Thermodynamics and contact geometry
Lecture 4: Introducing metrics

Eivind Schneider

Department of Mathematics, University of Hradec Králové

July 22, 2021

Metrics in thermodynamics

The contact geometry of thermodynamics seems to have been understood to a great extent by Gibbs 150 years ago. More recently others have introduced other geometric constructions to thermodynamics, in particular Riemannian metrics.

For example, Ruppeiner introduced in 1979 a metric on the 2-dimensional space of states, related it to fluctuation theory, and studied its curvature.

In fact, using the principle of maximal entropy (or minimal information gain), both the contact structure (see [Jaynes 1957]) and the metric (see [Mrugala 1990]) can be given statistical interpretations.

Riemannian metrics

Let us introduce a new type of structure on manifolds, namely Riemannian metrics.

Definition

A Riemannian metric g on a smooth manifold M is a collection of positive definite symmetric bilinear maps $g_p \colon T_pM \times T_pM \to M$ that depend smoothly on the point $p \in M$.

If $x^1, ..., x^n$ are coordinates on M, we can write g as

$$g = g_{ij}(x) dx^i \otimes dx^j$$

where the matrix $[g_{ij}(x)]$ is symmetric and positive definite at each point. Recall that a matrix A is symmetric if $A = A^T$ and that a symmetric matrix is positive definite if $v^T A v > 0$ for every nonzero column vector v. If $X = a^i \partial_{x^i}$ and $Y = b^i \partial_{y^i}$, we have

$$g(X, Y) = g_{ij} a^i b^j.$$

The Euclidean metric

Consider $\mathbb{R}^2$ with the metric $g = dx^2 + dy^2$. If $X = a_1\partial_x + a_2\partial_y$ and $Y = b_1\partial_x + b_2\partial_y$, then

$$g(X, Y) = a_1b_1 + a_2b_2.$$

We can also compute lengths of vectors:

$$\|X\|^2 = g(X, X) = a_1^2 + a_2^2.$$

Note: By $dx^i dx^j$ we mean $\frac{1}{2}(dx^i \otimes dx^j + dx^j \otimes dx^i)$. In particular $dx^2 = dxdx = dx \otimes dx$. We have $dx^i \otimes dx^j(X, Y) = dx^i(X)dx^j(Y)$.

Surfaces in Euclidean space

The first examples of nontrivial Riemannian manifolds were probably surfaces in Euclidean space.

Consider the space $\mathbb{R}^3$ with the Euclidean metric $g = dx^2 + dy^2 + dz^2$, and let $S \subset \mathbb{R}^3$ be a surface given by a function $z = f(x, y)$. Then we can take x and y to be coordinates on S. The metric "restricted to S" becomes

$$\begin{aligned}\tilde{g} &= dx^2 + dy^2 + df(x,y)^2 \\ &= dx^2 + dy^2 + (f_x dx + f_y dy)^2 \\ &= (1 + f_x^2)dx^2 + (1 + f_y^2)dy^2 + 2f_x f_y dxdy.\end{aligned}$$

Examples:

- Let $f(x, y) = C$. Then $\tilde{g} = dx^2 + dy^2$.
- Let $f(x, y) = h(x)$. Then $\tilde{g} = (1 + h_x^2)dx^2 + dy^2$.
- Let $f(x, y) = \sqrt{1 - x^2 - y^2}$. Then $\tilde{g} = \frac{(1-y^2)dx^2 + 2xydxdy + (1-x^2)dy^2}{1-x^2-y^2}$.

The curvature tensor

The previous examples were curved 2-dimensional manifolds in 3-dimensional space. The Riemannian metric g on a manifold M lets us talk about "curved" manifolds without embedding them into a higher-dimensional space.

