determinism free*

*except for conservation laws and Newton's 2nd law

A probability-based introduction to atmospheric thermodynamics

BETA 0.8

Demetris Koutsoyiannis

Department of Water Resources and Environmental Engineering

School of Civil Engineering

National Technical University of Athens, Greece

(dk@ntua.gr, http://itia.ntua.gr/dk/)

Entropy is uncertainty quantified: Definitions

- For a discrete random variable[*] $\underline{z}$ taking values z_j with probability mass function $P_j \equiv P(z_j) = P\{\underline{z} = z_j\}$, $j = 1,\ldots,w$, where

$$\sum_{j=1}^{w} P_j = 1 \qquad (1)$$

the entropy[†] is a dimensionless nonnegative quantity defined as (e.g., Papoulis, 1991):

$$\Phi[\underline{z}] := E[-\ln P(\underline{z})] = -\sum_{j=1}^{w} P_j \ln P_j \qquad (2)$$

- For a continuous random variable $\underline{z}$ with probability density function $f(z)$, where

$$\int_{-\infty}^{\infty} f(z)\, dz = 1 \qquad (3)$$

the entropy[†] is defined as:

$$\Phi[\underline{z}] := E[-\ln[\,f(\underline{z})/h(\underline{z})]] = -\int_{-\infty}^{\infty} \ln\,[f(z)/h(z)]\, f(z)\, dz \qquad (4)$$

The function $h(z)$ can be any probability density, proper (with integral equal to 1, as in (3)) or improper (meaning that its integral does not converge); typically it is an (improper) Lebesgue density, i.e. a constant with dimensions $[h(z)] = [f(z)] = [z^{-1}]$, so that $\Phi[\underline{z}]$ is again dimensionless.

---

[*] An underlined symbol denotes a random variable; the same symbol not underlined represents a value of the random variable.
[†] In case of risk of ambiguity, we will characterize $\Phi[\underline{z}]$ as *probabilistic entropy*.

The principle of maximum entropy (ME)

- The importance of the entropy concept springs from the principle of maximum entropy, an extremely powerful principle in physics and, at the same time, in logic.[*]

- Although this principle stands behind the Second Law of thermodynamics, which was first formulated in the mid-19th century by Clausius (who also coined the term entropy), it took 100 years to recognize its general applicability (Jaynes, 1957) also in logical inference (to infer unknown probabilities from known information) and to formalize it.

- The principle of maximum entropy postulates that the entropy of a random variable z should be at maximum, under some conditions, formulated as constraints, which incorporate the information that is given about this variable.

- In simple words, the principle advises us to express what we know for a variable z in the form of mathematical constraints, which are either equations or inequalities. What we do not know, we determine in probabilistic terms by maximizing entropy, i.e., uncertainty.

- The logic behind the principle is very simple and almost self-evident: If uncertainty is not maximized then there must be some more knowledge, which, however, should already have been incorporated in the constraints.

[*]In an optimistic view that our logic in making inference about natural systems could be consistent with the behaviour of the natural systems, we can regard the principle of maximum entropy both as a physical principle to determine thermodynamic states of natural systems and as a logical principle to make inference about natural systems. We will see later that this is reasonable.

Entropy maximization: The die example
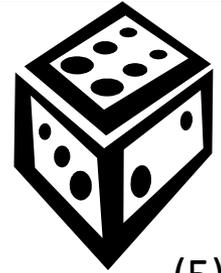
- What is the probability that the outcome of a die throw will be i?
- The entropy is:

$$\Phi := E[-\ln P(\underline{z})] = -P_1 \ln P_1 - P_2 \ln P_2 - P_3 \ln P_3 - P_4 \ln P_4 - P_5 \ln P_5 - P_6 \ln P_6 \qquad (5)$$

- The equality constraint is

$$P_1 + P_2 + P_3 + P_4 + P_5 + P_6 = 1 \qquad (6)$$

- The inequality constraint is $0 \le P_i \le 1$ (but in this case it is not necessary to include)
- Solution of the optimization problem (e.g. by the Lagrange method[*]) yields a single maximum:

$$P_1 = P_2 = P_3 = P_4 = P_5 = P_6 = 1/6 \qquad (7)$$

- The entropy is $\Phi = -6 \ (1/6) \ln (1/6) = \ln 6$. In general the entropy for w equiprobable outcomes is

$$\Phi = \ln w \qquad (8)$$

- In this case, the application of the ME principle (mathematically, an "extremization" form) is equivalent to the principle of insufficient reason (Bernoulli-Laplace; mathematically, an "equation" form).

---

[*] Simple calculations like this and other in the following pages should be regarded as homework.

Entropy maximization: The loaded die example

- What is the probability that the outcome of a die throw will be i if we know that it is loaded, so that $P_6 - P_1 = 0.2$?

- The principle of insufficient reason does not work in this case.

- The ME principle works. We simply pose an additional constraint:

$$P_6 - P_1 = 0.2$$

- The solution of the optimization problem (e.g. by the Lagrange method) is a single maximum as shown in the figure.

- The entropy is $\Phi = 1.732$ smaller than in the case of equiprobability, where $\Phi = \ln 6 = 1.792$.
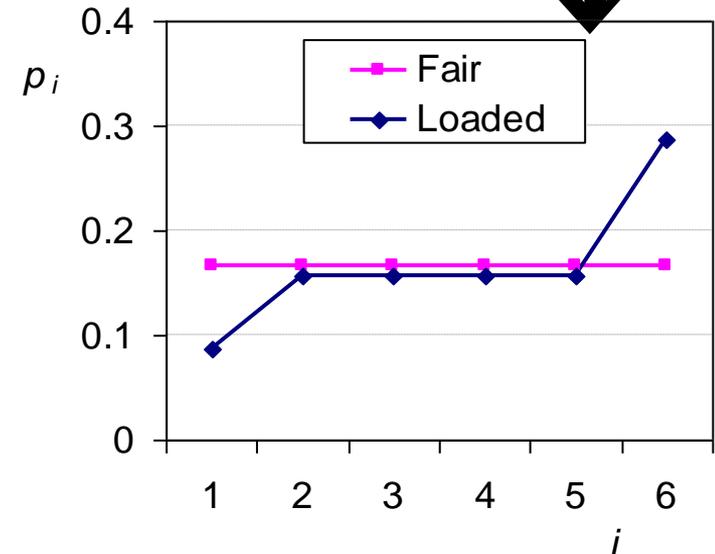
The decrease of entropy in the loaded die derives from the additional information incorporated in the constraints.

Entropy and information are complementary to each other.

When we know (observe) that the outcome is i ($P_i = 1$, $P_j = 0$ for $j \neq i$), the entropy is zero.

Expected values as constraints: General solution

- In the most typical application of the ME principle, we wish to infer the probability density function $f(z)$ of a continuous random variable z (scalar or vector) with constraints formulated as expectations of functions $g_j(z)$.

- In other words, the given information used in the ME principle is expressed as a set of constraints formed as

$$E[g_j(\underline{z})] = \int_{-\infty}^{\infty} g_j(z)\, f(z)\, \mathrm{d}z = \eta_j,\ j = 1, \dots, n \tag{9}$$

- The resulting maximum entropy distribution by maximizing entropy as given in (4) with constraints (9) and the obvious additional constraint (3) is (Papoulis, 1991, p. 571) is

$$f(z) = \exp\left[-\lambda_0 - \sum_{j=1}^{n} \lambda_j\, g_j(z) \right] \tag{10}$$

where λ_0 and λ_j are constants determined such as to satisfy (3) and (9), respectively.

- The resulting maximum entropy is

$$\Phi[\underline{z}] := \lambda_0 + \sum_{j=1}^{n} \lambda_j\, \eta_j \tag{11}$$

Typical results of entropy maximization

Constraints for the continuous variable z	Resulting distribution $f(z)$ and entropy Φ (for $h(z) = 1$)
z bounded within $[0, w]$, no equality constraint	$f(z) = 1/w$ (uniform) $\Phi = \ln w$
z unbounded from both below and above	
No constraint or constrained mean μ	not defined
Constrained mean μ and standard deviation σ	$f(z) = \exp\{-[(z - \mu)/\sigma]^2/2\} / (\sigma\sqrt{2\pi})$ (Gaussian) $\Phi = \ln (\sigma\sqrt{2\pi e})$
Nonnegative z unbounded from above	
No equality constraint	not defined
Constrained mean μ	$f(z) = (1/\mu) \exp(-z/\mu)$ (exponential) $\Phi = \ln (\mu e)$
Constrained mean μ and standard deviation σ with $\sigma < \mu$	$f(z) = A \exp\{-[(z - \alpha)/\beta]^2/2\}$ (truncated Gaussian tending to exponential as $\sigma \rightarrow \mu$). The constants A, α and β are determined from the constraints and Φ from (11) (the equations are involved and are omitted)
As above but with $\sigma > \mu$	not defined

A first application of the ME principle to uncertain motion of a particle: setup[*]

- We consider a motionless cube with edge a (volume $V = a^3$) containing spherical particles of mass m_0 (e.g. monoatomic molecules) in fast motion[†], in which we cannot observe the exact position and velocity.

- A particle's state is described by 6 variables, 3 indicating its position $\underline{x}_i$ and 3 indicating its velocity $\underline{u}_i$, with $i = 1, 2, 3$; all are represented as random variables, forming the vector $\underline{z} = (\underline{x}_1, \underline{x}_2, \underline{x}_3, \underline{u}_1, \underline{u}_2, \underline{u}_3)$.

- The constraints for position are:

$$0 \le \underline{x}_i \le a, \; i = 1, 2, 3 \tag{12}$$

- The constraints for velocity are (where the integrals are over feasible space Ω, i.e. $(0, a)$ for each x_i and $(-\infty, \infty)$ for each u_i):
  - Conservation of momentum: $E[m_0 \underline{u}_i] = m_0 \int_\Omega u_i f(\underline{z}) \, \mathrm{d}\underline{z} = 0$ (the cube is motionless), or:

$$E[\underline{u}_i] = 0, \; i = 1, 2, 3 \tag{13}$$

  - Conservation of energy[‡]: $E[m_0 \|\underline{u}\|^2/2] = (m_0/2) \int_\Omega \|\underline{u}\|^2 f(\underline{z}) \, \mathrm{d}\underline{z} = \varepsilon$, where ε is the energy per particle and $\|\underline{u}\|^2 = u_1^2 + u_2^2 + u_3^2$; thus, the constraint is

$$E[\|\underline{u}\|^2] = 2\varepsilon/m_0 \tag{14}$$

---

[*] This analysis (and those of the following pages up to p. 22) is explained in more detail in Koutsoyiannis (2013).
[†] See a 2D animation in en.wikipedia.org/wiki/File:Translational_motion.gif.
[‡] The expectation $E[\underline{u}_i]$ represents a macroscopic motion, while $\underline{u}_i - E[\underline{u}_i]$ represents fluctuation at a microscopic level. If $E[\underline{u}_i] \ne 0$, then the macroscopic and microscopic kinetic energies should be treated separately, the latter being $\varepsilon = E[m_0 (\|\underline{u} - E[\underline{u}]\|)^2/2]$; see also pp. 14-15.

