Problème des moments pour la théorie cinétique

Vincent PAVAN, J. Schneider

IUSTI - équipe DTF

Workshop Bordeaux 8-9/11/2018 - problèmes de moments - S. Brull organisateur

Le problème des moments originel

Soit $(m_k, k \in \mathbb{N})$ une liste de réels et I un intervalle. Existe-t-il une mesure $d\mu(x)$ (et une seule ?) telle que l'on ait

$$\forall k \in \mathbb{N}, \quad \int_I x^p d\mu(x) = m_k \tag{1}$$

On rapelle que:

1. pour $I = [0, 1]$ c'est le problème de Hausdorff
2. pour $I = \mathbb{R}$ c'est le problème de Hamburger
3. pour $I = \mathbb{R}^+$ c'est le problème de Stieltjes

De nombreuses variantes

- On peut ne retenir qu'un nombre fini de moment: $m_p, p \in [0, 2k]$. Il s'agit alors du problème dit tronqué (par opposition au problème complet)
- Lorsqu'il n'y a pas unicité de la mesure on peut imposer des restrictions sur la mesure: un cas particulier est celui des mesures atomiques $\sum_i a_i \delta_{x_i}$
- On peut imaginer toute sorte de variations sur l'ensemble d'intégration I. Le cas compact est particulièreent bien étudié

Le livre de la Jungle

Chaque variation dans la formulation du problème donne lieu à une littérature intensive et spécialisée utilisant un large spectre de techniques mathématiques (algèbre, analyse convexe, topologie, analyse complexe...)

Les résultats sont plus ou moins difficile à obtenir:

Theorem (Problème tronqué sur $\mathbb{R}$)

Soit $(m_k, k \in [0, 2p])$ *une suite de réels. Alors il existe une solution au problème tronqué si et seulement si la matrice de Hankel* $H_{ij} := m_{i+j}, 0 \leq i, j \leq p$ *est définie positive.*

Cette caractérisation utilise l'équivalence, pour les fonctions polynômes de la variables réelle, entre le polynômes positifs et les polyômes qui sont S.O.S (sum of squares). Pour le cas complet,

Theorem (Problème complet sur $\mathbb{R}$)

Soit $(m_k, k \in \mathbb{N})$ *une suite de réels. Alors il existe une solution au problème complet si et seulement pour tout* p *le problème tronqué d'ordre* $2p$ *possède une solution (autrement dit les matrices de Hankel correspondantes sont définies positives)*

Definition (représentation des polynômes multivariés)

Soit $\mathbf{v} \in \mathbb{R}^d$ l variable vectorielle de dimension d. On appelle

1. monome d'ordre n associé à $\mathbf{v}$ le tenseur symétrique $\mathrm{p}_n(\mathbf{v}) := \mathbf{v} \otimes \mathbf{v} \otimes \cdots \otimes \mathbf{v}$.
2. coefficient d'ordre n tout tenseur symétrique d'ordre n et taille d noté $\mathrm{A}_n \in \mathbb{S}(d, n)$
3. (fonction) polynôme de degré inférieur à n toute fonction de la forme

$$\pi(\mathbf{v}) = \sum_{k=0}^{k=n} \mathrm{A}_k : \mathrm{p}_k(\mathbf{v}), \ \mathrm{A}_k \in \mathbb{S}(d, k) \tag{2}$$

où le symbole : est le produit de contraction entre tenseur de même ordre et même dimension.

Pour une sous partie $I \subset \mathbb{R}^d$ et une mesure $d\mu(\mathbf{v})$ adaptée on peut définir le tenseur symétrique suivant

$$\mathrm{S}_k := \int_I \mathrm{p}_k(\mathbf{v})\, d\mu(\mathbf{v}) \tag{3}$$

on peut ainsi généraliser pour des tenseurs symétriques (d'odre k et de taille d) S_k le problème des moments (complet ou tronqué)

difficulté du cas multivarié

Dans le cas mono-varié (d=1) tous les A_k sont des coefficients réels: les sous espaces associés aux coefficients de degré k sont toujours de dimension 1. Dans le cas multi-varié, les sous espaces associés aux coefficients homogène de degré k ont une dimension $(k+d-1)!/\left(k!\,(d-1)!\right)$