On each Riemannian manifold (M, g) there exists a tensor called the curvature tensor. It is a 4-tensor

$$R = R_{ijkl} dx^i \otimes dx^j \otimes dx^k \otimes dx^l.$$

The coefficients R_{ijkl} can be given in terms of g_{ij} (here g^{ij} denotes the matrix inverse to g_{ij}):

$$\begin{aligned} R_{ijkl} = {} & \frac{1}{2}\frac{\partial}{\partial x^k}\left(\frac{\partial g_{ij}}{\partial x^l} + \frac{\partial g_{lj}}{\partial x^i} - \frac{\partial g_{il}}{\partial x^j}\right) - \frac{1}{2}\frac{\partial}{\partial x^l}\left(\frac{\partial g_{ij}}{\partial x^k} + \frac{\partial g_{kj}}{\partial x^i} - \frac{\partial g_{ik}}{\partial x^j}\right) \\ & + \frac{1}{4} g^{hm}\left(\frac{\partial g_{im}}{\partial x^k} + \frac{\partial g_{km}}{\partial x^i} - \frac{\partial g_{ik}}{\partial x^m}\right)\left(\frac{\partial g_{jh}}{\partial x^l} + \frac{\partial g_{lh}}{\partial x^j} - \frac{\partial g_{jl}}{\partial x^h}\right) \\ & - \frac{1}{4} g^{hm}\left(\frac{\partial g_{im}}{\partial x^l} + \frac{\partial g_{lm}}{\partial x^i} - \frac{\partial g_{il}}{\partial x^m}\right)\left(\frac{\partial g_{jh}}{\partial x^k} + \frac{\partial g_{kh}}{\partial x^j} - \frac{\partial g_{jk}}{\partial x^h}\right) \end{aligned}$$

The scalar curvature on 2-dimensional manifolds

For a 2-dimensional manifold the only nonvanishing part of the curvature tensor is

$$R_{1212} = -R_{2112} = -R_{1221} = R_{2112}.$$

From the curvature tensor, we can generate a scalar function, called the scalar curvature. If $g = a dx^2 + 2b dx dy + c dy^2$, then the scalar curvature is given by

$$\mathsf{Sc} = R_{ijkl} g^{ik} g^{jl} = -\frac{a_{yy} - 2b_{xy} + c_{xx}}{ac - b^2} - \frac{(-a a_y c_y + 2a b_x c_y - a c_x^2 - b a_x c_y + 2b a_y b_y + b a_y c_x - 4b b_x b_y + 2b b_x c_x + 2c a_x b_y - c a_x c_x - c a_y^2)}{2(ac-b^2)^2}$$

In two dimensions, the scalar curvature is essentially the only invariant scalar quantity that depends a, b, c and their first- and second-order partial derivatives. By "invariant" we mean invariant under coordinate transformations (local diffeomorphisms): If we change coordinates, the functions a, b, c change but the expression of Sc in terms of a, b, c and their derivatives remains the same.

The scalar curvature on submanifolds in Euclidean space

In the case when

$$g = (1 + f_x^2)dx^2 + 2f_x f_y dxdy + (1 + f_y^2)dy^2$$

the scalar curvature simplifies a lot:

$$\text{Sc} = 2\frac{f_{xx}f_{yy} - f_{xy}^2}{(f_x^2 + f_y^2 + 1)^2}$$

This is called (up to a constant factor) the Gaussian curvature of the surface. Let us compute it for the examples we considered earlier.

Examples:

- Let $f(x, y) = C$. Then $\tilde{g} = dx^2 + dy^2$ and $\text{Sc} = 0$.
- Let $f(x, y) = h(x)$. Then $\tilde{g} = (1 + h_x^2)dx^2 + dy^2$ and $\text{Sc} = 0$.
- Let $f(x, y) = \sqrt{r^2 - x^2 - y^2}$. Then $\tilde{g} = \frac{(r^2-y^2)dx^2+2xydxdy+(r^2-x^2)dy^2}{r^2-x^2-y^2}$ and $\text{Sc} = \frac{2}{r^2}$.