A first application of the ME principle to uncertain motion of a particle: results

- We form the entropy of $\underline{z}$ as in (4) recognizing that the constant density $h(\underline{z})$ in ln $[f(\underline{z})/h(\underline{z})]$ should have units $[\underline{z}^{-1}] = [x^{-3}] [u^{-3}] = [L^{-6} T^3]$. To make this, we utilize a universal constant, i.e. the Planck constant h = 6.626 × 10^{-34} J·s; its dimensions are $[L^2 M T^{-1}]$. If we combine it with the particle mass m_0, we observe that the quantity $(m_0/h)^3$ has the required dimensions $[L^{-6} T^3]$, thereby giving the entropy as

$$\Phi[\underline{z}] := E[-\ln[(h/m_0)^3 f(\underline{z})]] = -\int_\Omega \ln [(h/m_0)^3 f(\underline{z})] \, f(\underline{z}) \, d\underline{z} \qquad (15)$$

- Application of the principle of maximum entropy with constraints (3), (12), (13) and (14) will give the distribution of $\underline{z}$ (see proof below) as:

$$f(\underline{z}) = (1/a)^3 (3m_0 / 4\pi\varepsilon)]^{3/2} \exp(-3m_0 \|\boldsymbol{u}\|^2 / 4\varepsilon), \ \ 0 \le \underline{x}_i \le a \qquad (16)$$

- The marginal distribution of each of the location coordinates x_i is uniform in [0, a], i.e.,

$$f(x_i) = 1/a, \ \ i = 1, 2, 3 \qquad (17)$$

- The marginal distribution of each of the velocity coordinates u_i is derived as

$$f(u_i) = (3m_0 / 4\pi\varepsilon)^{1/2} \exp(-3m_0 u_i^2 / 4\varepsilon), \ \ i = 1, 2, 3 \qquad (18)$$

This is Gaussian with mean 0 and variance $2\varepsilon / 3m_0$ = 2 × energy per unit mass per degree of freedom.

A first application of the ME principle to uncertain motion of a particle: results (2)

- The marginal distribution of the velocity magnitude $\|u\|$ results as:

$$f(\|u\|) = (2/\pi)^{1/2}(3m_0 / 2\varepsilon)^3 \|u\|^2 \exp(-3m_0\|u\|^2/ 4\varepsilon) \qquad (19)$$

  This is known as the Maxwell–Boltzmann distribution.

- The entropy is then calculated as follows, where e is the base of natural logarithms:

$$\Phi[\underline{z}] = \frac{3}{2} \ln\left(\frac{4\pi e}{3} \frac{m_0}{h^2} \varepsilon\, V^{2/3} \right) = \frac{3}{2} \ln\left(\frac{4\pi e}{3} \frac{m_0}{h^2} \right) + \frac{3}{2} \ln \varepsilon + \ln V \qquad (20)$$

  An extended version of (20) (for many particles), but with some differences, is known as the Sackur-Tetrode equation (after H. M. Tetrode and O. Sackur, who developed it independently at about the same time in 1912).

- From (16) we readily observe that the joint distribution $f(z)$ is a product of functions of z's coordinates x_1, x_2, x_3, u_1, u_2, u_3. This means that all six random variables are jointly independent. The independence results from entropy maximization.

- From (16) and (18) we also observe a symmetry with respect to the three velocity coordinates, resulting in uniform distribution of the energy ε into $\varepsilon/3$ for each direction or degree of freedom. This is known as the **equipartition** principle and is again a result of entropy maximization.

- From (20) we can verify that the entropy $\Phi[\underline{z}]$ is a dimensionless quantity.

Sketch of proof of equations (16)-(20)

According to (10) and taking into account the equality constraints (3), (13) and (14), the ME distribution will have density $f(z) = \exp[-\lambda_0 - \lambda_1 u_1 - \lambda_2 u_2 - \lambda_3 u_3 - \lambda_4(u_1^2 + u_2^2 + u_3^2)]$. This proves that the density will be an exponential function of a second order polynomial of (u_1, u_2, u_3) involving no products of different u_i. The $f(z)$ in (16) is of this type, and thus it suffices to show that it satisfies the constraints.

Note that the inequality constraint (12) is not considered at this phase but only in the integration to evaluate the constraints. That is, the integration domain will be $\Omega := \{(0 \le x_1 \le a, 0 \le x_2 \le a, 0 \le x_3 \le a, -\infty < u_1 < \infty, -\infty < u_2 < \infty, -\infty < u_3 < \infty)\}$. We denote by $\int_\Omega dz$ the integral over this domain. It is easy then to show (the integrals are trivial) that:

$\int_\Omega f(z)\, dz = 1$; $\int_\Omega u_1 f(z)\, dz = 0$; $\int_\Omega u_2 f(z)\, dz = 0$; $\int_\Omega u_3 f(z)\, dz = 0$; $\int_\Omega (u_1^2 + u_2^2 + u_3^2) f(z)\, dz = 2\varepsilon/m_0$.

Thus, all constraints are satisfied.

To find the marginal distribution of each of the variables we integrate over the entire domain of the remaining variables; due to independence this is very easy and the results are given in (17) and (18). To find the marginal distribution of $\|u\|$ (eqn. (19)), we recall that the sum of squares of n independent $N(0, 1)$ random variables has a $\chi^2(n)$ distribution (Papoulis, 1990, p. 219, 221) and then we use known results for the density of a transformation of a random variable (Papoulis, 1990, p. 118) to obtain the distribution of the square root, thus obtaining (19).

To calculate the entropy, we observe that $-\ln[f(z)] = (3/2) \ln[4\pi\varepsilon / 3m_0] + \ln a^3 + 3m_0 (u_1^2 + u_2^2 + u_3^2) / 4\varepsilon$ and $\ln[h(z)] = 3 \ln(m_0/h)$. Thus, the entropy, whose final value is given in (20) is derived as follows:

$\Phi[\underline{z}] = \int_\Omega \{-\ln[f(z)] + \ln[h(z)]\} f(z)\, dz = (3/2) \ln[(4\pi\varepsilon / 3m_0 (m_0/h)^2)] + \ln a^3 + (3m_0 / 4\varepsilon)(2\varepsilon/m_0)$

$= (3/2) \ln(4\pi m_0 / 3 h^2) + (3/2) \ln \varepsilon + \ln V + (3/2)$.

In a similar manner we can prove the results of the cases discussed below, whose proofs are omitted.

The ME principle applied to one diatomic molecule

- Analysis of diatomic gases is important for atmospheric physics because the dominant atmospheric gases are diatomic (N_2, O_2).

- In a diatomic gas, in addition to the kinetic energy, we have rotational energy at two axes x and y perpendicular to the axes defined by the two molecules; these are $L_x^2 / 2I$ and $L_y^2 / 2I$, where L denotes angular momentum and I denotes rotational inertia and has dimensions $[M\,L^2]$ (due to symmetry, $I_x = I_y = I$).

- We consider again a motionless cube with edge a (volume $V = a^3$) containing identical diatomic molecules, each one with mass m_0 and kinetic and rotational energy ε.

- Each molecule is described by 8 variables, 3 indicating its position $\underline{x}_i$, 3 indicating its velocity $\underline{u}_i$ ($i = 1, 2, 3$) and two indicating its rotation, $\underline{u}_4 := \underline{L_x}/\sqrt{I\,m_0}$ and $\underline{u}_5 := \underline{L_y}/\sqrt{I\,m_0}$; all are represented as random variables, forming the vector $\underline{z} = (\underline{x}_1, \underline{x}_2, \underline{x}_3, \underline{u}_1, \underline{u}_2, \underline{u}_3, \underline{u}_4, \underline{u}_5)$.

- The constraints are the same as before:
  - position: $0 \le \underline{x}_i \le a$;
  - momentum/angular momentum: $E[\underline{u}_i] = 0$ (the cube is not in motion);
  - energy: $E[\|\underline{u}\|^2] = 2\varepsilon/m_0$.

- The constant density $h(z)$ in $\ln[f(z)/h(z)]$ in (4) should have units $[z^{-1}] = [x^{-3}][u^{-5}] = [L^{-8} T^5]$. Combining the Planck constant h with the particle mass m_0 and rotational inertia I, we observe that the required dimensions are attained by the quantity $m_0^4 I/h^5$, so that

$$\Phi[\underline{z}] := E[-\ln[(h^5/m_0^4 I)\,f(\underline{z})]] = -\int_\Omega \ln[(h^5/m_0^4 I)\,f(\underline{z})]\,f(\underline{z})\,\mathrm{d}\underline{z} \tag{21}$$

The ME principle applied to one diatomic molecule (2)

- Application of the ME principle with the above constraints will give the density function as:

$$f(\mathbf{z}) = (1/a)^3 \, (5m_0 / 4\pi\varepsilon)^{5/2} \exp(-5m_0 \|\mathbf{u}\|^2 / 4\varepsilon), \ \ 0 \le \underline{x}_i \le a \tag{22}$$

which is again uniform for the location components and Gaussian for the translational and rotation components, and indicates independence of all 8 components and equipartitioning of energy (1/5 for each degree of freedom).

- Note that the energy per degree of freedom i is $E[m_0 u_i^2 / 2]$, which for the rotational components becomes $E[m_0 u_4^2 / 2] = E[L_x^2 / 2I]$ and $E[m_0 u_5^2 / 2] = E[L_y^2 / 2I]$.

- The entropy is then calculated as

$$\Phi[\underline{z}] = \frac{5}{2} \ln\left(\frac{4\pi e}{5} \frac{m_0^{3/5} I^{2/5}}{h^2} \varepsilon \, V^{2/5} \right) = \frac{5}{2} \ln\left(\frac{4\pi e}{5} \frac{m_0^{3/5} I^{2/5}}{h^2} \right) + \frac{5}{2} \ln \varepsilon + \ln V \tag{23}$$

- Generalizing (20) and (23) for β degrees of freedom we obtain

$$\Phi[\underline{z}] = \frac{\beta}{2} \ln\left(\frac{4\pi e}{\beta} \frac{m_0^{3/\beta} I^{1-3/\beta}}{h^2} \varepsilon \, V^{2/\beta} \right) = \frac{\beta}{2} \ln\left(\frac{4\pi e}{\beta} \frac{m_0^{3/\beta} I^{1-3/\beta}}{h^2} \right) + \frac{\beta}{2} \ln \varepsilon + \ln V \tag{24}$$

Q: Why the kinetic energy is equally distributed among the different degrees of freedom?
A: Because this maximizes entropy, that is, uncertainty.

