On interplay between jet and information geometries

Valentin Lychagin
IPU RAS,lab6

October 12, 2022

1 Plan of the talk

1. Measurment and Information Geometry
2. Equations of State, EoS. Partition functions and Lagrangian manifolds. Relations with Jet geometry and Integral manifolds of the Cartan distribution.
3. Moments, positive symmetric differential forms, applicable domains and geometrical structures on Lagrangian manifolds.
4. Phase transitions. Two types of singularities on integral manifolds: (1) singularities of projections and (2) singularities of prolongations.
5. Example. Phase transitions in real (Van der Waals) gases.

2 Measurment and Information Geometry

1. Let (Ω, q) be a probability space ,$\mathbf{V}$ be a vector space $(\dim \mathbf{V} = n)$, q be a probability measure, and

   $$X : \Omega \to \mathbf{V}$$

   be a random vector.
2. By result of measurement we mean the avarage

   $$\mathbf{E}\left(X, q\right) = \int_{\Omega} X\left(\omega\right) dq,$$

   and consider q as a measurement device.
3. For given vector $x \in \mathbf{V}$, we a looking for a device (=probability measure p) such, that $\mathbf{E}\left(X, p\right) = x$.

4. Let ρ be the density of p wrt q, $dp = \rho\, dq$ and let
$$I(p) = \int_{\Omega} \rho \ln(\rho)\, dq$$
be the Information gain (or Kullback–Leibler divergence).

It is a "quasi distance": $I(p) \geq 0$ and $I(p) = 0$ iff $\rho = 1$ almost everywhere.

5. Variational problem: find a non-negative function ρ such that
$$\int_{\Omega} \rho dq = 1,\ x = \int_{\Omega} \rho(\omega) X(\omega)\, dq$$
and $I(p)$ is minimal for probability measures satisfying the condition $\mathbf{E}(X, p) = x$.

6. Auxiliary functional:
$$\rho \to \int_{\Omega} \rho \ln(\rho)\, dq - \lambda_0 \left(\int_{\Omega} \rho(\omega)\, dq - 1 \right) - \lambda \left(\int_{\Omega} \rho(\omega) X(\omega)\, dq - x \right)$$
where $\lambda_0 \in \mathbb{R}, \lambda \in \mathbf{V}^*$ are the Lagrange multipliers.

7. Solution $\frac{\delta \mathcal{F}}{\delta \rho} = 0$,
$$\begin{aligned} \rho_\lambda &= \frac{1}{Z(\lambda)} \exp \langle \lambda, X \rangle, \\ Z(\lambda) &= \int_{\Omega} \exp \langle \lambda, X \rangle\, dq, \\ x &= \frac{1}{Z} \frac{\partial Z}{\partial \lambda}. \end{aligned}$$

8. Here $Z(\lambda)$ is the so-called *partition function.* The probability measures p_λ, where $dp_\lambda = \rho_\lambda dq$ we call *extreme measures* and function
$$H = \ln Z$$
we call Plank potential (or *Gibbs free entropy).*
These names are taken from thermodynamics.
Between the information gain $I(\lambda) = I(p_\lambda)$ and the Plank potential, there is the following relation:
$$I(\lambda) = \left\langle \lambda, \frac{\partial H}{\partial \lambda} \right\rangle - H(\lambda).$$

9. Geometrical interpretation of the above calculations:

   $\mathbf{B} = \mathbf{V}^*$ the manifold of the extreme probabulity measures (λ). $\mathbf{T}^*(\mathbf{B}) = \mathbf{V} \times \mathbf{V}^*$ the measurment phase manifold. Lagrangian

$$L_H = \left\{ x = \frac{\partial H}{\partial \lambda} \right\} \subset \mathbf{T}^*(\mathbf{B}),$$

   or Legendrian

$$L_H^{(1)} = \left\{ x = \frac{\partial H}{\partial \lambda}, u = H(\lambda) \right\} \subset \mathbf{J}^1(\mathbf{B}),$$

   manifolds are the measurment Equations of State.

10. We'll accept more general picture: by EoS for measurment we mean a Lagrangian (or Legendrian) manifold $L \subset \mathbf{T}^*(\mathbf{B})$ or $L \subset \mathbf{J}^1(\mathbf{B})$, or an n-dimentional integral manifold of the Cartan distribution: $L \subset \mathbf{J}^k(\mathbf{B})$, where $k \geq 1$. By $L^{(i)} \subset \mathbf{J}^{k+i}(\mathbf{B})$ we denote the i-th prolongation of L.

