TENSOR, N. S.
VOL. 39 (1982)

INFORMATION GEOMETRY OF CLASSICAL THERMODYNAMICAL SYSTEMS.

Dedicated to Professor Akitsugu Kawaguchi
on the occasion of his 80th birthday.

By Roman S. INGARDEN, Michiaki KAWAGUCHI and Yoshiharu SATO.

A systematic introduction of the local Riemannian geometrical structure into the parameter space of classical thermodynamical systems (statistically defined) is given. The Riemannian metric has the physical meaning of the entropy gain of Rényi-Kullback, and therefore is closely connected with the Glansdorff-Prigogine principle of irreversible thermodynamics as interpreted, e.g., by Schlögl. Two simplest classical examples are discussed in detail (the ideal gas and generalized oscillator), both giving the Euclidean metric of the space of thermodynamical parameters. For the ideal gas the normal coordinates (the Euclidean representation) are introduced and discussed.

§1. Introduction. The present paper has two independent origins: attempts of geometrical formulation of information thermodynamics [4][1], and developments of geometrical methods in statistics, especially, in multivariative analysis [20]. It was a happy circumstance that, because of the invitation of one of the authors (R.S.I.) to the Hokkaido University, these two lines of the research crossed. In 1977 the first results of the common investigations have been presented (by R.S.I.) in 3 lectures in different universities in Japan (in the Hokkaido University on Sept. 8, in the Sagami Institute of Technology on Sept. 16, and in the Tsukuba University on Sept. 24), cf. [5], and at the 10th Symposium on Math. Phys. in Torun on Dec. 5. The present publication gives the results in a more elaborated and developed form, together with some new results. The main scope of the investigations is the local Riemannian structure in the parameter space of thermodynamics, where the latter is constructed as a statistical theory by means of the method of information thermodynamics ([10], [9], [8]). The metric is derived from the Rényi "information gain" or "conditional expectation". Here only general theory and simplest classical examples are discussed. In the forthcoming publication [6] a discussion of geometry of thermodynamics in the quantum case, while in [7] an investigation of another classical examples will be given.

§2. Information gain. The concept of information gain has a rather long history and here we are not able to present it in detail (cf. [14], [25], [3], [1]), where further references can also be found. The term "gain of information" has been proposed by Rényi [19] who gave also the most penetrating analysis of the meaning of this concept. The other authors used the terms "relative entropy", "conditional entropy", "Kullback (or Kullback-Leibler) entropy" (cf. [15]), "information between operators" (in the quantum case, Umegaki [24]), "directed divergence" [14]. The

---

Received May 11, 1982.

[1] Numbers in brackets refer to the references at the end of the paper.

term "divergence" (without the adjective "directed") was applied by Kullback [**14**] to the symmetrized gain of information, cf. below, introduced by Jeffreys ([**11**], [**12**]). In physics the concept of information gain was extensively used, especially, by F. Schlögl ([**21**], [**22**]) who applied it for the foundation of thermodynamics, classical and quantum (cf. also [**18**] where further references are given).

It seems that the simplest method of introduction the information gain is to begin with a more general concept, that of the *Kerridge inaccuracy* [**13**]

$$(1) \qquad S(q\|p)=-\sum_{i=1}^{\nu} p_i \ln q_i \geqslant 0 \qquad (\nu=1, 2, \cdots <\infty \text{ or } \nu=+\infty)$$

in the discrete classical case (e.g., the Ising model of a spin system [**7**]), where

$$p_i,\ q_i \geqslant 0 \qquad (i=1, 2, \cdots, \nu), \qquad \sum_{i=1}^{\nu} p_i=\sum_{i=1}^{\nu} q_i=1\,.$$

In the continuous classical case of a system of r identical particles ($r=1, 2, \cdots<\infty$)

$$(2) \qquad S(g\|f)=-\int_{\Omega} f \ln g d\mu=-\frac{1}{r!h^{3r}}\int_{\Omega} f(x, y) \ln g(x, y) d^{3r}x d^{3r}y\,,$$

where h is the Planck constant (introduced here only because of the dimensional reasons, it comes from the semiclassical theory), Ω is the phase space of the system (a 6-dimensional simplectic manifold), and $x=(x^1, \cdots, x^{3r})$, coordinates, $y=(y_1, \cdots, y_{3r})$, momenta, are canonical symplectic coordinates, while

$$(3) \qquad \forall (x, y) \in \Omega\,, \qquad f(x, y),\ g(x, y) \geqslant 0\,, \qquad \int_{\Omega} f d\mu=\int_{\Omega} g d\mu=1\,.$$