The ME principle applied to *N* molecules

- The N molecules are assumed to be of the same kind and thus each one has the same mass m_0, rotational inertia I, and degrees of freedom β (for monoatomic and diatomic molecules, $\beta = 3$ and $\beta = 5$, respectively).

- Their coordinates form a vector $\underline{\mathbf{Z}} = (\underline{\mathbf{z}}_1,\ldots, \underline{\mathbf{z}}_N)$ with $3N$ location coordinates and βN velocity coordinates; this could be rearranged as $\underline{\mathbf{Z}} = (\underline{\mathbf{X}}, \underline{\mathbf{U}})$, with $\underline{\mathbf{X}} = ((\underline{x}_1, \underline{x}_2, \underline{x}_3)_1, \ldots, (\underline{x}_1, \underline{x}_2, \underline{x}_3)_N)$ and $\underline{\mathbf{U}} = ((\underline{u}_1, \ldots, \underline{u}_\beta)_1, \ldots, (\underline{u}_1, \ldots, \underline{u}_\beta)_N)$.

- If E is the total kinetic energy of the N molecules and $\varepsilon = E/N$ is the energy per particle, then conservation of energy yields

$$E[\|\underline{\mathbf{U}}\|^2] = 2E/m_0 = 2N\varepsilon/m_0 \tag{25}$$

- Application of the ME principle with constraints (3), (12), (13) and (25) gives:

$$f(\underline{\mathbf{Z}}) = (1/a)^{3N} (\beta m_0 / 4\pi\varepsilon)^{\beta N/2} \exp(-\beta m_0 \|\underline{\mathbf{U}}\|^2 / 4\varepsilon), \quad 0 \le \underline{x}_i \le a \tag{26}$$

- The entropy for N particles is:

$$\Phi[\underline{\mathbf{Z}}] = \frac{\beta N}{2} \ln\left(\frac{4\pi e}{\beta} \frac{m_0^{3/\beta} I^{1-3/\beta}}{h^2} \varepsilon\, V^{2/\beta} \right) = \frac{\beta N}{2} \ln\left(\frac{4\pi e}{\beta} \frac{m_0^{3/\beta} I^{1-3/\beta}}{h^2} \right) + \frac{\beta N}{2} \ln \varepsilon + N \ln V \tag{27}$$

$$\Phi[\underline{\mathbf{Z}}] = \frac{\beta N}{2} c + \frac{\beta N}{2} \ln \varepsilon + N \ln V = \frac{\beta N}{2} c + \frac{\beta N}{2} \ln \frac{E}{N} + N \ln V \tag{28}$$

where c incorporates mathematical and physical constants.

Is the entropy subjective or objective?

- In physics most quantities are subjective in the sense that they depend on the *observer*. There may be also some objective quantities that are unaltered if the observer's choices change.
  - Thus, the location coordinates (x_1, x_2, x_3) depend on the observer's choice of the coordinate frame and change if this frame is translated or rotated; however the distance between two points remains constant if the frame changes.
  - Also, the velocity depends on the relative motion of the frame of reference; the velocity of a car whose speedometer indicates 100 km/h is zero for an observer moving with the car, 100 km/h for an observer sitting at the road and 107 000 km/h for a coordinate system attached to the sun. The kinetic energy, as well as changes thereof, depend on the reference frame, too.
- Surprisingly, however, the entropy $\Phi[\underline{Z}]$ of the gas in a container of a fixed volume V, whose general form is given in (24), does not change with the change of the reference frame (see demonstration below), provided that the kinetic energy per gas molecule ε is defined based on the difference of velocity $\underline{u}$ from its mean $E[\underline{u}]$, i.e., $\varepsilon = E[m_0 (\|\underline{u} - E[\underline{u}]\|)^2/2]$; in this case ε is also invariant, despite that $\underline{u}$ changes with the reference frame. The invariance extends to the entropy maximizing distribution.
- Therefore, despite that entropy is based on probabilities, it is an objective quantity that can be measured and its magnitude does not depend on the reference frame.

Demonstration of invariance properties of entropy

We consider again the container with a gas with spherical particles (3 degrees of freedom). We assume that an observer is moving with velocity u_0 parallel to the horizontal axis x_1 as shown in figure. According to this observer each particle has location coordinates x'_i and velocity coordinates u'_i related to those of the fixed frame, x_i and u_i, by the relationships shown in figure.

$$x'_1 = x_1 - x_0 \qquad u'_1 = u_1 + u_0$$
$$x'_2 = x_2 \qquad u'_2 = u_2$$
$$x'_3 = x_3 \qquad u'_3 = u_3$$

For the moving frame, $E[\underline{u}'_1] = u_0$ while $E[\underline{u}'_i] = 0$ for $i = 2, 3$. Thus, $E[m_0 (\|\underline{u}' - E[\underline{u}']\|)^2/2] = E\{m_0 [(\underline{u}'_1 - u_0)^2 + \underline{u}'^2_2 + \underline{u}'^2_3]\}/2 = \varepsilon$, the same as for the fixed frame.

For one molecule, the density function will be

$f(\textbf{z}) = (1/a)^3 (3m_0 / 4\pi\varepsilon)]^{3/2} \exp\{-3m_0[(\underline{u}'_1 - u_0)^2 + \underline{u}'^2_2 + \underline{u}'^2_3]/ 4\varepsilon\}$, $-x_0 \leq x'_1 \leq -x_0 + a$, $0 \leq x'_2 \leq a$, $0 \leq x'_2 \leq a$

It can be verified, using the same method as in p. 10, that it satisfies all constraints.

To calculate the entropy we observe that $-\ln[f(\textbf{z})] = (3/2) \ln [4\pi\varepsilon / 3m_0)] + \ln a^3 + 3m_0 [(\underline{u}'_1 - u_0)^2 + \underline{u}'^2_2 + \underline{u}'^2_3] / 4\varepsilon$ and $\ln [h(\textbf{z})] = 3 \ln (m_0/h)$. Thus, the entropy is calculated as $\Phi[\underline{z}] = (3/2) \ln (4\pi e m_0 /3 h^2) + (3/2) \ln \varepsilon + \ln V$, which is the same as in (20).

Likewise, the entropy of N molecules with respect to the moving frame will be the same as that with respect to the fixed frame.
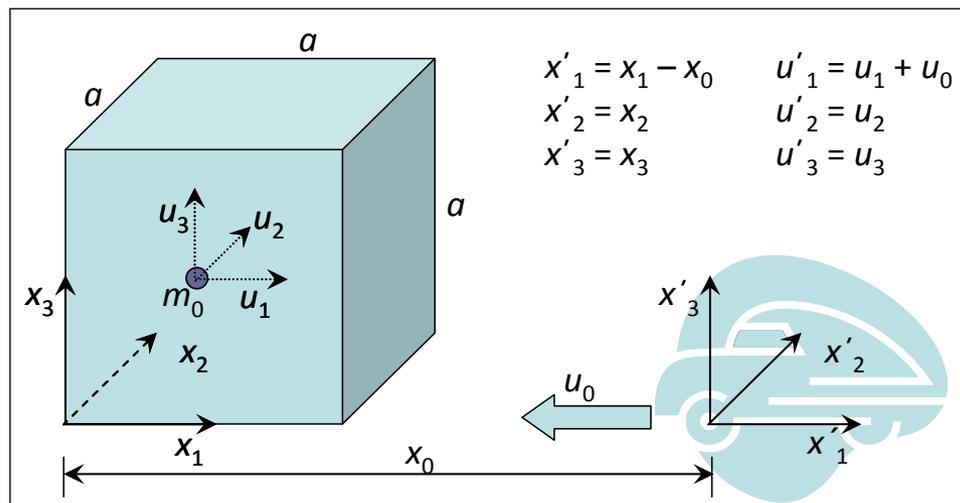
The extensive entropy

- In very large systems, such as the atmosphere as a whole, all physical quantities change with location and the homogeneity (independence of expected values from location) assumed in our "gas container" example does not hold.

- Still the equations we have derived are valid but at a local scale, i.e. at a small volume V for which homogeneity can be assumed. It looks convenient to use as a possibly objective quantity the entropy for a single particle $\varphi(\varepsilon, V) := \Phi[\underline{z}] = (\beta/2)c + (\beta/2) \ln \varepsilon + \ln V = \Phi[\underline{Z}]/N$, where N is the number of particles contained in volume N.

- However, $\varphi(\varepsilon, V)$ is not an objective/invariant quantity, as it depends on the selection of the volume V. To make it objective, we observe that the quantity $\varphi(\varepsilon, V) - \ln N = (\beta/2)c + (\beta/2) \ln \varepsilon + \ln v =: \varphi^*(\varepsilon, v)$ where $v := V / N$, is invariant under change of V, provided that the density of particles N/V is fairly uniform.

- This leads to the definition of two derivative quantities, which we call *standardized entropies* and more specifically, *intensive entropy* and *extensive entropy* respectively:

$$\varphi^*(\varepsilon, v) := \varphi(\varepsilon, v) = \varphi(\varepsilon, V) - \ln N = \frac{\beta}{2} \ln\left(\frac{4\pi e}{\beta} \frac{m_0^{3/\beta} I^{1 - 3/\beta}}{h^2} \right) + \frac{\beta}{2} \ln \varepsilon + \ln v \qquad (29)$$

$$\Phi^*(E, V, N) := N \varphi^*(\varepsilon, v) = \Phi[\underline{Z}] - N \ln N = \frac{\beta N}{2} \ln\left(\frac{4\pi e}{\beta} \frac{m_0^{3/\beta} I^{1 - 3/\beta}}{h^2} \right) + \frac{\beta N}{2} \ln \frac{E}{N} + N \ln \frac{V}{N} \qquad (30)$$

Properties of intensive and extensive entropy

- Like the energy per particle, ε, and the volume per particle, v, the standardized entropy per particle $\varphi^*(\varepsilon, v)$, is an *intensive* property (hence its name) in the sense that it does not depend on the size of system that an observer, justifiably or arbitrarily, considers.