A metric on thermodynamic systems

Thermodynamic systems (in the form of Legendrian submanifolds in $\mathbb{R}^5(T, p, e, v, s)$) can also be equipped with a metric. The one we will focus on here is given by

$$g = -(s_{ee}(e,v)de^2 + 2s_{ev}(e,v)dedv + s_{vv}(e,v)dv^2)$$

in (e, v)-coordinates (see for example [Mrugala et. al. 1990] or [Lychagin 2020]). It is required to be positive definite.

G. Ruppeiner relates this metric to the variance of fluctuations ([Ruppeiner 1979]) and investigated the physical meaning of the scalar curvature and computed it for several thermodynamic systems (see [Ruppeiner 1995]).

A metric on the ideal gas

For an ideal gas, given by $pv = aT$, $e = cT$ and $s = c \ln e + a \ln v + s_0$, the metric is given by

$$g = \frac{c}{e^2} de^2 + \frac{a}{v^2} dv^2.$$

The scalar curvature Sc for this metric is 0. This means that there exists coordinates $x(e, v), y(e, v)$ such that $g = dx^2 + dy^2$.

From the intrinsic viewpoint (allowing arbitrary coordinate changes on L) this would mean that the ideal gas is the unique "flat" thermodynamic system.

Some concepts from statistics

The following is based on [Lychagin 2020] (see also [Mrugala et. al. 1990] and [Lychagin, Roop 2021]). First we need some statistical notions.

Let Ω be a set which we call the sample space. A σ-algebra $\mathcal{A}$ on Ω is a collection of subsets of Ω such that

- $\Omega \in \mathcal{A}$,
- $\Omega \setminus S \in \mathcal{A}$ for each $S \in \mathcal{A}$,
- $\cup_i S_i \in \mathcal{A}$ for all countable subsets $\{S_i\} \subset \mathcal{A}$.

The σ-algebra lets us define a measure on Ω, and pair $(\Omega, \mathcal{A})$ is called a measurable space. A measure μ is a map $\mu \colon \mathcal{A} \to [0, \infty)$ with the properties

- $\mu(\emptyset) = 0$,
- $\mu(\cup_i S_i) = \sum_i \mu(S_i)$ for all countable subsets $\{S_i\} \subset \mathcal{A}$ such that $S_i \cap S_j = \emptyset$ for every $i \neq j$.

The triple $(\Omega, \mathcal{A}, \mu)$ is called a measure space. If $\mu(\Omega) = 1$, we call μ a probability measure and $(\Omega, \mathcal{A}, \mu)$ a probability space.

More statistics

For the space $V = \mathbb{R}^n$ we can define the (Borel) σ-algebra $\mathcal{B}$ as the smallest σ-algebra which contains all open sets of V. Then a random vector is a map

$$X\colon (\Omega, \mathcal{A}, \mu) \to V,$$

satisfying $X^{-1}(U) \in \mathcal{A}$ for all sets $U \in \mathcal{B}$ and

$$\begin{aligned} X^{-1}(U_1 \cup U_1) &= X^{-1}(U_1) \cup X^{-1}(U_2), \\ X^{-1}(U_1 \cap U_1) &= X^{-1}(U_1) \cap X^{-1}(U_2), \\ X^{-1}(V \setminus U) &= \Omega \setminus X^{-1}(U). \end{aligned}$$

Then a measure on V is induced by the formula

$$X_*(\mu)(U) = \mu(X^{-1}(U)).$$

In this lecture all measures will be probability measures, and measure spaces will be probability spaces.

Measurements

Consider a random vector

$$X \colon (\Omega, \mathcal{A}, \mu_0) \to V$$

where Ω is the sample space, $\mathcal{A}$ is the σ-algebra of events, and μ_0 is a probability measure.

We interpret X as a measurement of $x_0 \in V$ if

$$E_{\mu_0}(X) = \int_\Omega X d\mu_0 = x_0.$$

Let us choose an affine frame such that $x_0 = 0$. The measurement of a vector $x \in V$ is given by a probability measure μ different from μ_0, satisfying $E_\mu(X) = \int_\Omega X d\mu = x$. Assuming that μ is absolutely continuous with respect to μ_0, we have by the Radon-Nikodym theorem $d\mu = \rho d\mu_0$.