Now assuming all the conditions of existence and regularity (e.g., in the continuous case that functions f and g are mutually and with respect to μ absolutely continuous (cf. [**14**], p. 3), we may connect the Kerridge inaccuracy with other "inaccuraties" ("uncertainties", "informations" or "entropies"; for the sake of simplicity we shall use here the same symbol S for all these quantities, distinguishing them only by the arguments which cannot be dropped if any ambiguity is possible). Namely, in discrete case we may write

$$(4) \qquad S(q\|p)=S(p)+S(q\,|\,p)\,,$$

where $S(p)=-\sum_{i=1}^{\nu} p_i \ln p_i \geqslant 0$ is the *Shannon information* (*entropy*) and

$$(5) \qquad S(q\,|\,p)=\sum_{i=1}^{\nu} p_i(\ln p_i-\ln q_i)=\sum_{i=1}^{\nu} p_i \ln \frac{p_i}{q_i} \geqslant 0$$

is the *Rényi gain of information.* In the classical continuous case we obtain respectively

$$(6) \qquad S(g\|f)=S(f)+S(g\,|\,f)\,,$$

where

$$(7) \qquad S(f)=-\int_{\Omega} f \ln f d\mu \quad \text{is the \textit{Boltzman-Gibbs entropy}},$$

$$(8) \qquad S(g\,|\,f)=\int_{\Omega} f(\ln f-\ln g) d\mu \quad \text{is the \textit{Kullback entropy}}.$$

To get a physical interpretation let us now imagine that we like to fix experimentally a definite statistical state p or f, respectively (we shall call it simply *state* and denote by one symbol a, say). Measuring a, we actually obtain, in view of unavoidable errors of measurement and instrumentation, some other near state b (denoted above by q or g). Looking at formulae (4), (6) we may say that $S(a)$ is the *absolute uncertainty* of state a, while $S(b\,|\,a)$ is the *relative uncertainty* of a with respect to b. The Kerridge entropy $S(b\|a)$ is then the *total uncertainty* of a with respect to b, taking also into account the "inner" or "own" uncertainty, its absolute entropy, in the situation where a is actually represented by b. Therefore, $S(b\,|\,a)$ can be considered as an "error" connected with position of a by measuring entropy, or an "information distance" of a and b.

The discrete classical case is, of course, a special case of the general classical case, and in mathematics there is no need of considering it separately. (cf. **[14]** or [16], [17], in the latter papers the entropy is called the *Boltzman-Gibbs-Shannon entropy*). The discrete case may be also obtained from arbitrary continuous case by introducing "quantum cells" $A_i \subset \Omega$, $i=1, \cdots, \nu$ such that

$$A_i \cap A_j = \varnothing \qquad (i \neq j), \qquad \bigcup_{i=1}^{\nu} A_i = \Omega, \qquad \int_{A_i} d\mu = 1$$

and putting $p_i = \int_{A_i} f\,d\mu$ $(i=1, \cdots, \nu)$.

§3. Symmetrization and local metric properties. The Rényi-Kullback entropy $S(b\,|\,a)$ is not, in general, a distance function in the sense of the Fréchet *metric*. Indeed, of the three Fréchet axioms

$$d(b, a) \geqslant 0 \quad \text{and} \quad d(b, a)=0 \quad \text{iff} \quad b=a\,, \tag{9}$$

$$d(b, a) = d(a, b)\,, \tag{10}$$

$$d(c, a) \leqslant d(c, b) + d(b, a)\,, \tag{11}$$

only the first one is always satisfied (cf. **[14]** and **[25]**). Instead of (11) we have the so-called *joint convexity*, i.e. (cf. **[25]**),

$$S(b\,|\,a) \leqslant \lambda S(b_1\,|\,a_1) + (1-\lambda) S(b_2\,|\,a_2)\,,$$

where $b=\lambda b_1 + (1-\lambda) b_2$, $a=\lambda a_1 + (1-\lambda) a_2$, $0 \leqslant \lambda \leqslant 1$. To realize the second axiom (10) Jeffreys (**[11]**, **[12]**) proposed the symmetrization of $S(b\,|\,a)$, as is mentioned above,

$$I(a, b) = \tfrac{1}{2}[S(b\,|\,a) + S(a\,|\,b)]\,. \tag{12}$$

This definition, however, does not guarantee the axiom (11). On the other hand, when we are interested in local properties only, in the spirit of differential geometry (as Riemannian geometry, etc.) we obtain (10) and (11) without (12) in exactly the same form for $S(b\,|\,a)$ as for $I(b, a)$. (The difference by factor 1/2 in **[14]** is caused by omitting this factor in (12).)

In order to speak about differential manifolds with local metric defined by $S(b\,|\,a)$, we have to assume that $S(b\,|\,a)$ is of the class C^∞ (or C^ω, or at least C^3) with respect to the local coordinates.

Let us now assume that we have a differentiable manifold of thermodynamically regular states **[9]**. The manifold, called a *thermodynamical space* and denoted by $M=M^K$, may be infinite- or finite-dimensional ($K=1, 2, \cdots$ or $K=\infty$), and in the limit may also contain all the thermodynamically regular states, i.e., in the classical case we take the set $\Delta(\Omega)$ of all probability measures on Ω with finite entropy, or a differentiable submanifold of $\Delta(\Omega)$. But, as a rule, in thermodynamics we usually consider thermodynamical spaces of a finite and rather small number of dimensions defined by independent thermodyamical parameters such as temperature, volume, mass, number of particles etc. Some simple concrete examples of M will be discussed below. Introducing in M a local or global coordinate system $u=(u^1, \cdots, u^K)$ (we assume below $K<\infty$) we have for any $m\in M$ such a neighbourhood $V\subseteq M$ of m that ($u\in U\subseteq \boldsymbol{R}^K$)

$$f=f(u)=f(u^1, \cdots, u^K), \qquad f: U\rightarrow V\subseteq \Delta(\Omega),$$

where f is a function of C^∞ class.

