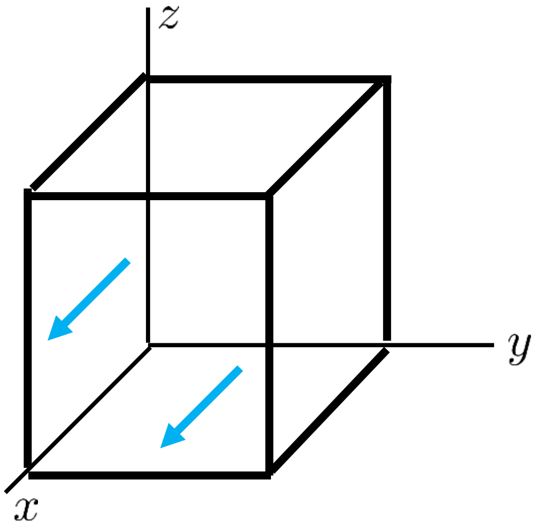
- You get the Laplacian spectrum $\rho(\omega)$ for your cavity ! (For the boundary conditions relevant for the electromagnetic field inside.)

Example of boundary conditions
Free electromagnetic field in a conducting box

Free electromagnetic field in a conducting box (like microwave) satisfies wave equation inside the box. The more difficult part of the problem are the boundary condition.

Let us investigate the x-component of the electric field $E_x(x, y, z)$ in a cube-like conducting box $a \times a \times a$.

Tangential components on the boundary have to be zero



$$E_x(x,0,z) = E_x(x,a,z) = E_x(x,y,0) = E_x(x,y,a) = 0$$

According to **Fourier**: function $E_x(x, y, z)$ satisfying such conditions can be constructed from the set of functions (cyclically for other components)

$$\tilde{E}_{x;n,l} = f_{x;n,l}(x)\sin(\frac{n\pi y}{a})\sin(\frac{l\pi z}{a})$$ where f-s are so far unknown functions

The electric field has to satisfy the condition div $\vec{E} = 0$ also in points with $x = 0, x = a$. There the y and z field components do not contribute since their sinuses are 0 there, therefore we get the condition for the derivative of f

$$f'_{x;n,l}(0) = f'_{x;n,l}(a) = 0$$

so the derivatives of f can be constructed from sinuses, so f from cosines.

Free electromagnetic field in a conducting box

We have found that the electric field automatically satisfying the boundary conditions can be looked for in the form

$$E_x(x,y,z,t) = \sum_{mnl} E_{x;mnl}(t)\cos(\frac{m\pi x}{a})\sin(\frac{n\pi y}{a})\sin(\frac{l\pi z}{a})$$

$$E_y(x,y,z,t) = \sum_{mnl} E_{y;mnl}(t)\sin(\frac{m\pi x}{a})\cos(\frac{n\pi y}{a})\sin(\frac{l\pi z}{a})$$

$$E_z(x,y,z,t) = \sum_{mnl} E_{z;mnl}(t)\sin(\frac{m\pi x}{a})\sin(\frac{n\pi y}{a})\cos(\frac{l\pi z}{a})$$

where $E_{x;mnl}(t), E_{y;mnl}(t), E_{z;mnl}(t)$ are so far unknown functions. Actually not all these functions are independent, because of the zero divergence condition, which gives the conditions

$$\frac{m\pi}{a}E_{x;mnl}(t) + \frac{n\pi}{a}E_{y;mnl}(t) + \frac{l\pi}{a}E_{z;mnl}(t) = 0$$

so for any mnl **only two of these functions are independent** (for example the x and y component). The third one can be expressed from the other two. **We have got the Laplacian spectrum**: $j = (m,n,l)$, $k_j = \left(\frac{m\pi}{a}, \frac{n\pi}{a}, \frac{l\pi}{a}\right)$

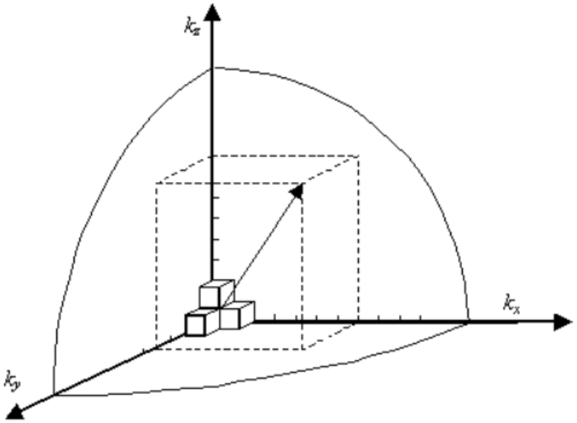
$$\boxed{\omega_j = c\sqrt{\frac{\pi^2}{a^2}(m^2+n^2+l^2)}}$$

Laplacian spectral density

$$\omega_j = c\sqrt{\frac{\pi^2}{a^2}(m^2+n^2+l^2)}$$

The states (m, n, l) can be visualized in an abstract (m, n, l)-space as points **with integer coordinates in the positive octant.** To each unit cube belongs on average one its vertex, so number of integer-coordinate points for which $\omega_j < \omega$ denoted as $\varphi(\omega)$is 1/8 of the volume of a sphere with radius $r_\omega = \sqrt{m^2+n^2+l^2} = \frac{a}{c\pi}\omega$

$$\varphi(\omega) = 2\frac{1}{8}\frac{4}{3}\pi r_\omega^3 = 2\frac{1}{8}\frac{4}{3}\pi(\frac{a}{c\pi}\omega)^3$$



The pre-factor 2 is for 2 “polarizations” for each (m, n, l).

$$\boxed{\rho(\omega) = \varphi'(\omega) = \frac{a^3\omega^2}{\pi^2c^3} = V\frac{\omega^2}{\pi^2c^3}}$$

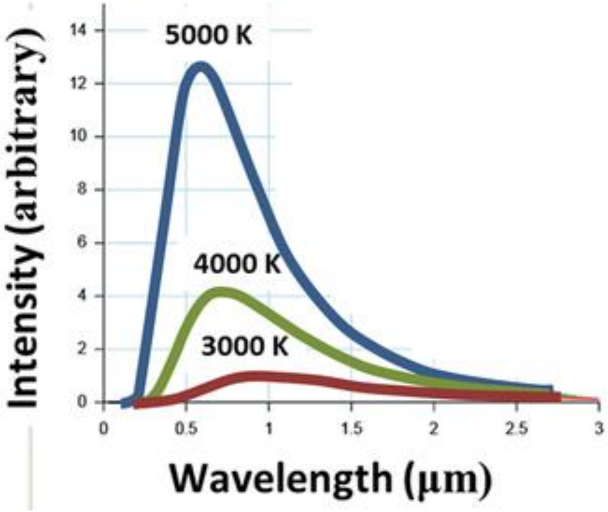
Energy spectral density

$$u(\omega) = \frac{1}{V}\bar{\varepsilon}(\omega, T)\rho(\omega)$$

$$\bar{\varepsilon}(\omega, T) = \frac{\hbar\omega}{\exp\left(\frac{\hbar\omega}{kT}\right) - 1}$$

$$\rho(\omega) = V\frac{\omega^2}{\pi^2 c^3}$$

$$u(\omega) = \frac{\hbar\omega^3}{\pi^2 c^3}\frac{1}{\exp\left(\frac{1}{kT} - 1\right)}$$



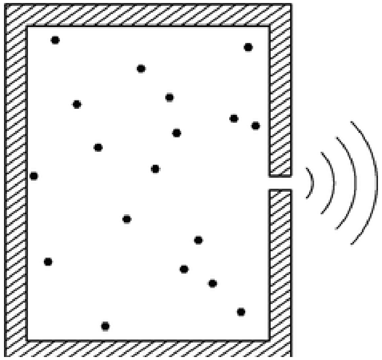
We have got this for a special case cube-like box cavity with perfectly conducting walls

How to measure $u(\omega)$: make hole into the cavity

You can experimentally determine the Laplacian spectrum for your cavity:

- Measure the energy spectral density
- Divide by $$\frac{\hbar\omega}{\exp\left(\frac{\hbar\omega}{kT}\right)-1}$$
- You get the Laplacian spectrum $\rho(\omega)$ for your cavity ! (For the boundary conditions relevant for the electromagnetic field inside.)

$$u(\omega)=\frac{1}{V}\frac{\hbar\omega}{\exp\left(\frac{\hbar\omega}{kT}\right)-1}\rho(\omega)$$



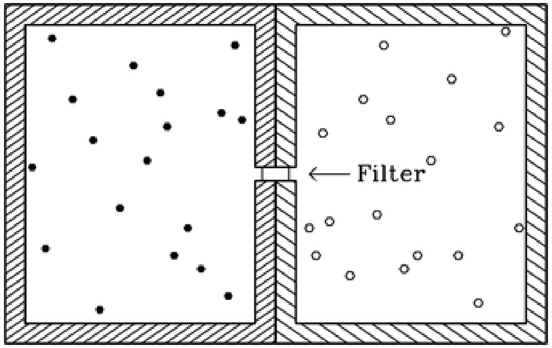
$$j(\omega)=\frac{1}{4}cu(\omega)$$

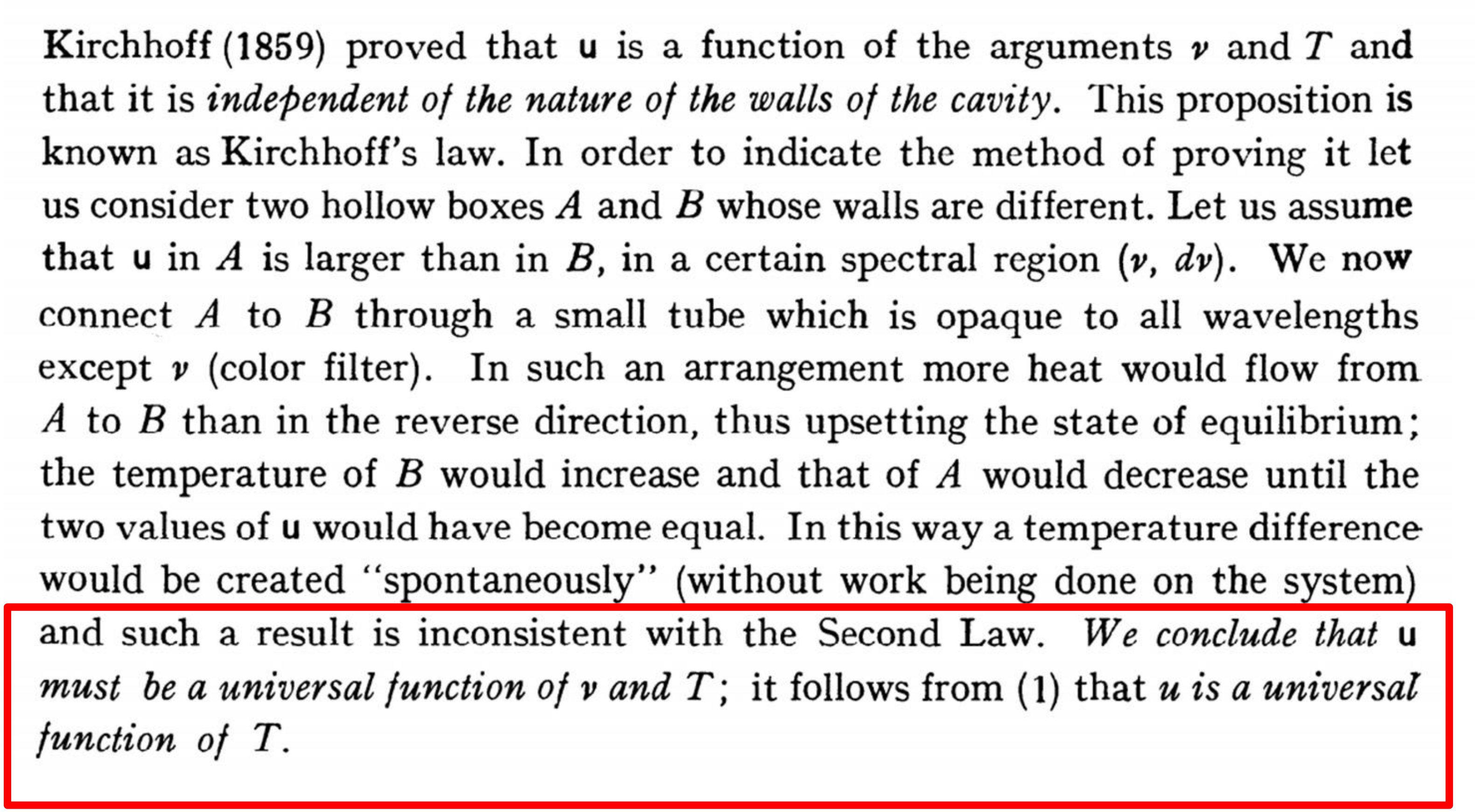
spectral energy density flow (energy per unit of the hole area per unit of time per unit of frequency

Easy argument using photons: photons have velocity c, averaging over angles gives 1/4.

How photons enter the game ? Be patient, we come to it soon

Kirchhoff Sommerfeld thermodynamic universality proof





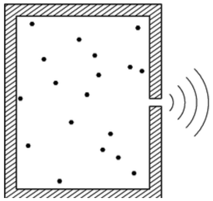
Kirchhoff (1859) proved that **u** is a function of the arguments ν and T and that it is *independent of the nature of the walls of the cavity*. This proposition is known as Kirchhoff's law. In order to indicate the method of proving it let us consider two hollow boxes A and B whose walls are different. Let us assume that **u** in A is larger than in B, in a certain spectral region (ν, $d\nu$). We now connect A to B through a small tube which is opaque to all wavelengths except ν (color filter). In such an arrangement more heat would flow from A to B than in the reverse direction, thus upsetting the state of equilibrium; the temperature of B would increase and that of A would decrease until the two values of **u** would have become equal. In this way a temperature difference would be created "spontaneously" (without work being done on the system) and such a result is inconsistent with the Second Law. *We conclude that* **u** *must be a universal function of* ν *and* T; it follows from (1) that *u is a universal function of* T.

Total $u(\omega)$ universality cannot be true

$$u(\omega) = \frac{1}{V}\frac{\hbar\omega}{\exp\left(\frac{\hbar\omega}{kT}\right) - 1}\rho(\omega)$$

If $u(\omega)$ is universal, then $\rho(\omega)$ is universal as well, that is **Laplacian spectrum is universal for a cavity of arbitrary shape.** **This cannot be true for small ω.**

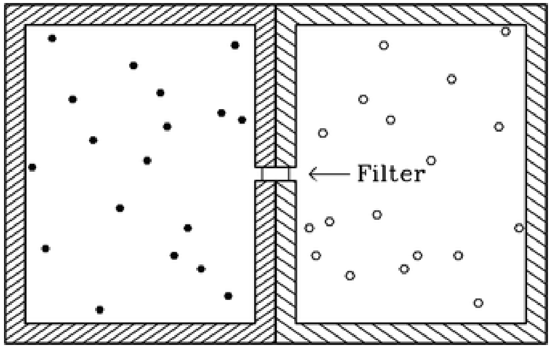
So there is a flow in the thermodynamic argument as was presented .



$$j(\omega) = \frac{1}{4}cu(\omega)$$

Small frequencies mean **large wave lengths, which do not get through a small hole**. Everybody knows that: you can look into running microwave oven through net full of small holes. Visible light gets out, 15 cm microwaves do not: our eyes are not burned.

Kirchhoff Sommerfeld thermodynamic universality proof



The thermodynamic universality proof survives only for frequencies for which the formula

$$j(\omega) = \frac{1}{4} c u(\omega)$$

is a good approximation, that is for large frequencies. Conclusion:

Laplacian spectral density $\rho(\omega)$ is asymptotically universal.

This is roughly what Weyl's theorem says.

We have presented a thermodynamic proof of the Weyl theorem

We have presented a thermodynamic proof of the Weyl theorem

Well, no.

There is another flow in the thermodynamic argument

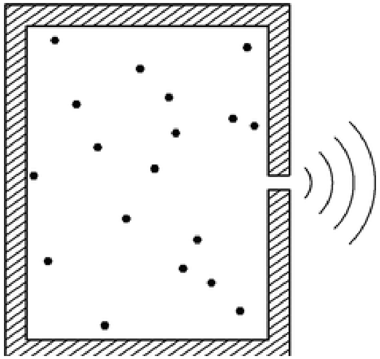
Actually the thermodynamic conclusion is correct, just the argumentation is not clean enough.

We have first to learn more about photons.

What is electromagnetic field?

We have retold the whole story about our physics animal: the electromagnetic field in a box.

- its state is given as a set of states of independent harmonic oscillators with frequencies ω_j



Electromagnetic field in a box

is

a set of harmonic oscillators

Stupid naive question: the oscillators run out from the hole?

What **is** electromagnetic field? Alternative formulation

$$E_x(x,y,z,t)=\sum_{mnl}E_{x;mnl}(t)\cos(\frac{m\pi x}{a})\sin(\frac{n\pi y}{a})\sin(\frac{l\pi z}{a})$$

$$E_y(x,y,z,t)=\sum_{mnl}E_{y;mnl}(t)\sin(\frac{m\pi x}{a})\cos(\frac{n\pi y}{a})\sin(\frac{l\pi z}{a})$$

$$E_z(x,y,z,t)=\sum_{mnl}E_{z;mnl}(t)\sin(\frac{m\pi x}{a})\sin(\frac{n\pi y}{a})\cos(\frac{l\pi z}{a})$$

Electromagnetic field in a box

is

a set of stationary waves

Stupid naive question: why standing waves run out from the hole?

Easy answer: a superposition of standing waves can give you a running wave, which gets out. **Il si muove**.

Never formulate statements of ontological character in physics.

?

Easy answer: a superposition of standing waves can give you a running wave, which gets out.

But

Where are the superposition microstates in the statistical physics calculation of the statistical sum?

$$Z(V,T) = \sum_i \exp\left(-\frac{E_i}{kT}\right)$$

$$\overline{E}(T) = \sum_j \bar{\varepsilon}(\omega_j, T)$$

Free electromagnetic field: still another alternative description - photons

The complete (micro)state of the electromagnetic field we have represented by the table where the excitation numbers are any integers. If you see the table do you recall that you have already met a similar table in a very different context? Here it is: The state of non-interacting indistinguishable spinless bosons was given by the list of one-particle states each with its occupational number which could be any integer. The total energy was

mnl,s	excit.number
111,x	n_{111x}
111,y	n_{111y}
112,x	n_{112x}
112,y	n_{112y}
121,x	n_{121x}
121,y	n_{121y}
…	…

one particle state	occup. number
1,1,1	1
1,1,2	3
1,2,1	0
1,2,2	2
2,1,1	0
2,1,2	0
2,2,1	12
2,2,2	4
⋮	⋮

$$E = \sum_{n_1 n_2 n_3} n_{n_1 n_2 n_3} \varepsilon_{n_1 n_2 n_3}$$

For the electromagnetic field the total energy is

$$E = \sum_{mnl} n_{mnl} \hbar\omega_{mnl} = \sum_{mnl} n_{mnl} \varepsilon_{mnl}$$

The same (almost) description of the states, the same formula for the total energy.

Free electromagnetic field: still another alternative description - photons

mnl,s	excit.number
111,x	n_{111x}
111,y	n_{111y}
112,x	n_{112x}
112,y	n_{112y}
121,x	n_{121x}
121,y	n_{121y}
…	…

one particle state	occup. number
1,1,1	1
1,1,2	3
1,2,1	0
1,2,2	2
2,1,1	0
2,1,2	0
2,2,1	12
2,2,2	4
⋮	⋮

Now to the **problem of "spin labels".** The particles we want to "invent and introduce" must be bosons, since the excitation numbers interpreted as occupational numbers are any integers. The bosons should not be spinless since we have to interpret the labels x, y as some spin-state labels. So far we do not know any bosons which would have just two spin states. Spin 1 bosons should have 3 "spin projections". Two spin projections suggest spin 1/2, but spin 1/2 particles are fermions!

However, the above mentioned classification is valid if we consider only nonrelativistic particles. Relativistic particles with the mass equal to zero with the total spin 1 have only two spin-projection states +1 and -1. Projection 0 is not possible. The reason is hidden deeply in relativistic quantum mechanics (better: field theory). We skip the detailed argumentation, just take it as a fact.

So there is a chance for a particle-like alternative description of the states of the electromagnetic field introducing relativistic mass=0 particles with spin 1 named photons. The "spin-projection labels" for photons do not necessarily mean true spin projections: other quantum mechanical base of "spin states" can be used equally well. For photons we most other use the "polarization basis".

Free electromagnetic field: photons

We found that the energy and momentum formulas for photons should be

$$\varepsilon_{mnl} = \frac{c\hbar\pi}{a}\sqrt{m^2+n^2+l^2} \qquad \vec{p}_{mnl} = \hbar\left(\frac{m\pi}{a}, \frac{n\pi}{a}, \frac{l\pi}{a}\right)$$

This gives us the following relation between energy and momentum

$$\boxed{\varepsilon = c|\vec{p}|}$$

Such a relation is expected for relativistic massless particle because of the general relativistic formula for the particle rest mass

$$\boxed{m_0^2 = \frac{1}{c^4}\varepsilon^2 - \frac{1}{c^2}|\vec{p}|^2}$$

For photons this gives the rest mass 0. The general formula for the relation between the energy, rest mass and velocity of a particle is $\varepsilon = \frac{m_0c^2}{\sqrt{1-\frac{v^2}{c^2}}}$

To get non-zero energy, the zero-mass particle velocity must always be $v = c$. So photons always move with the velocity c.