- In contrast, the total energy, E, the volume, V, and the number of particles, N, are *extensive* properties in the sense that depend on the observer's selection of the system and are proportional one another; that is, a system of volume αV, where α is any positive number contains αN particles with a total energy αE. Likewise, the extensive entropy $\Phi^*(E, V, N)$ is indeed an extensive property, as it is easily seen that

$$\Phi^*(\alpha E, \alpha V, \alpha N) := \alpha \, \Phi^*(E, V, N) \tag{31}$$

- It is noted that the probabilistic entropy per particle $\varphi(\varepsilon, V) = \Phi[\underline{z}]$ is not intensive as it depends on the system volume V. Likewise, the probabilistic total entropy $\Phi(E, V, N) = \Phi[\underline{Z}]$ is not extensive: it can be seen that

$$\Phi(\alpha E, \alpha V, \alpha N) - \alpha \, \Phi(E, V, N) = \alpha N \ln \alpha \neq 0 \tag{32}$$

Interpretation of extensive entropy

- An interpretation of standardized entropies $\varphi^*(\varepsilon, v)$ and $\Phi^*(E, V, N)$ is that they are not strictly (probabilistic) entropies, but differences of entropies, taken with the aim to define quantities invariant under change of the observer's choices (as in taking differences of linear coordinates to make them invariant under translation of the frame of reference).

- The reference entropies, from which these differences are taken are $\ln N$ and $N \ln N = \ln N^N$ for φ and Φ^*, respectively. Thus, φ^* or Φ^* measures how much larger the entropy $\Phi[\underline{z}]$ or $\Phi[\underline{Z}]$, respectively, is from the entropy of a simplified reference system, in which only the particle location, discretized into N bins, counts (with the number N of bins here representing a discretization of the volume V that is not a subjective choice of an observer). Clearly, in gases (and fluids in general) there are N and N^N ways of placing one and N particles, respectively, in the N bins, so that the reference entropies are $\ln N$ and $N \ln N$, respectively[*].

- An easy perception of $\varphi^*(\varepsilon, v)$ is that it is identical to the probabilistic entropy of a system with a fixed volume equal to v. Also $\Phi^*(E, V, N)$, is identical to the probabilistic entropy of a system of N particles, each of which is restricted in a volume v.

- A more common interpretation of φ^* and Φ^* is that they in fact represent the probabilistic entropies of the gas under study, under the assumption that the particles are indistinguishable. This interpretation has several problems.

---

[*] Notably in solids the locations of particles are fixed (only one possible way) and thus the reference entropy is $\ln 1 = 0$. Thus, φ^* and Φ^* in solids become identical to φ and Φ, respectively, which agrees with the classical result for solids.

Equivalence of descriptions by the two entropy measures

- We assume that N molecules are in motion in a container of volume V and entropy per particle φ.

- We make an arbitrary partition of the container into two parts A and B with volumes V_A and V_B, respectively, with $V_A + V_B = V$. The partition is only mental—no material separation was made. Therefore, at any instance any particle can be either in part A with probability π, or in part B with probability $1 - \pi$.

- We assume that we are given the information that a particle is in part A or B. We denote the conditional entropy, for each of the two cases as φ_A and φ_B, respectively.

- The unconditional entropy (for the unpartitioned volume) can be calculated from the conditional entropies as (see proof in box below)

$$\varphi = \pi\, \varphi_A + (1 - \pi)\, \varphi_B + \varphi_\pi, \quad \text{where } \varphi_\pi := -\pi \ln \pi - (1 - \pi) \ln (1 - \pi) \tag{33}$$

- Substituting N_A/N for π and N_B/N for $(1 - \pi)$, where N_A and N_B are the expected number of particles in parts A and B respectively, we get (see proofs in box below)

$$\Phi = \Phi_A + \Phi_B + N \ln N - N_A \ln N_A - N_B \ln N_B \tag{34}$$

$$\Phi^* = \Phi^*_A + \Phi^*_B \tag{35}$$

- Equations (34) are (35) precisely equivalent and describe the same thing, using either probabilistic entropies Φ or extensive entropies Φ^*. Apparently, (35) is simpler and therefore the description using Φ^* is more convenient when the number of particles N matters.

Proof of equations (33)-(35)

To show that (33) holds true (cf. also Papoulis, 1991, p. 544) we observe that the unconditional density $f(z)$ is related to the conditional ones $f(z|A)$ and $f(z|B)$ by

$$f(z) = \begin{cases} \pi\, f(z|A), & z \text{ in } A \\ (1-\pi)\, f(z|B), & z \text{ in } B \end{cases} \tag{36}$$

Denoting $\int_A g(z)\,dz$ and $\int_B g(z)\,dz$ the integral of a function $g(z)$ over the intersection of the domain of z with the volume A and volume B, respectively, the unconditional entropy will be

$$\begin{aligned} \varphi = \Phi[z] \ \ &= -\int_A f(z)\ln[f(z)/h(z)]dz - \int_B f(z)\ln[f(z)/h(z)]dz = \\ &= -\int_A \pi\, f(z|A)\ln[\pi\, f(z|A)/h(z)]dz - \int_B (1-\pi)\, f(z|B)\ln[(1-\pi)\, f(z|B)/h(z)]dz = \\ &= -\pi \int_A f(z|A)\ln[f(z|A)/h(z)]dz - \pi\ln\pi \int_A f(z|A)\,dz \\ &\quad - (1-\pi)\int_B f(z|B)\ln[f(z|B)/h(z)]dz - (1-\pi)\ln(1-\pi)\int_B f(z|B)\,dz \end{aligned}$$

We observe that $\int_A f(z|A)\ln[f(z|A)/h(z)]dz = \varphi_A$ and $\int_B f(z|B)\ln[f(z|B)/h(z)]dz = \varphi_B$, whereas $\int_A f(z|A)\,dz = \int_B f(z|B)\,dz = 1$. Evidently, then, (33) follows directly.

From (33) we get:

$$\varphi = (N_A/N)\,\varphi_A + (N_B/N)\,\varphi_B - (N_A/N)\ln(N_A/N) - (N_B/N)\ln(N_B/N)$$

$$N\varphi = N_A\,\varphi_A + N_B\,\varphi_B - N_A\ln N_A - N_B\ln N_B + N\ln N$$

so that (34) follows directly. In turn, from (34) we obtain

$$\varphi - \ln N = (N_A/N)\,(\varphi_A - \ln N_A) + (N_B/N)\,(\varphi_B - \ln N_B)$$

$$N\varphi^* = N_A\varphi^*_A + N_B\,\varphi^*_B$$

from which (35) follows directly.

Definition of temperature

- Temperature is defined to be the inverse of the partial derivative of entropy with respect to energy, i.e.,

$$\frac{1}{\vartheta} := \frac{\partial \Phi}{\partial E} \tag{37}$$

Since entropy is dimensionless and E has dimensions of energy, temperature has also dimensions of energy (joules). This contradicts the common practice of using different units of temperature, such as kelvins or degrees Celsius. To distinguish from the common practice, we use the symbol ϑ (instead of T which is used for temperature in kelvins[*]) and we call ϑ the *natural* temperature (instead of absolute temperature for T).

- From (27), (24), (30), (29) we obtain

$$\frac{1}{\vartheta} = \frac{\partial \Phi}{\partial E} = \frac{\partial \Phi^*}{\partial E} = \frac{\partial \varphi}{\partial \varepsilon} = \frac{\partial \varphi^*}{\partial \varepsilon} = \frac{\beta}{2\varepsilon} \tag{38}$$

or

$$\vartheta = \frac{2\varepsilon}{\beta} \tag{39}$$

That is, the temperature is proportional to the kinetic energy per particle; in fact it equals twice the particle's kinetic energy per degree of freedom.

---

[*] We can regard the unit of kelvin an energy unit, a multiple of the joule, like the calorie and the Btu, i.e. 1 K = 0.138 06505 yJ .

The law of ideal gases

- We consider again the cube of edge a containing N identical molecules of a gas, each with mass m_0 and $β$ degrees of freedom.

- We consider a time interval dt; any particle at distance from the bottom edge $dx_3 \le -u_3 \, dt$ will collide with the cube edge ($x_3 = 0$).

- From (22), generalized for $β$ degrees of freedom, the joint distribution function of (x_3, u_3) of a single particle is

$$f(x_3, u_3) = (1/a)(β \, m_0 / 4π \, ε)^{1/2} \exp(−β m_0 u_3^2 / 4ε) \tag{40}$$

- Thus, the expected value of the momentum $\underline{g}(dt)$ of molecules colliding at the cube edge ($x_3 = 0$) within time interval dt is

$$E[\underline{g}(dt)] = N \int_0^\infty dx_3 \int_{-\infty}^{-x_3/dt} m_0 \, u_3 \, f(x_3, u_3) \, du_3 = N \, ε \, dt/(β \, a) \tag{41}$$

- According to Newton's 2nd law, the force exerted on the edge is $F = 2 \, E[\underline{g}(dt)]/dt$ and the pressure is $p = F / a^2 = 2 \, N \, ε /(β \, V)$, or finally (by using (39)),

$$p = N \, ϑ / V = ϑ / v \Leftrightarrow p \, V = N \, ϑ \Leftrightarrow p \, v = ϑ \tag{42}$$

This is the well-known law of ideal gasses written for natural temperature.

Q: Is the **law of ideal gases** an empirical relationship or can it be deduced and how?
A: It can be easily derived by maximizing entropy.

Alternative expression of entropy

The definition of temperature ϑ and its relationship with kinetic energy per particle ε (equation (39)) along with law of ideal gasses (equation (42)) allows expressing the intensive entropy φ^* (equation (29)) in terms of temperature and pressure as follows

$$\varphi^* = \frac{\beta}{2} \ln\left(2\pi e \, \frac{m_0^{3/\beta} l^{1-3/\beta}}{h^2} \right) + \left(1 + \frac{\beta}{2} \right) \ln \vartheta - \ln p \tag{43}$$

Differential form of entropy

- From (27), (24), (30), (29), the partial derivatives of entropy with respect to V and v are

$$\frac{\partial \Phi}{\partial V} = \frac{\partial \Phi^*}{\partial V} = \frac{\partial \varphi^*}{\partial v} = \frac{1}{v} = \frac{p}{\vartheta}, \quad \frac{\partial \varphi}{\partial V} = \frac{1}{V} \tag{44}$$

  where the term p/ϑ was obtained from the ideal gas law.

- From the same equations, the partial derivatives with respect to N are

$$\frac{\partial \Phi}{\partial N} = N \frac{\partial \varphi}{\partial N}, \quad \frac{\partial \Phi^*}{\partial N} = \varphi^* + N \frac{\partial \varphi^*}{\partial N} \tag{45}$$

- The latter can be simplified by using of the so called *chemical potential*, μ (cf., Wannier, 1987, p. 139)[*],

$$-\frac{\mu}{\vartheta} := \frac{\partial \Phi^*}{\partial N} = \varphi^* + N \frac{\partial \varphi^*}{\partial N} \tag{46}$$

- Based on the above, the probabilistic and the extensive entropy can be written, respectively, in differential form as

$$d\Phi = \frac{1}{\vartheta} dE + \frac{p}{\vartheta} dV + \left(-\frac{\mu}{\vartheta} + 1 + \ln N \right) dN, \quad d\Phi^* = \frac{1}{\vartheta} dE + \frac{p}{\vartheta} dV - \frac{\mu}{\vartheta} dN \tag{47}$$

- The latter can be written in the equivalent form

$$\vartheta\, d\Phi^* = dE + p\, dV - \mu\, dN \tag{48}$$

---

[*] Note that the definition applies when E is the internal energy, which for gases is identical to the thermal energy.