Denoting for simplicity by T the integration over all the phase space M we may write (3) and (8) (denoting $f\rightarrow f(u)$, $g\rightarrow f(v)$) as

$$T[f(w)]=T[f(v)]=1, \tag{13}$$

$$S(v, u)=S[f(v)\,|\,f(u)]=T[f(u)(\ln f(u)-\ln f(v))].$$

From (13) we directly obtain by differentiation with respect to u the identities (it is sufficient to write only those for u)

$$T(\partial f/\partial u^\alpha)=0, \qquad T(\partial^2 f/\partial u^\alpha \partial u^\beta)=0 \qquad (\alpha, \beta=1, \cdots, K). \tag{14}$$

Using (14) we obtain finally for

$$\left.\frac{\partial S}{\partial u^\alpha}\right|_{u=v}=0, \qquad \left.\frac{\partial^2 S}{\partial u^\alpha \partial u^\beta}\right|_{u=v}=T\left[f\cdot\left(\frac{\partial \ln f}{\partial u^\alpha}\cdot\frac{\partial \ln f}{\partial u^\beta}\right)\right]=\left\langle\frac{\partial \ln f}{\partial u^\alpha}\,\frac{\partial \ln f}{\partial u^\beta}\right\rangle=2g_{\alpha\beta}=2g_{\beta\alpha}. \tag{15}$$

Therefore, we can write down to be second order of accuracy in $du=(du^1, \cdots, du^K)$, a "small" vector in the tangent space $T_m M$,

$$S(u+du, u)=\sum_{\alpha,\beta=1}^{K} g_{\alpha\beta}(u)du^\alpha du^\beta+\cdots\geqslant 0. \tag{16}$$

Because of the positive definiteness property (8) of $S(u, v)$ we obtain the very important result that the symmetric matrix $g=(g_{\alpha\beta})$ is positive definite, so in particular det $(g)>0$, or rank $(g)=K$. Since by diffeomorphisms of coordinates g behaves as a tensor, as is directly seen, we have all the conditions required for the Riemannian metric. Thus we may write locally $ds^2=g_{\alpha\beta}du^\alpha du^\beta\geqslant 0$ (since now we adapt the Einsteinian summation convention for $\alpha, \beta=1, \cdots, K$) where ds is by definition an *information distance* between the states parametrized by u and $u+du$ in the K-dimensional state manifold M^K. In such a way we transform M locally into a Riemannian manifold. Of course, for ds, "in the small", locally the axioms (10) and (11) are fulfilled since by a suitable coordinate transformation we obtain locally just a Euclidean metric. For larger du the higher order terms in du have to be considered in (16), and globally M is not Riemannian.

Since the first derivatives of $S(u, v)$ disappear identically in any point of the parameter space (cf. (15)), and the next leading term of the Taylor expansion is symmetric in du, the symmetrization (16) cannot change anything locally, and here we are not interested in the non-local distance. That (16) gives the same first and second derivatives can be also checked directly by calculation.

§ 4. **The classical ideal gas of identical particles.** The first example which we shall discuss is, naturally, the classical ideal gas of r identical particles of mass m in volume V. Applying the method of information thermodynamics, as mentioned above, we take the total energy in V as

$$H=H(x, y)=K(y)+P(x)=K(y)=\frac{y^2}{2m}=\sum_{i=1}^{3r}\frac{y_i^2}{2m}. \tag{17}$$

Then the maximum entropy solution for f is in the form of (7)

$$f = Z^{-1}(T, V) \exp\left(-\frac{y^2}{2mT}\right), \qquad Z(T, V) = \frac{v^r}{h^{3r} r!}(2\pi m T)^{3r/2}, \tag{18}$$

T being the absolute temperature connected with the mean energy $U = \langle H \rangle$ by

$$U = T^2 \frac{\partial \ln Z(T, V)}{\partial T} = \frac{3r}{2} T \quad \text{or} \quad T = \frac{2}{3r} U \tag{19}$$

(T is here expressed in the energy units, so the Boltzmann constant $k=1$). Calculating the Boltzmann-Gibbs entropy (7) for the equilibrium state (18) we obtain

$$S(f) = r \ln\,[h^{-3}(2\pi m T)^{3/2} V] - \ln\,(r!) + 3r/2\,. \tag{20}$$

Assuming, as usually, that $r \gg 1$ we use the approximation

$$\ln\,(r!) \approx r \ln r \qquad (r \gg 1) \tag{21}$$

following from the Stirling formula. Then

$$S(f) = r \ln\,[h^{-3}(2e\pi m T)^{3/2} V \cdot r^{-1}] = r \ln\,[h^{-3}(\tfrac{4}{3} e m \pi U/r)^{3/2} V/r]\,. \tag{22}$$