3 Moments and symmetric differential forms

1. Let m_k be the k-th moment of random vector X :

$$m_k(\lambda) = E\left(X^{\otimes k}, p_\lambda\right) = \int_\Omega X(\omega)^{\otimes k} \rho_\lambda(\omega)\, dq \in \mathbf{S}^k(\mathbf{V}),$$

   and

$$c_k(\lambda) = E\left((X - m_1(\lambda))^{\otimes k}, p_\lambda\right) \in \mathbf{S}^k(\mathbf{V})$$

   be the central k-th moment.

2. Let $\{e_1, ..., e_n\}$ be a basis in vector space $\mathbf{V}$ and $\{e_1^*, ..., e_n^*\}$ be the dual basis in $\mathbf{V}^*$. Then

$$X(\omega) = \sum_{i=1}^n X_i(\omega)\, e_i$$

   and

$$X^{\otimes k}(\omega) = \sum_{k_1+..+k_n=k} \binom{k}{k_1...k_n} X_1^{k_1} \cdots X_{k_n}^{k_n} e_1^{k_1} \cdot \cdots \cdot e_n^{k_n},$$

   where $\cdot$ stands for the symmetric product of symmetric tensors.

3. If we diffirantiate the formula

$$Z(\lambda) = \int_\Omega \exp\langle \lambda, X\rangle\, dq$$

   we get

$$\frac{\partial Z}{\partial \lambda_i} = \int_\Omega \exp\langle \lambda, X\rangle\, X_i dq = Z\ E(X_i)$$

and

$$\frac{\partial^k Z}{\partial\lambda_{i_1}..\partial\lambda_{i_k}} = \int_\Omega \exp\langle\lambda, X\rangle X_{i_1}...X_{i_k} dq = Z\ E\left(X_{i_1}...X_{i_k}\right).$$

4. Finally,

$$m_k(\lambda) = \frac{1}{Z}\sum_{k_1+..+k_n=k}\binom{k}{k_1...k_n}\frac{\partial^k Z}{\partial\lambda_1^{k_1}...\partial\lambda_1^{k_n}} e_1^{k_1}\cdot\cdots\cdot e_n^{k_n}.$$

5. Remark, that at each point $\lambda = (\lambda_1, ..., \lambda_n) \in \mathbf{B} = \mathbf{V}^*$ we have isomorphism between frames $(e_1^*, ..., e_n^*)$ and $(\partial_{\lambda_1}, ..., \partial_{\lambda_n})$ in $\mathbf{T}_\lambda(\mathbf{B}) = \mathbf{V}^*$, and $(e_1, ..., e_n)$ and $(d\lambda_1, ..., d\lambda_n)$ in $\mathbf{T}_\lambda^*(\mathbf{B}) = \mathbf{V}$.

6. Thus, the moments $m_k(\lambda), \lambda \in \mathbf{B}$, define a symmetric k-form on EoS L_H :

$$m_k = \frac{1}{Z}\sum_{k_1+..+k_n=k}\binom{k}{k_1...k_n}\frac{\partial^k Z}{\partial\lambda_1^{k_1}...\partial\lambda_1^{k_n}} d\lambda_1^{k_1}\cdot\cdots\cdot d\lambda_1^{k_n} = \frac{d_k Z}{Z},$$

   where $d_k Z$ is the k-th order differential of the partition function (wrt the standard affine connection on $\mathbf{B}$!).

4 Partition functions and Lagrangian manifolds

1. Let $L \subset \mathbf{T}^*(\mathbf{B})$ be a Lagrangian manifold and let $\varrho = \sum x_i d\lambda_i$ be the universal Liouville-Poincare differential $1-$form. Then, as we have seen, the Plank potential equals

$$H = \int_\gamma \varrho.$$

   Therefore,

$$Z = \exp\left(\int_\gamma \varrho\right).$$

2. For general Lagrangian manifolds these formulae give us multivalued functions on L.

3. Remark, that the Born quantization condition means that L shoud be such that "energy" H multivalued with step $\hbar$.