Dans la pratique, on travaille avec des tenseurs $\mathrm{t}_k^l(\mathbf{v})$ mais qui représentent des sous espaces de polynôme (homogènes) de degré k. Principe de construction: l'invariance par rotation et par translation:

$$\forall\Theta,\ \mathrm{A} : \mathbf{t}_k(\Theta\mathbf{v}) = L_\theta(\mathrm{A}) : \mathrm{t}_k(\mathbf{v}) \tag{4}$$

$$\forall\mathbf{u},\ \sum_{k=0}^{k=n} \mathrm{A}_k : \mathrm{t}_k(\mathbf{v}-\mathbf{u}) = \sum_{k=0}^{k=n} L_{\mathbf{u},k}(\mathrm{A}_0,\cdots,\mathrm{A}_n) : \mathrm{t}_k(\mathbf{v}) \tag{5}$$

Example (Exemple de tenseur de degré deux invariant par rotation)

$$\mathrm{t}_2^1(\mathbf{v}) = \frac{1}{d}\mathbf{v}^2\mathbb{I},\ \ \mathtt{dim} = 1 \tag{6}$$

$$\mathrm{t}_2^2(\mathbf{v}) = \mathbf{v}\otimes\mathbf{v} - \frac{1}{d}\mathbf{v}^2\mathbb{I},\ \mathtt{dim} = \frac{d(d+1)}{2} - 1 \tag{7}$$

Definition

Le problème des moments multivariés est dit ***tronqué partiel***, lorsque les monomes $\mathrm{t}_k(\mathbf{v})$ (homogènes de degré k) retenus ne sont pas choisis égaux aux $\mathrm{p}_k(\mathbf{v})$, autrement dit la suite des tenseurs symétriques $\mathrm{s}_k \in \mathbb{S}(d,k)$ doit être résolue sur les expressions de la forme:

$$\mathrm{S}_k = \int_I \mathrm{t}_k(\mathbf{v})\, d\mu(\mathbf{v}) \tag{8}$$

Example

les bases suivantes sont utilisées en théorie cinétique:

1. la base d'Euler $\mathbf{m}(\mathbf{v}) = \left(1, \mathbf{v}, \mathbf{v}^2\right)$
2. La base de Gauss: $\mathbf{m}(\mathbf{v}) = (1, \mathbf{v}, \mathbf{v} \otimes \mathbf{v})$
3. La base de Grad: $\mathbf{m}(\mathbf{v}) = \left(1, \mathbf{v}, \mathbf{v} \otimes \mathbf{v}, \mathbf{v}^2\mathbf{v} \vee \mathbb{I}\right)$
4. La base de Levermore: $\mathbf{m}(\mathbf{v}) = \left(1, \mathbf{v}, \mathbf{v} \otimes \mathbf{v}, \mathbf{v}^2\mathbf{v} \vee \mathbb{I}, \mathbf{v}^2\mathbb{I} \vee \mathbb{I}\right)$

Toutes sont invariantes par transformation galiléennes (translation rotation).
Le fait d'avoir un problème tronqué partiel augmente encore la difficulté de caractérisation des solutions.

Definition (cone convexe)

Un cone convexe est toute partie $\mathcal{C}$ telle que

1. On a $\mathcal{C} + \mathcal{C} \subset \mathcal{C}$
2. on a $\forall \lambda > 0,\ \lambda \mathcal{C} \subset \mathcal{C}$

Pour tout cone, le cone polaire associé (en tant que sous partie du dual) est tel que

$$l \in \mathcal{C}^\circ \Leftrightarrow \forall \mathbf{x} \in \mathcal{C}, \quad l(\mathbf{x}) \leq 0 \tag{9}$$

Dans certains ca particulier, on peut caractériser l'intérieur d'un cone convexe solid (i.e. d'intérieur non vide)