Notice that E_μ behaves well under affine transformations:

$$E_\mu(AX + B) = AE_\mu(X) + B.$$

Constraints on ρ

We require

$$\int_\Omega \rho d\mu_0 = 1, \qquad E_\mu(X) = \int_\Omega \rho X d\mu_0 = x.$$

This is an underdetermined set of conditions on ρ. We define the information gain

$$I(\mu, \mu_0) = \int_\Omega \rho \ln \rho d\mu_0,$$

and require ρ to minimize $I(\mu, \mu_0)$. Jaynes noted that this (or more precisely the the maximation of entropy) is "the only unbiased assignment we can make". The three conditions imply

$$\rho = \frac{1}{Z(\lambda)} e^{\langle \lambda, X \rangle}$$

with $\lambda \in V^*$. Here λ is the (multidimensional) Lagrange multiplier for the proposed optimization problem. $Z(\lambda) = \int_\Omega e^{\langle \lambda, X \rangle} d\mu_0$ is called the partition function.
Choosing basis on V gives us coordinates x^i on V, and dual coordinates λ_i on V^*.

Symplectic and contact structures

If we define $H(\lambda) = -\ln Z(\lambda)$, we get

$$H_{\lambda_i}(\lambda) = -\frac{1}{Z(\lambda)} Z_{\lambda_i}(\lambda) = -\int_\Omega X^i \frac{e^{\langle \lambda, X \rangle}}{Z(\lambda)} d\mu_0 = -x^i.$$

These n equations define an n-dimensional submanifold $L \subset V \times V^*$ which is Lagrangian with respect to symplectic form $dx^i \wedge d\lambda_i$. Restricting $I(\mu, \mu_0)$ to L gives

$$I(\lambda) = H(\lambda) + \langle \lambda, x \rangle = H(\lambda) - \lambda_i H_{\lambda_i}(\lambda).$$

Or if we can solve $H_{\lambda_i} = -x^i$ for $\lambda(x)$, we may write

$$I(x) = H(\lambda(x)) + \langle \lambda(x), x \rangle.$$

And we notice that $I_{x^i} = \lambda_i$. This determines a Legendrian submanifold

$$\tilde{L} = \{u = I(x), \lambda_i = I_{x^i}(x)\} \subset V \times \mathbb{R} \times V^*$$

with respect to the contact form $du - \lambda_i dx^i$.

Contact and symplectic structures

The measurement of a point in V leads us to consider Legendrian submanifolds in the contact space $V \times \mathbb{R} \times V^*(x^i, u, \lambda_j)$ with contact structure given by $du - \lambda_i dx^i$. They are locally given by $u = I(x), \lambda_i = I_{x^i}(x)$.

It may be convenient to remove information gain from the picture, and consider only Lagrangian submanifolds in $(V \times V^*, dx^i \wedge d\lambda_i)$. Locally they are given by n functions $x^i(\lambda)$, satisfying $x^i_{\lambda_j} = x^j_{\lambda_i}$ (similar to the Maxwell relations).

Example

In classical thermodynamics the fundamental thermodynamic relation is given by

$$de - Tds + pdv = 0 \quad \text{or} \quad -ds - (-T^{-1})de - (-pT^{-1})dv = 0.$$

Geometrically this means that systems in thermodynamic equilibrium are Legendrian submanifolds in a five-dimensional contact manifold. We can connect this to the above discussion by setting $du = -ds, x^1 = e, x^2 = v, \lambda_1 = -T^{-1}, \lambda_2 = -pT^{-1}$. We notice that we can identify the thermodynamic identity with $du - \lambda_i dx^i$ if we "measure" $(e, v) \in V$.