Gas mixtures: probability distribution

- We assume that our cube of edge a contains a mixture of two gases: N_A molecules of a gas A with particle mass m_{0_A} and β_A degrees of freedom and N_B molecules of a gas B with particle mass m_{0_A} and β_B degrees of freedom. We split the vector of velocities into two sub-vectors, i.e $\underline{U} = (\underline{U}_A, \underline{U}_B)$, with $\underline{U}_A = ((\underline{u}_1, ..., \underline{u}_{\beta_A})_1, ..., (\underline{u}_1, ..., \underline{u}_{\beta_A})_{N_A})$ and $\underline{U}_B = ((\underline{u}_1, ..., \underline{u}_{\beta_B})_1, ..., (\underline{u}_1, ..., \underline{u}_{\beta_B})_{N_B})$.

- The characteristic quantities of the mixture are total number of are given in the table:

	Total	Average
Number of particles	$N = N_A + N_B$	
Mass	$M = m_{0_A}N_A + m_{0_B}N_B$	$m_0 = (m_{0_A}N_A + m_{0_B}N_B)/N$
Energy	$E = E_A + E_B$	$\varepsilon = E / N$
Degrees of freedom	$\beta N = \beta_A N_A + \beta_B N_B$	$\beta = (\beta_A N_A + \beta_B N_B)/N$

- Conservation of energy yields:

$$E[m_{0_A} \|\underline{U}_A\|^2 + m_{0_B} \|\underline{U}_B\|^2] = 2E = 2N\varepsilon \tag{49}$$

- Maximization of entropy will give the following probability density function:

$$f(\underline{Z}) = (1/a)^{3N} (\beta/4\pi\varepsilon)^{\beta N/2} m_{0A}{}^{\beta_B N_A/2} m_{0B}{}^{\beta_B N_B/2} \exp[-m_{0A}\|\underline{U}_A\|^2 - m_{0B}\|\underline{U}_B\|^2)(\beta/4\varepsilon)], \ 0 \le \underline{x}_i \le a \tag{50}$$

Gas mixtures: entropy

- The maximized entropy of the gas mixture is then calculated as:

$$\Phi[\underline{Z}] = \frac{\beta N}{2} \ln c + \frac{\beta N}{2} \ln \varepsilon + N \ln V \qquad (51)$$

  where c incorporates mathematical and physical constants.

- The interpretation of the above results is that a mixture of gases, statistically behaves like a hypothetical single gas with molecular mass, energy per particle and degrees of freedom equal to the corresponding averages in the mixture of gases.

- However, a mixture is not identical to a single gas as demonstrated with the following example (Q&A).

Q: Why in the **composition of Earth's atmosphere**, nitrogen and oxygen are present and hydrogen is absent, while in planets far from the Sun hydrogen is present?

A: Because, to maximize entropy, the kinetic energy is equally distributed among different molecules; hence, hydrogen, which has molecular mass lower than oxygen and nitrogen, moves faster (~4 times) and escapes to space, while nitrogen and oxygen cannot reach the escape velocity.

Planets far from the sun have lower temperature, which is proportional to the kinetic energy, and thus hydrogen cannot reach the escape velocity.

Bringing two systems in contact

- We consider two systems initially isolated to each other. System A with volume V_A contains N_A molecules with average energy ε_A; system B with volume V_B contains N_B molecules with average energy ε_B; both systems contain molecules of the same type (equal molecular mass m_0 and degrees of freedom β).

- The extensive entropies of the two systems are:

$$\Phi^*_A = \frac{\beta N_A}{2} \ln c + \frac{\beta N_A}{2} \ln \varepsilon_A + N_A \ln \frac{V_A}{N_A}, \quad \Phi^*_B = \frac{\beta N_B}{2} \ln c + \frac{\beta N_B}{2} \ln \varepsilon_B + N_B \ln \frac{V_B}{N_B} \tag{52}$$

- As far as the systems are isolated, the total entropy will be the sum of the two partial ones:

$$\Phi^{*(0)} = \frac{\beta N}{2} \ln c + \frac{\beta N_A}{2} \ln \varepsilon_A + \frac{\beta N_B}{2} \ln \varepsilon_B + N_A \ln \frac{V_A}{N_A} + N_B \ln \frac{V_B}{N_B} \tag{53}$$

- Now let us bring the systems in contact, but keep them isolated from the environment. There are two kinds of system interactions. In a *closed* interaction the two systems exchange energy but not mass, reaching at an *equilibrium* state (1) with energies per particle $\varepsilon_A^{(1)}$ and $\varepsilon_B^{(1)}$, respectively. Due to conservation of energy,

$$N_A \varepsilon_A^{(1)} + N_B \varepsilon_B^{(1)} = N_A \varepsilon_A + N_B \varepsilon_B = E \tag{54}$$

- In an *open* interaction, the systems can also exchange mass reaching at an *equilibrium* state (2), where in addition to energy, mass conservation should also be considered, i.e.,

$$N_A^{(2)} + N_B^{(2)} = N_A + N_B = N \tag{55}$$

Closed interaction

- After a closed interaction, $\varepsilon_A^{(1)}$ and $\varepsilon_B^{(1)}$ can be determined by maximizing the extensive entropy:

$$\Phi^{*(1)} = \frac{\beta N}{2}\ln c + \frac{\beta N_A}{2}\ln \varepsilon_A^{(1)} + \frac{\beta N_B}{2}\ln \varepsilon_B^{(1)} + N_A \ln \frac{V_A}{N_A} + N_B \ln \frac{V_B}{N_B} \qquad (56)$$

- Maximization of $\Phi^{(1)}$ with respect to $\varepsilon_A^{(1)}$ and $\varepsilon_B^{(1)}$ subject to (54), results in $\varepsilon_A^{(1)} = \varepsilon_B^{(1)} = \varepsilon^{(1)} = E/N$. Thus,

$$\Phi^{*(1)} = \frac{\beta N}{2}\ln c + \frac{\beta N}{2}\ln \varepsilon^{(1)} + N_A \ln \frac{V_A}{N_A} + N_B \ln \frac{V_B}{N_B} \qquad (57)$$

- We observe that the entropy change is

$$\Phi^{(1)} - \Phi^{(0)} = \Phi^{*(1)} - \Phi^{*(0)} = \frac{\beta N}{2}\ln \frac{N_A \varepsilon_A + N_B \varepsilon_B}{N} - \frac{\beta N_A}{2}\ln \varepsilon_A - \frac{\beta N_B}{2}\ln \varepsilon_B \qquad (58)$$

- It is readily understood that $\Phi^{(1)} - \Phi^{(0)} \geq 0$, where the equality sign applies to the case that the initial entropy was already at maximum. The spontaneous increase of entropy, due to entropy maximization, constitutes a version of the 2nd law of thermodynamics.

- It can be easily verified that if the two systems have different degrees of freedom β_A and β_B, then maximum entropy is achieved for $\varepsilon_A^{(1)}/\beta_A = \varepsilon_B^{(1)}/\beta_B$ or for equal temperature:

$$\vartheta_A^{(1)} = \vartheta_B^{(1)} = \vartheta^{(1)} \qquad (59)$$

Open interaction

- After mass and energy exchange, the compound system A + B will reach at a state (2), with the boxes A and B containing $N_A^{(2)}$ and $N_B^{(2)}$ particles with energies per particle $\varepsilon_A^{(2)}$ and $\varepsilon_B^{(2)}$, respectively; the entire volume will be $V = V_A + V_B$.

- Consequently, the new entropy will be

$$\Phi^{*(2)} = \frac{\beta N}{2} \ln c + \frac{\beta N_A^{(2)}}{2} \ln \varepsilon_A^{(2)} + \frac{\beta N_B^{(2)}}{2} \ln \varepsilon_B^{(2)} + N_A^{(2)} \ln \frac{V_A}{N_A^{(2)}} + N_B^{(2)} \ln \frac{V_B}{N_B^{(2)}} \qquad (60)$$

- Maximization of $\Phi^{*(2)}$ with respect to $\varepsilon_A^{(2)}$, $\varepsilon_B^{(1)}$, $N_A^{(2)}$, $N_B^{(2)}$, subject to (54) and (55), results in $\varepsilon_A^{(2)} = \varepsilon_B^{(2)} = \varepsilon^{(2)} = E/N$ and $V_A/N_A^{(2)} = V_B/N_B^{(2)} = V/N$. Thus,

$$\Phi^{*(2)} = \frac{\beta N}{2} \ln c + \frac{\beta N}{2} \ln \varepsilon^{(2)} + N \ln \frac{V}{N} \qquad (61)$$

- We observe that the entropy change is

$$\Phi^{*(2)} - \Phi^{*(0)} = \frac{\beta N}{2} \ln \frac{N_A \varepsilon_A + N_B \varepsilon_B}{N} - \frac{\beta N_A}{2} \ln \varepsilon_A - \frac{\beta N_B}{2} \ln \varepsilon_B + N \ln \frac{V}{N} - N_A \ln \frac{V_A}{N_A} - N_B \ln \frac{V_B}{N_B} \quad (62)$$

$$\Phi^{(2)} - \Phi^{(0)} = \frac{\beta N}{2} \ln \frac{N_A \varepsilon_A + N_B \varepsilon_B}{N} - \frac{\beta N_A}{2} \ln \varepsilon_A - \frac{\beta N_B}{2} \ln \varepsilon_B + N \ln V - N_A \ln V_A - N_B \ln V_B \quad (63)$$

- Clearly, $\Phi^{*(2)} - \Phi^{*(0)} \geq 0$ and $\Phi^{(1)} - \Phi^{(0)} > 0$, which again constitutes a version of the 2nd law of thermodynamics. In particular, even if $\Phi^{*(2)} - \Phi^{*(0)} = 0$, the difference $\Phi^{(1)} - \Phi^{(0)}$ is always positive, reflecting the larger uncertainty due the mixing of the two systems.

Constrained open interaction

- We have seen that if two systems of same composition are brought in contact and are allowed to exchange energy (closed interaction) then they reach to equal energy per particle, or equal temperature. If they are also allowed to exchange mass (open interaction), then they reach to full uniformity, with equal temperature and volume per particle, thus practically forming a single system.