Formula (22) gives the correct macroscopic behaviour of entropy of perfect gases (for sufficient high temperatures, of course. since for $T \to 0$, $S(f) \to -\infty$ against the Nernst-Planck principle) cf. e.g. **[23]** p. 53, **[2]**, which can be also checked immediately. Indeed, (22) gives (when S is expressed by U, V and r) the so-called *fundamental equation* **[2]** for an ideal gas in the entropy representation,

$$U = \{3h^2 r^{5/3}/4\pi e m V^{2/3}\} \exp\,(2S/3r) \tag{23}$$

and hence by differentiation the three so-called *equations of state* **[2]** for temperature

$$T = \partial U/\partial S = (2/3r)U\,, \tag{24}$$

for pressure

$$P = \partial U/\partial V = 2U/3V = rT/V\,, \tag{25}$$

and for chemical potential

$$\mu = \frac{\partial U}{\partial r} = \frac{5U}{3r} - \frac{2SU}{3r^2} = \frac{5}{2} T - \frac{ST}{r} = \frac{U + pV - ST}{r}\,.$$

Eq. (24) is identical with (19); (25) gives the well-known equation of state of the ideal gas expressing the experimental laws of Boyle, Charles and Avogadro; and (26) is equivalent to the Euler equation corresponding to the first order positive homogeneity of the fundamental equation (23) with respect to S, U, V and r (expressing the extensive character of entropy, internal energy and volume in thermodynamical equilibrium, cf. **[2]**, p. 48), $S = U/T + pV/T - \mu r/T$.

As is well-known, from (23) or (26) it follows that the equations for T, p and μ are zero order positive homogeneous and these three quantities are intensive ones (actually, only asymptotically for large r, cf. (20)). On the other hand, the homonogeneity properties allow the elimination of r and μ from the macroscopic theory and consideration of r as a scaling factor, cf. **[2]**. Therefore, only T and V (or any two quantities from the five ones: S, U, V, T, p) are considered as independent *thermodynamical degrees of freedom*. Strictly speaking, for small values r cannot be a coordinate of a differentiable manifold, being not a continuous variable, and therefore one cannot differentiate with respect to it. But for very large r, $r \gg 1$, the difference between the real continuous and the positive integer parameter is not physically perceptible, especially, when we go over to the macroscopically measurable *mole numbers* [2]

$$N = r/R\,, \qquad R = \text{Avogadro's number} = 6.0225 \times 10^{23}\,. \tag{27}$$

In our units ($k=1$) R is also the *gas constant*, as is seen from the equation of state (25) $pV = RNT$. In our geometrical theory we consider N (or r) as a fixed parameter, while T and V as (continuous) positive real variables

$$0 < T < \infty\,, \qquad 0 < V < \infty\,. \tag{28}$$

In order to eliminate N we introduce, as usually in thermodynamics [2], *mole densities*

$$s = S/N\,, \qquad u = U/N\,, \qquad v = V/N\,, \tag{29}$$

while T and p, as intensive quantities, remain the same. We remark that, s, u, v are no more extensive, and are formally intensive. For convenience, however, in our calculations we shall use N as a fixed macroscopical parameter, not necessarily large, but definitely not very small, i.e., $RN = r \gg 1$ or $N \gg 1/R \approx 10^{-24}$. Nevertheless, to be precise, the obtained formulae are valid only in the *thermodynamical limit*, i.e., for

$$\begin{aligned} &\lim U = \infty\,, \qquad \lim V = \infty\,, \qquad \lim N = \infty\,, \\ &\text{but} \qquad \lim U/N = u = \text{const.} < \infty\,, \\ &\text{and} \qquad \lim V/N = v = \text{const.} < \infty\,, \end{aligned}$$

then by (22) also

$$\lim S/N = R \ln\left[(V/Rh^3)(4\pi emU/3R)^{3/2}\right] = s = \text{const.} < \infty\,.$$

Only at the end of our calculation we shall use (29) and express our results in densities.

Calculating g according to the formula (15) for f given by (18) and $u^1 = T$, $u^2 = V$, $r = RN$, we obtain finally

$$g = \begin{pmatrix} 3RN2^{-1}T^{-2}\,, & 0 \\ 0 & ,\; R^2N^2V^{-2} \end{pmatrix} = \begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{pmatrix},$$

and, in consequence,

$$g^{-1} = \begin{pmatrix} 2T^2 3^{-1}(RN)^{-1}\,, & 0 \\ 0 & ,\; V^2(RN)^{-2} \end{pmatrix} = \begin{pmatrix} g^{11} & g^{12} \\ g^{21} & g^{22} \end{pmatrix}.$$

Thus

$$\det(g) = \tfrac{3}{2}R^3N^3/T^2V^2 > 0\,. \tag{30}$$

Now we obtain for Christoffel symbols as coefficients of a Riemannian connection

$$\begin{aligned} &\Gamma^1_{12} = \Gamma^1_{21} = 0\,, \qquad \Gamma^1_{11} = -1/T\,, \qquad \Gamma^2_{11} = 0\,, \\ &\Gamma^2_{22} = -1/V\,, \qquad \Gamma^1_{22} = 0\,, \qquad \Gamma^2_{21} = \Gamma^2_{12} = 0\,. \end{aligned}$$