We have discovered photons !!!

Thermal radiation in a cavity as a photon gas

We have arrived at a conclusion that we can consider the electromagnetic field in a cavity as a photon gas. So we can try to calculate the properties of the thermal radiation by calculating the properties of a photon gas in equilibrium with the cavity walls at the temperature T.

How to do that? we certainly do not know a priori the total number of photons in the cavity so we try to work within the grand canonical technology pretending to know the chemical potential of photons, calculate with it as with a symbolical abstract number μ and finally find the correct value of it by adjusting this value to get the correct thermal radiation results which we already know.

So we have a the photon gas with one-particle-states mnls. The mean occupational number a one-particle-state will be (Bose-Einstein)

$$n_{mnls} = \frac{1}{\exp\left(\frac{\hbar\omega_{mnl}-\mu}{kT}\right) - 1}$$

The total energy of the photon gas will be

$$\overline{E} = \sum_{mnls} \frac{\hbar\omega_{mnl}}{\exp\left(\frac{\hbar\omega_{mnl}-\mu}{kT}\right) - 1}$$

Thermal radiation in a cavity as a photon gas

Let us compare the formula we have just got with the formula we have got for "equivalent oscillators" where we had "mean excitation numbers" instead of "mean occupational numbers"

$$\overline{E} = \sum_{mnls} \frac{\hbar\omega_{mnl}}{\exp\left(\frac{\hbar\omega_{mnl}-\mu}{kT}\right) - 1} \quad \longleftrightarrow \quad \overline{E} = \sum_{mnls} \frac{\hbar\omega_{mnl}}{\exp\left(\frac{\hbar\omega_{mnl}}{kT}\right) - 1}$$

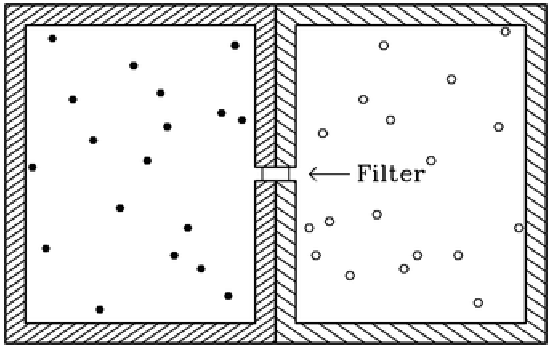
The formula for ω_{mnl} is the same in both cases, the only difference is that there is μ in the formula on the left. The conclusion is: we get exactly the same results for the thermal radiation if calculated as for oscillators and if calculated as for photons **if we set $\boldsymbol{\mu = 0}$**.

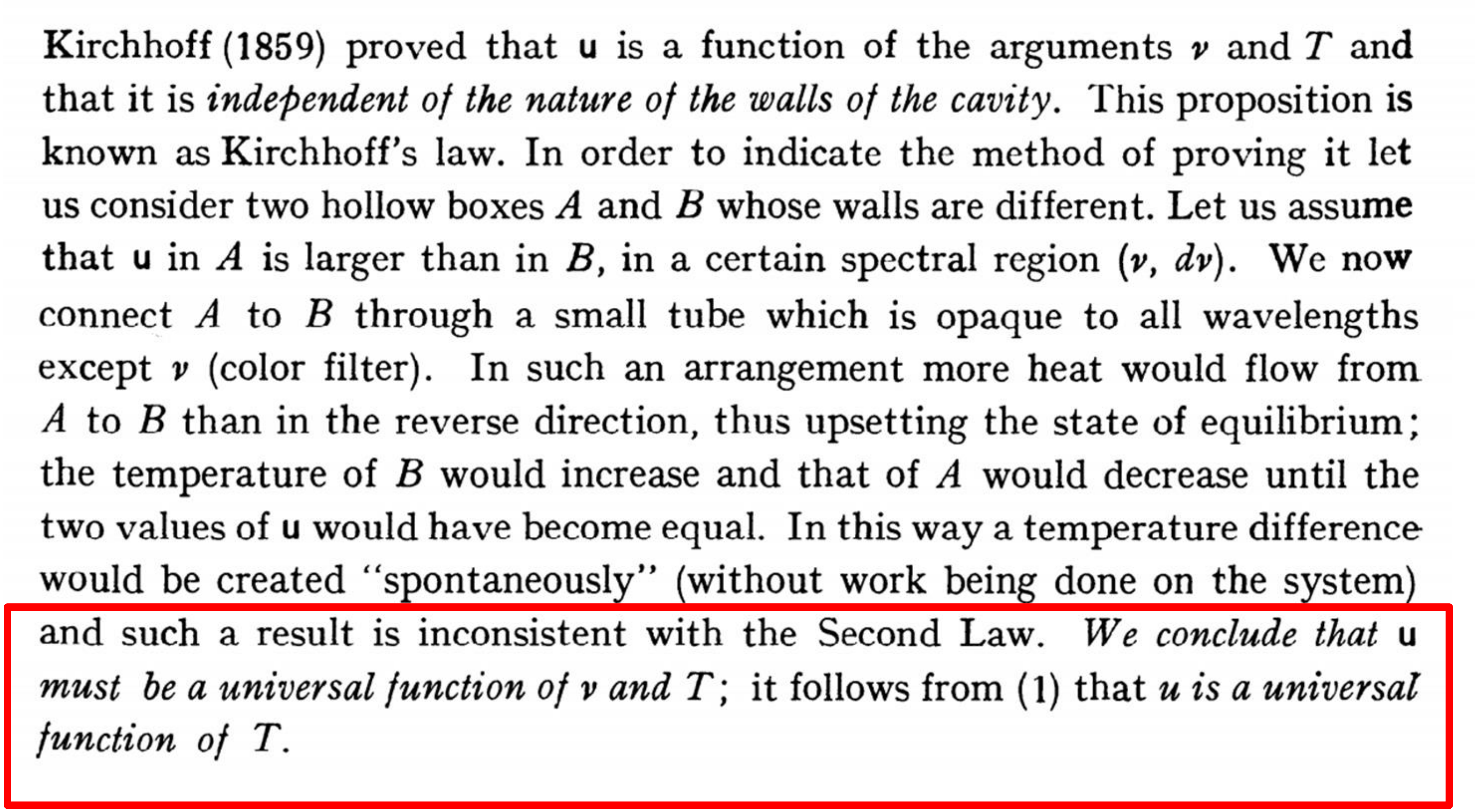
We discovered a new law of nature: chemical potential of photons is (always) zero.

Putting $\mu = 0$ we continue in exactly the same way as we did for the "equivalent oscillators": we replace the sum by an integral using formula for the **density of one-particle-states which would be for photons the same as was the formula for the density of oscillators** and we will get the same result

$$u(\omega) = \frac{\hbar\omega^3}{\pi^2 c^3} \frac{\hbar\omega}{\exp\left(\frac{1}{kT} - 1\right)}$$

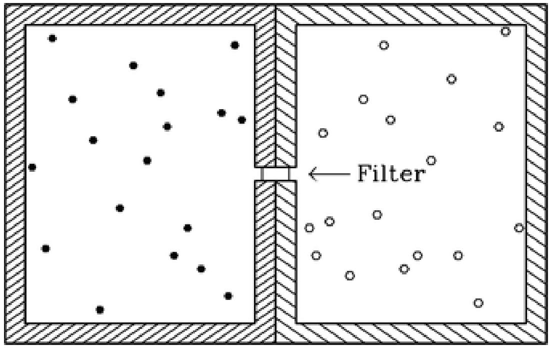
Kirchhoff Sommerfeld thermodynamic universality proof





Kirchhoff (1859) proved that $\mathbf{u}$ is a function of the arguments ν and T and that it is *independent of the nature of the walls of the cavity*. This proposition is known as Kirchhoff's law. In order to indicate the method of proving it let us consider two hollow boxes A and B whose walls are different. Let us assume that $\mathbf{u}$ in A is larger than in B, in a certain spectral region $(\nu, d\nu)$. We now connect A to B through a small tube which is opaque to all wavelengths except ν (color filter). In such an arrangement more heat would flow from A to B than in the reverse direction, thus upsetting the state of equilibrium; the temperature of B would increase and that of A would decrease until the two values of $\mathbf{u}$ would have become equal. In this way a temperature difference would be created "spontaneously" (without work being done on the system) and such a result is inconsistent with the Second Law. *We conclude that* $\mathbf{u}$ *must be a universal function of* ν *and* T; it follows from (1) that *u is a universal function of* T.

Kirchhoff Sommerfeld thermodynamic universality proof applied to ordinary ideal gas. Why not, particles like particles



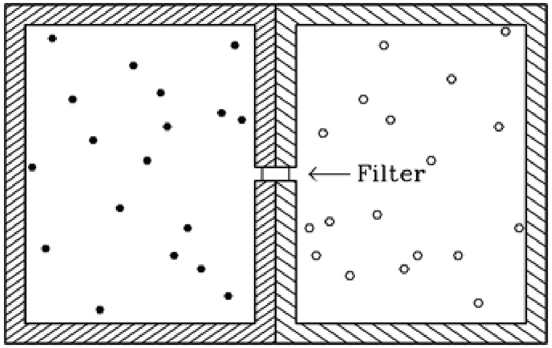
Two containers with monoatomic ideal gas, same temperatures, common hole with filter allowing to go through only to particles with specified $v = |\vec{v}|$ that is with specified kinetic energy.

Let the energy density per unit of absolute velocity will be denoted as $u(v)$ than the spectral energy flow through the hole would be

$$\frac{1}{4} v u(v)$$

If the energy density is not the same on both sides, them net energy flow would result, **changing temperatures originally equal ! Forbidden by thermodynamics. Conclusion: the energy density in an ideal gas is a universal function of temperature. That is grossly wrong! But the logical syntax is the same as in Kirchhoff Sommerfeld proof, just semantics is different.**

Kirchhoff Sommerfeld thermodynamic universality proof applied to ordinary ideal gas



Where is the error?

$$u(v) = n\frac{1}{2}mv^2 \qquad\qquad n = \frac{N}{V}$$

For standard ideal gas there is **additional independent parameter: particle density**. The containers can have the same temperature but different particle densities (different pressures). After making the hole pressures will be equalized, the temperatures will not change.

The energy flow does not necessarily change the temperature for ordinary gas

For radiation Sommerfeld calls the energy flow as **heat flow (wärmestrahlung).**

What is the difference between the ordinary gas and the photon gas?

Why is there no independent parameter like photon density for the photon gas?

Because photons have zero chemical potential.

Actually the logic is reverse: photon have zero chemical potential because there is **no independent parameter** like photon (spatial) density

Ordinary gas

one particle state	occup. number
1,1,1	1
1,1,2	3
1,2,1	0
1,2,2	2
2,1,1	0
2,1,2	0
⋮	⋮

$$\sum = N$$

Oscillators

mnl,s	excit.number
111,x	n_{111x}
111,y	n_{111y}
112,x	n_{112x}
112,y	n_{112y}
121,x	n_{121x}
…	…

$$\overline{E} = \sum_{mnls} \frac{\hbar\omega_{mnl}}{\exp\left(\frac{\hbar\omega_{mnl}-\mu}{kT}\right)-1}$$

$$\overline{E} = \sum_{mnls} \frac{\hbar\omega_{mnl}}{\exp\left(\frac{\hbar\omega_{mnl}}{kT}\right)-1}$$

No state independent sum here

What is the role of the number of photons in a box?

Parameter of non-equilibrium

Start with empty box at $T = 0$. Heat the walls suddenly to T.
No photons inside, so the system is not in equilibrium.
Observe relaxation to equilibrium: $N(t)$
Number of photons finally reaches the equilibrium value $\bar{N}$.

How to recognize equilibrium at a given temperature:
Free energy is minimal

$$F_{noneq}(T, V, N) \to F_{eq}(T, V)$$

Close to equilibrium

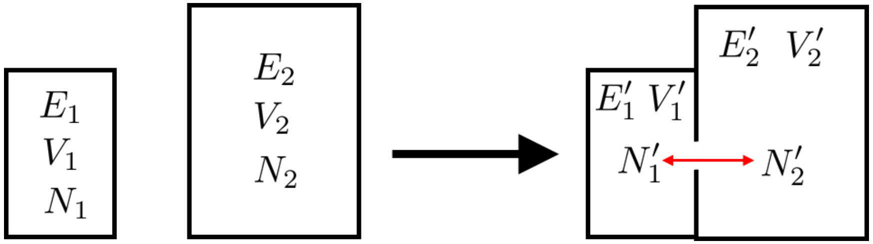
$$\frac{\partial F_{noneq}(T, V, N)}{\partial N} = 0$$

$$\boxed{\bar{N} = \mathrm{argmin}_N F_{noneq}(T, V, N)}$$

Landau:

$$\frac{\partial F(T, V, N)}{\partial N} = \mu \implies \mu = 0$$

Chemical potential definition



After the final common equilibrium is established the total entropy will be

$$S' = S_1(E_1', V_1, N_1') + S_2(E_2', V_2, N_2')$$

The final values of energies will be the same as in the case when we have considered just a thermal contact. The only unknown value in this relation is N_1', since N_2' is given by the total **number of particles conservation** $N_2' = N_1 + N_2 - N_1'$.

$$\frac{dS'(N_1')}{dN_1'} = \left.\frac{\partial S_1}{\partial N}\right|_{N=N_1'} - \left.\frac{\partial S_2}{\partial N}\right|_{N=N_1+N_2-N_1'} = 0$$

$$\left.\frac{\partial S_1}{\partial N}\right|_{N=N_1'} = \left.\frac{\partial S_2}{\partial N}\right|_{N=N_2'}$$

$$\boxed{\frac{\mu}{T} = -\frac{\partial S}{\partial N}}$$

Photon strange stoichiometry

$$2H_2 + O_2 \leftrightarrow 2H_2O$$

$$2\mu_{H_2} + \mu_{O_2} = 2\mu_{H_2O}$$

$$\boxed{\begin{gathered}\text{nothing} \leftrightarrow \gamma \\[1em] 0 = \mu_\gamma\end{gathered}}$$

What is the difference between the ordinary gas and the photon gas?

Ordinary gas

one particle state	occup. number
1,1,1	1
1,1,2	3
1,2,1	0
1,2,2	2
2,1,1	0
2,1,2	0
⋮	⋮

$$\sum = N$$

Oscillators

mnl,s	excit.number
111,x	n_{111x}
111,y	n_{111y}
112,x	n_{112x}
112,y	n_{112y}
121,x	n_{121x}
…	…

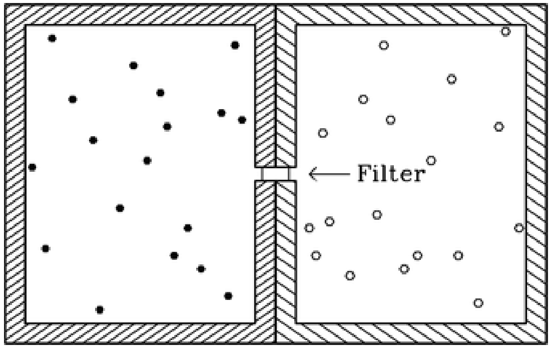
Bose-Einstein distribution is derived by grand canonical technology. Needs conservation of the total number of particles!

No grand canonical calculations for photon gas!

Well, you can do grand canonical calculations but then you have to understand well why you can do it.

No state independent sum here

Kirchhoff Sommerfeld thermodynamic universality proof



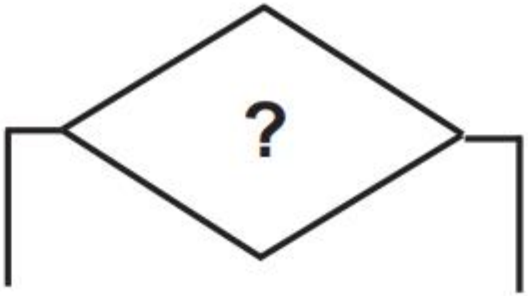
We conclude that u must be asymptotically a universal function of ω and T .

$$u(\omega) = \frac{1}{V} \frac{\hbar\omega}{\exp\left(\frac{\hbar\omega}{kT}\right) - 1} \rho(\omega)$$

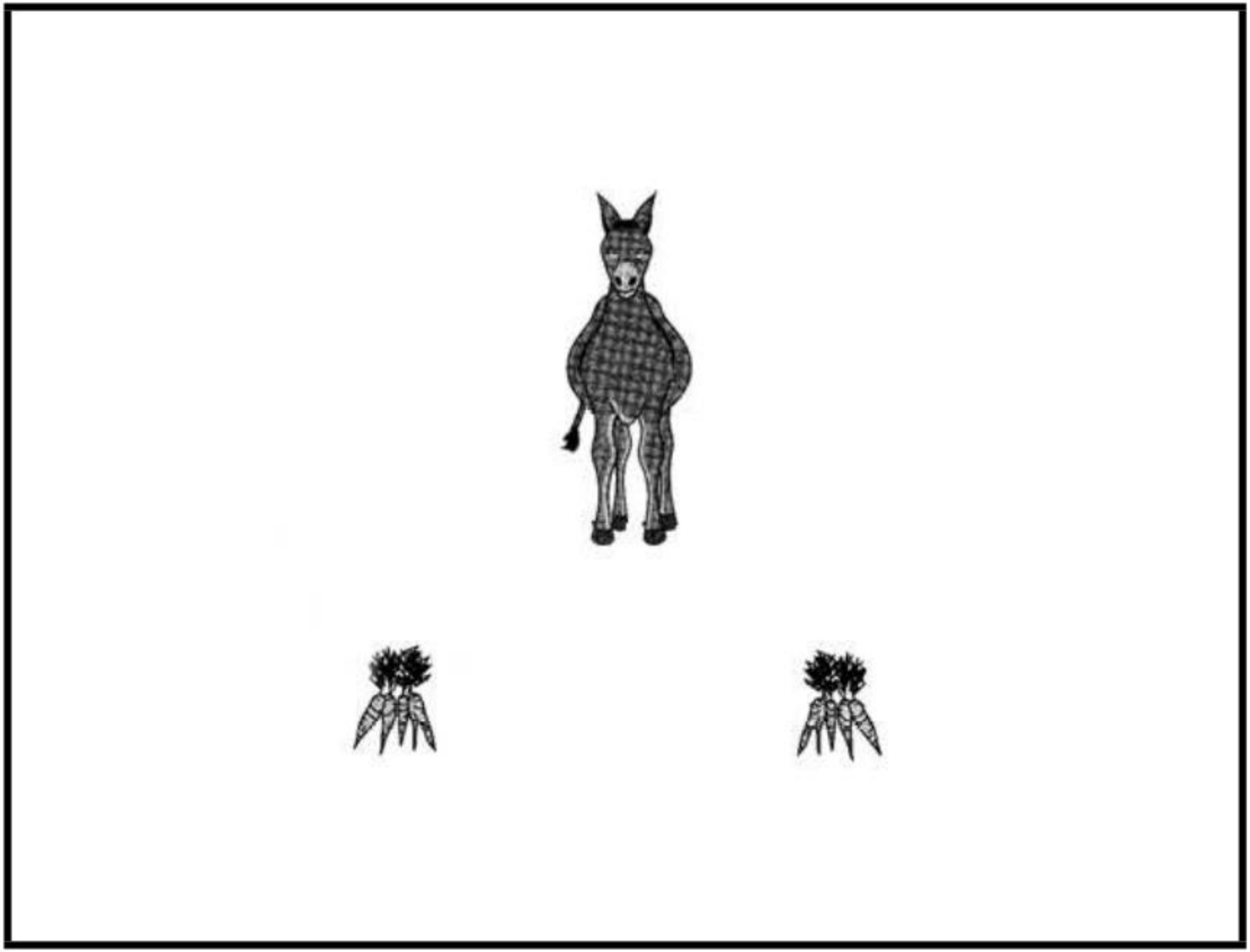
If $u(\omega)$ is asymptotically universal, then $\rho(\omega)$ is asymptotically universal as well, that is **Laplacian spectrum is asymptyotically universal for a cavity of arbitrary shape.**