5 Applicable domains

1. Consider symmetric k-forms $\tau \in \mathbf{S}^k(\mathbf{V})$ as homogeneous polynomials of degree k on the vector space $\mathbf{V}^*$. We say that τ is positive- definite,$\tau > 0$, if this polynomial is positive.

2. In addition, symmetric $2k$-forms $\tau \in \mathbf{S}^{2k}(\mathbf{V})$ we consider as quadrics Q_τ on $\mathbf{S}^k(\mathbf{V}^*)$, where
$$Q_\tau(a,b) = \langle \tau, a \cdot b \rangle .$$

   Theorem 1 *A symmetric $2k$-form $\tau \in \mathbf{S}^{2k}(\mathbf{V})$ is positive-definite if and only if the quadric Q_τ is positive-definite.*

   Example 2 *A symmetric 4-form*
$$\tau = a_4x^4 + 4a_3x^3y + 6a_2x^2y^2 + 4a_1xy^3 + a_0y^4$$
   on the plane (x,y) is positive-definite if and only if the matrix :
$$M_\tau = \begin{Vmatrix} a_4 & a_3 & a_2 \\ a_3 & a_2 & a_1 \\ a_2 & a_1 & a_0 \end{Vmatrix}$$
   positive or if the following inequalities
$$a_4 > 0,\ a_4a_2 - a_3^2 > 0,$$
$$a_0a_2a_4 - a_0a_3^2 - a_1^2a_4 + 2a_1a_2a_3 - a_2^3 > 0,$$
   hold.

3. We say that a point $x_{2k} \in L \subset J^{2k}(\mathbf{B}), \lambda = \pi_{2k}(x_{2k})$, is *applicable* if for any positive-definite polynomial
$$P(X) = a_{2k}X^{\otimes 2k} + \binom{2}{2k} a_{2k-2}X^{\otimes 2k-2} + \binom{3}{2k} a_{2k-3}X^{\otimes 2k-3} + \cdots + \binom{l}{2k} a_{2k-l}X^{\otimes 2k-l} + \cdots + a_0 \in \mathbf{S}^{2k}(\mathbf{V})$$
   where $a_{2k-i} \in \mathbf{S}^i(\mathbf{V})$ and $X \in \mathbf{V}$, the mean of the random symmetric tensor $\omega \in \Omega \to P(X(\omega))$, $\mathbf{E}(P(X), p_\lambda)$ is positive-definite too, $\mathbf{E}(P(X), p_\lambda) > 0$.

4. $k = 1$.
   For example, for $k = 1$, the positive polynomial $P(X)$ has the form $P(X) = X^{\otimes 2} + a_0$, where $a_2 \in \mathbf{S}^2(\mathbf{V})$, $a_2 > 0$.

   We have
$$E(P(X)) - P(E(X)) = c_2,$$
   and $P(E(X)) > 0$. Therefore, $E(P(X)) > 0$ for any positive polynomial P if and only if $c_2 > 0$.

   Theorem 3 *A point $x_4 \in L^{(4)}$, is applicable iff $c_2 > 0$ and $\left(3c_2c_4 - 2c_3^2\right) >$ 0 at the point.*

Remark 4 *So, in the domain of applicable points in* $L \subset J^{2k}(\mathbf{B}), k \geq 1$*, we have Riemannian structure given by the second central moment* $c_2 = d_2H$*, and, in addition, positive symmetric 6-form* $3c_2c_4 - 2c_3^2$*.*

Remark 5 *Lagrangian manifolds* $L \subset T^*(\mathbf{B})$ *play the crucial role in physics. They do establish the relations between intensive* $(\mathbf{V}^*)$ *and extensive* $(\mathbf{V})$ *quantities. Like relations between velocity and momentum in classical mechanics, or relations between stress and deformation tensors in continuous mechanics. Often, these relations are linear, i.e. correspond to Lagrangian manifolds that are linear subspaces in* $\mathbf{V}^* \oplus \mathbf{V}$*. From the probabilistic point of view they correspond to the normal distribution law. They have the constant central moment* c_2 *or the trivial Riemannian structure. The next interesting relations we may get from higher central moments, like* $c_3 = const$*, etc.*

Remark 6 *We have*

$$\begin{aligned}
c_2 &= d_2H, \\
c_3 &= m_3 - 3m_1m_2 + 2m_1^3 \ = d_3H, \\
c_4 &= m_4 - 4m_1m_3 + 6m_1^2m_2 - 3m_1^4 = d_4H + 3d_2H^2. \\
3c_2c_4 - 2c_3^2 &= 3d_2H \cdot d_4H - 2d_3H^2 + 9d_2H^3.
\end{aligned}$$