This gives for all the components of the curvature tensor $R_{\alpha\beta\gamma\delta} = 0$, $\forall \alpha, \beta, \gamma, \delta = 1, 2$. Thus, the Riemannian and Gauss curvature disappear $R = K = 0$. In such a way we obtain geometry of a space developable into a Euclidean space, in this case a Euclidean plane. Introducing the "*normal* coordinates" (x_1, x_2)

$$T=\exp[(2/3RN)^{1/2}x_1], \qquad V=\exp[x_2/RN],$$
$$x_1=(\tfrac{3}{2}RN)^{1/2}\ln T, \qquad x_2=RN\ln V,$$

we have from (24) and (30)

$$-\infty<x_1<+\infty, \qquad -\infty<x_2<+\infty,$$
$$\forall x_1,x_2\in\boldsymbol{R}^2, \qquad \forall\alpha,\ \beta=1,2, \qquad g_{\alpha\beta}(x_1,x_2)=\delta_{\alpha\beta}=\text{Kronecker's delta}. \tag{31}$$

If we use the mole densities (29), we obtain

$$g=\begin{pmatrix}3RN/2T^2, & 0\\ 0 & ,R^2/V^2\end{pmatrix}, \qquad \det(g)=\tfrac{3}{2}R^3N/T^2V^2>0.$$

(T is an intensive parameter, so it does not change), and the normal coordinates $(\xi_T, \xi_v)\in\boldsymbol{R}^2$ (we call them the *Euclidean coordinates* of T and V or T and v in the *Euclidean representation*)

$$T=\exp[(2/3RN)^{1/2}\xi_T], \qquad V=\exp(\xi_V/R), \tag{32}$$

$$\xi_T=(\tfrac{3}{2}RN)^{1/2}\ln T=x_1, \qquad \xi_V=R\ln V=x_2/N-R\ln N, \tag{33}$$
$$-\infty<\xi_T<+\infty, \qquad -\infty<\xi_V<+\infty,$$

in which the metric is Euclidean as in (31)

$$\forall(\xi_T,\xi_V)\in\boldsymbol{R}^2, \qquad \forall\alpha,\ \beta=1,2, \qquad g_{\alpha\beta}(\xi_T,\xi_V)=\delta_{\alpha\beta}. \tag{34}$$

The most general coordinate or point transformation preserving (34) is an inhomonogeous orthogonal mapping (Euclidean motion in the plane) u^1, u^2 being some thermodynamical parameters,

$$\begin{aligned}\xi_{u^1}&=a_{11}\xi_T+a_{12}\xi_v+a_1, \qquad a_{11}^2+a_{12}^2=1, \qquad a_{11},a_{12},a_1\in\boldsymbol{R}^1,\\ \xi_{u^2}&=a_{21}\xi_T+a_{22}\xi_v+a_2, \qquad a_{21}^2+a_{22}^2=1, \qquad a_{21},a_{22},a_2\in\boldsymbol{R}^1;\end{aligned} \tag{35}$$

$$a_{11}a_{21}+a_{12}a_{22}=0. \tag{36}$$

Hence we may easily calculate the following coefficients for the respective Euclidean coordinates for entropy density s, internal energy density u and pressure P, cf. (22), (19), (25), (29), (27):

$$\begin{aligned}\xi_s&=c_s[(3R/2N)^{1/2}\xi_T+\xi_V]+c_sR\ln[(3e\pi m)^{3/2}/Rh^3]=c_ss,\\ \xi_u&=c_u(2R/3N)^{1/2}\xi_T+c_uR\ln\tfrac{3}{2}RN=c_uR\ln u,\\ \xi_P&=c_P[(2R/3N)^{1/2}\xi_T-\xi_V]+c_PR\ln R=c_PR\ln P.\end{aligned} \tag{37}$$

We obtain from (35)

$$c_s=(2N/(3R+2N))^{1/2}, \qquad c_u=(3N/2R)^{1/2}, \qquad c_P=(3N/(2R+3N))^{1/2},$$

therefore,

$$\text{a)}\quad \xi_s=\left(\frac{3R}{3R+2N}\right)^{1/2}\xi_T+\left(\frac{2N}{3R+2N}\right)^{1/2}\xi_V+\left(\frac{2N}{3R+2N}\right)^{1/2}\ln\frac{(2e\pi m)^{3/2}}{Rh^3},$$

$$\text{(38)}\quad \text{b)}\quad \xi_u=\xi_T+(\tfrac{3}{2}NR)^{1/2}\ln(\tfrac{3}{2}RN),$$

$$\text{c)}\quad \xi_P=\left(\frac{2R}{2R+3N}\right)^{1/2}\xi_T-\left(\frac{3N}{2R+3N}\right)^{1/2}\xi_V+\left(\frac{3N}{2R+3N}\right)^{1/2}R\ln R.$$