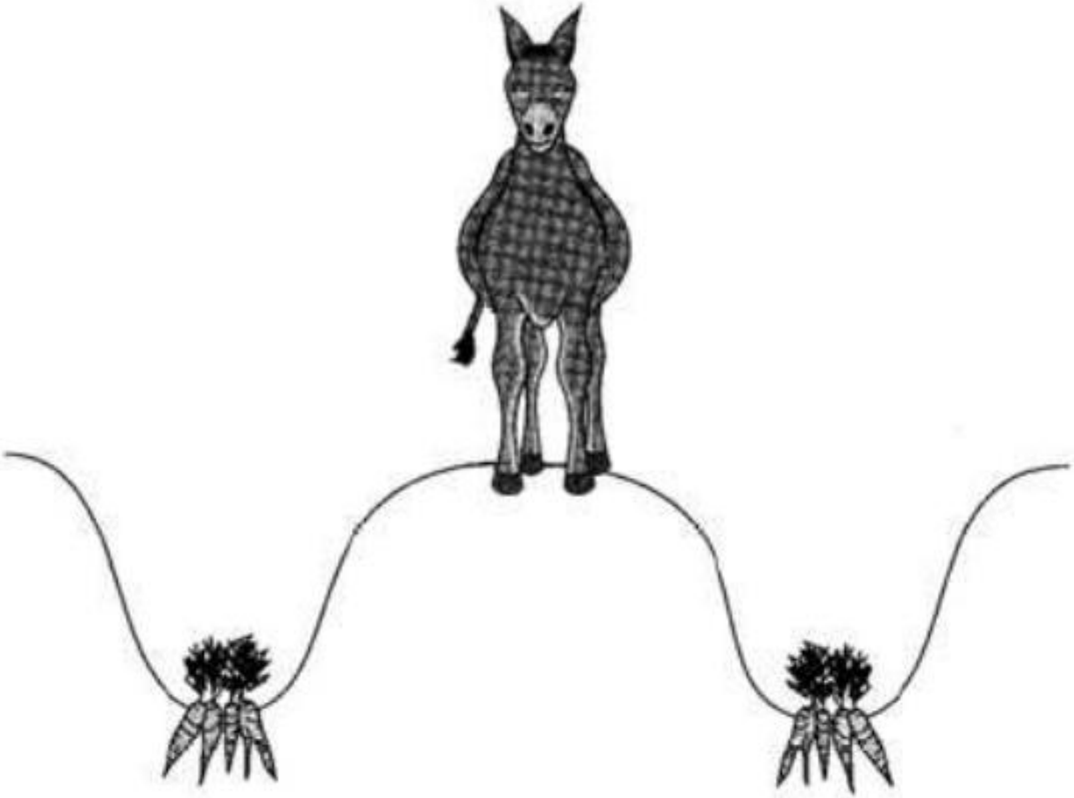
The end

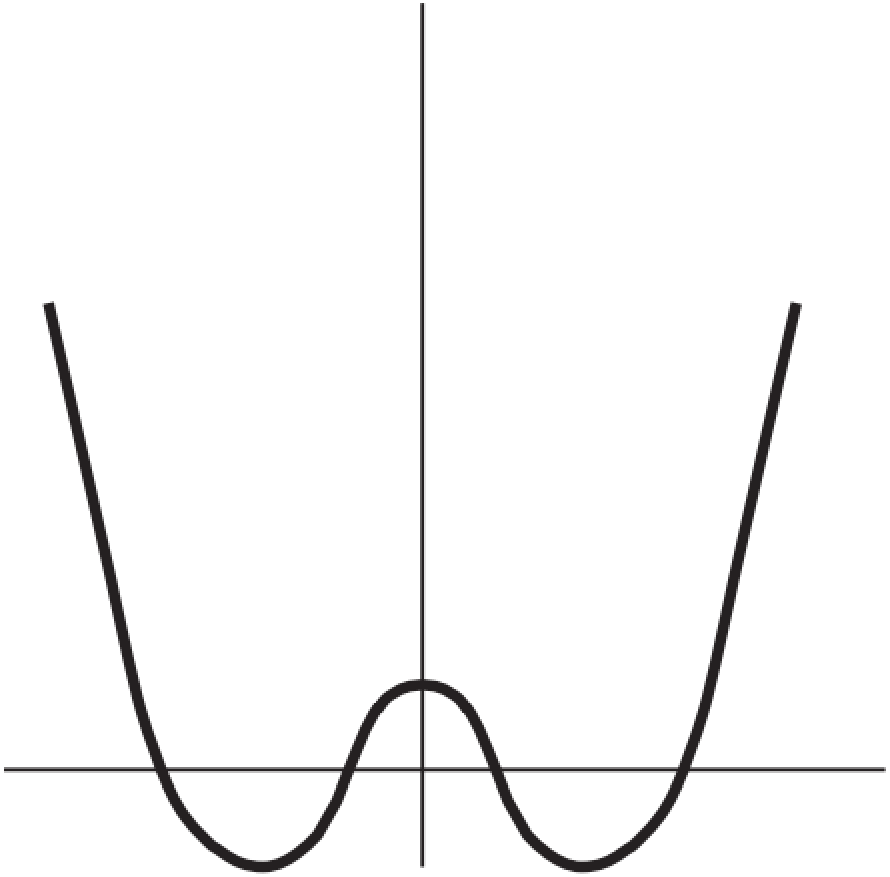
Energia, entropia, informácia a inteligencia

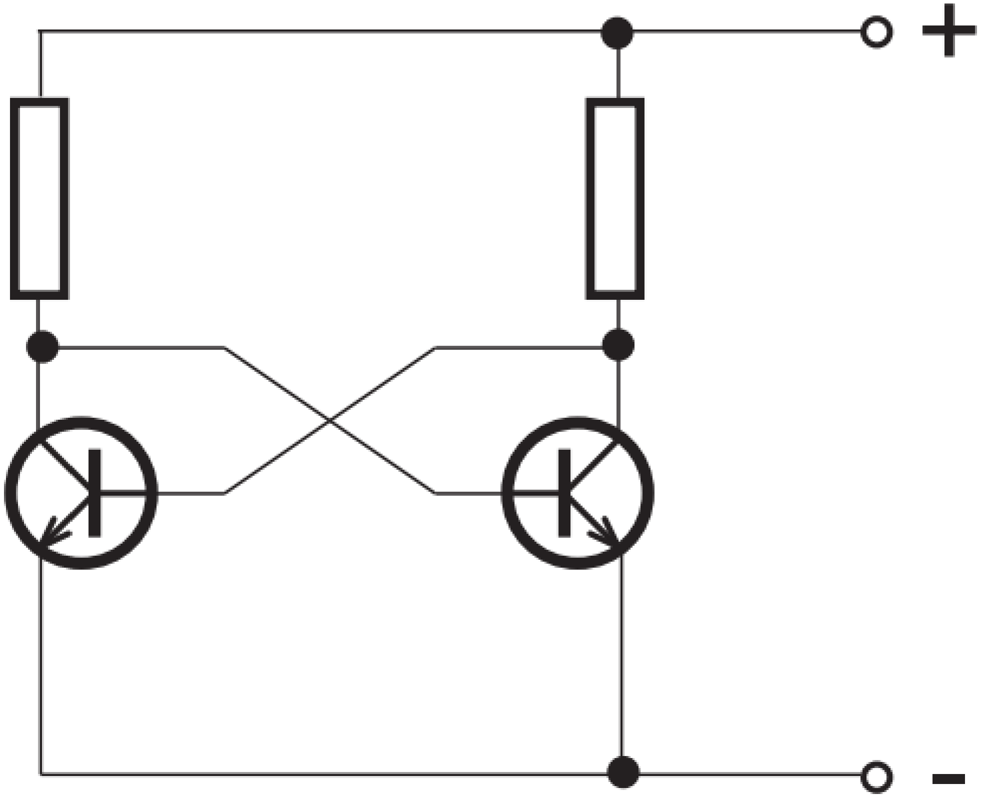


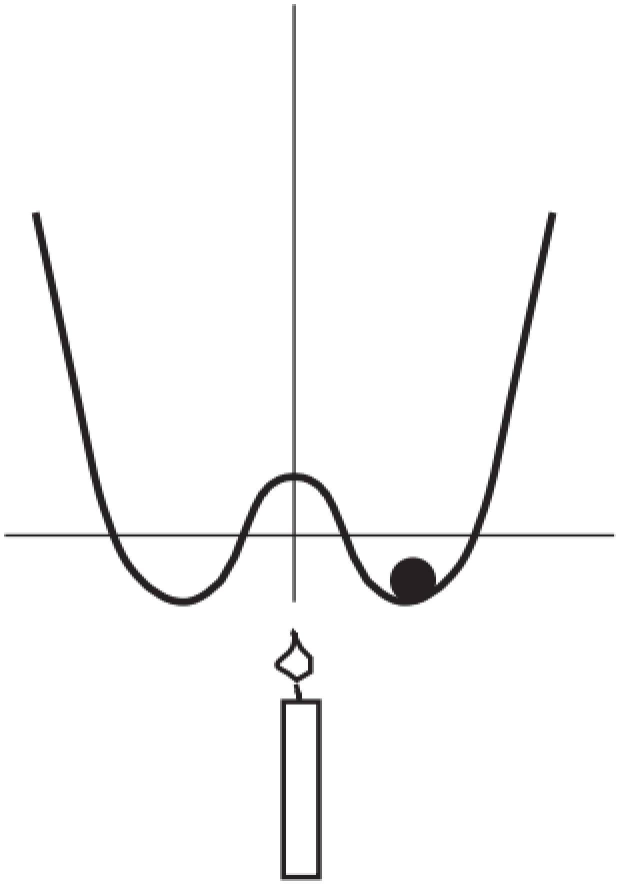
?



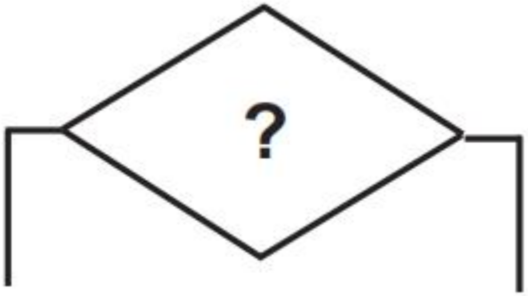








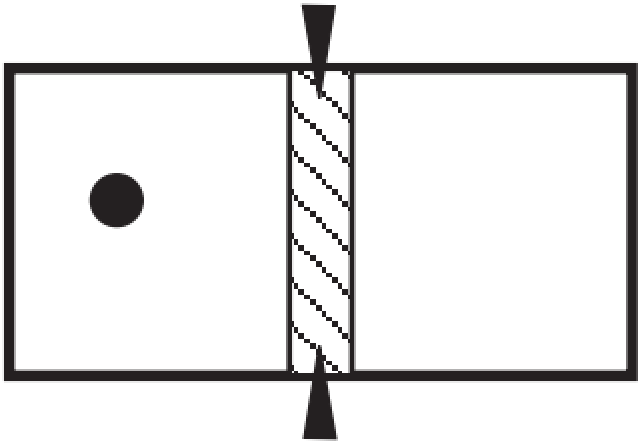
Pri T = 300 K, energia molekúl je rádove 0.02 eV.



Agent, ktorý sa rozhoduje, musí mať pamäťový register, aby vedel realizovať konštrukciu

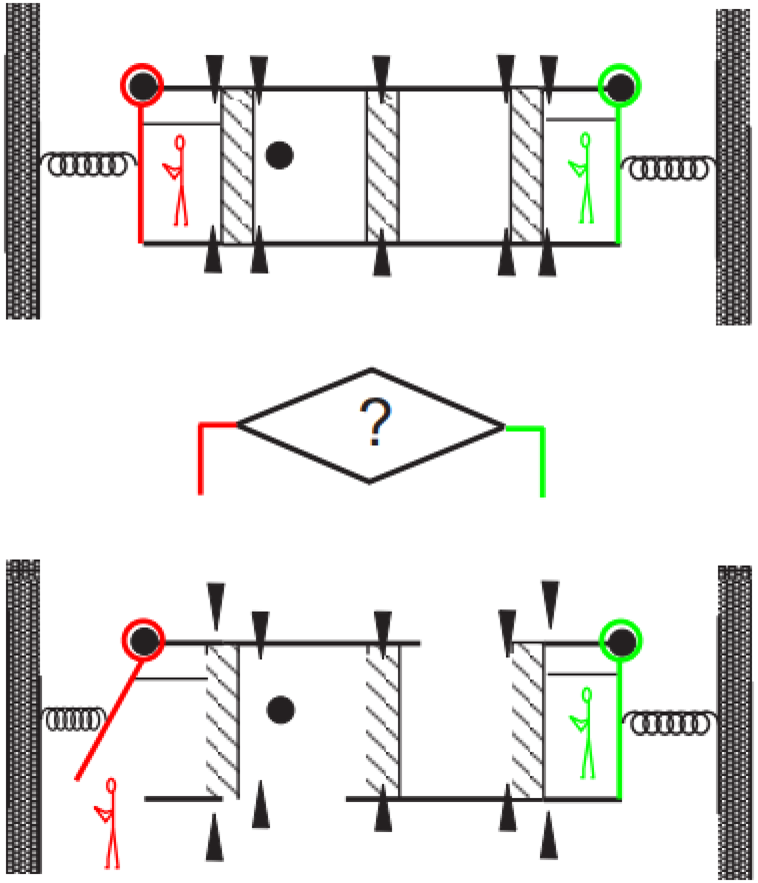
if – then – else

Feynmanov register

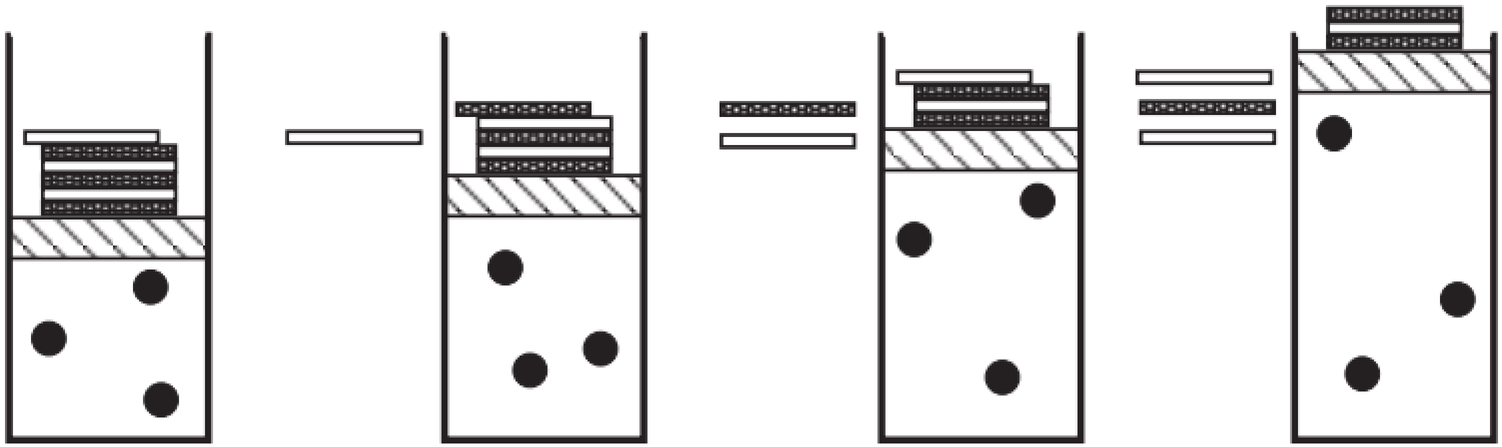


$$pV = NkT$$

CPU register



Vratná izotermická expanzia plynu



$$pV = NkT$$

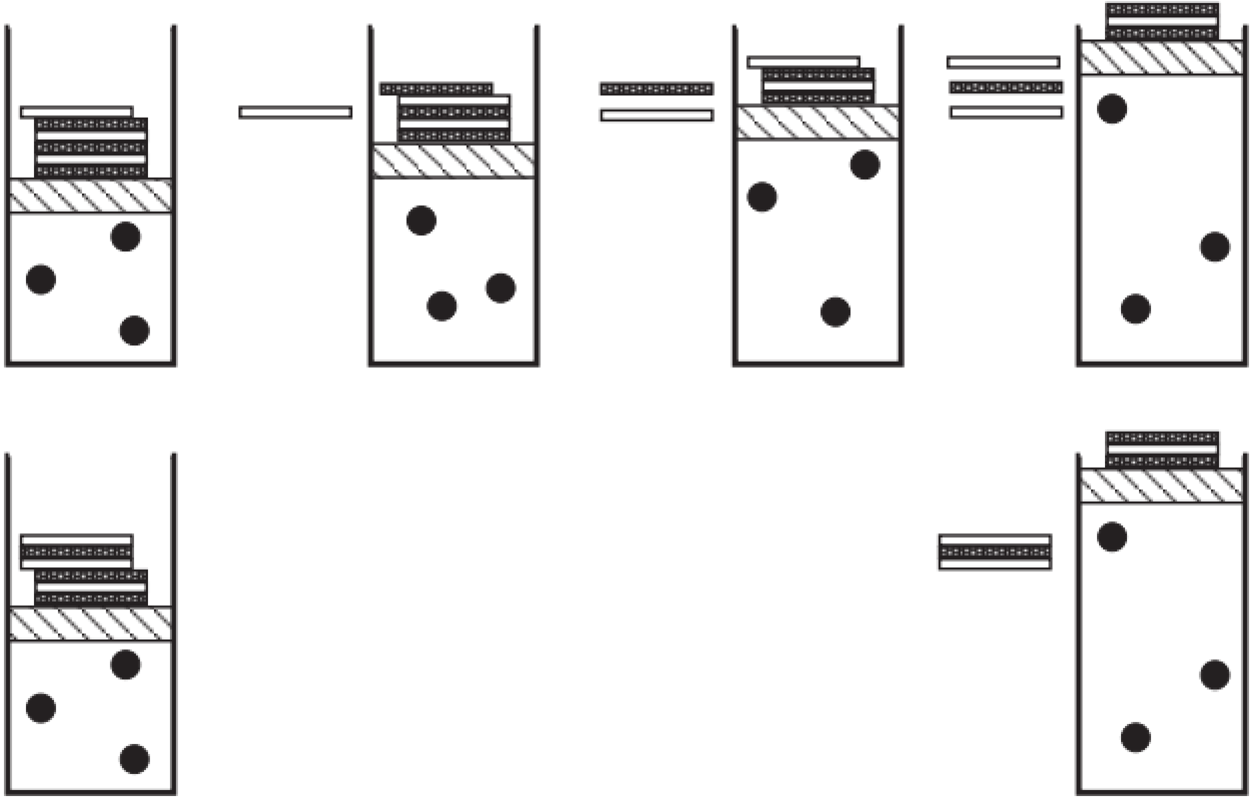
Práca pri posune piesta o dx je

$$\delta A = F\,dx = pS\,dx = p\,dV$$

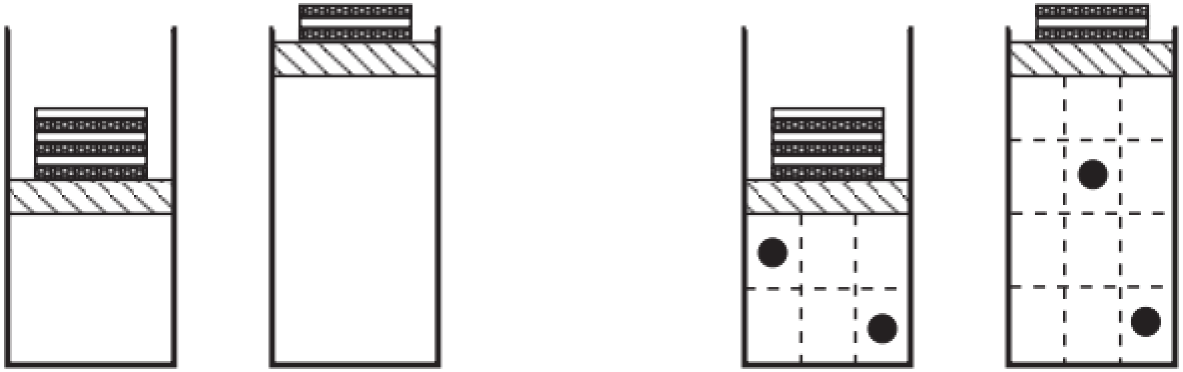
Celková práca

$$A = \int p\,dV = NkT \int \frac{dV}{V} = NkT \ln(2)$$

Nevratná expanzia plynu



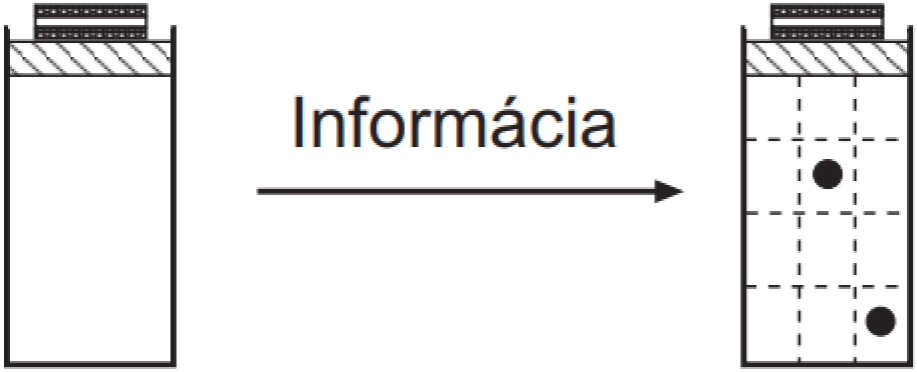
Zmena počtu mikrostavov pri expanzii



Počet možných mikrostavov pred expanziou: 15, teda na špecifikáciu konkrétneho stavu treba približne 4 bity

Počet možných mikrostavov po expanzii: 66, teda na špecifikáciu konkrétneho stavu treba približne 6 bitov