Theorem (intérieur d'un cône solide dans $\mathbb{R}^n$)

Soit $\mathcal{C}$ un cone convexe, solide de $\mathbb{R}^n$. Alors on a

$$\mathbf{x} \in \mathit{int}(\mathcal{C}) \Leftrightarrow \forall \mathbf{y} \in \mathcal{C}^\circ, \quad [\mathbf{y} \neq \mathbf{0} \Rightarrow \mathbf{y} \cdot \mathbf{x} < 0] \tag{10}$$

remarque

Ce théorème de caractérisation de l'intérieur des cônes solides n'est pas si facile à trouver, même si c'est un résultat classique d'analyse convexe.

Definition

Soit $\mathbf{m}(\mathbf{x}) := (\mathrm{h}_0(\mathbf{v}), \cdots, \mathrm{h}_k(\mathbf{v}), \cdots, \mathrm{h}_n(\mathbf{v}))$ une liste de fonctions (organisée en tenseur) définie Ω, muni de la mesure de Lebesgues $d\lambda(\mathbf{v})$. On dit qu'elle est pseudo-Haar lorsque:

$$\forall \boldsymbol{\alpha} = (\alpha_0, \ldots, \alpha_n), \quad [\boldsymbol{\alpha} \neq \mathbf{0} \Rightarrow \boldsymbol{\alpha} \cdot \mathbf{m}(\mathbf{v}) \neq 0], \quad \lambda.a.e\ \mathbf{v} \in \mathbb{R}^d \tag{11}$$

$$\boldsymbol{\alpha} \cdot \mathbf{m}(\mathbf{v}) := \sum_k \alpha_k : \mathrm{h}_k(\mathbf{v}) \tag{12}$$

Example

Les fonctions polynomiales sont pseudo-Haar

Definition

Soit $\mathbf{m}(\mathbf{x})$ pseudo-Haar. On pose $\mathbb{L}^1(\mathbf{m})$ l'ensemble des fonctions mesurables $\mathbb{R}^d$, $f : \mathbb{R}^d \mapsto \mathbb{R}$ telles que:

$$\forall k \in [0, n], \quad \int \|\mathrm{h}_k(\mathbf{v})\| \, |f(\mathbf{v})| \, d\lambda(\mathbf{v}) < +\infty \tag{13}$$

Definition (cone des moments d'une famille pseudo-haar)

Soit $\mathbf{m}(\mathbf{x})$ pseudo-Haar soit q la dimension totale de l'espace généré par $(m_0, \cdots, \mathbf{m}_n)$. Ond édinit l'application de moment $\mathtt{R} : \mathbb{L}^1(\mathbf{m}) \mapsto \mathbb{R}^q$ selon:

$$\forall f \in \mathbb{L}^1(\mathbf{m}), \quad \mathtt{R}[f] = \int \mathbf{m}(\mathbf{x}) f(\mathbf{x}) d\lambda(\mathbf{x}) \tag{14}$$

$$\mathcal{R}_{\mathbf{m}}^{+} = \left\{ \mathtt{R}[f], f \in \mathbb{L}^{1,+}(\mathbf{m}) \right\}, \quad \mathcal{R}_{\mathbf{m}}^{*+} = \left\{ \mathtt{R}[f], f \in \mathbb{L}^{1,*+}(\mathbf{m}) \right\} \tag{15}$$

Definition (cone généré par une famille pseudo-haar)

Soit $\mathbf{m}(\mathbf{x})$ pseudo-Haar on Ω et soit q la dimension de l'espace généré par les $(\mathrm{m}_0(\mathbf{v}), \cdots, \mathrm{m}_n(\mathbf{v}))$. On définit C_Ω son cone associé dans $\mathbb{R}^q$:

$$C_\Omega = \left\{ \sum_i \lambda_i \mathbf{m}(\mathbf{v}_i), \lambda_i \geq 0, \quad \mathbf{v}_i \in \Omega \right\} \tag{16}$$

Theorem (struture du cone $\mathcal{R}$)