We see that in this representation all the essential physical properties of the gas, entropy, energy, temperature, volume and pressure are linearly connected, and the theory becomes rather simple and nice. We have not only rotations, but also translations (shifts) of coordinates. Of course, if we like to take any two of quantities (38), as a new coordinate system, we have yet to check (36). E.g., we see that if we take the pair (ξ_P, ξ_s), the condition (36) gives with (38)b) and (38)c)

$$\sqrt{6}\,(RN-1)/(3RN+2)(3RN+3)^{1/2}=0\,,\qquad \text{so}\quad RN=r=1\,,$$

against our assumption (21) that $r\gg 1$. Since ξ_u is only translated but non rotated with respect to ξ_T, we have also an orthogonal system (ξ_u, ξ_v), while the systems

$$(\xi_u, \xi_s)\,,(\xi_v, \xi_s)\,,\quad (\xi_T, \xi_s)\,,\quad (\xi_p, \xi_u)\,,\quad (\xi_p, \xi_v)\,,\quad (\xi_p, \xi_T)$$

are not orthogonal (so the metric is not of the form (34) in these coordinate systems), and the system (ξ_T, ξ_u) is impossible since ξ_u is parallel to ξ_T and the Jacobian of the transformation is 0. In such a way we obtained a nice geometrical classification of the different thermodynamical coordinate systems in which the systems (ξ_T, ξ_v) and (ξ_u, ξ_v) appeared to be distinguished as orthogonal, and therefore especially convenient. Of course, there is an infinite number of other orthogonal systems.

Parameters (33), (37) are the Euclidean representations of the 5 basic thermodynamical properties of the ideal gas. We also say that these properties are then represented in the *logarithmical scale*, except for the entropy which contains the logarithm in the original definition (7) (and this was just the reason why in the Euclidean representation all the parameters are linearly connected and can be treated on the same footing and with the same unit). In our above formulae the unit of all the parameters in the Euclidean representation is that of R (27), i.e., a pure number. For completeness we have to add, however, that in (32) we tacitly used two dimensional constants of numerical value 1. Namely, in an explicitly unit-covariant way we actually should write instead of (32)

$$T=T_0\exp\left[(2/3RN)^{1/2}\xi_T/R\right],\qquad v=v_0\exp(\xi_v/R)\,,$$

and therefore instead of (38)a) (formulae for ξ_s)

$$\xi_S=\left(\frac{3R}{3R+2N}\right)^{1/2}\xi_T+\left(\frac{2N}{3R+2N}\right)^{1/2}\xi_v+\left(\frac{2N}{3R+2N}\right)^{1/2}R\ln\left[\frac{(2e\pi mT_0)^{3/2}v_0}{Rh^3}\right],$$

where we have in SI units ($[A]$ denotes the physical dimension or unit of a quantity A)

$$[T_0]=1\,\mathrm{kg\,m^2\,s^{-2}}\,,\qquad [v_0]=1\,\mathrm{m^3}\,,\qquad [m]=1\,\mathrm{kg}\,,\qquad [h]=1\,\mathrm{kg\,m^2\,s^{-1}}\,.$$

Needless to say that the expression of all these physical properties of the gas by the dimensionless Euclidean parameters enables their geometrical treatment on the same footing, as well as introduction of other possible coordinates in the Euclidean representation for all the thermodynamical (equilibrium) states of the gas.

§ 5. A classical model of the solid state. We have above discussed the example of the ideal gas with considerable detail in view of the importance of this case for thermodynamics, in spite of its idealisation. Now we present rather shortly the results of similar calculation for some other simple examples which allow a rigorous treatment. All these cases will be also "ideal", i.e., without mutual interactions of particles, but we assume more general forms of Hamiltonian than (17), as well as more general types of equilibrium. First of all we shall consider particles in an external field with the potential quadratic (and then with other powers) in x. The most important ones are those which are easy to integrate. In other words, they will be the cases of harmonic and then generalized oscillators, or of a set of non-interacting oscillators. The latter set may be considered as a (classical) model of the solid state (the oscillators correspond to the normal vibrations of an elastic body). For simplicity, however, we shall put here $r=1$ (one particle).

In general, if we have a particle with an electric charge e in an external electromagnetic field with an electric (scalar) potential $\Phi(x)$, $x=(x^1, x^2, x^3)$, and a magnetic (vector) potential $A(x)=(A^1(x), A^2(x), A^3(x))$, we obtain in place of (17) the Hamiltonian

$$H(x, y)=(1/2m)[p-(e/c)A(x)]^2+e\Phi(x)=(1/2m)y^2+e\Phi(x)\,, \tag{39}$$

where $c=$ the velocity of light, a constant connected with chosen units, $p=(p^1, p^2, p^3)$ is the generalized (canonical) momentum, and

$$y=p-(e/c)A(x)=mv \tag{40}$$

is the kinetical momentum ($v=(v^1, v^2, v^3)$ is the velocity of the particle).

From (39) we see that H depends only on the *square* of the velocity $v=y/m$ which according to the Lorentz equation is independent of the magnetic field. Thus we obtain the same results with and without magnetic field if we consider the kinetic phase space (x, y) and not the canonical phase space (x, p). Since the transformation between them has the Jacobian equal to 1 (cf. (40)), the partition function Z (which fixes thermodynamics) and the metric g (which fixes geometry) can be calculated over (x, y) as below. We remark that only y has a direct physical meaning, while p is fixed only up to an arbitrary gauge of $A(x)$.