Ω: počet mikrostavov



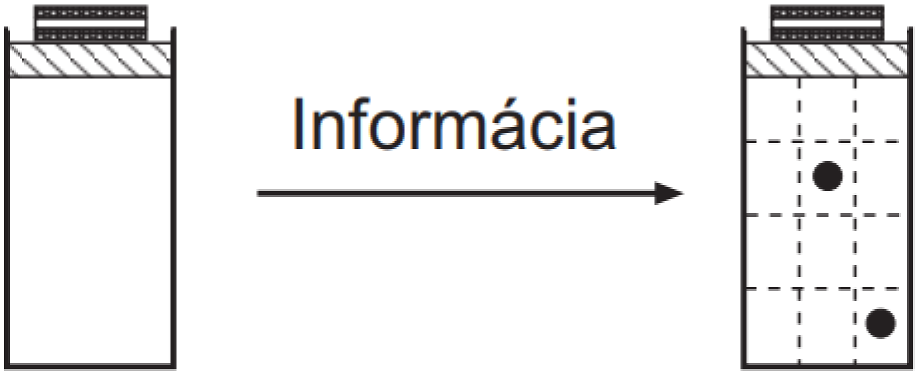
Počet častíc: N, **počet priehradiek:** C

$$\ln\Omega = \ln\frac{C!}{N!(C-N)!} \approx N\ln C$$

počet bitov čísla $\Omega = \log_2(e) \times \ln\Omega$

počet tritov čísla $\Omega = \log_3(e) \times \ln\Omega$

Entropia



Počet častíc: N, počet priehradiek: C

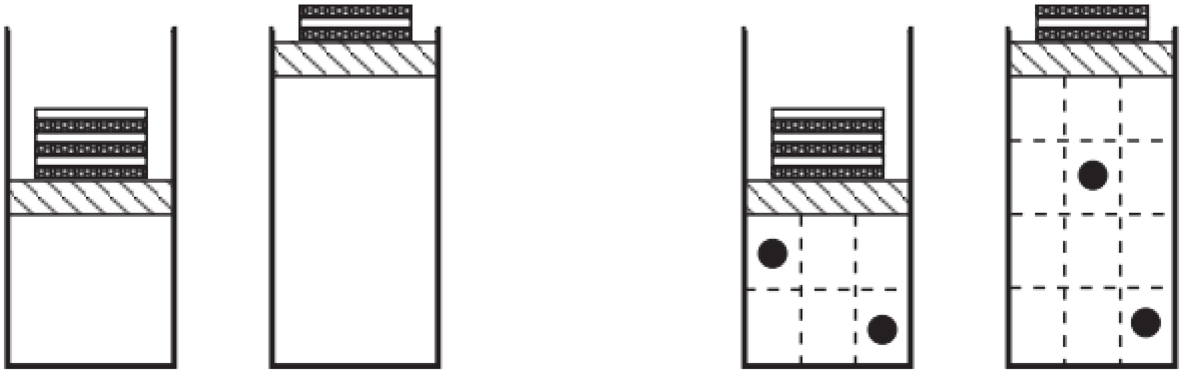
$$\ln \Omega \approx N \ln C$$

Chýbajúca informácia sa nazýva

entropia

$$S = N \ln C$$

Zmena entropie pri expanzii



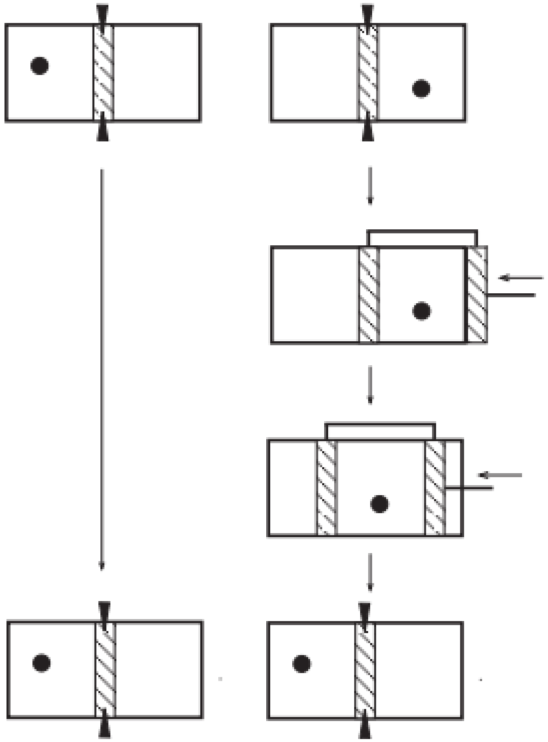
$$\Delta S = S_2 - S_1 = N \ln(2C) - N \ln C = N \ln 2$$

Práca vykonaná pri expanzii $A = kTN \ln 2$

Okolie udržiavalo konštantnú teplotu, dodalo množstvo tepla $Q = A$

$$Q = kT\,\Delta S$$

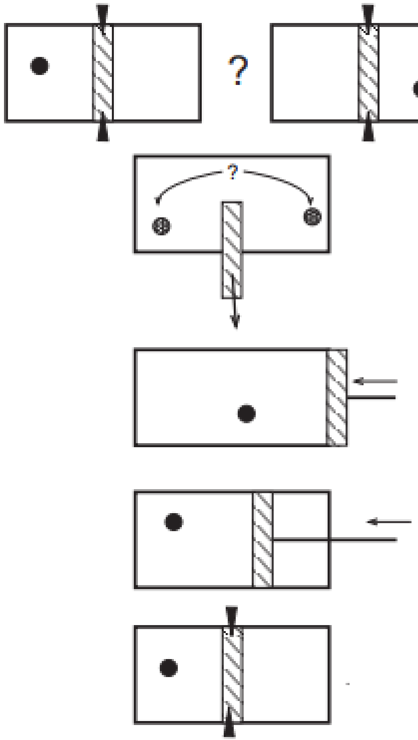
Manipulácie s registrom



Vychádzajúc so známeho stavu nastav na 0

Procedúra beží bez nároku na vykonanie práce.

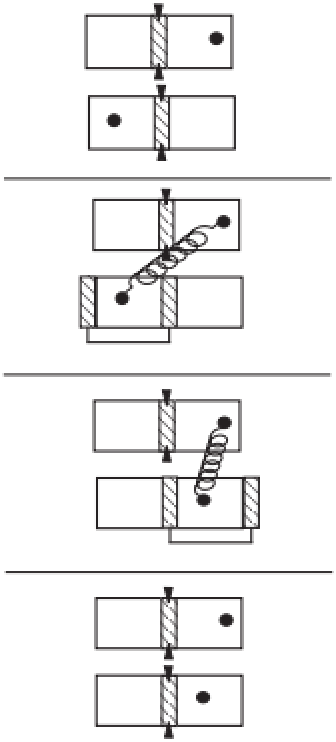
Manipulácie s registrom



Vychádzajúc s neznámeho stavu nastav na 0

Treba vykonať prácu $A = kT \ln 2$

Manipulácie s registrom



Okopíruj register v neznámom stave do registra v defaultovom stave

Manipulácie s registrom

Ak je register v známom stave, môžem s ním urobiť akúkoľvek operáciu bez nároku na prácu.

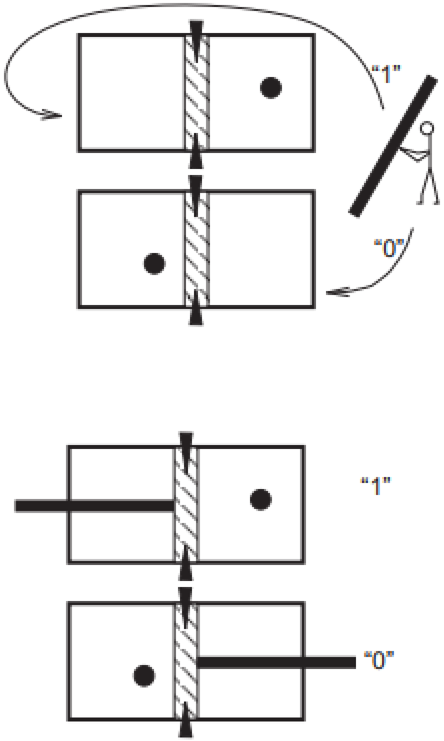
Operácie typu "set" však vyžadujú proces typu if-then-else a teda pamäťový register, s ktorým mám vykonať želanú operáciu musí byť okopírovaný do registra procesora. Po vykonaní želanej operácie v registri procesora ostane kópia pôvodného stavu.

Register môžem resetovať do defaultového stavu univerzálnou procedúrou (bez if-then-else, teda bez potreby kópie v registri procesora), ale na takú operáciu treba vykonať prácu:

$$A = kT \ln 2$$

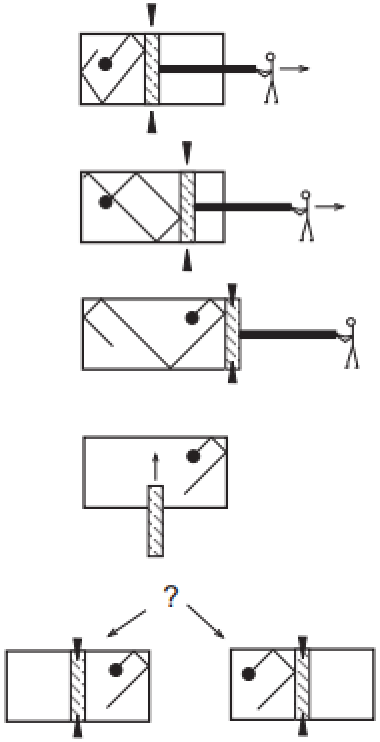
Prácu (energiu) treba na jedinú procedúru: na likvidovanie informácie, teda na resetovanie registrov

Extrakcia práce na základe informácie



Príprava na extrakciu užitočnej práce

Extrakcia práce



Extrakcia užitočnej práce za cenu privedenia registra do neznámeho stavu, teda za cenu zvýšenia jeho entropie.

Extrakcia práce

Dodací list: 0,1,1,0,1,...

Induktívna inferencia

$$2, 4, 6, 8, \ldots$$

Induktívna inferencia

$$2, 4, 6, 8, 34$$

Induktívna inferencia

$$2, 4, 6, 8, 34$$

$$n^4 - 10n^3 + 35n^2 - 48n + 24$$

Wiliam of Ockham

1285–1349

Ockhamova britva

Pluralitas non est ponenda sine necessitate.

Induktívna inferencia

$$2, 4, 6, 8, \ldots$$

$$n^4 - 10n^3 + 35n^2 - 48n + 24$$

$$2n$$

Kolmogorovova zložitosť

Dátové reťazce ("dodacie listy")

- **0,1,0,1,0,1,0,1,0,1,0,1,0,1,0,1,0,1,0,1**
- **0,1,0,0,1,0,0,0,0,1,0,0,0,0,0,1,0,0,0,0**
- **0,0,1,0,1,1,0,1,0,0,0,1,1,0,1,0,1,0,0,1**

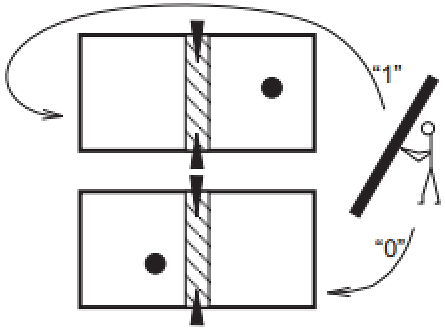
$p_U(x_n)$ **je program v jazyku** U **ktorý vygeneruje postupnosť** x_n **a zastaví sa.**

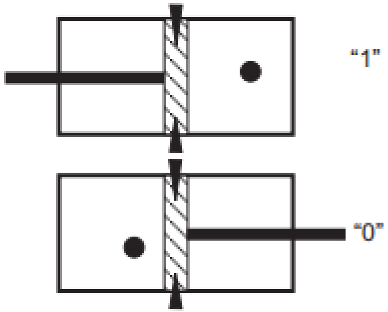
$$K_U(x_n) = \min |p_U(x_n)|$$

Extrakcia práce

Dodací list: ”samé nuly”

Extrakcia práce na základe informácie



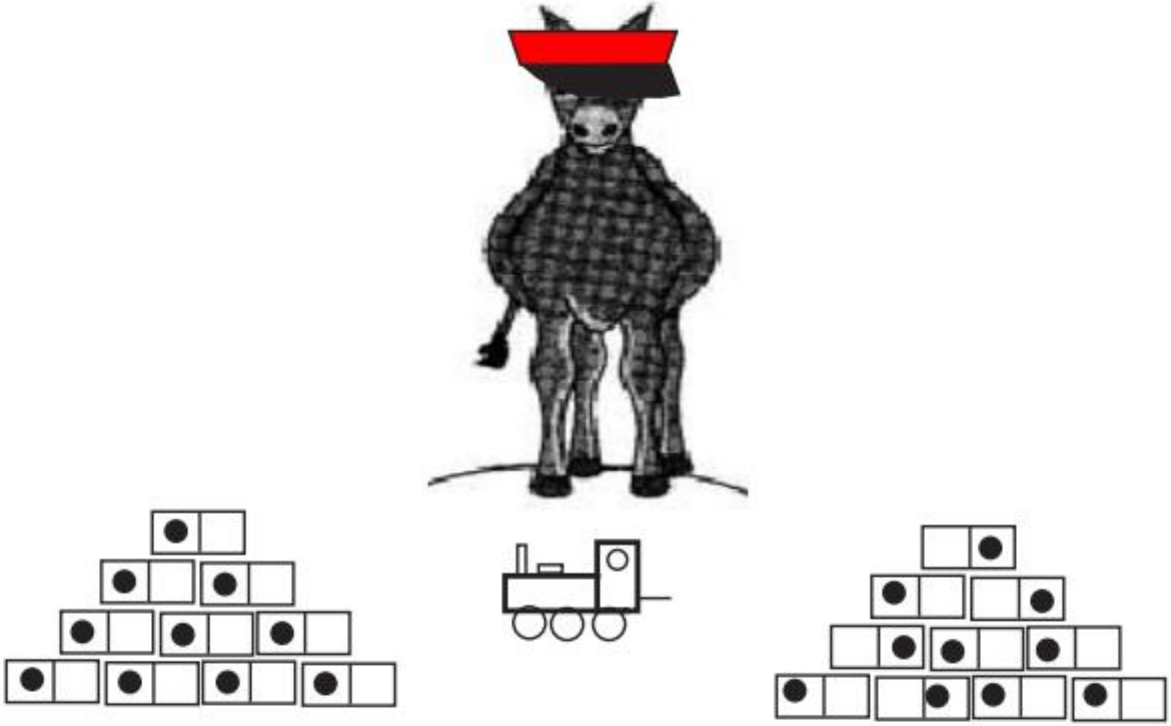


Príprava na extrakciu užitočnej práce

Ako sa nakladá takýto vlak

Aby sa nestalo, že výpravca vlaku potrebuje na jeho naloženie rovnakú prácu, ako príjemca získa

Výpravca a induktívna inferencia

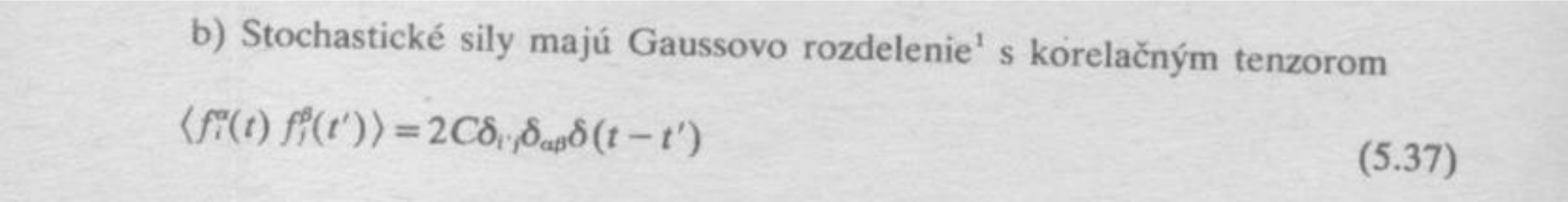


Heureka!

$2, 4, 6, 8, \ldots$

Induktívna inferencia nedáva istotu, poskytuje len šancu.

Prežijú iba inteligentní hazardéri.



b) Stochastické sily majú Gaussovo rozdelenie[1] s korelačným tenzorom

$$\langle f_i^{\alpha}(t)\, f_j^{\beta}(t') \rangle = 2C\delta_{ij}\delta_{\alpha\beta}\delta(t-t') \qquad (5.37)$$

$$m\frac{d\dot{x}}{dt} = -\alpha\dot{x} + f'$$

For strong dumping we can neglect in the Langevin equation the inertial term and we get a Aristotle type of equation

$$0 = -\alpha\dot{x} + f'$$
$$\dot{x} = \frac{1}{\alpha}f'$$
$$x(t) = x(0) + \int_0^t d\tau\frac{f'(\tau)}{\alpha}$$

For the initial condition $x(0) = 0$ we get

$$\langle x^2(t)\rangle = \frac{1}{\alpha^2}\int_0^t d\tau_1 d\tau_2\langle f'(\tau_1)f'(\tau_2)\rangle$$

Now the correlation depends only on the time difference and we get

$$\langle x^2(t)\rangle = \frac{1}{\alpha^2}\int_0^t d\frac{\tau_1+\tau_2}{2}d(\tau_1-\tau_2)\langle f'(0)f'(\tau_1-\tau_2)\rangle$$

$$\langle x^2(t)\rangle = \frac{1}{\alpha^2}\int_0^t d\frac{\tau_1+\tau_2}{2} d(\tau_1-\tau_2)\langle f'(0)f'(\tau_1-\tau_2)\rangle$$

$$\langle x^2(t)\rangle = \frac{t}{\alpha^2}\int_{-\infty}^{\infty} d\tau \langle f'(0)f'(\tau)\rangle$$

where we extended the region of integration formally to $(-\infty, \infty)$ since the correlation function is different from zero only on a very small interval and so the integration region does not play any role.

We, however, know from previous calculation that

$$\langle x^2\rangle = \frac{2kT}{\alpha}t$$

and comparing the two results we get

$$\alpha = \frac{1}{2kT}\int_{-\infty}^{\infty} d\tau \langle f'(0)f'(\tau)\rangle$$

So the dissipation coefficient α is given by the correlations of the fluctuation force. This is so called fluctuation-dissipation theorem.

Integral runs only through a very short correlation time interval, so to get finite α the correlation function must be very large at $\tau = 0$, it is approximately a delta function.

Let us consider a cylinder with a gas of particles with a piston in the middle. (See Fig.1.) To make the calculations simple let us consider the gas as one dimensional, the molecules moving only along the cylinder axis. The molecules strike against the piston from both sides. If the pressure on both sides is equal, then the mean net force on the piston is zero. However the net force on the cylinder is fluctuating around the mean zero value. Our aim is to estimate the fluctuations. The situation is similar, but not equivalent, to the case of brownian particle in the gas as described by the Langevin equation with a fluctuating force at the right hand side.

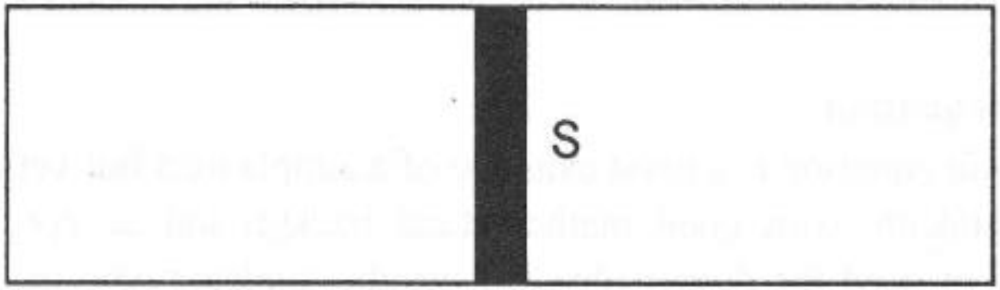


Fig. 1

If we denote the fluctuating force at the time t as $f(t)$ we can define the autocorrelation function as

$$\left\langle f(0)f(t)\right\rangle$$

If we denote the fluctuating force at the time t as $f(t)$ we can define the autocorrelation function as

$$\langle f(0)f(t)\rangle$$

where the angle brackets denote the average over the ensemble of equivalent cylinders with pistons. If the time difference t in the autocorrelation function is large, the values of $f(0)$ and $f(t)$ are independent and therefore

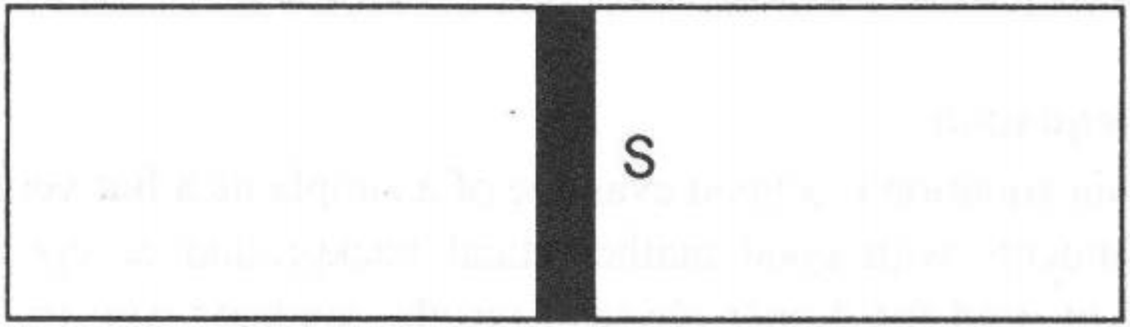
$$\langle f(0)f(t)\rangle = \langle f(0)\rangle\langle f(t)\rangle = 0$$

Large t means t larger then a certain correlation time τ_c describing how long the molecules remember their state.

For the short times $t \approx \tau_c$ we expect large correlation between *f(0)* and *f(t)*. To show how large, we try to estimate the mean square value of the fluctuation force.