6 Phase transitions

1. Let $L \subset \mathbf{T}^*(\mathbf{B})$ be a Lagrangian manifold. The first type of singularity (so-called *first order phase transitions*) that we study is the singularity of the projection $\pi : L \to \mathbf{B}$, that is the set

$$\mathrm{Sing}(L) = \{x \in L, \dim \mathrm{Ker}\, d_x\pi \neq 0\}.$$

Removing this singularity and consider only applicable points we get a Lagrangian manifold

$$L_{\mathrm{reg}} = \left\{x \in L, \dim \mathrm{Ker}\, d_x\pi = 0, c_2 > 0, 3c_2c_4 - 2c_3^2 > 0, ...\right\}$$

that can be seen as the graph of the differential of a multivalued Plank function H (Gibbs energy).

The image of the singular set

$$L_{\mathrm{sing}} = \left\{x \in L, \dim \mathrm{Ker}\, d_x\pi \neq 0, \mathrm{or}\ c_2 \leq 0, \mathrm{or}\ 3c_2c_4 - 2c_3^2 \leq 0, ...\right\}$$

$\pi(L_{\mathrm{sing}})$ is called *coexistence surface,* and branches of the Plank potential are called *phases*.

2. Taking now prolongations of L_{reg} we get integral manifolds $L_{\mathrm{reg}}^{(k)} \subset J^k(B)$ that are prolongations of EoS. We say that at a point $x_k \in L_{\mathrm{reg}}^{(k)}$ the phase transition of order $(k+1)$ is occurs if $x_k \notin \pi_{k+1,k}\left(L_{\mathrm{reg}}^{(k+1)}\right)$.To understand this phenomena we need the following

Proposition 7 *Let X, Y be connected manifolds and $\alpha : Y \to X$ be a covering of degree m. Then any function $f \in C^\infty(Y)$ satisfies the polynomial equation*

$$f^m + a_{m-1}(x) f^{m-1} + \cdots a_1(x) f + a_0(x) = 0,$$

where $a_i \in C^\infty(X)$.

Proof. Let $\pi_1(X)$ and $\pi_1(Y)$ be fundamental groups of the manifolds X and Y respectively. Then, it is known, that the manifold Y is a $\pi_1(X)$-manifold, and $\pi_1(X)$-orbits are exactly the fibres of the projection α and they are isomorphic to the coset $\pi_1(X)/\pi_1(Y)$ of the group $\pi_1(X)$ by the subgroup $\pi_1(Y)$.Moreover, if we identify functions on X with functions on Y, then the subgroup $\pi_1(Y)$ acts in the trivial way on $C^\infty(X)$ and if function $h \in C^\infty(Y)$ is a fixed element of $\pi_1(Y)$-action, then $h \in C^\infty(X)$.

Using this observation, we take a point $x \in X$ and let the fibre $\alpha^{-1}(x)$ consist of points $\{y_1, ..., y_m\}$. Take a function $f \in C^\infty(Y)$ and consider the following polynomial in a parameter z :

$$p(z, x) = \prod_i (z - f(y_i)) = z^m + a_{m-1} z^{m-1} + \cdots a_1 z + a_0.$$

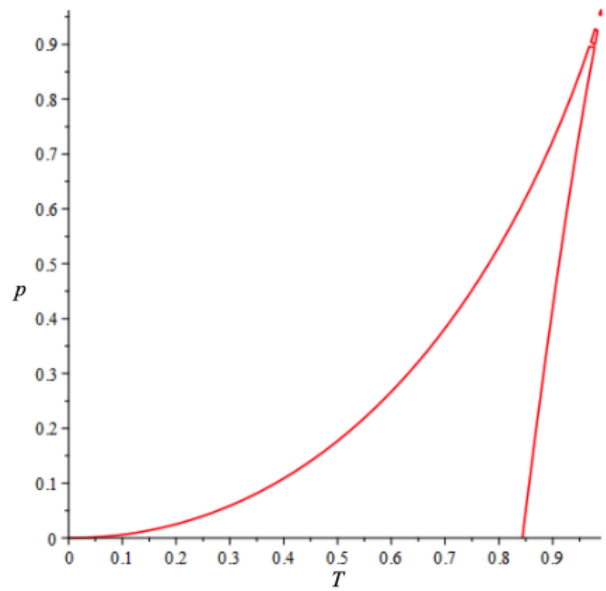
Assume, that the fundamental group $\pi_1(Y)$ acts in the trivial way on z. Then polynomial is $p(z, x)$ $\pi_1(X)$-invariant. Therefore, $a_i \in C^\infty(x)$. ■