$$\begin{aligned} f(x, y)&=Z(T, V)^{-1}\exp(-H(x, y)/T)=Z^{-1}(T, V)\ \exp(-y^2/2mT-e\Phi(x)/T)\,,\\ Z(T, V)&=\frac{1}{h^3}\int_\Omega \exp\left(-\frac{H(x, y)}{T}\right)d^3x\,d^3y=h^{-3}\int_V d^3x\exp\left(-\frac{e\Phi(x)}{T}\right)\left[\int_{-\infty}^{+\infty}dy\exp\left(-\frac{y^2}{2Tm}\right)\right]^3\\ &=\frac{(2\pi mT)^{3/2}}{h^3}\int_V \exp\left(-\frac{e\Phi(x)}{T}\right)d^3x\,. \end{aligned} \tag{41}$$

Now we obtain from (15)

$$\begin{gathered} g_{11}=g_{TT}=\left\langle\left(\frac{\partial \ln f}{\partial T}\right)^2\right\rangle=\frac{1}{T^4}\left\langle\left[-T^2\frac{\partial \ln Z}{\partial T}+\frac{y^2}{2m}+e\Phi(x)\right]^2\right\rangle\,,\\ g_{12}=g_{21}=0\,,\qquad g_{22}=g_{VV}=\langle[-\partial \ln Z/\partial V]^2\rangle\,. \end{gathered} \tag{42}$$

Confining ourselves now only to $\Phi(x)$, we first consider the most important integrable case

$$\Phi(x)=m\omega^2x^2/2e\,,\qquad H(x, y)=y^2/2m+m\omega^2x^2/2\,, \tag{43}$$

where we replaced the coupling constant, electric charge, e by the circular frequency ω of the resulting (3-dimensional isotropic) *harmonic oscillator* (which may occur not necessarily from the direct charge interaction, but rather indirectly, through elastic forces in the solid state, e.g., where also quantum exchange forces etc. may be active). Thus we obtain from (41) where we put $V \to \infty$[2])

$$Z(T, \omega)=\frac{(2\pi mT)^{3/2}}{h^3}\left[\int_{-\infty}^{+\infty} \exp\left(-\frac{m\omega^2 x^2}{2T}\right)dx\right]^3=\left(\frac{2\pi T}{h\omega}\right)^3. \tag{44}$$

We now remark that the potential (43) can be generalized to infinite many other cases which are also integrable for calculation of metric (42), but which do not have so important physical interpretation and application as (43). Starting from the Euler formula for the Γ-function

$$\Gamma(z)=\int_0^\infty \exp(-t)t^{z-1}dt, \qquad z\in \mathbf{C}, \qquad \mathrm{Re}\, z>0,$$

we put

$$z=(p+1)/q, \qquad t=qx^2, \qquad dt=aqx^{q-1}dx, \qquad a>0, \qquad p>-1, \qquad q>0,$$

and obtain

$$\int_0^\infty \exp(-ax^q)x^p dx=a^{-(p+1)/q}q^{-1}\Gamma((p+1)/q).$$

If we introduce the potential

$$\Phi(x)=\frac{a}{e}|x|^q, \qquad |x|^q=\sum_{i=1}^{3} |x^i|^q, \qquad a>0, \qquad q>0,$$

we may call the potential of a *generalized oscillator of order* $q/2$ $(q>0)$ (the usual oscillator being of order 1; if $q/2=1, 2, 3, \cdots$, the oscillator is regular, analytic for $x=0$, otherwise the derivatives of the potential have singularities at $x=0$). Then we obtain in place of (44) (for the latter case $q=2$, $a=m\omega^2/2$)

$$Z_q(T, a)=\frac{(2\pi mT)^{3/2}}{h^3}\left[2\int_0^\infty \exp\left(-\frac{ax^q}{T}\right)dx\right]^3=\left[\frac{2(2\pi m)^{1/2}T^{1/2+1/q}}{hqa^{1/q}}\Gamma\left(\frac{1}{q}\right)\right]^3,$$

and for the metric tensor components we have, finally,

$$g=\begin{pmatrix} 3/T^2q+3/T^2, & 3/aTq \\ 3/aTq, & 3/a^2q \end{pmatrix}, \qquad \det(g)=9/a^2T^2q>0, \quad \text{and} \tag{45}$$

$$g^{-1}=\begin{pmatrix} T^2/3, & -aT/3 \\ -aT/3, & a^2(1+q)/3 \end{pmatrix}, \qquad \det(g^{-1})=a^2T^2g/9>0.$$

Calculating the curvature we finally obtain $(u^1=T,\ u^2=a)$

$$\Gamma^1_{11}=-1/T, \qquad \Gamma^2_{11}=\Gamma^1_{12}=\Gamma^1_{21}=\Gamma^2_{12}=\Gamma^2_{21}=\Gamma^1_{22}=0, \qquad \Gamma^2_{22}=-1/a,$$

$$R_{1212}=0, \qquad R=K=0.$$

So, as for the ideal gas, we obtained a flat Euclidean space as the thermodynamical space. Similarly as in the previous section, we may then introduce normal coordinates and Euclidean

---

[2] This is a rare thermodynamically regular case in the infinite space, just because of the sufficiently strong repulsive potential at infinity which actually makes the volume of the motion finite, but depending on the energy, as in magnetic bottle. Parameter V is therefore replaced by as a parameter of the repulsive potential.

representation of thermodynamical parameters, as well as generalize this theory for $r>1$. We resign here from these developments because of the lack of space.