.Within the time interval τ_c the molecules which are in small volumes within the distance *l* from the piston strike at the piston. Here *l* denotes the mean free path of the molecules. The molecules within that distance from the piston typically do not collide between themselves, those which have velocities in the direction towards the piston strike at the piston. The next "bunch" of molecules coming after the time τ_c has suffered collisions between themselves and therefore their state does not remember what was the state of the molecules in the previous bunch. Very roughly we can therefore represent the value *f(0)* as the value of the force during the time interval τ_c averaged to smear out the impacts of individual molecules. The net impulse on the piston is therefore given as

$$f(0)\tau_c = \left(\frac{1}{2}n_{left} - \frac{1}{2}n_{right}\right)2m|v|$$



$$f(0)\tau_c = \left(\frac{1}{2}n_{left} - \frac{1}{2}n_{right}\right)2m|v|$$

where n_{left} and n_{right} denote the number of molecules which strike at the piston from the left and from the right during the time τ_c. From now on we use the approximation that all the molecules have the same absolute value of velocity $|v|$. If we now average over the ensemble of equivalent gas systems we get

$$\langle f(0)\rangle\tau_c = \langle n_{left} - n_{right}\rangle m|v| = 0\,.$$

$$\left\langle f^2(0)\right\rangle \tau_c^2 = \left\langle \left(n_{left} - n_{right}\right)^2 \right\rangle \left(m|v|\right)^2$$

$$\left\langle f^2(0)\right\rangle \tau_c^2 = \left\langle n_{left}^2 + n_{right}^2 - 2n_{left}n_{right} \right\rangle \left(m|v|\right)^2$$

$$\left\langle f^2(0)\right\rangle \tau_c^2 = 2\left(\left\langle n^2 \right\rangle - \left\langle n \right\rangle^2\right)\left(m|v|\right)^2$$

We see that size of the net force fluctuations is given by the fluctuation of the number of molecules striking at the piston. We stress that we do not consider here the fluctuations in molecule velocities which would contribute as well to the fluctuation of force. We have to estimate the fluctuation of the number of molecules in the small volume of gas as given by the surface S of the piston and the mean free path l. The mean number of molecules (over the ensemble) in this volume is

$$\left\langle n \right\rangle = \frac{lS}{V} N$$

where V denotes the volume of the cylinder on one side of the piston and N is the total number of molecules again on one side of the piston.

To calculate the fluctuation one has to use the binomial distribution. Since the volume lS is very small wit respect to the total volume V the variance of the number of particles within the small volume can be approximated as

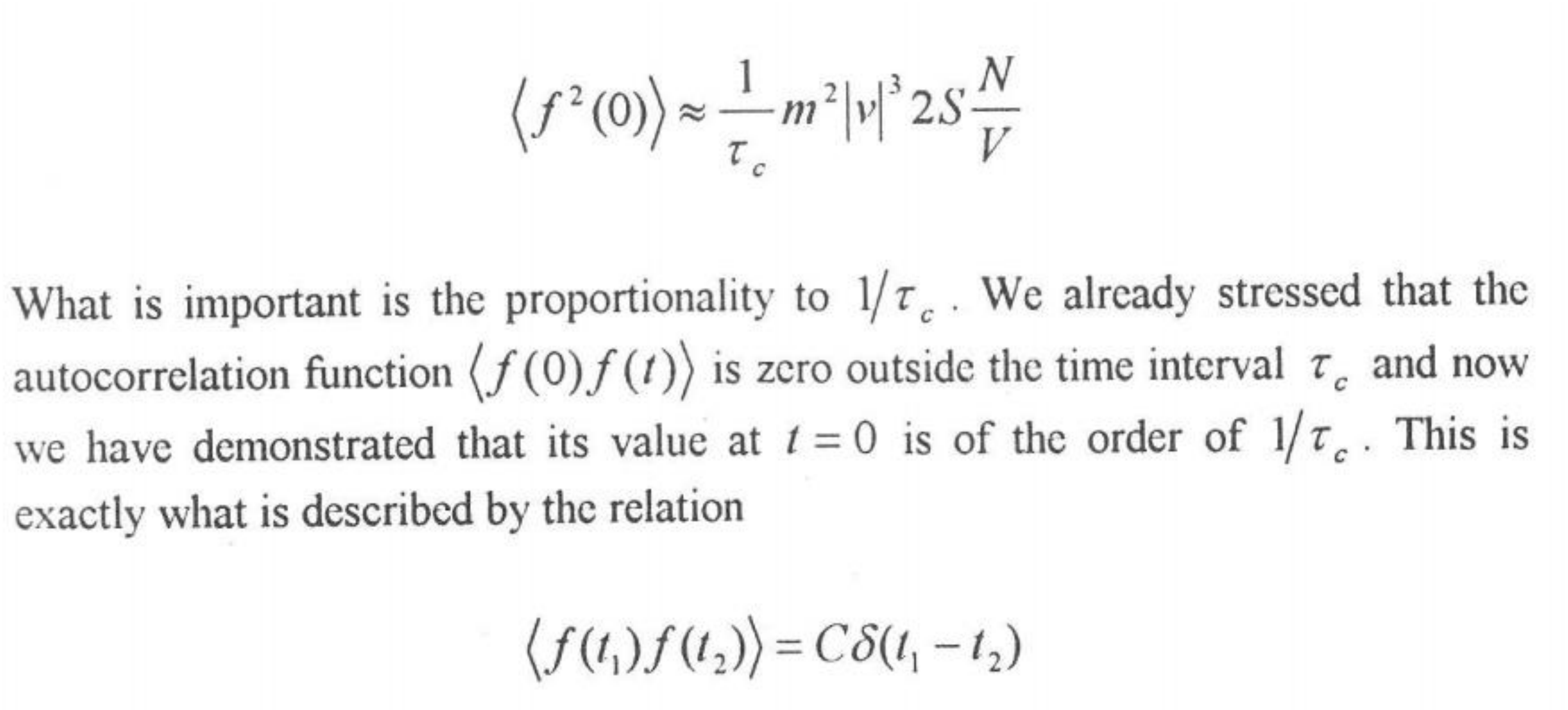
$$\left\langle \left(n-\langle n\rangle\right)^2 \right\rangle = \langle n\rangle$$

$$\left\langle f^2(0)\right\rangle \tau_c^2 = 2\left(\langle n^2\rangle - \langle n\rangle^2\right)(m|v|)^2$$

$$\left\langle f^2(0)\right\rangle \approx \frac{1}{\tau_c^2}(m|v|)^2\, 2lS\frac{N}{V}$$

$$l = |v|\tau_c$$

$$\left\langle f^2(0)\right\rangle \approx \frac{1}{\tau_c} m^2|v|^3\, 2S\frac{N}{V}$$



$$\left\langle f^2(0)\right\rangle \approx \frac{1}{\tau_c} m^2 |v|^3 2S \frac{N}{V}$$

What is important is the proportionality to $1/\tau_c$. We already stressed that the autocorrelation function $\left\langle f(0)f(t)\right\rangle$ is zero outside the time interval τ_c and now we have demonstrated that its value at $t=0$ is of the order of $1/\tau_c$. This is exactly what is described by the relation

$$\left\langle f(t_1)f(t_2)\right\rangle = C\delta(t_1 - t_2)$$

The Wiener-Khintchine theorem states a relationship between two important characteristics of a random process: the power spectrum of the process and the correlation function of the process. ll

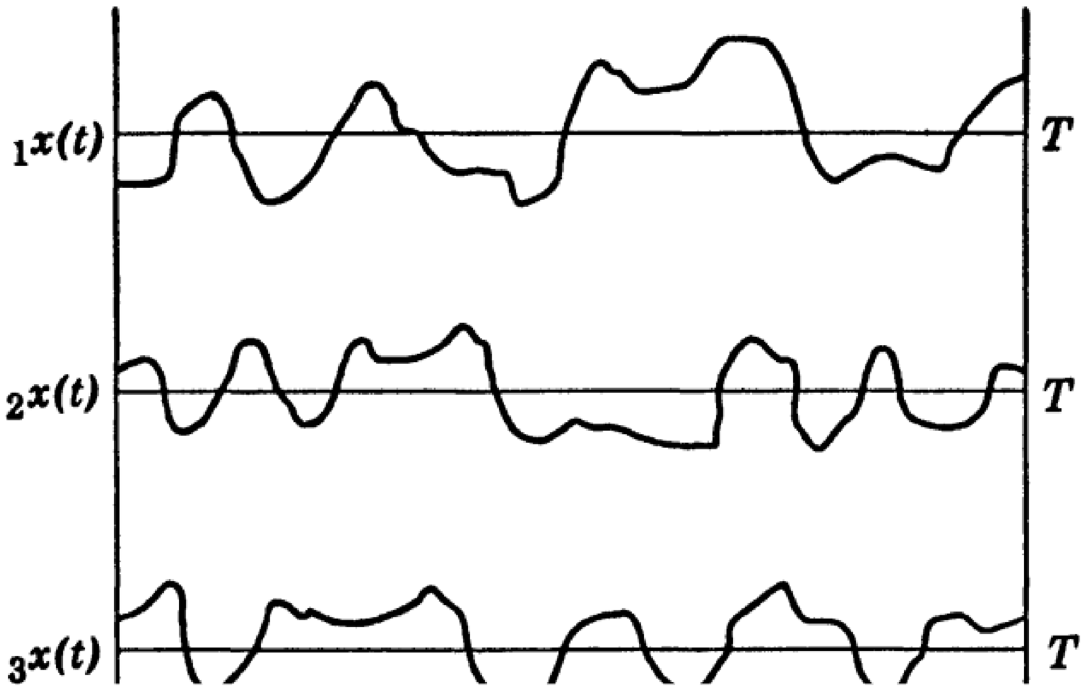


Fig. 27.2. Recordings of $x(t)$ versus t for three systems of an ensemble, as simulated by taking three intervals of duration T from a single long recording. Time averages are taken in a horizontal direction in such a display; ensemble averages are taken in a vertical direction. (After S. O. Rice.)

Suppose we develop one of the records in Fig. 27.2 of $x(t)$ for $0 < t < T$ in a Fourier series:

$$x(t) = \sum_{n=1}^{\infty} (a_n \cos 2\pi f_n t + b_n \sin 2\pi f_n t), \tag{28.1}$$

where $f_n = n/T$. We assume that $\langle x(t) \rangle = 0$, where the angular parentheses $\langle\,\rangle$ denote time average; because the average is assumed zero there is no constant term in the Fourier series. The Fourier coefficients are highly variable from one record of duration T to another. For many types of noise the a_n, b_n have Gaussian distributions. When this is true the process (28.1) is said to be a Gaussian random process.

Let us now imagine that $x(t)$ is an electric current flowing through unit resistance. The instantaneous power dissipation is $x^2(t)$. Each Fourier component will contribute to the total power dissipation. The power in the nth component is

$$\mathcal{P}_n = (a_n \cos 2\pi f_n t + b_n \sin 2\pi f_n t)^2. \tag{28.2}$$

We do not consider cross product terms in the power of the form

$$(a_n \cos 2\pi f_n t + b_n \sin 2\pi f_n t)(a_m \cos 2\pi f_m t + b_m \sin 2\pi f_m t)$$

because for $n \neq m$ the time average of such terms will be zero. The time average of $\mathcal{P}_n$ is

$$<\mathcal{P}_n> = <a_n{}^2 + b_n{}^2>/2, \tag{28.3}$$

We now turn to ensemble averages, denoted here by a bar over the quantity. We recall from Sec. 27 that in the present context an ensemble average is an average over a large set of independent records of the type shown in Fig. 17.2, each record running in time from 0 to T. For a random process we will have

$$\overline{a_n} = 0; \quad \overline{b_n} = 0; \quad \overline{a_n b_m} = 0; \tag{28.5}$$

$$\overline{a_n a_m} = \overline{b_n b_m} = \sigma_n{}^2 \delta_{nm}, \tag{28.6}$$

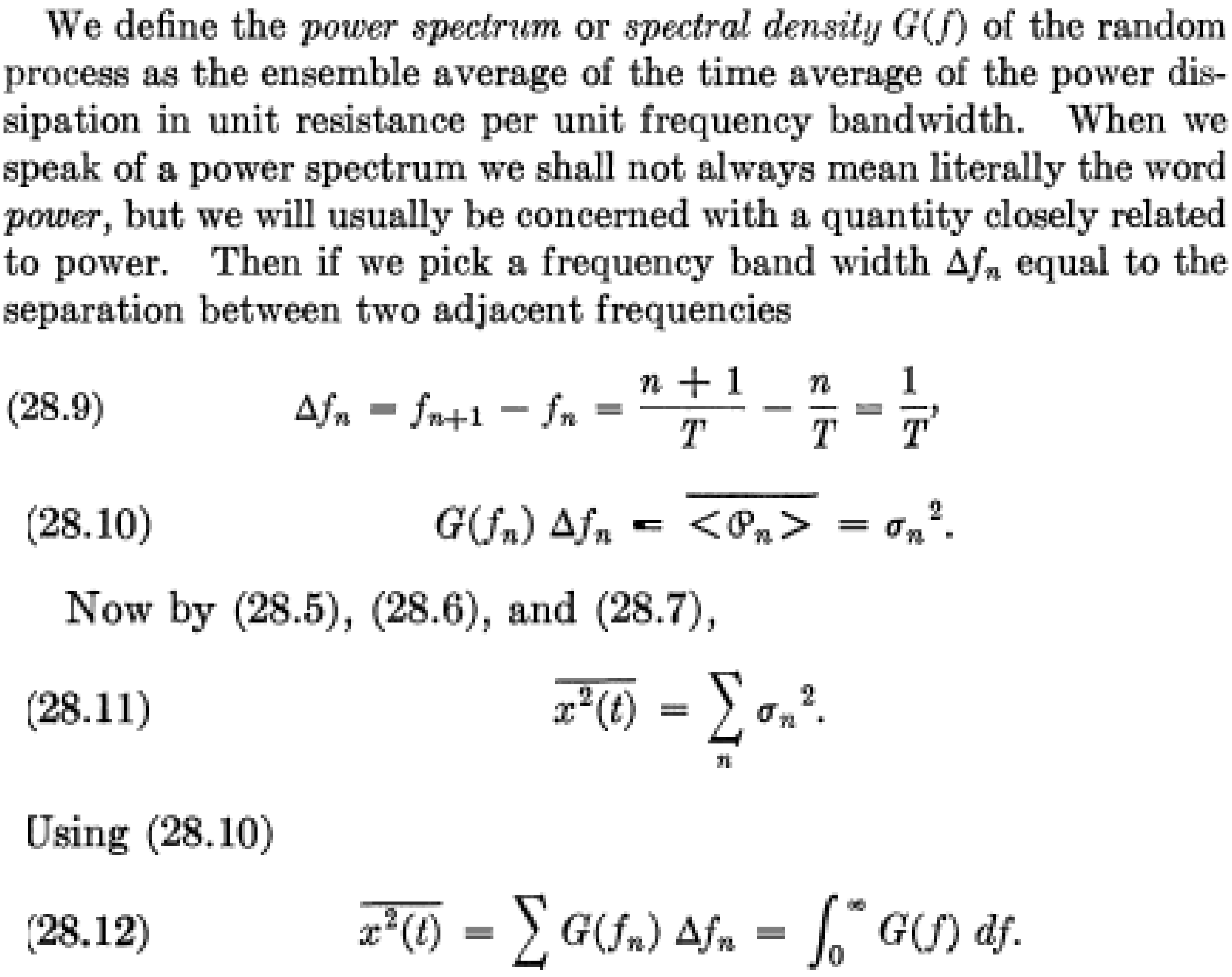
where for a Gaussian random process σ_n is just the standard deviation, as in (27.15). Thus

$$\begin{aligned} \overline{(a_n \cos 2\pi f_n t + b_n \sin 2\pi f_n t)^2} &= \sigma_n{}^2(\cos^2 2\pi f_n t + \sin^2 2\pi f_n t) \\ &= \sigma_n{}^2. \end{aligned} \tag{28.7}$$

Thus, from (28.3) the ensemble average of the time average power dissipation associated with the nth component of $x(t)$ is

$$\overline{<\mathcal{P}_n>} = \sigma_n{}^2. \tag{28.8}$$

$$\overline{x^2(t)} = \sum_n \sigma_n{}^2.$$



We define the *power spectrum* or *spectral density* $G(f)$ of the random process as the ensemble average of the time average of the power dissipation in unit resistance per unit frequency bandwidth. When we speak of a power spectrum we shall not always mean literally the word *power*, but we will usually be concerned with a quantity closely related to power. Then if we pick a frequency band width Δf_n equal to the separation between two adjacent frequencies

$$\Delta f_n = f_{n+1} - f_n = \frac{n+1}{T} - \frac{n}{T} = \frac{1}{T}, \tag{28.9}$$

$$G(f_n)\,\Delta f_n = \overline{<\mathcal{P}_n>} = \sigma_n{}^2. \tag{28.10}$$

Now by (28.5), (28.6), and (28.7),

$$\overline{x^2(t)} = \sum_n \sigma_n{}^2. \tag{28.11}$$

Using (28.10)

$$\overline{x^2(t)} = \sum G(f_n)\,\Delta f_n = \int_0^\infty G(f)\,df. \tag{28.12}$$

$$C(\tau) = <x(t)\, x(t+\tau)>$$

$$(28.14) \quad C(\tau) = \overline{<x(t)\, x(t+\tau)>}$$

$$= \overline{< \sum_{n,m} [a_n \cos 2\pi f_n t + b_n \sin 2\pi f_n t][a_m \cos 2\pi f_m (t+\tau)}$$

$$\overline{+ b_m \sin 2\pi f_m (t+\tau)]>}$$

$$= \tfrac{1}{2} \sum_n \overline{({a_n}^2 + {b_n}^2)} \cos 2\pi f_n \tau = \sum_n {\sigma_n}^2 \cos 2\pi f_n \tau.$$

Using (28.10)

$$(28.15) \qquad C(\tau) = \int_0^\infty G(f) \cos 2\pi f \tau \, df.$$

Thus the correlation function is the Fourier cosine transform of the power spectrum. We can use our previous formulas (27.5) and (27.6)

$$\boxed{G(f) = 4 \int_0^\infty C(\tau) \cos 2\pi f \tau \, d\tau.}$$

The last two formulas are known as Wiener-Khinchin theorem

$$C(\tau) = e^{-\tau/\tau_c}, \tag{28.17}$$

we may say that τ_c is a measure of the average time the system exists without changing its state, as measured by $x(t)$, by more than e^{-1}. We may think of τ_c as a persistence time or *correlation time*. We then expect physically that frequencies much higher than, say, $1/\tau_c$ will not be represented in an important way in the power spectrum. Now if $C(\tau)$ is given by (28.17), the Wiener-Khintchine theorem tells us that

$$G(f) = 4\int_0^\infty e^{-\tau/\tau_c} \cos 2\pi f\tau \, d\tau = \frac{4\tau_c}{1 + (2\pi f\tau_c)^2}. \tag{28.18}$$

Thus, as shown in Fig. 28.1, the power spectrum is flat (on a log frequency scale) out to $2\pi f \approx 1/\tau_c$, and then decreases as $1/f^2$ at high frequencies. We say roughly that the noise spectrum for the correlation function $e^{-\tau/\tau_c}$ is "white" out to a cutoff $f_c \approx 1/2\pi\tau_c$.

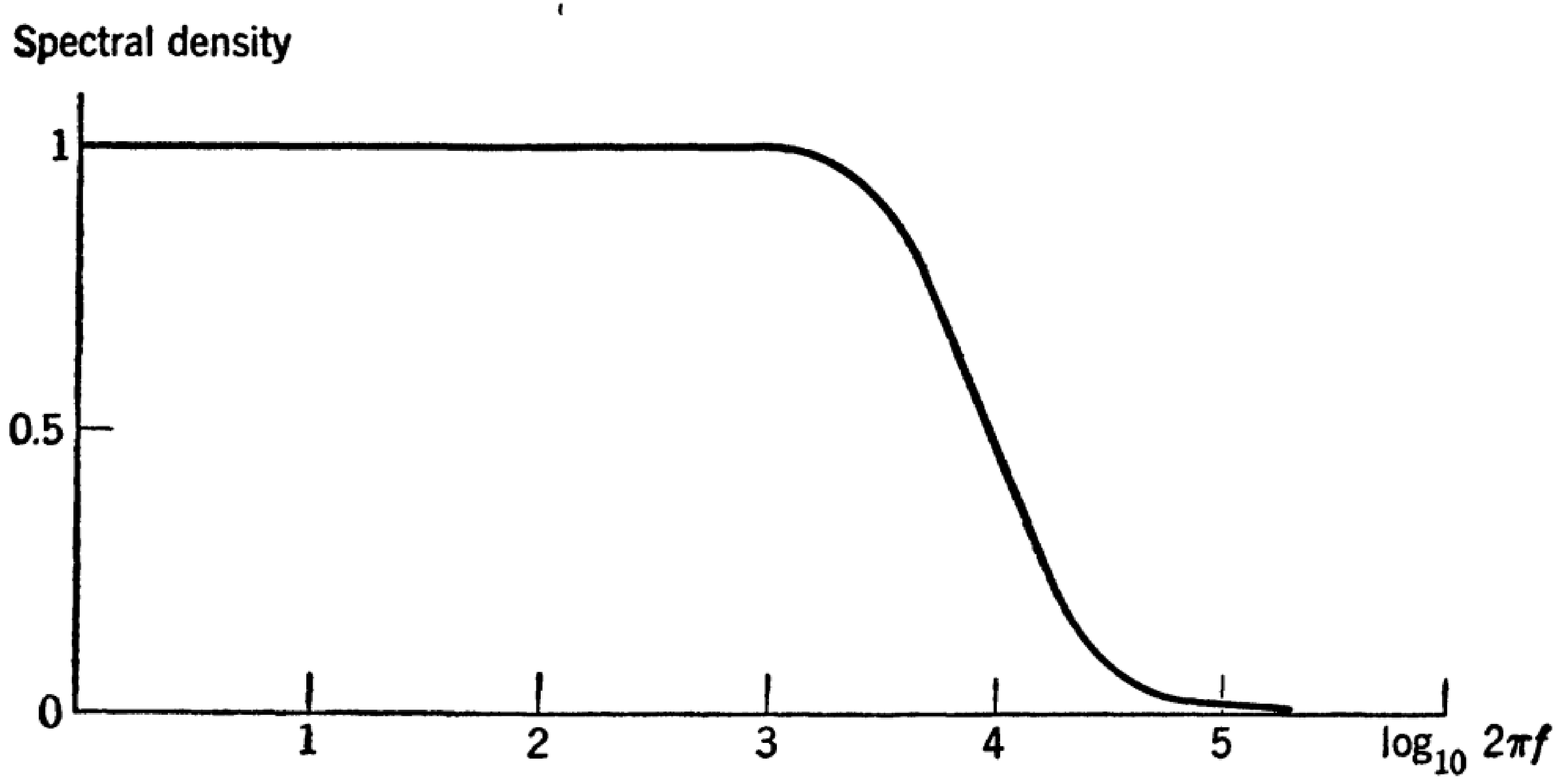
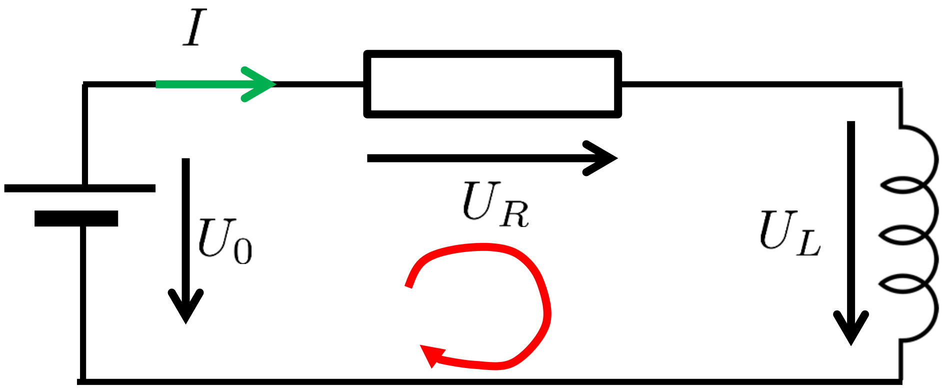


Fig. 28.1. Plot of spectral density versus $\log_{10} 2\pi f$ for an exponential correlation function with $\tau_c = 10^{-4}$ sec.