1. Let us consider now the case when $Y \subset L_{\mathrm{reg}}^{(k)}$, $X = \pi_k(Y)$ and $\pi_k : Y \to X$ is a covering of degree m. Let (λ, u, u_σ) be the standard coordinates in $J^{(k)}$ and π_k and let $u_\sigma, |\sigma| = k$, satisfies the above algebraic equation $p(u_\sigma) = 0$ of degree m. Taking total derivatives we get the following equations

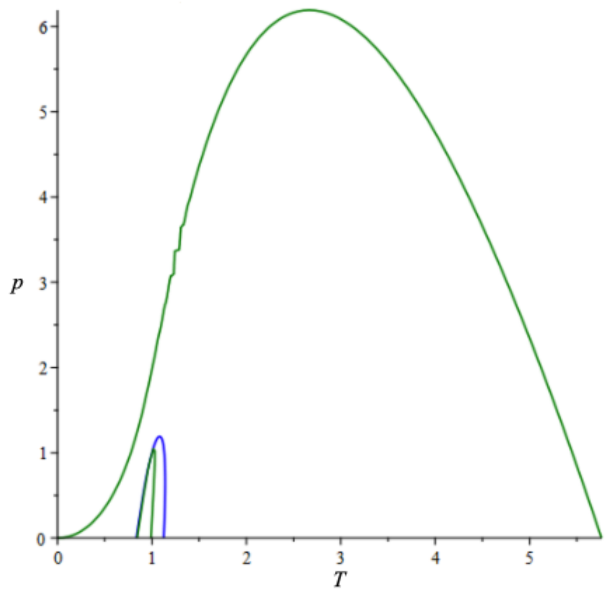
$$u_{\sigma+1_i} \frac{\partial p}{\partial z} + \frac{dp}{d\lambda_i} = 0,$$

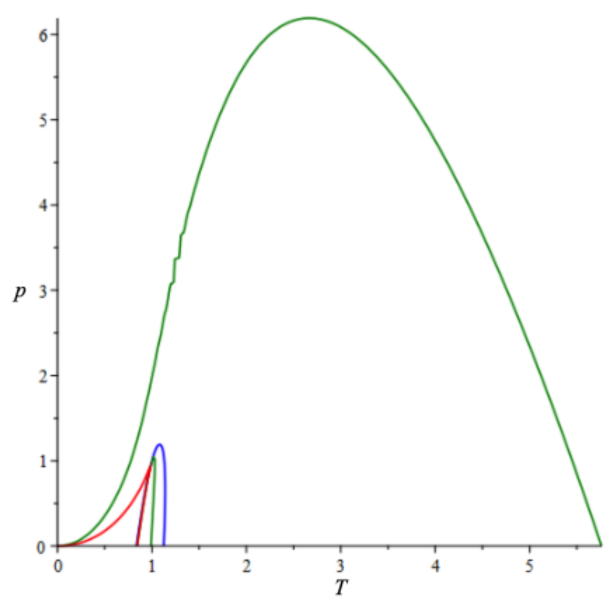
that could be solved wrt $u_{\sigma+1_i}$ if $\frac{\partial p}{\partial z} \neq 0$.

2. Let's call points on Y *prolongable*, the all above equations have smooth roots $u_{\sigma+1_i}$, i.e. points where the discriminants of all polynomials p for all $|\sigma| = k$ do not vanish or if a discriminant of some polynomial p equals zero then common roots of polynomials p and $\frac{\partial p}{\partial z}$ are also roots of polynomials $\frac{dp}{d\lambda_i}$.Going back to the probabilistic interpretation, we will say that non-prolongable points correspond to states where (central) moments of order $k+1$ are blowing up i.e. tends to infinity in some directions.



7 Example van der Waals gas


p
T
0
1
2
3
4
5
6


p
T
0
1
2
3
4
5
6