A metric on Legendrian submanifolds

On a Legendrian submanifold $L \subset V \times \mathbb{R} \times V^*$ (and also on the corresponding Lagrangian submanifold in $\subset V \times V^*$), we can define a metric σ_2. If we use λ_i as coordinates on L, we have

$$\sigma_2 = -H_{\lambda_i \lambda_j} d\lambda_i \otimes d\lambda_j = x^i_{\lambda_j} d\lambda_i \otimes d\lambda_j.$$

In coordinates x^i it is given by

$$\sigma_2 = u_{x^i x^j} dx^i \otimes dx^j.$$

The metric encodes (as we will see soon) the variance of the random vector X with respect to the extremal measure ρ corresponding to λ. As mentioned, this metric has received attention in the fields of information geometry and geometric thermodynamics. For instance, Ruppeiner has suggested a physical meaning of the scalar curvature.

Action by the affine group

- Recall that $E_{\mu_0}(AX+B) = AE_{\mu_0}(X) + B$.
- Affine transformations preserve the contact and the metric structure, and they are the only transformations on V that does that.

$$(x, u, \lambda) \mapsto (Ax + B, u, (A^{-1})^T \lambda), \qquad A \in GL(V), B \in V$$

Conclusion: The principle of minimal information gain leads us to consider submanifolds in $V \times V^* \times \mathbb{R}$ which are Legendrian with respect to the contact form $du - \lambda_i dx^i$. In addition it gives a metric $u_{x^i x^j} dx^i \otimes dx^j$ on Legendrian submanifolds. The group Aff acts on the space of Legendrian manifolds as the largest group acting on V that preserves both the contact structure and the metric.

Alternatively, we may consider submanifolds in $V \times V^*$ which are Lagrangian with respect to the symplectic form $d\lambda_i \wedge dx^i$. The Lagrangian submanifolds come equipped with a metric $x^i_{\lambda_j} d\lambda_i \otimes d\lambda_j$. The group Aff acts on the space of Lagrangian submanifolds.

More invariant symmetric tensors

Central moments (one of which is σ_2) gives additional structure on Legendrian submanifolds.

We write $X = X^i d\lambda_i$. The kth moment corresponds to a symmetric k-form:

$$\begin{aligned} m_k &= \int_\Omega X^{\otimes k} \rho d\mu_0 = \left(\int_\Omega X^{i_1} \cdots X^{i_k} \rho d\mu_0 \right) d\lambda_{i_1} \otimes \cdots \otimes d\lambda_{i_k} \\ &= \frac{Z_{\lambda_{i_1} \cdots \lambda_{i_k}}}{Z} d\lambda_{i_1} \otimes \cdots \otimes d\lambda_{i_k} \end{aligned}$$

These are $GL(V)$-invariant, but not invariant under translations. The kth central moment (i.e. the moment of $X - m_1$) is given by

$$\sigma_k = \sum_{i=0}^{k} \binom{k}{i} m_i \cdot m_1^{(k-i)}.$$

They define Aff-invariant symmetric k-forms on Lagrangian manifolds, for $k \geq 2$.

To be continued

Tomorrow we will use the central moments to find scalar differential invariants with respect to Aff. We will see that some of the invariants are well-known physical quantities.

Sources

- Mrugala, Nulton, Schön, Salamon, *Statistical approach to the geometric structure of thermodynamics*, Phys. Rev. A (1990).
- Ruppeiner, *Thermodynamics: A Riemannian geometric model*, Phys. Rev. A (1979).
- Ruppeiner, *Riemannian geometry in thermodynamic fluctuation theory*, Review of Modern Physics **67** (1995).
- Jaynes, *Information Theory and Statistical Mechanics*, Phys. Rev. **106** (1957),
- Lychagin, *Contact Geometry, Measurement, and Thermodynamics* in Nonlinear PDEs, Their Geometry, and Applications, Birkhäuser (2020)
- Lychagin, Roop, *Critical Phenomena in Darcy and Euler Flows of Real Gases*, in Differential Geometry, Differential Equations, and Mathematical Physics: The Wisla 19 Summer School, Birkhäuser (2021).
- Schneider, *Differential Invariants of Measurements, and Their Relation to Central Moments*, Entropy **22** (2020).