Indukčnosť v elektrickom obvode



$$-U_0 + IR + L\frac{dI}{dt} = 0$$

Skontrolujme ešte, či tento postup dáva správne znamienko „prídavného napätia“ samoindukovaného cievkou. Ak prúd rastie, prídavné napätie ako sme ho napísali bude kladné a bude mať smer k nemu nakreslenej šípky. Zjavne pôsobí „proti snahám batérie“ a teda bude sa snažiť zmenšovať prúd pretláčaný batériou. To je OK. Ak by prúd klesal, bude nami napísané „prídavné napätie“ záporné. Bude mať smer opačný ako nakreslená šípka a bude teda „pomáhať batérii“ pretláčať prúd, teda snažiť sa ho zväčšovať. To je znovu OK.

lem. Indeed, if $\mathcal{V}$ denotes an applied emf and $V(t)$ the effective fluctuating emf representing the interaction of the conduction electrons with all the other degrees of freedom, the current I satisfies the equation

$$L\frac{dI}{dt} = \mathcal{V}(t) + V(t) \tag{15·8·12}$$

This is analogous to (15·5·1) with $L = m$, $I = v$, and $V(t) = F(t)$. If $\mathcal{V}$ is a relatively slowly varying function of time, an analysis similar to that leading to (15·8·9) then gives the equation

$$L\frac{d\bar{I}}{dt} = \mathcal{V} - R\bar{I} \tag{15·8·13}$$

$$m\frac{d\dot{x}}{dt} = -\alpha\dot{x} + f'$$

$$\alpha = \frac{1}{2kT}\int_{-\infty}^{\infty} d\tau \langle f'(0)f'(\tau)\rangle$$

$$R = \frac{1}{2kT}\int_{-\infty}^{\infty} \langle V(0)V(s)\rangle_0\, ds$$

$$R = \frac{1}{2kT} \int_{-\infty}^{\infty} \langle V(0)V(s)\rangle_0 \, ds$$

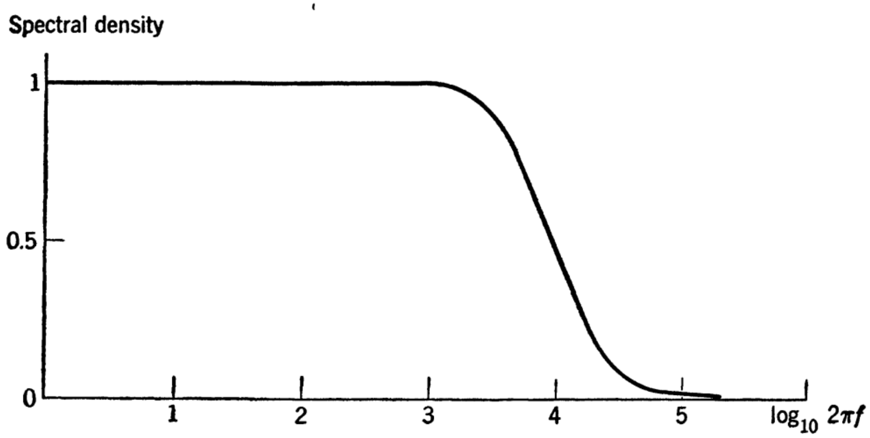
$$R = \frac{1}{2kT} 2 \int_0^{\infty} < V(0)V(s) > \cos(2\pi 0 s) ds$$

$$G(f) = 4 \int_0^{\infty} C(\tau) \cos 2\pi f \tau \, d\tau.$$

$$G(0) = 4kTR$$

$$G(f) \approx G(0)$$

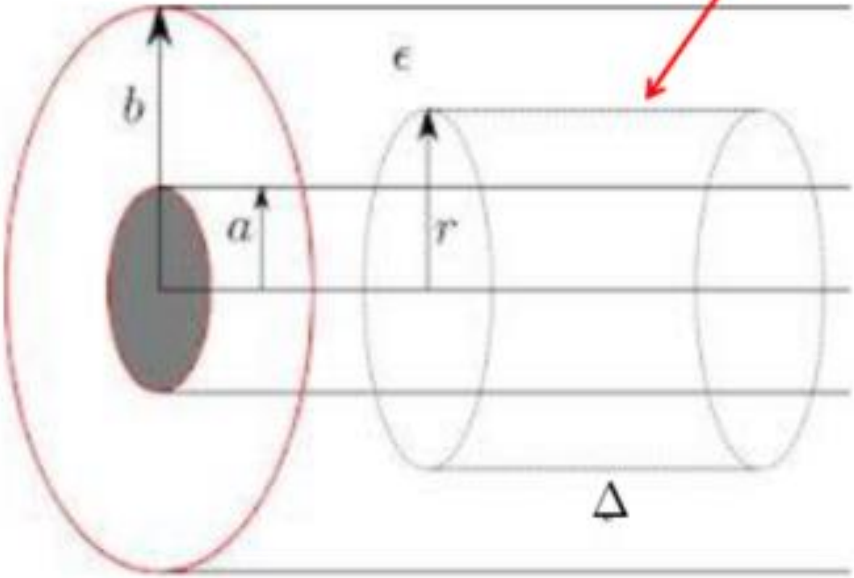
$$G(f) = 4RkT$$



Nyquist theorem

Koaxiálny kábel

Kapacita na meter dĺžky



Plocha pre Gaussovu vetu

$$2\pi r \Delta E(r) = \frac{1}{\varepsilon_0} Q$$

$$E(r) = \frac{1}{2\pi\varepsilon_0 r}\frac{Q}{\Delta}$$

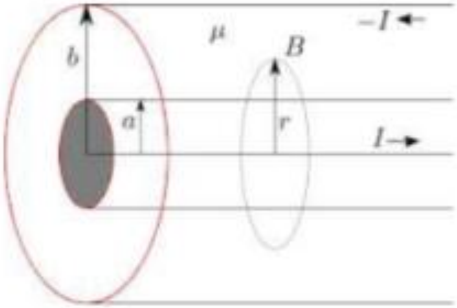
$$U = \int_a^b \frac{1}{2\pi\varepsilon_0 r}\frac{Q}{\Delta} dr =$$

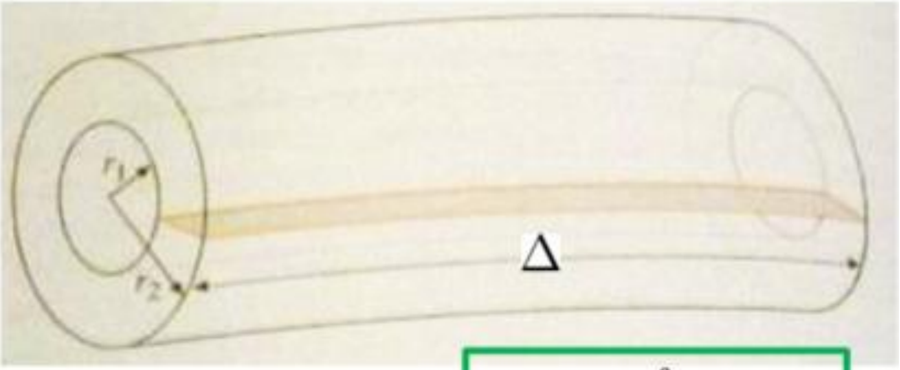
$$= \frac{1}{2\pi\varepsilon_0}\frac{Q}{\Delta}\ln\frac{b}{a} = \frac{Q}{C}$$

$$C = 2\pi\varepsilon_0 \Delta \frac{1}{\ln\frac{b}{a}}$$

$$\boxed{\tilde{C} = 2\pi\varepsilon_0 \frac{1}{\ln\frac{b}{a}}}$$

Indukčnosť na meter dĺžky





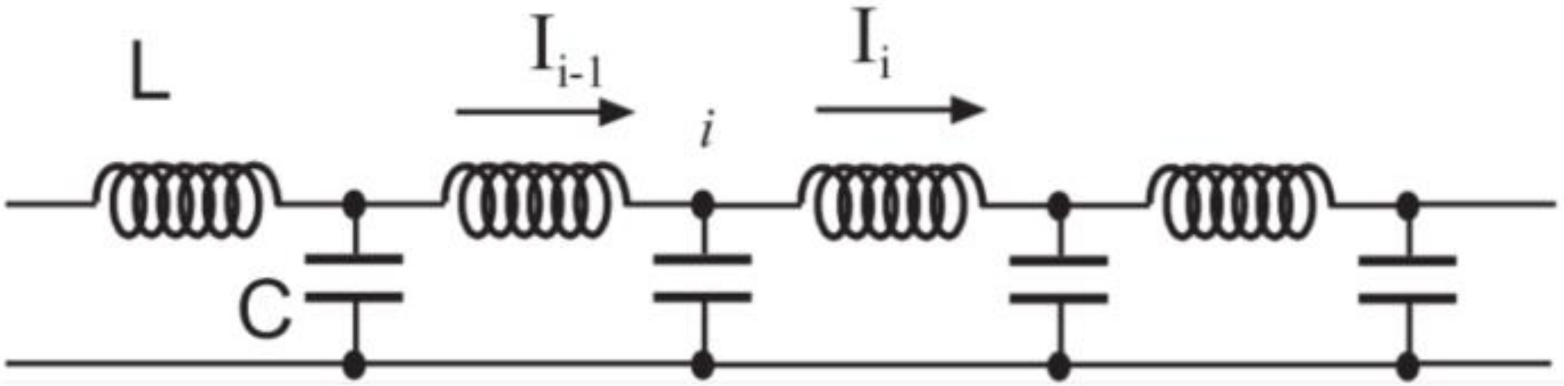
$$\boxed{LI = \int \vec{B}.d\vec{S}}$$

$$2\pi r B(r) = \mu_0 I$$

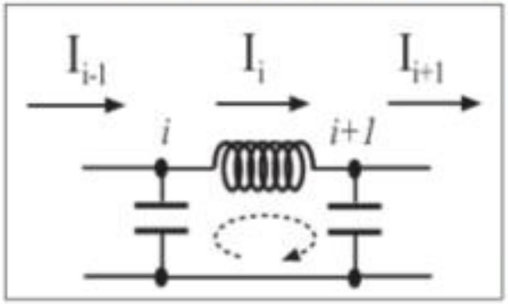
$$\Phi = \int_a^b B(r)\Delta dr = \frac{\mu_0 \Delta I}{2\pi}\int_a^b \frac{1}{r}dr = \frac{\mu_0\Delta}{2\pi} I \ln\frac{b}{a} = LI$$

$$L = \frac{\mu_0\Delta}{2\pi}\ln\frac{b}{a}$$

$$\boxed{\tilde{L} = \frac{\mu_0}{2\pi}\ln\frac{b}{a}}$$



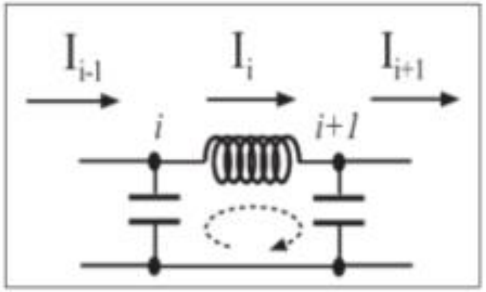
Vedenie je poskladané reťazením elementárnych buniek: bunku tvorí kondenzátor a cievka vpravo od neho. Bunky očíslujeme: nad i-tym kondenzátorom sa nachádza i-ty prúdový uzol. Cievka vpravo od kondenzátora je označená rovnakým číslom i a prúd ňou pretekajúci je I_i



Prúd pritekajúci do i-teho kondenzátora je,

$$I_{i-1} - I_i$$

náboj na i-tom kondenzátore je Q_i



$$\frac{dQ_i}{dt} = I_{i-1} - I_i$$
$$U_i = \frac{1}{C} Q_i$$

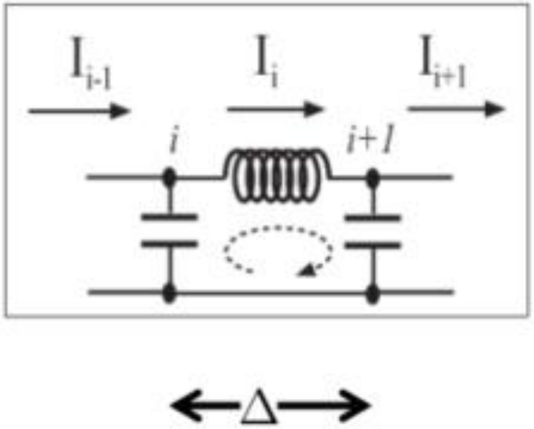
Pre súčet napätí okolo naznačenej slučky dostaneme

$$-U_i + L\frac{dI_i}{dt} + U_{i+1} = 0$$
$$-\frac{1}{C}Q_i + L\frac{dI_i}{dt} + \frac{1}{C}Q_{i+1} = 0$$

Zderivujme poslednú rovnicu podľa času a dostaneme

$$-\frac{1}{C}(I_{i-1} - I_i) + L\frac{d^2 I_i}{dt^2} + \frac{1}{C}(I_i - I_{i+1}) = 0$$

$$\frac{d^2 I_i}{dt^2} = \frac{1}{LC}(I_{i-1} - 2I_i + I_{i+1})$$



$$\frac{d^2 I_i}{dt^2} = \frac{1}{LC}(I_{i-1} - 2I_i + I_{i+1})$$

$$\frac{d^2 I_i}{dt^2} = \frac{\Delta^2}{LC}\frac{I_{i-1} - 2I_i + I_{i+1}}{\Delta^2}$$

$$I_i(t) \rightarrow I(t, x = i\Delta)$$

$$\frac{\partial^2 I(t,x)}{\partial t^2} = \frac{1}{\frac{L}{\Delta}\frac{C}{\Delta}}\frac{\partial^2 I(t,x)}{\partial x^2}$$

Indukčnosť aj kapacita kúska vedenia dĺžky Δ je úmerná (uvidíme to) dĺžke Δ, teda $L = \tilde{L}\Delta \quad C = \tilde{C}\Delta$, kde $\tilde{L},\ \tilde{C}$ sú indukčnosť a kapacita vedenia na jednotku dĺžky. Pre prúd vo vedení dostaneme teda vlnovú rovnicu

$$\frac{\partial^2 I(t,x)}{\partial t^2} = \frac{1}{\tilde{L}\tilde{C}}\frac{\partial^2 I(t,x)}{\partial x^2}$$

$$\frac{\partial^2 I(t,x)}{\partial t^2} = \frac{1}{\tilde{L}\tilde{C}} \frac{\partial^2 I(t,x)}{\partial x^2}$$

$$\frac{\partial^2 I(t,x)}{\partial t^2} = c^2 \frac{\partial^2 I(t,x)}{\partial x^2}$$

Vzruch sa teda šíri po vedení rýchlosťou

$$c = \frac{1}{\sqrt{\tilde{L}\tilde{C}}}$$

$$\frac{\partial^2 I(t,x)}{\partial t^2} = c^2 \frac{\partial^2 I(t,x)}{\partial x^2} \qquad c = \frac{1}{\sqrt{\tilde{L}\tilde{C}}}$$

$$\tilde{C} = 2\pi\varepsilon_0 \frac{1}{\ln\frac{b}{a}} \qquad \tilde{L} = \frac{\mu_0}{2\pi} \ln\frac{b}{a}$$

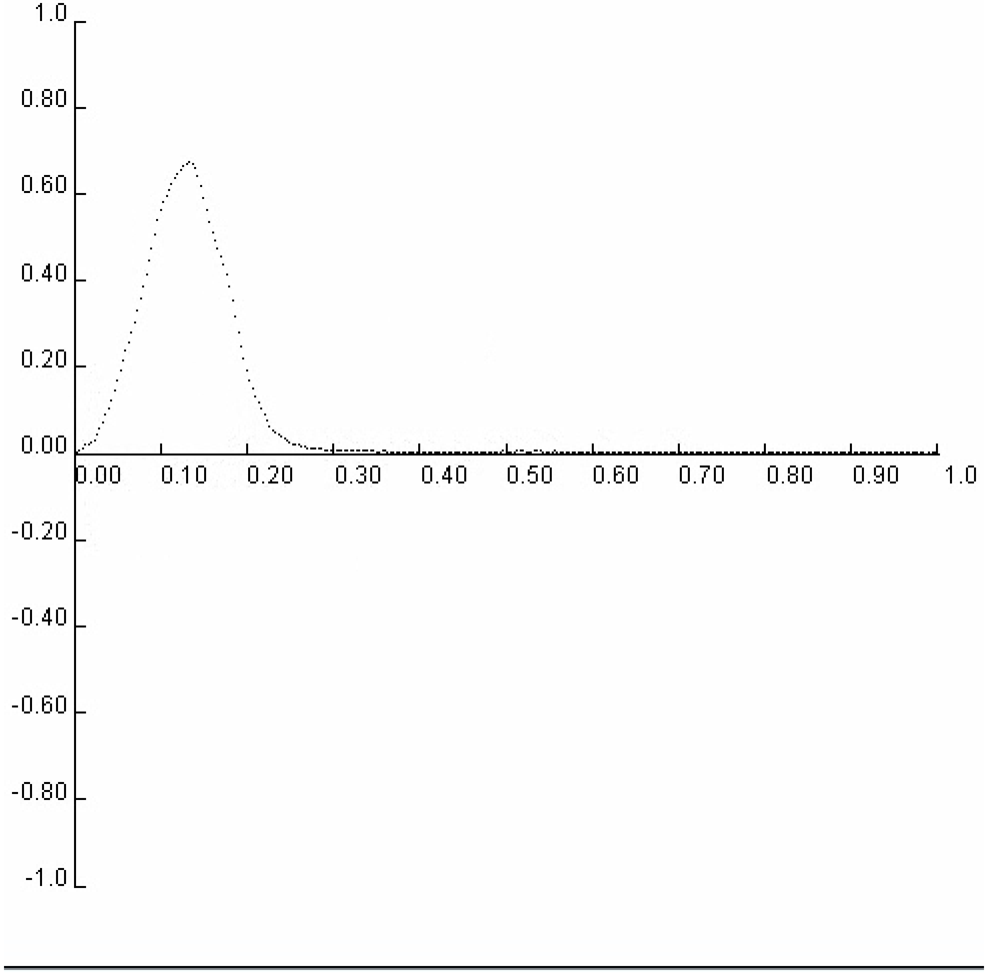
$$c = \frac{1}{\sqrt{\tilde{L}\tilde{C}}} = \frac{1}{\sqrt{2\pi\varepsilon_0 \frac{1}{\ln\frac{b}{a}} \frac{\mu_0}{2\pi} \ln\frac{b}{a}}} = \frac{1}{\sqrt{\mu_0\varepsilon_0}}$$

c = 299 792 458 m/s,   čo je rýchlosť svetla vo vákuu.

Vzruchy sa takýmto káblom (kde izolátorom medzi centrálnym a obvodovým vodičom je vákuum) šíria rýchlosťou svetla. Reálny vodič má ako izolátor nejaké dielektrikum. Tvar vzorcov sa nemení, len namiesto permitivity vákua ε_0 bude všade vystupovať permitivita dielektrika $\varepsilon > \varepsilon_0$. Preto v kábli s dielektrikom sa bude signál šíriť rýchlosťou menšou ako je rýchlosť svetla vo vákuu.