Soit $\mathbf{m}(\mathbf{v})$ *pseudo-haar, soit* $\mathtt{R} : \mathbb{L}^1(\mathbf{m}) \mapsto \mathcal{R}$ *l'application de moment et soit* C_Ω *le cône associé à* $\mathbf{m}(\mathbf{v})$. *Alors on a:*

1. $$Im(R) = C_\Omega - C_\Omega = \mathbb{R}^q, \quad \mathcal{R}^{*+}_{\mathbf{m}} = int(C_\Omega) \tag{17}$$

   où $int(\cdot)$ *est l'intérieur topologique*

2. *Deplus, si les* $\mathrm{h}_k(\mathbf{v})$ *sont continues, pour tout élément* $\boldsymbol{\rho}$ *of* $\mathcal{R}^{*+}$ *il existe une fonction positive non nulle* $\phi : \Omega \mapsto \mathbb{R}^+$, *qui est régulière à support compact* $K \subset \Omega$ *telle que:*

   $$\boldsymbol{\rho} = \int_\Omega \mathbf{m}(\mathbf{x})\,\phi(\mathbf{x})\,d\lambda(\mathbf{x}) \tag{18}$$

remarque

Le fait que le cone soit ouvert est lié à deux choses:

- la famille $\mathbf{m}(\mathbf{v})$ est pseudo-haar
- on prend $f \neq 0 \in \mathbb{L}^{1,+,*}(\mathbf{m})$

et donc si l'on a $\boldsymbol{\alpha} \cdot \mathbf{m} \leq 0$ (et donc $\boldsymbol{\alpha} \in C^\circ_\Omega$) on peut montrer que $\boldsymbol{\alpha} \cdot \mathrm{R}_{\mathbf{m}}[f] < 0$ (et on est forcément dans l'intérieur d'un cone)

Remarque

Si C_Ω est le cone associé à la famille pseudo-haar, on a immédiatement $C_\Omega^\circ = \{\boldsymbol{\alpha},\ \forall \mathbf{v}, \boldsymbol{\alpha} \cdot \mathbf{m}(\mathbf{v}) \leq 0\}$. Comme $\mathcal{R}_{\mathbf{m}}^{+,*} = int\,(C_\Omega)$ on a

$$\boldsymbol{\rho} \in \mathcal{R}_{\mathbf{m}}^{+,*} \Leftrightarrow \forall \boldsymbol{\alpha} \in C_\Omega^\circ, \boldsymbol{\alpha} \neq \mathbf{0} \Rightarrow \boldsymbol{\alpha} \cdot \boldsymbol{\rho} < 0 \tag{19}$$

17 ième problème de Hilbert

Les moments réalisables sont caractérisables grâce au cone polaire de C_Ω qui regroupe l'ensemble des fonctions négatives écrites sur la base des $\mathbf{m}(\mathbf{v})$. La description des moments réalisables est donc équivalente à la descriotion des fonctions négatives. Pour $\mathbf{m}(\mathbf{v})$ polynomiale, on retrouve le 17ième problème de Hilbert.

Une des questions importantes du problème de Hilbert est la suivante:

- Tout polynôme qui s'écrit comme somme de carré de polynôme (S.O.S sum of squares) est positif.
- Réciproquement: si un polynôme est positif, s'écrit-il nécessairement comme un polynôme S.O.S

Les résultats connus sur l'équivalence entre polynôme positif et S.O.S ont été donné par Hilbert pour les degrés suivants (d est la dimension de la variable vectorielle et n le degré du polynôme)

1. Si $d = 3$,les polynômes homogènes positifs de degré $n = 4$ sont S.O.S
2. Si $d = 2$ tous les polynômes de degré $n = 4$ positifs sont SOS
3. Si $n = 2$ pour tout $d \in \mathbb{N}$ les polynômes positifs sont SOS
4. Si $d = 1$, tous les polynômes positifs sont SOS

Le problème de caractériser les polynômes positifs via les SOS peut se généraliser sur les structure "quadratiques"

Definition (structure quadratique)