- However, due to external effects, sometimes the two systems cannot reach full uniformity. Assuming that the intensive entropies per particle, when the system will reach a maximum entropy state (3) (not an equilibrium state in the classical sense) will be $\varphi^*_A{}^{(3)}$ and $\varphi^*_B{}^{(3)}$ for parts A and B, respectively, a general expression for the extensive entropy of the compound system will be (omitting the superscript (3) for simplicity):

$$\Phi^* = N_A \varphi^*_A + N_B \varphi^*_B \tag{64}$$

- To maximize Φ^* under constraint (55) (constraint (54) is not used, assuming fixed ε_A and ε_B), we form the function Ψ incorporating the constraint with a Langrage multiplier λ:

$$\Psi = N_A \varphi^*_A + N_B \varphi^*_B + \lambda (N_A + N_B - N) \tag{65}$$

- Equating the derivatives with respect to N_A and N_B to 0 to maximize Ψ, we obtain

$$\frac{\partial \Psi}{\partial N_A} = \varphi^*_A + N_A \frac{\partial \varphi^*_A}{\partial N_A} + \lambda = 0, \quad \frac{\partial \Psi}{\partial N_B} = \varphi^*_B + N_B \frac{\partial \varphi^*_B}{\partial N_B} + \lambda = 0 \tag{66}$$

which yields

$$\varphi^*_A - \varphi^*_B = -N_A \frac{\partial \varphi^*_A}{\partial N_A} + N_B \frac{\partial \varphi^*_B}{\partial N_B} \tag{67}$$

Constrained open interaction (2)

- We recall the expression of φ^* from (29), which can be combined with (39) and (42) to give various different expressions as seen in the table below, from which the quantity $N\,\partial\varphi^*/\partial N$ is derived as follows:

Expression	$\dfrac{\partial\varphi^*}{\partial N}$	$-N\dfrac{\partial\varphi^*}{\partial N}$
φ^*	0	0
$\varphi^* = c + (β/2)\ln ε + \ln v = c' + (β/2)\ln ϑ + \ln v$ $= c' + (β/2)\ln ϑ + \ln(ϑ/p) = c' + (1 + β/2)\ln ϑ - \ln p$ [where $c' = c + (β/2)\ln(β/2)$]	0	0
$\varphi^* = c + (β/2)\ln (E/N) + \ln (V/N)$	$-(1 + β/2)/N$	$1 + β/2$
$\varphi^* = c + (β/2)\ln ε + \ln (V/N) = c' + (β/2)\ln ϑ + \ln (V/N)$	$-1/N$	1
$\varphi^* = c + (β/2)\ln (E/N) + \ln v = c + (β/2)\ln (E/N) + \ln(ϑ/p)$	$-(β/2)/N$	$β/2$

- In all cases contained in the table, the quantity $N\,\partial\varphi^*/\partial N$ proves to be constant, independent of N, so that the right hand side of (67) equals 0. Hence,

$$\varphi^*_A = \varphi^*_B = \varphi^* \qquad (68)$$

- This defines an isentropic state, which is reached when full uniformity is not possible.

Isentropic state and the law of adiabatic change

- It is easy to show that at the isentropic state the following equalities follow:

$$\vartheta_A^{\,\theta/2} V_A/N_A = \vartheta_B^{\,\theta/2} V_B/N_B \Leftrightarrow \vartheta_A^{\,\theta/2} v_A = \vartheta_B^{\,\theta/2} v_B \Leftrightarrow \vartheta_A^{\,1+\theta/2}/p_A = \vartheta_B^{\,1+\theta/2}/p_B \qquad (69)$$

- The latter equations are known as the *law of isentropic or (reversible) adiabatic change*.

- In particular, a reversible process for which the equality of intensive entropies in the two states (as implied by (68)) implies zero change in entropy ($d\Phi^* = 0$). As we will see below (equation (89)) this implies zero heat transfer ($\delta Q = 0$); a process with zero heat transfer is called an adiabatic process.

Q: Why the temperature in a vertical cross section across the troposphere varies substantially (decreases with increasing elevation), while the entropy per unit mass is fairly uniform (an **isentropic** state)?

A: Because this maximizes entropy, i.e., uncertainty. Note that, due to absorption of solar energy by the ground, the temperature close to the ground level is not the same as at higher elevation in the atmosphere. Also, due to the hydrostatic law, the pressure cannot be uniform in a vertical cross section and therefore uniformity cannot hold. Thus, both pressure and temperature vary according to (69).

Phase change: setup

- When the two systems that are brought into contact are in different phases, e.g. system A is gas (vapour) and system B is liquid (water), then there is a difference in the equation of energy conservation, which should include the *phase change energy*, i.e. the amount of energy per molecule ξ to break the bonds between molecules of the liquid phase in order for the molecule to move to the gaseous phase.

- Using subscripts A and B for the gaseous and liquid phase, respectively, the total entropy will be

$$\Phi^* = N_A \varphi^*_A + N_B \varphi^*_B \tag{70}$$

  with

$$\varphi^*_A = c_A + (\beta_A/2) \ln (E_A/N_A) + \ln (V/N_A), \quad \varphi^*_B = c_B + (\beta_B/2) \ln (E_B/N_B) \tag{71}$$

  where in the liquid phase we neglected the volume per particle, which is by several orders of magnitude smaller than that of the gaseous phase.

- The two systems (phases) are in open interaction and the constraints are:

$$E_A + E_B + N_A\xi = E \tag{72}$$
$$N_A + N_B = N \tag{73}$$

- We wish to find the conditions which maximize the entropy Φ^* in (70) under constraints (72) and (73) with unknowns E_A, E_B, N_A, N_B. We form the function Ψ incorporating the total entropy Φ^* as well as the two constraints with Langrage multipliers κ and λ:

$$\Psi = N_A\varphi^*_A + N_B \varphi^*_B + \kappa (E_A + E_B + N_A\xi - E) + \lambda (N_A + N_B - N) \tag{74}$$

Phase change: results

- To maximize Ψ, equating to 0 the derivatives with respect to E_A and E_B, we obtain

$$\frac{\partial \Psi}{\partial E_A} = \frac{N_A \, \beta_A}{2E_A} + \kappa = 0, \quad \frac{\partial \Psi}{\partial E_B} = \frac{N_B \, \beta_B}{2E_B} + \kappa = 0 \tag{75}$$

and since $E/N = \varepsilon$ and by virtue of (39), this obviously results in equal temperature ϑ, i.e.

$$\kappa = -1/\vartheta_A = -1/\vartheta_B = -1/\vartheta \tag{76}$$

- Equating to 0 the derivatives with respect to N_A and N_B, we obtain

$$\frac{\partial \Psi}{\partial N_A} = \varphi^*_A - \frac{\beta_A}{2} - 1 + \kappa \xi + \lambda = 0, \quad \frac{\partial \Psi}{\partial N_B} = \varphi^*_B - \frac{\beta_B}{2} + \lambda = 0 \tag{77}$$

and after eliminating λ, substituting κ from (76), and making algebraic manipulations,

$$\varphi^*_A - \varphi^*_B = \xi/\vartheta - (\beta_B/2 - \beta_A/2 - 1) \tag{78}$$

- On the other hand, from (71), but expressed in terms of ϑ and p (equation (43)), the entropy difference is

$$\varphi^*_A - \varphi^*_B = -(\beta_B/2 - \beta_A/2 - 1) \ln \vartheta - \ln p + \text{constant} \tag{79}$$

- Combining (78) and (79), and eliminating $\varphi^*_A - \varphi^*_B$, we find

$$p = \text{constant} \times e^{-\xi/\vartheta} \, \vartheta^{-(\beta_B/2 - \beta_A/2 - 1)} \tag{80}$$

Assuming that at some temperature ϑ_0, $p(\vartheta_0) = p_0$, we write (80) in a more convenient and dimensionally consistent manner as:

$$p = p_0 \, e^{\xi/\vartheta_0 - \xi/\vartheta} \, (\vartheta_0/\vartheta)^{(\beta_B/2 - \beta_A/2 - 1)} \tag{81}$$

For application of (81) to the change phase of water see p. 41.

The Clausius-Clapeyron equation

Equations (78) and (79) can be written in differential form as

$$d(\varphi^*_A - \varphi^*_B) = -\xi\, d\vartheta/\vartheta^2 = -(\varphi^*_A - \varphi^*_B)\, d\vartheta/\vartheta - (\beta_B/2 - \beta_A/2 - 1)\, d\vartheta/\vartheta \tag{82}$$

$$d(\varphi^*_A - \varphi^*_B) = -(\beta_B/2 - \beta_A/2 - 1)\, d\vartheta/\vartheta - dp/p \tag{83}$$

respectively. Equating the right-hand sides of the two, after algebraic manipulations we find

$$(\varphi^*_A - \varphi^*_B)\, d\vartheta/\vartheta = dp/p \tag{84}$$

or

$$dp/\, d\vartheta = (\varphi^*_A - \varphi^*_B)\, (p/\vartheta) \tag{85}$$

This is the well known Clausius-Clapeyron equation, a differential equation whose solution, obviously, is (80). Here we derived it, as a result of entropy maximization, just for the completeness of the presentation. In fact, the differential form is not necessary because in application only the closed solution is actually needed.

Warning: Classical and statistical thermodynamics books typically integrate the Clausius-Clapeyron equation using an incorrect assumption, that $\varphi^*_A - \varphi^*_B = \xi/\vartheta$, with constant ξ, which results in the incorrect, albeit quite common, solution:

$$p = \text{constant} \times e^{-\xi/\vartheta} \tag{86}$$

Q: What determines how much water is evaporated and condensed, thus providing the physical basis of the hydrological cycle?

A: A combination of entropy maximization (eqn. (70)) with energy availability (eqn. (72)).

Laws of thermodynamics: the zeroth law

- In classical (non-statistical) thermodynamics the zeroth law states that if two systems are in thermal equilibrium with a third system, then they are in thermal equilibrium with each other; this law defines the notion of thermal equilibrium. In turn, this is necessary to define, temperature as two systems that are in equilibrium have the same temperature.

- In statistical thermodynamics this law looks not necessary. Two systems are in equilibrium if they are put in contact and the entropy of the compound system has been maximized. Besides, the temperature is defined through (37).

- As we have seen already, (59) implies that two systems put in contact, in which entropy has been maximized, will have the same temperature. This is a consequence of entropy maximization and does not presuppose an axiomatic introduction of the zeroth law.

Laws of thermodynamics: the first law

- We have already used several times the principle of conservation of energy, according to which energy can be neither created nor destroyed but can only change forms.

- In classical thermodynamics, this principle is implemented as the first law of thermodynamics, which is typically stated as: the heat supplied to a system (δQ) equals the increase in internal energy of the system (dE) plus the work done by the system (δW).

- While the internal energy of a gas in a motionless container is kinetic energy, the expected value of the velocity of a molecule is zero and thus macroscopically it cannot produce work. In contrast, the forces exerted on the walls of the container by the colliding molecules are all of the same direction and give a resultant force (pressure times area) whose expected value is not zero. Macroscopically, this can produce work.