Všimnite si na videu, ako sa vlna na koncoch odráža

$$\frac{\partial^2 u(t,x)}{\partial t^2} = c^2 \frac{\partial^2 u(t,x)}{\partial x^2}, \text{ s okrajovou podmienkou } \; u(t,0) = u(t,L) = 0$$



Pre potreby bezproblémovej komunikácie by sme potrebovali zariadiť, aby **na konci kábla nedochádzalo k odrazom.**

Ako to urobiť?

Myšlienka je takáto: **„signál si musí myslieť, že vedenie je nekonečné“.** Na konci vedenia musíme „zahrať nejaké divadlo“, že je to tam také isté, ako keby vedenie pokračovalo ďalej až do nekonečna.

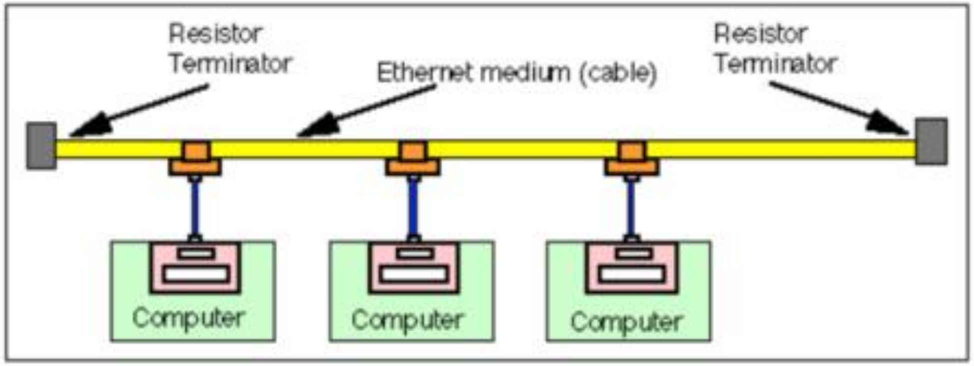
Ako to „zahrať“ najlepšie naznačí energetická úvaha.

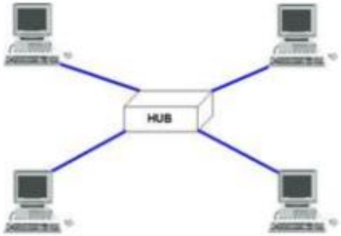
Pohybujúca sa vlna súčasne prenáša aj energiu. Pri odraze sa tá energia zasa vracia, nepokračuje ďalej. Idea je, že energia sa nemôže odraziť, **ak ju niekto alebo niečo na konci vedenia „zožerie“**. Lebo to je to isté, z hľadiska pozorovateľa vo vedení, ako keby energia pokračovala ďalej až do nekonečna. V nekonečnom kábli tú energiu „zožerie nekonečno“. Lenže **energia je nezničiteľná, nedá sa „zožrať“**. Ale možno ju premeniť na nejakú inú formu, ktorá **už nepotrebuje na „odchod do nekonečna“ kábel**. Napríklad ju môžeme **premeniť na teplo**, ktoré sa potom rozptýli do okolia.

Takže bezodrazová idea:

,

na ktorom sa celá dopadajúca energia bude meniť na teplo a odraz nenastane

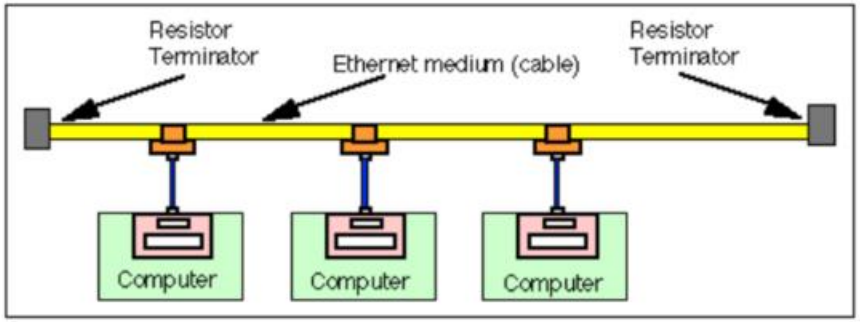




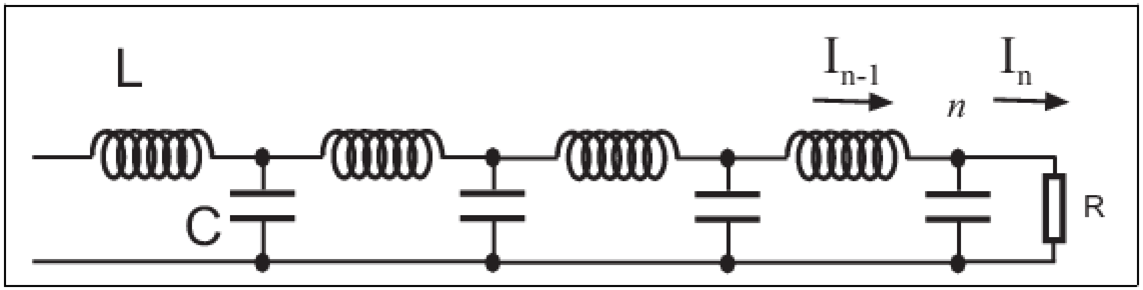
hub

twisted pair kábel

štrukturovaná kabeláž: zakončovacie odpory sú dnu za konektormi



daisy chain



$$\frac{1}{C}\int (I_{n-1} - I_n)dt = RI_n$$

$$\frac{1}{C}(I_{n-1} - I_n) = R\frac{\partial I_n}{\partial t}$$

Prechodom do spojitej limity dostaneme odtiaľ okrajovú podmienku

$$\boxed{-\frac{1}{\tilde{C}}\frac{\partial I}{\partial x} = R\frac{\partial I}{\partial t}}$$

mimo okrajov kábla platí :

$$\frac{\partial^2 I(t,x)}{\partial t^2} = \frac{1}{\tilde{L}\tilde{C}} \frac{\partial^2 I(t,x)}{\partial x^2}$$

pre x=L (na konci kábla) platí:

$$-\frac{1}{\tilde{C}} \frac{\partial I}{\partial x} = R \frac{\partial I}{\partial t}$$

Vráťme sa teraz k otázke ako treba voliť veľkosť R, aby na pravom konci kábla nenastal odraz. Uvedomme si že ľubovoľné riešenie vlnovej rovnice môžme napísať v tvare

$$I(t,x) = f(t - \frac{x}{c}) + g(t + \frac{x}{c})$$

kde f a g musia byť volené tak, aby boli splnené okrajové podmienky, inak majú ľubovoľný tvar. Je pritom zrejmé, že funkcia f popisuje vzruchy šíriace sa zľava doprava, funkcia g popisuje vzruchy šíriace sa sprava doľava. dosaďme teraz do tohto vyjadrenia okrajovú podmienku pravého konca kábla. Dostaneme

$$\frac{1}{c\tilde{C}} f' - \frac{1}{c\tilde{C}} g' = R f' + R g'$$

$$(\frac{1}{c\tilde{C}} - R) f' - (\frac{1}{c\tilde{C}} + R) g' = 0$$

$$\frac{1}{c\tilde{C}}f' - \frac{1}{c\tilde{C}}g' = Rf' + Rg'$$

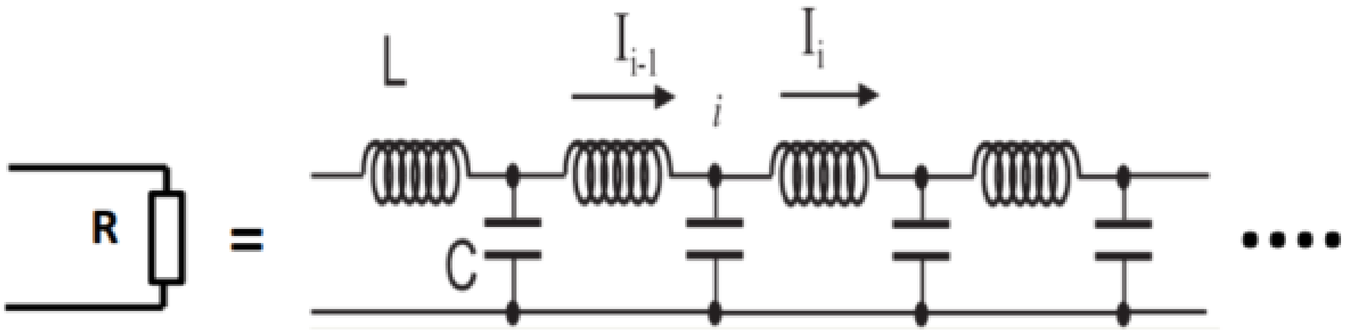
$$(\frac{1}{c\tilde{C}} - R)f' - (\frac{1}{c\tilde{C}} + R)g' = 0$$

Zvoľme: $$\boxed{R = \frac{1}{c\tilde{C}} = \sqrt{\frac{\tilde{L}}{\tilde{C}}}}$$

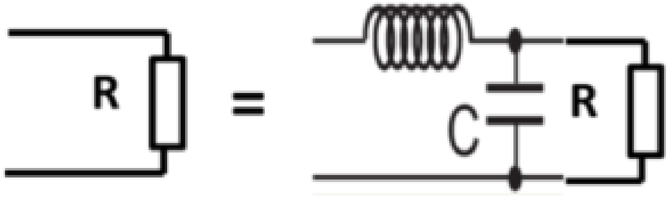
Dostaneme g = const, takže žiadne šírenie zľava doprava

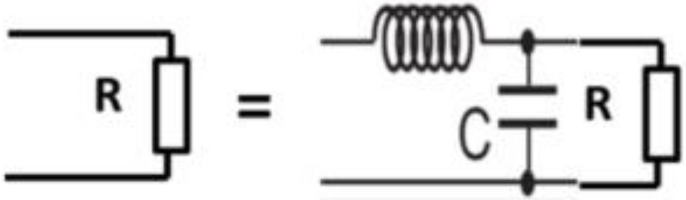
Po dosadení typických hodnôt pre ethernetový koaxiálny kábel dostaneme R = 50 Ω, televízny koaxiálny kábel dáva R = 70 Ω.

Správnu hodnotu zakončovacieho odporu môžeme vyrátať i iným trikom:



nekonečný kábel sa musí javiť z ľavého konca ako odpor R. Nekonečne veľa buniek je to isté ako nekonečno plus jedna bunka navyše, takže dostaneme rovnicu, ktorú schematicky môžeme vyjadriť ako





Paralelné spojenie impedancií C a R dáva

$$\frac{1}{X} = \frac{1}{R} + \frac{1}{X_C}$$

A to spojené sériovo s impedanciou cievky musí dať R, teda dostaneme rovnicu

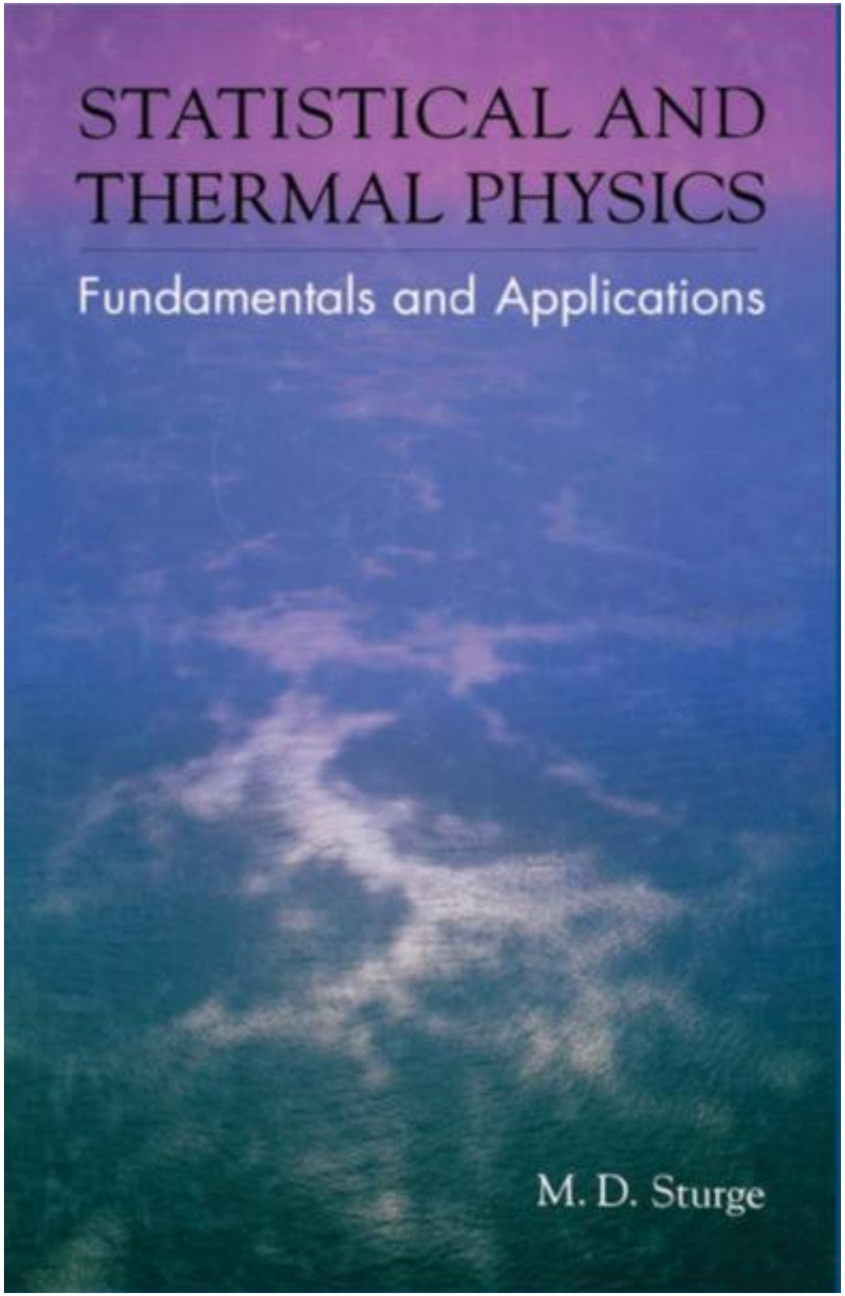
$$R = X_L + X = X_L + \frac{1}{\frac{1}{R} + \frac{1}{X_C}}$$

$$R = \hat{i}\omega L + \frac{1}{\frac{1}{R} + \hat{i}\omega C}$$

$$1 + \hat{i}\omega CR = \hat{i}\omega\frac{L}{R} - \omega^2 LC + 1$$

$$R^2\frac{C}{L} = 1 + \hat{i}\omega CR = 1 + \hat{i}R\tilde{C}\Delta \approx 1$$

$$R = \sqrt{\frac{L}{C}} = \sqrt{\frac{\tilde{L}}{\tilde{C}}}$$



STATISTICAL AND THERMAL PHYSICS

Fundamentals and Applications

M. D. Sturge

These thermal fluctuations are known as *thermal* or *Johnson noise.*[18]) "Noise" in this context means random fluctuations of the voltage across a resistor, and thermal noise is the electrical analog of Brownian motion. There are other, nonthermal, sources of electrical noise which we ignore here.

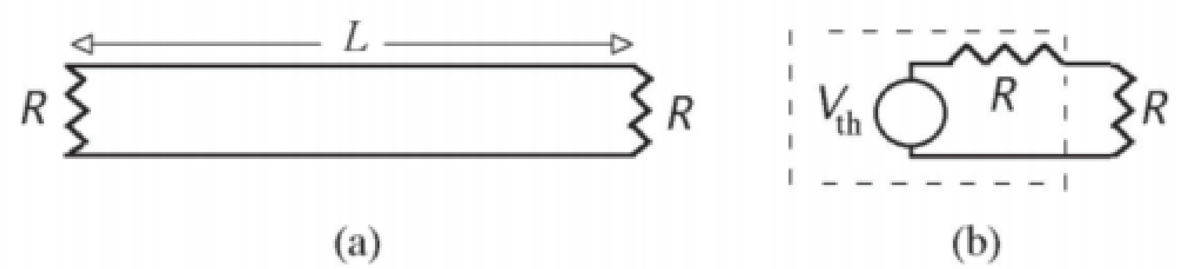
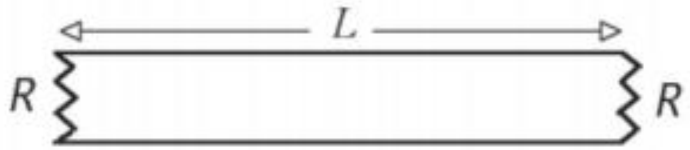
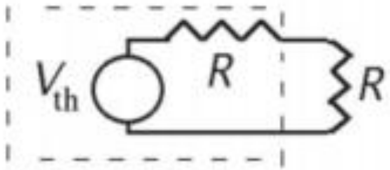


Figure 6.8. (a) Transmission line of length L, terminated at each end by a matching resistor. (b) Equivalent circuit of terminating resistor (within dashed lines), feeding a transmission line of characteristic impedance R.



(a)



(b)

Consider the transmission line shown in Figure 6.8(a). Its length is L and each end is terminated by resistors whose resistance R (not to be confused with the gas constant R_0) is equal to the characteristic impedance of the line. Electromagnetic waves travel along the line with velocity c, and if the ends were short-circuited, the normal modes would be standing waves like the waves on a string discussed in Section 6.2. The resistors, on the other hand, absorb all the energy that comes to them along the line, but we assume L to be so large that the resistors at the ends have negligible effect on the density of normal modes. We can then use equipartition and the results of Section 6.2 to find the thermal energy contained in the normal modes.

From Equation (6.6), the number of modes[19] per unit length with κ less than k is $G(k) = k/\pi$. Substituting $\nu = c/2\pi k$ and differentiating with respect to ν gives $g(\nu) = 2/c$. Each mode is a harmonic oscillator, and since frequencies in electrical circuits are low relative to $\frac{k_{\mathrm{B}}T}{h}$, except at extremely low temperatures, we can use the classical equipartition result Equation (6.4) for the mean energy of an oscillator, so that the thermal energy in each mode is $k_{\mathrm{B}}T$. A mode is a standing wave and is the sum of two counterpropagating waves, so that the energy in each propagating wave is $k_{\mathrm{B}}T/2$. The energy per unit length on the line in frequency interval (bandwidth) $\Delta\nu$, moving (say) to the right, is thus $\frac{1}{2}k_{\mathrm{B}}T g(\nu)\,\Delta\nu = \frac{k_{\mathrm{B}}T\,\Delta\nu}{c}$. This energy propagates at the velocity of light c, so that the energy delivered to one of the resistors per unit time (that is, the power) in this frequency interval, is $k_{\mathrm{B}}T\,\Delta\nu$. Since the system is in thermal equilibrium, on average the net power transfer between the line and the resistor must be zero; hence, the resistor must deliver an equal power to the line. Figure 6.8(b) shows the equivalent circuit of the resistor, enclosed within the dashed lines. It consists of a noise generator (which of course does not exist, but the resistor behaves as if it did) whose open circuit voltage is V_{th}, in series with a resistance R, feeding into the transmission line whose characteristic impedance is also R. The voltage across the transmission line is $V_{th}/2$, so that the resistor delivers a power $V_{th}^2/4R$ to the line. In equilibrium, this must equal the power delivered by the line to the resistor, $\frac{k_{\mathrm{B}}T\,\Delta\nu}{c}$, so that the Johnson noise voltage is given by

$$V_{th} = (4Rk_{\mathrm{B}}T\,\Delta\nu)^{1/2}\,. \tag{6.52}$$

If $T = 300$ K, $R = 1$ MΩ, and $\Delta\nu = 1$ MHz, $V_{th} \approx 130\ \mu$V, a significant input voltage in a high gain amplifier.