§ 6. **Conclusions and acknowledgements.** Concluding, we may say that the method of information geometry gives a natural mathematical frame for statistical thermodynamics. In the simplest cases, as the both discussed above, we obtain the simplest geometry possible, the Euclidean one. In this geometry the usual thermodynamical parameters, as temperature, volume, pressure etc., appear to be curvilinear, and only entropy as information becomes rectilinear from the beginning. But it is easy to introduce special "scales" for parameters (e.g., logarithmic) in which they appear as Euclidean or Cartesian coordinates (we call this the Euclidean representation). As is seen from (45), in the case of oscillators there are more involved combinations of the original parameters, than in case of the ideal gas, as the natural coordinates. Then we may rotate and translate the natural coordinates according to the Euclidean group of motions.

In Euclidean or curved Riemannian spaces we obtain the principle of the shortest way (the movement along geodesics) as the simplest principle for thermodynamical processes, in general, irreversible (when entropy is changed). Because of the meaning of our information metric, we naturally obtain the principle of minimum information gain or the Glansdorff-Prigogine principle of irreversible thermodynamics, as was also formulated by Schlögl ([**21**], [**22**]) and others, cf. [**18**], although without our geometrical interpretation.

The authors thank Dr. Henryk Janyszek for his valuable remarks and discussion which help to improve some of our results.

Institute of Physics
N. Copernicus University
87-100 Toruń, Poland
and
Division of Information Engineering
Hokkaido University
Sapporo 060, Japan

REFERENCES

[1] J. Aczél and Z. Daróczy: On measures of information and their characterizations, *Academic Press, New York*, (1975).
[2] B. H. Callen: Thermodynamics, *Wiley, New York*, (1960).
[3] S. Guiasu: Information theory with applications, *McGraw-Hill, New York*, (1977).
[4] R. S. Ingarden: Differential geometry and physics, *Tensor, N. S.*, **30** (1976), 201–209.
[5] R. S. Ingarden, Y. Sato, K. Sugawa and M. Kawaguchi: Information thermodynamics and differential geometry, *Tensor, N. S.*, **33** (1979), 347–353.
[6] R. S. Ingarden, H. Janyszek, A. Kossakowski and T. Kawaguchi: Information geometry of quantum statistical systems, *Tensor, N. S.*, **37** (1982), 105–111.
[7] R. S. Ingarden and H. Janyszek: On the local Riemannian structure of the state space of classical information thermodynamics, *Tensor, N. S.*, **39** (1982), 279–285.
[8] R. S. Ingarden and A. Kossakowski: On the connection of non-equilibrium information thermodynamics with non-hamiltonian quantum mechanics of open systems, *Annal of Physics*, **89** (1975), 451–485.
[9] R. S. Ingarden and K. Urbanik: Quantum informational thermodynamics, *Acta Physica Polonica*, **21** (1962), 281–304.
[10] E. T. Jaynes: Information theory and statistical mechanics, *Physical Review*, **106** (1957), 620–630.
[11] H. Jeffreys: An invariant form for the prior probability in estimation problems, *Proceedings of the Royal Society, London*, **A186** (1946), 453–461.
[12] H. Jeffreys: Theory of probability, *Oxford Univ. Press, Oxford*, (1948).

[13] D. F. Kerridge: Inaccuracy and inference, *Journal of the Royal Statistical Society*, **B23** (1961), 184–194.
[14] S. Kullback: Information theory and statistics, *Wiley, New York*, (1959).
[15] S. Kullback and R. A. Leibler: On information and sufficiency, *Annals of Mathematical Statistics*, **22** (1951), 79–86.
[16] W. Ochs: Basic properties of the generalized Boltzmann-Gibbs-Shannon entropy, *Reports on Mathematical Physics*, **9** (1976), 135–155.
[17] W. Ochs: A unique characterization of the generalized Boltzmann-Gibbs-Shannon entropy, *Reports on Mathematical Physics*, **9** (1979), 331–354.
[18] E. Pfaffenhuber: Information-theoretical stability and evolution criteria in irreversible thermodynamics, *Journal of Statistical Physics*, **16** (1977), 69–90.
[19] A. Rényi: Wahrscheinlichkeitsrechnung, *VEB Verlag d.Wiss., Berlin*, (1962), Engl. transl.: Probability theory, *Acad. Kiádo, Budapest*, (1970).
[20] Y. Sato, K. Sugawa and M. Kawaguchi: The geometrical structure of the parameter space of the two-dimensional normal distribution, *Reports on Mathematical Physics*, **16** (1979), 111–124.
[21] F. Schlögl: Zur statistichen Theorie der Entropieproduktion in nicht abgeschlossenen Systemen, *Zeitschrift für Physik*, **191** (1966), 81–90.
[22] F. Schlögl: On the statistical foundation of the thermodynamic evolution cryterion of Glansdorff and Prigogine, *Annals of Physics*, **45** (1967), 155–163.
[23] Ya. P. Terletskii: Statistical physics, *Wiley, New York*, (1960).
[24] H. Umegaki: Conditional expectation in an operator algebra, IV (entropy and information), *Kodai Mathematical Seminar*, **14** (1962), 59–85.
[25] A. Wehrl: General properties of entropy, *Reviews of Modern Physics*, **50** (1978), 221–260.