- This is demonstrated in the figure; if the piston whose area is A is moved by dx, then the work produced is $\delta W = p\, A\, dx = p\, dV$, or per particle $\delta w = p\, A\, dx\, /\, N = p\, dv$.

- If energy δQ, called heat, is supplied to the system then, because of energy conservation,

$$\delta Q = dE + \delta W = dE + p\,dV \qquad (87)$$

or per particle

$$\delta q = d\varepsilon + p\,dv \qquad (88)$$

These are the mathematical expressions of the first law.
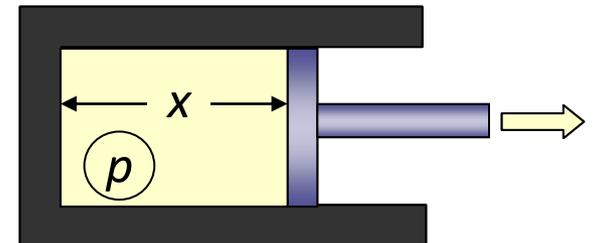
Laws of thermodynamics: the second law

- In classical thermodynamics, the second law states that in the process of reaching a thermodynamic equilibrium, the total entropy of a system increases, or at least does not decrease.

- In this respect, the second law is none other than the principle of maximum entropy applied to thermodynamic systems. Since, starting from any condition and approaching the equilibrium, the entropy change $\mathrm{d}\Phi^*$ of the total system can only be positive.

- In classical thermodynamics, the entropy change is defined as $\delta Q/T$ for a *reversible process*, where T is the absolute temperature (see below), which in our formalism should equivalently be written as $\delta Q/\vartheta$.

- Taking the piston example in the previous page, and comparing (87) with (48), we may see that, since in our example $\mathrm{d}N = 0$, their right-hand sides are equal. Thus

$$\delta Q = \vartheta \, \mathrm{d}\Phi^* \tag{89}$$

  which indicates the equivalence of the classical entropy with the extensive entropy of our framework. Note that (89) holds true only for reversible systems.

- However, we have seen (Closed interaction in p. 28, Open interaction in p. 29) that the entropy can increase ($\mathrm{d}\Phi^* > 0$) without energy gain or work production ($\delta Q = \delta W = 0$). Processes in which this happens are known as *irreversible* and are more generally characterized by

$$\delta Q < \vartheta \, \mathrm{d}\Phi^* \tag{90}$$

Classical (Boltzmann) formalism

- We have seen that the statistical framework provided here can produce the classical thermodynamic framework, yet it is more general as the principle of maximum entropy can be applied in other systems.

- In order for the statistical system to be able to reproduce classical thermodynamics also numerically, some numerical adaptations are necessary.

- In the classical formalism the temperature is (unnecessarily) regarded as an independent fundamental unit (kelvin, K) and the (unnecessary) constants $k = 0.138\,065$ yJ/K (yocto-joules per kelvin) and

$$R_* = k\,N_a = 8\,314.472 \text{ J K}^{-1} \text{ kmol}^{-1} \tag{91}$$

are used, where N_a (= 6.022×10^{23} mol^{-1}) is the Avogadro constant (the number of particles per mole of substance).

- Accordingly, the so-called absolute (or thermodynamic) temperature T and the classical entropy S are defined as follows, so that the relationship $1/T = \partial S/\partial E$ is retained:

$$T := \vartheta/k \text{ [units: K]}, \quad S := k\,\Phi^* \text{ [units: J/K]} \tag{92}$$

- The classical entropy per unit mass becomes:

$$s = k\,\varphi^* / m_0 = [(1 + \beta/2) \ln(T/T_0) - \ln(p/p_0)] (R_*/M_0) \tag{93}$$

where T_0 and p_0 designate an arbitrary macroscopic state to which we assign zero entropy (usually in atmospheric thermodynamics this is $T_0 = 200$ K and $p_0 = 1000$ hPa) and $M_0 := N_a\,m_0$ is the molecular mass.

Classical (Boltzmann) formalism (2)

- By setting

$$R_*/M_0 =: R, \quad (1 + β/2)\, R =: c_p, \quad (β/2)\, R =: c_v \tag{94}$$

(which are, respectively, the gas constant and the specific heat for constant pressure and constant volume, respectively), we obtain the final classical entropic formula:

$$s = c_p \ln(T/T_0) - R \ln(p/p_0) \tag{95}$$

- The law of ideal gases becomes

$$p\,V = n\,R_*\,T \Leftrightarrow p\,v = R\,T \tag{96}$$

where now $v := V/m = V/(n\,M_0)$ is the volume per unit mass (= 1/density), m is the mass of the gas in volume V and n is the number of moles.

- All of the above equations, after being adapted, substituting T for ϑ and S for Φ^* based on (92), are valid also in classical thermodynamics.

Q: Is the **specific heat** (heat capacity per unit mass) of a gas an experimental quantity or can it be derived theoretically?

A: The entropy maximization framework can derive heat capacities theoretically.

As an example we consider the atmospheric air mixture which is mostly composed by diatomic molecules (N_2, O_2) so that the degrees of freedom per molecule are $β = 5$. In this mixture, $M_0 = 28.96$ kg/kmol, so that $R = 8314.472/28.96 = 287.1$ J K^{-1} kg^{-1}. Thus, $c_p = (1 + 5/2) \times 287.1 = 1004.8$ J K^{-1} kg^{-1}, while the experimental value is $c_p = 1004$ J K^{-1} kg^{-1}!

The phase change using classical formalism

- Equation (78) can be easily transformed into the classical formalism using the above definitions, thus obtaining

$$s_A - s_B = \frac{\alpha}{T} - (c_L - c_p) \qquad (97)$$

  where $\alpha := \xi / m_0$.

- We choose as reference point the triple point of water, for which it is known with accuracy that $T_0 = 273.16$ K (= 0.01°C) and $p_0 = 6.11657$ hPa (Wagner and Pruss, 2002).

- The specific gas constant of water vapour is $R = 461.5$ J kg^{-1} K^{-1}.

- The specific heat of water vapour for constant pressure, again determined at the triple point, is $c_p = 1884.4$ J kg^{-1} K^{-1} and that of liquid water is $c_L = 4219.9$ J kg^{-1} K^{-1} (Wagner and Pruss, 2002), so that $c_L - c_p = 2335.5$ J kg^{-1} K^{-1} and $(c_L - c_p)/R = 5.06$.

- The latent heat, defined as $L = (s_A - s_B) T$, at T_0 is $L_0 = 2.501 \times 10^6$ J kg^{-1} so that $\alpha = L_0 + (c_L - c_p) T_0 = 3.139 \times 10^6$ J kg^{-1} and $\xi / kT_0 = \alpha / RT_0 = 24.9$. This results in the functional form

$$L \text{ [J kg}^{-1}] = \alpha - (c_L - c_p)T = 3.139 \times 10^6 - 2336 \, T \text{ [K]} \qquad (98)$$

The phase change using classical formalism (2)

- It can be verified (see figure) that equation (98) is very close to tabulated data from Smithsonian Meteorological Tables (List, 1951), as well as a commonly suggested empirical linear equation for latent heat (e.g. Shuttleworth, 1993):

$$L \text{ [J kg}^{-1}\text{]} = 3.146 \times 10^6 - 2361\, T \text{ [K]} \; (= 2.501 \times 10^6 - 2361\, T_C \text{ [}^{\circ}\text{C]}) \tag{99}$$

- It is important to know that the entropic framework which gives the saturation vapour pressure is the same framework that predicts the relationship of the latent heat of vaporization with temperature.
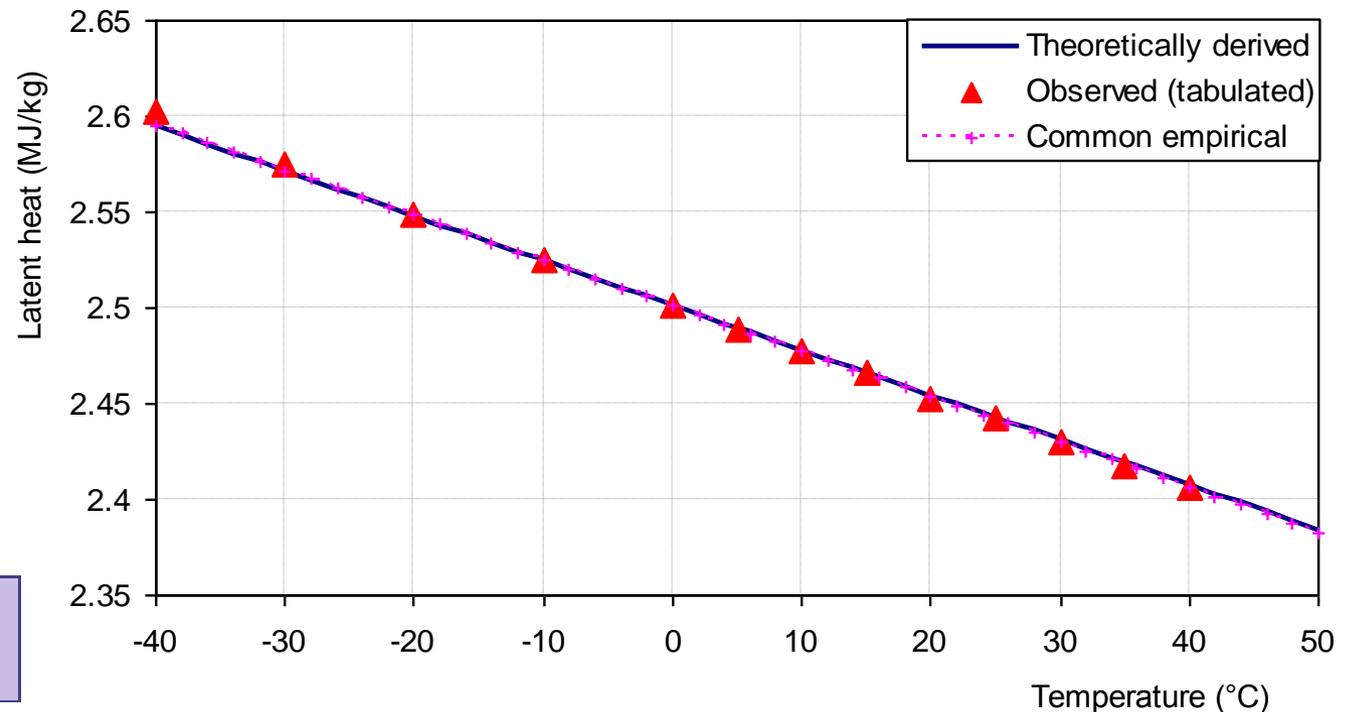
Source of figure: Koutsoyiannis (2012)

The phase change using classical formalism (3)

- Equation (81), if converted to the classical formalism, takes the form:

$$p = p_0 \exp\left[\frac{\alpha}{RT_0}\left(1 - \frac{T_0}{T}\right)\right]\left(\frac{T_0}{T}\right)^{(c_L - c_p)/R} \tag{100}$$

- With the above values of thermodynamic properties, (100) becomes

$$p = p_0 \exp\left[24.921\left(1 - \frac{T_0}{T}\right)\right]\left(\frac{T_0}{T}\right)^{5.06}, \text{ with } T_0 = 273.16 \text{ K}, p_0 = 6.11657 \text{ hPa.} \tag{101}$$

where we have slightly modified the last two decimal digits of the constant α / RT_0 to optimize its fit to the data (see below).