The Equipartition Theorem

Consider the quantity

$$\begin{aligned}\left\langle q_i \frac{\partial E}{\partial q_i} \right\rangle &\equiv \int\int \ldots \int q_i \frac{\partial E}{\partial q_i} P(q_1, q_2, \ldots)\, dq_1\, dq_2\, \ldots \\ &= Z^{-1} \int\int \ldots \int q_i \frac{\partial E}{\partial q_i} e^{-\beta E}\, dq_1\, dq_2\, \ldots \\ &= -(\beta Z)^{-1} \int\int \ldots \int q_i \frac{\partial}{\partial q_i} \left(e^{-\beta E}\right)\, dq_1\, dq_2\, \ldots .\end{aligned}$$

Now integrate with respect to q_i by parts:

$$\int q_i \frac{\partial}{\partial q_i} \left(e^{-\beta E}\right)\, dq_i = -\int e^{-\beta E}\, dq_i$$

$$\left\langle q_i \frac{\partial E}{\partial q_i} \right\rangle = -(\beta Z)^{-1} \int\int \ldots \int e^{-\beta E}\, dq_1\, dq_2\, \ldots$$

$$\boxed{\left\langle q_i \frac{\partial E}{\partial q_i} \right\rangle = \beta^{-1} = k_{\mathrm{B}} T.}$$

In many cases, a system of coordinates exists in which E can be written as a sum of quadratic functions of the coordinates:

$$E = \sum_i a_i q_i^2, \tag{G.10}$$

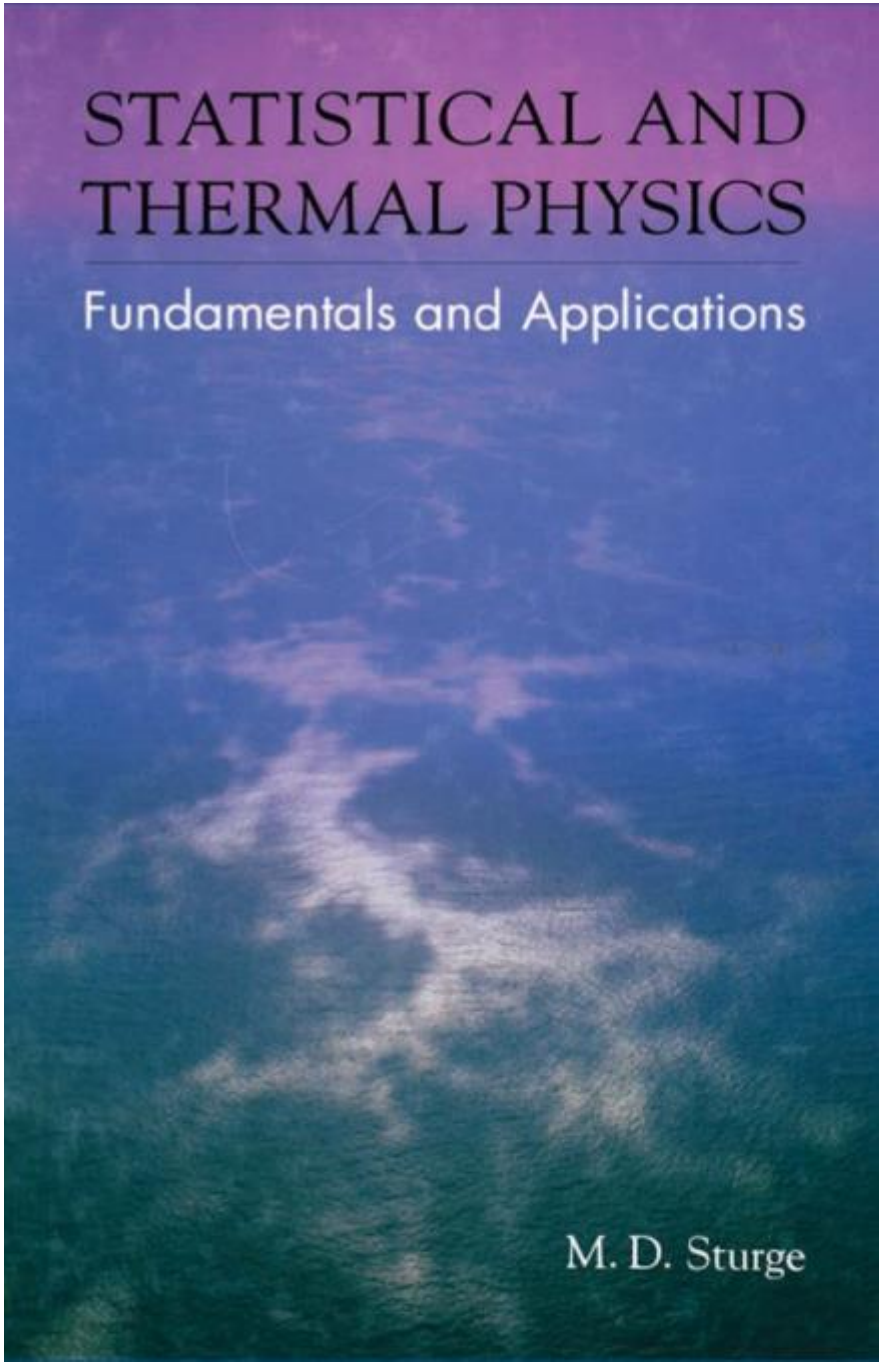
so that

$$\frac{\partial E}{\partial q_i} = 2a_i q_i.$$

$$\boxed{\langle a_i q_i^2 \rangle = \frac{1}{2}\left\langle q_i \frac{\partial E}{\partial q_i} \right\rangle = \frac{1}{2} k_\mathrm{B} T.}$$

The fact that the contribution to the internal energy of the i'th degree of freedom is independent of the magnitude of the coefficient a_i has a curious and (in classical terms) paradoxical consequence. Suppose that a_i tends to zero. However small a_i may be, so long as it is nonzero, the equipartition theorem says that the contribution to U is $Nk_\mathrm{B}T/2$. On the other hand, if a_i actually *is* zero, that degree of freedom does not appear in the energy expression and cannot contribute to U.

This nonanalytic behavior is a consequence (like Gibbs' paradox; see Section 9.4) of the internal inconsistency of classical statistical mechanics. Quantum theory resolves this paradox, since as a_i becomes smaller, the corresponding quantum levels become further and further apart, so that a higher and higher temperature is needed for the assumption of continuous energy levels to be valid and for equipartition to hold. As a_i tends to zero, the required temperature tends to infinity.



STATISTICAL AND THERMAL PHYSICS

Fundamentals and Applications

M. D. Sturge

4.6 Entropy of Mixing

The fact that $dS = \frac{đQ}{T}$ in a reversible transfer of heat might suggest that entropy change is exclusively associated with heat transfer, but the irreversible processes illustrated in Figure 2.2 and discussed in the previous section shows that this is not the case. In general, there are contributions to the entropy that have nothing directly to do with heat. For example, suppose we have two pure crystals[7] of different elements A and B at a very low temperature, where thermal vibrations can be neglected. Every atom is in its place, so that there is only one microstate and the entropy is zero.

Now suppose that we melt these two crystals (containing N_a and N_b atoms, respectively) together and allow the resulting alloy to recrystallize, forming a crystal of $N = N_a + N_b$ atoms. We then return the crystal to its original temperature. In many cases, it is found that the different atoms are distributed at random over the possible sites in the new crystal, so that the probability of finding an A atom in any arbitrarily chosen site is $x = \frac{N_a}{N}$, and the probability of finding a B atom is $1 - x$. If we neglect thermal vibrations, the multiplicity of this macrostate is, from Equation (D.1)

$$\Omega = \frac{N!}{(xN)![(1-x)N]!}$$

If $N \gg 1$ the entropy is, from Stirling's formula in the form given in Equation (D.6):

$$S = -Nk_{\mathrm{B}}[x \ln x + (1-x)\ln(1-x)] \tag{4.16}$$

(note that $0 < x < 1$, so that both logarithms are negative and the entropy is positive).

If $N_a = N_b$, so that $x = \frac{1}{2}$, S has its maximum value

$$S = Nk_{\mathrm{B}} \ln 2. \tag{4.17}$$

Thus, the entropy is greater than that of the separate pure crystals by an amount of order Nk_{B}. This entropy increase is called the *entropy of mixing* and an increase of this order occurs whenever two distinct substances mix irreversibly; see Problem 4.5(c) for the case of two ideal gases mixing.

The entropy of mixing is not small; for $x \sim 0.5$ it is of the same order as the increase in the entropy of an ideal gas with N molecules when its absolute temperature is doubled (see Chapter 9).

energy (which is negative) is reduced in magnitude. Interfaces between different substances also have energy and contribute to G; for example, the large interfacial energy of the oil-water boundary is responsible for the fact that oil and water do not mix even though their entropy would be increased by mixing. A detergent reduces the interfacial energy, thus lowering the free energy of the mixed state and allowing mixing to occur. Further exam-

Consider the process illustrated by Figure 2.2, where an ideal gas, initially contained in a volume V_1, is allowed (by the removal of a partition) to expand irreversibly to a volume $V_2 = V_1 + V_0$. No work is done and no heat flows, so that U, and therefore T and n_{q}, remain constant. For simplicity, we take $V_2 = 2V_1$. The number of gas molecules N is fixed, so that the density is halved, and Equation (9.14) gives for the increase in entropy

$$\Delta S = N k_{\mathrm{B}} \ln 2. \tag{9.17}$$

Now suppose that instead of a vacuum, we start with a different gas in the righthand volume, at the same pressure, also with N molecules. When the partition is removed, this gas also expands to twice its volume, and its entropy increases by the same amount, so long as the gases do not interact with each other. Thus, the total entropy of the system has increased by $2N k_{\mathrm{B}} \ln 2$. This extra entropy is an example of entropy of mixing (see (4.17) and Problem 4.5 (c)). Note that this increase in entropy involves no heat flow; as we have said, mixing is an irreversible process so that the relation $dS = \frac{đQ}{T}$ (Equation (4.14)) does not apply.

sures, is unchanged. What if both gases are the same? Then nothing is changed by the removal of the partition; the entropy before and after its removal must be the same. In classical theory this result is a paradox, since one could imagine the two gases being made more and more alike, until it became impossible to distinguish them by their properties; yet the increase in entropy due to mixing is independent of how different the two gases are.

9.9 Ideal Solutions and Osmotic Pressure

When you drop an ionic solid such as potassium chloride into water, it dissolves; that is, the positive and negative ions that make up the solid separate and move freely in the liquid. If you measure the temperature of the water carefully while the solid is dissolving you will find that it drops; this is because energy (the heat of solution) is needed to separate the ions.[12] However, although the energy of the ions is higher in solution, their entropy is much greater, so that their free enthalpy is lower and they dissolve.

An *ideal solution* is one in which the dissolved molecules or ions do not interact with each other, so that the only contributions to the free enthalpy difference between the solid and solution are the heat of solution and the entropy of mixing, which is from Equation (4.16)

$$S = -Nk_{\mathrm{B}}\left[x \ln x + (1-x)\ln(1-x)\right],$$

where x is the concentration of the solute. We label the solute x and solvent s, writing $N_{\mathrm{x}} = Nx$ for the number of dissolved particles and $N_{\mathrm{s}} = N(1-x)$

---

[12] The heat of solution of KCl is about 0.18 eV per $\mathrm{K}^+ - \mathrm{Cl}^-$ pair; this energy is less than the cohesive energy of the crystal because the dissolved ions attract the highly polarizable H_2O molecules.

for the number of solvent particles. The free enthalpy difference between the solution and the pure materials is then

$$\Delta G = N_{\mathrm{x}}\epsilon + k_{\mathrm{B}}T\left[N_{\mathrm{x}}\ln x + N_{\mathrm{s}}\ln(1-x)\right] \tag{9.43}$$

where ϵ is the heat of solution per solute atom and the quantity in square brackets is negative since $x < 1$. Note that however large ϵ is, ΔG is negative for sufficiently small x, since $\ln x \to -\infty$ as $x \to 0$. It follows that a small amount of material will always dissolve. In this limit of large ϵ, the solubility (defined as the value of x for which $\Delta G = 0$) increases with temperature according to Arrhenius' law, with activation energy ϵ.

We will now examine the remarkable phenomenon of osmosis. If a solution is separated from the pure solvent by a membrane permeable to the solvent but not to the solute, the solvent will diffuse through the membrane until a sufficient pressure difference, called the osmotic pressure Δp, has built up in the solution to prevent any further diffusion. This happens because the chemical potential of the *solvent*, ζ_{s}, is decreased by the presence of the solute, but increases with pressure. The solvent will diffuse until the pressure difference equalizes ζ_{s} on both sides of the membrane.

To make this quantitative, we assume that the solution is dilute, that is $x \ll 1$, so that $\ln(1-x) \approx -x$. Then

$$\Delta G \approx N_{\mathrm{x}}\epsilon + k_{\mathrm{B}}T\left[N_{\mathrm{x}}\ln x - N_{\mathrm{s}}x\right]. \tag{9.44}$$

The presence of the solute thus reduces the chemical potential of the solvent by

$$\Delta\zeta_{\mathrm{s}} = \left(\frac{\partial\Delta G}{\partial N_{\mathrm{s}}}\right)_{T,p,N_{\mathrm{x}}} = -k_{\mathrm{B}}Tx. \tag{9.45}$$

There is a Maxwell relation (see Problem 9.12(a))

$$\left(\frac{\partial\zeta_i}{\partial p}\right)_{T,N_i} = \left(\frac{\partial V}{\partial N_i}\right)_{T,p} = n_i^{-1}, \tag{9.46}$$

so that an increase in pressure Δp increases ζ_{s} by $\frac{\Delta p}{n}$, where n is the particle density of the solvent.

Hence, equilibrium is achieved when

$$\Delta\zeta_{\mathrm{s}} = -k_{\mathrm{B}}Tx + \frac{\Delta p}{n} = 0.$$

Hence,

$$\Delta p = n_{\mathrm{x}}k_{\mathrm{B}}T \tag{9.47}$$

where $n_{\mathrm{x}} = nx$ is the particle density of the solute and we have neglected x relative to 1.

Equation (9.47) is known as van't Hoff's[13] law, and is identical to the ideal gas law (Equation (9.11)), with the osmotic pressure substituted for the actual pressure of the gas. This is not a coincidence, since the density dependence of the chemical potential of the solute has the same form as that of an ideal gas (see Problem 9.12(b)). Since kinetic theory cannot possibly apply to a solution, where motion of the solute molecules is severely restricted by the solvent, this result shows that it is a mistake to think of the ideal gas law as merely a consequence of kinetic theory.

Example 9.4. (Lowering of the freezing point by a solute.) Solubilities are usually much lower in solids than in liquids, and for the purpose of this example, we will assume that the solute is not soluble in the solid solvent. Then, while the chemical potential of the liquid solvent containing solute is lower than that of pure liquid by $\Delta\zeta_s$, given by Equation (9.45), that of the solid is unaffected. If we start at a temperature above the freezing point and reduce it, freezing becomes possible when the chemical potential of the solid drops below that of the liquid. Since the presence of the solute reduces the chemical potential of the liquid, the liquid remains at the stable phase to a lower temperature and the freezing point is depressed. Find the change ΔT, in temperature, to first order in x.

$$H = G - T\left(\frac{\partial G}{\partial T}\right)$$

$$\frac{H}{N} = \zeta - T(\frac{\partial \zeta}{\partial T})$$

$$\frac{\partial}{\partial T}(\frac{\zeta}{T}) = -\frac{1}{T^2}\zeta + \frac{1}{T}\frac{\partial \zeta}{\partial T}$$

Besides the reduction $\Delta\zeta_s$ due to the solute, we also need to know how the chemical potential difference between the pure liquid and the solid varies with temperature. If this difference is $\Delta\zeta_0(T)$, freezing will occur when

$$\Delta\zeta_s + \Delta\zeta_0 = 0. \tag{9.48}$$

$\Delta\zeta_0$ can be calculated from the Gibbs-Helmholtz relation for G (Equation (4.45)), which can be written

$$\frac{\partial}{\partial T}\left(\frac{\zeta}{T}\right) = -\frac{H}{N_A T^2},$$

where H is the kmolar enthalpy, the derivative is at constant pressure, and we have used $G = N\zeta$ (Equation (7.10)).

The change in temperature ΔT is small, and $\Delta\zeta_0 = 0$ when $\Delta T = 0$, so that

$$\Delta\zeta_0 \approx T\frac{\partial}{\partial T}\left(\frac{\Delta\zeta_0}{T}\right)\Delta T = -\frac{L\,\Delta T}{N_A T}, \tag{9.49}$$

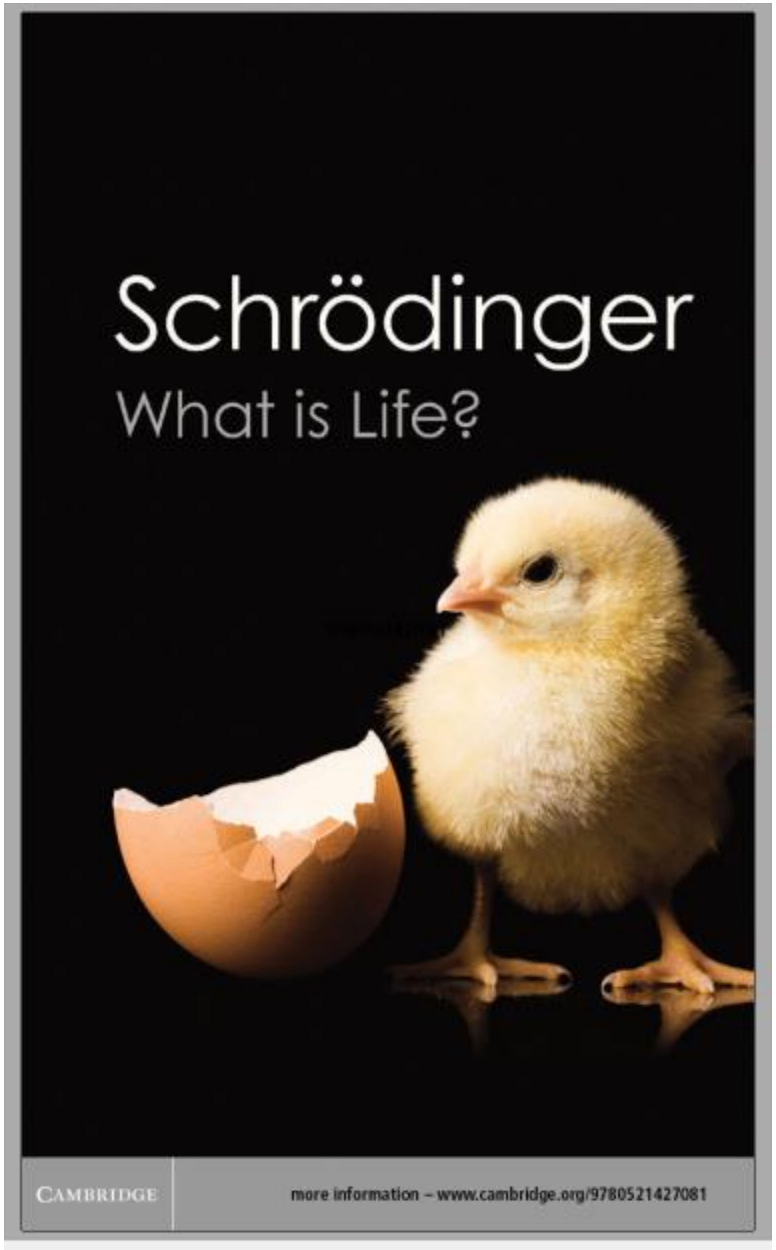
$$\Delta\zeta_s = \left(\frac{\partial \Delta G}{\partial N_s}\right)_{T,p,N_x} = -k_B T x.$$

where L is the latent heat (defined as the difference in kmolar enthalpy between the pure liquid solvent and the solid).

Substituting Equations (9.45) and (9.49) into Equation (9.48) we find

$$\Delta T = -\frac{R_0 T^2}{L} x. \tag{9.50}$$

Note that, so long as the assumption of ideality holds, $\frac{\Delta T}{x}$ is the same for all solutes, depending only on the latent heat of the pure solvent. An analogous calculation shows that the presence of a nonvolatile solute raises the boiling point (see Problem 9.13(b)). ◇



Schrödinger

What is Life?

CAMBRIDGE

more information – www.cambridge.org/9780521427081

WHAT IS LIFE?

THE PHYSICAL ASPECT OF THE LIVING CELL

Based on lectures delivered under the auspices of the Dublin Institute for Advanced Studies at Trinity College, Dublin, in February 1943

Like so many works that have had a great impact on human thinking, it makes points that, once they are grasped, have a ring of almost self-evident truth; yet they are still blindly ignored by a disconcertingly large proportion of people who should know better. How often do we still hear that quantum effects can have little relevance in the study of biology, or even that we eat food in order to gain energy? This serves to emphasize the continuing relevance that Schrödinger's *What is Life?* has for us today. It is amply worth rereading!

Roger Penrose
8 August 1991

WHY ARE THE ATOMS SO SMALL?

Clearly, the question is an evasion. For it is not really aimed at the size of the atoms. It is concerned with the size of organisms, more particularly with the size of our own corporeal selves. Indeed, the atom is small, when referred to our civic unit of length, say the yard or the metre. In atomic physics one is accustomed to use the so-called Ångström (abbr. Å), which is the 10^{10}th part of a metre, or in decimal notation 0.0000000001 metre. Atomic diameters range between 1 and 2Å. Now those civic units (in relation to which the atoms are so small) are closely related to the size of our bodies. There is a story tracing

THE WORKING OF AN ORGANISM REQUIRES EXACT PHYSICAL LAWS

PHYSICAL LAWS REST ON ATOMIC STATISTICS AND ARE THEREFORE ONLY APPROXIMATE

THEIR PRECISION IS BASED ON THE LARGE NUMBER OF ATOMS INTERVENING.

THE $\sqrt{n}$ RULE

The Hereditary Mechanism

THE HEREDITARY CODE-SCRIPT (CHROMOSOMES)

THE NECESSITY OF MUTATION BEING A RARE EVENT

PERMANENCE UNEXPLAINABLE BY CLASSICAL PHYSICS

faced with the question: How can we, from the point of view of statistical physics, reconcile the facts that the gene structure seems to involve only a comparatively small number of atoms (of the order of 1,000 and possibly much less), and that nevertheless it displays a most regular and lawful activity – with a durability or permanence that borders upon the miraculous?

EXPLICABLE BY QUANTUM THEORY

QUANTUM THEORY – DISCRETE STATES – QUANTUM JUMPS

THE APERIODIC SOLID

THE VARIETY OF CONTENTS COMPRESSED IN THE MINIATURE CODE

Order, Disorder and Entropy

LIVING MATTER EVADES THE DECAY TO EQUILIBRIUM

IT FEEDS ON 'NEGATIVE ENTROPY'

What then is that precious something contained in our food which keeps us from death? That is easily answered. Every

On Determinism and Free Will

that in my opinion, and contrary to the opinion upheld in some quarters, *quantum indeterminacy* plays no biologically relevant role in them, except perhaps by enhancing their