- For comparison, (86) (based on the inconsistent integration of the Clausius-Clapeyron equation for constant L) is

$$p = p_0 \exp\left[19.84\left(1 - \frac{T_0}{T}\right)\right], \text{ with } T_0 = 273.16 \text{ K}, p_0 = 6.11657 \text{ hPa} \tag{102}$$

- Empirical equations based on observations are in common use; among these we mention the so-called Magnus-type equations, two of which are:

$$p = 6.11 \, e^{17.27 \, T_C / (237.3 + T_C)} \quad [T_C \text{ in } ^{\circ}\text{C}, p \text{ in hPa}] \text{ (Tetens, 1930)} \tag{103}$$

$$p = 6.1094 \, e^{17.625 \, T_C / (243.04 + T_C)} \quad [T_C \text{ in } ^{\circ}\text{C}, p \text{ in hPa}] \text{ (Alduchov \& Eskridge, 1996)} \tag{104}$$

The latter is the most recent and is regarded to be the most accurate—but it is less accurate than (101) (see Koutsoyiannis, 2012).

Saturation vapour pressure: comparisons

- The figure on the right compares the two theoretical equations (101) and (102) also with the empirical (104). All seem indistinguishable.

- However, the figure below, which compares relative differences from measurements, clearly indicates the inappropriateness of
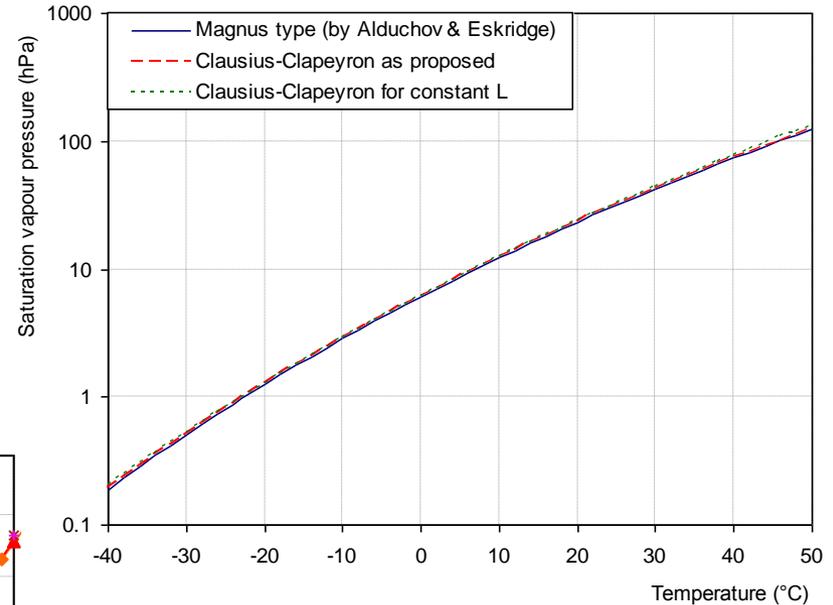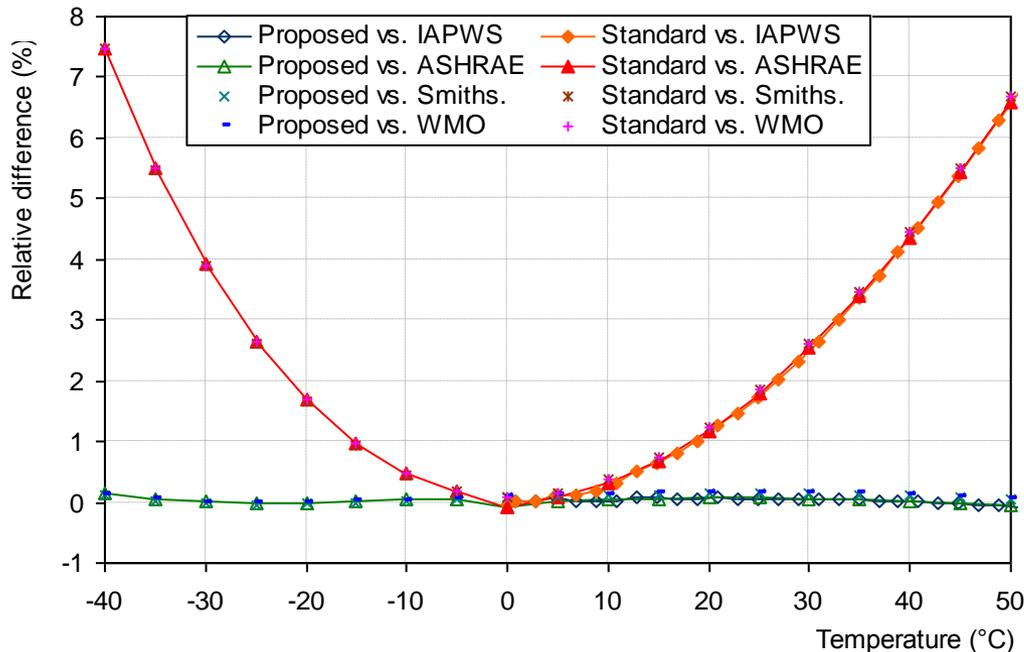
(102).

Source of figures: Koutsoyiannis (2012)

Calculation of dew point for known vapour pressure

- The simplicity of (101) makes numerical calculations easy. For known T, (101) provides p directly. The inverse problem (to calculate T, i.e. the saturation temperature, also known as dew point, for a given partial vapour pressure p) cannot be solved algebraically. However, the Newton-Raphson numerical method at an origin $T_0/T = 1$ gives a first approximation T' of temperature by

$$\frac{T_0}{T'} = 1 + \frac{1}{24.921 - 5.06} \ln\left(\frac{p_0}{p}\right) \tag{105}$$

Notably, this is virtually equivalent to solving (102) for T_0/T.

- This first approximation can be improved by re-applying (101) solved for the term T_0/T contained within the exponentiation, to give

$$\frac{T_0}{T} = 1 + \frac{1}{24.921} \ln\left(\frac{p_0}{p}\right) + \frac{5.06}{24.921} \ln\left(\frac{T_0}{T'}\right) \tag{106}$$

- A single application of (106) suffices to provide a value of T with a numerical error in T_0/T less than 0.1%, while a second iteration (setting the calculated T as T') reduces the error to 0.02%.

Statistical vs. classical thermodynamics

- In classical thermodynamics, first we define temperature T, as well as a basic unit for it (kelvins) and then the entropy S by $dS := \delta Q / T$, where Q denotes heat; the definition is perhaps affected by circularity, as it is valid for reversible processes, which are those in which $dS = \delta Q / T$, while irreversible are those in which $dS > \delta Q / T$.

- In statistical thermodynamics, the entropy Φ is just the uncertainty, as defined in probability theory, and is dimensionless; the temperature is then defined by $1/\vartheta := \partial\Phi/\partial E$, where E is the internal energy of the system; the natural unit for temperature ϑ is then identical with that of energy (joules).

- Classical thermodynamic equations can be derived from statistical thermodynamics by simple linear transformations ($S = k\,\Phi^{*}$ and $T = \vartheta/k$ where k is Boltzmann's constant).

- The essential difference is in interpretation:
  - Awareness of maximum uncertainty in statistical thermodynamics.
  - Delusion of deterministic laws in classical thermodynamics.

- The statistical thermodynamic framework is clearly superior to the classical framework for many reasons:
  - It is more parsimonious.
  - It provides explanations for the processes described by classical thermodynamics.
  - It can be generalized to other processes, involving higher levels of macroscopization.

Concluding remarks

- Heat is not a caloric fluid, as used to be thought of up to the 19th century.
- Thermodynamic laws e.g. the ideal gas law are statistical laws (typically relationships of expectations of random variables).
- While these laws are derived by maximizing entropy, i.e. uncertainty, they express near certainties and are commonly misinterpreted as deterministic laws.
- The explanation of near certainty relies on these two facts:
  - Typical thermodynamic systems are composed of hugely many identical elements: $N \sim 10^{24}$ per kilogram of mass.
  - The random motion of each of the system elements is practically independent to the others'.
- As a consequence, a random variable $\underline{x}$ expressing a macroscopic state, will have a variation $\text{std}[\underline{x}]/E[\underline{x}] \sim 1/\sqrt{N} \sim 10^{-12}$ (for a kilogram of mass).
- The fact that the macroscopic variability is practically zero should not mislead us to interpret the laws in deterministic terms.

References

- Alduchov, O. A., and R. E. Eskridge, Improved Magnus form approximation of saturation vapour pressure, *J. Appl. Meteor.*, 35, 601–609, 1996.

- Jaynes, E. T., Information theory and statistical mechanics, *Phys. Rev.*, 106(4), 620–630, 1957.

- Koutsoyiannis, D., Clausius-Clapeyron equation and saturation vapour pressure: simple theory reconciled with practice, *European Journal of Physics*, 33 (2), 295–305, 2012.

- Koutsoyiannis, D., Physics of uncertainty, the Gibbs paradox and indistinguishable particles, *Studies in History and Philosophy of Modern Physics*, 44, 480–489, 2013

- List, R. J., Smithsonian Meteorological Tables, Smithsonian Institution, Washington, DC, 527 pp., 1951.

- Papoulis, A., *Probability and Statistics*, Prentice-Hall, 1990.

- Papoulis, A., *Probability, Random Variables, and Stochastic Processes*, 3rd ed., McGraw-Hill, New York, 1991.

- Shuttleworth, W. J., Evaporation, ch.4 in *Handbook of Hydrology* (D. R. Maidment ed.), McGraw-Hill, New York, pp. 4.1-4.53, 1993.

- Tetens, O., Über einige meteorologische Begriffe, *Z. Geophys.*, 6, 297–309, 1930.

- Wagner, W., and A. Pruss, The IAPWS formulation 1995 for the thermodynamic properties of ordinary water substance for general and scientific use, *J. Phys. Chem. Data*, 31, 387-535, 2002.

- Wannier, G. H., *Statistical Physics*, Dover, New York, 1987.