Soit $\mathbf{r}(\mathbf{v}) := (\mathrm{r}_0, \cdots, \mathrm{r}_k)$ une base polynomiale multivariée. La structure quadratique associée est par définition

$$\mathbf{q}(\mathbf{v}) = (\mathrm{r}_0, \cdots, \mathrm{r}_k) \vee (\mathrm{r}_0, \cdots, \mathrm{r}_k) \tag{20}$$

Example

Pour $\mathbf{r}(\mathbf{v}) = (1, \mathbf{v})$ et $\mathbf{q}(\mathbf{v}) = (1, \mathbf{v}, \mathbf{v} \otimes \mathbf{v})$. Pour $\mathbf{r}(\mathbf{v}) = \left(1, \mathbf{v}, \mathbf{v}^2\mathbb{I}\right)$ on a $\mathbf{q}(\mathbf{v}) = \left(1, \mathbf{v}, \mathbf{v} \otimes \mathbf{v}, \mathbf{v}^2\mathbf{v} \vee \mathbb{I}, \mathbf{v}^4\mathbb{I} \vee \mathbb{I}\right)$

Les espaces ayant une structure quadratique sont (pour la théorie cinétique)

1. Espace de Gauss $\mathbf{g}(\mathbf{v}) = (1, \mathbf{v}, \mathbf{v} \otimes \mathbf{v})$
2. Espace de Levermore $\mathbf{l}(\mathbf{v}) = \left(1, \mathbf{v}, \mathbf{v} \otimes \mathbf{v}, \mathbf{v}^2\mathbf{v} \vee \mathbb{I}, \mathbf{v}^4\mathbb{I} \vee \mathbb{I}\right)$

Difficulté a priori pour l'espace de Grad

L'espace de Grad $\mathbf{G}r(\mathbf{v}) = \left(1, \mathbf{v}, \mathbf{v} \otimes \mathbf{v}, \mathbf{v}^2\mathbf{v} \vee \mathbb{I}\right)$ n'a pas de structure quadratique

Construction des matrices de Hankel pour les structures quadratiques

Lorsque la base de polynôme est quadratique, on peut construire la matrice de Hankel associée à la base de la manière suivante:

$$\begin{array}{cccc} \vee & 1 & \mathbf{v} & \mathbf{v}^2\mathbb{I} \\ 1 & 1 & \mathbf{v} & \mathbf{v}^2\mathbb{I} \\ \mathbf{v} & \mathbf{v} & \mathbf{v} \otimes \mathbf{v} & \mathbf{v}^2\mathbf{v} \vee \mathbb{I} \\ \mathbf{v}^2\mathbb{I} & \mathbf{v}^2\mathbb{I} & \mathbf{v}^2\mathbf{v} \vee \mathbb{I} & \mathbf{v}^2\mathbb{I} \vee \mathbb{I} \end{array} \tag{21}$$

Par intégration pour une fonction $f \in \mathbb{L}^{1,+,*}(\mathbf{m})$ on peut alors présenter les moments $\mathrm{R}_{\mathbf{m}}[f]$ dans une matrice de Hankel. De façon analogue un polynôme SOS peut se représenter via la matrice de Hankel

Example (Representation dans l'espace de Gauss)

On a ainsi

1. Pour les polynômes positifs:

$$\pi(\mathbf{v}) = \sum_i \lambda_i (a_i + \mathbf{v}_i \cdot \mathbf{v})^2 = \sum_i \lambda_i [a_i, \mathbf{b}_i] \begin{bmatrix} 1 & \mathbf{v} \\ \mathbf{v} & \mathbf{v} \otimes \mathbf{v} \end{bmatrix} \begin{bmatrix} a_i \\ \mathbf{b}_i \end{bmatrix} \quad (22)$$

2. Pour la représentation des moments $\boldsymbol{\rho} = (\rho, \mathbf{u}, \mathbb{D})$

$$\begin{bmatrix} \rho & \mathbf{u} \\ \mathbf{u} & \mathbb{D} \end{bmatrix} \quad (23)$$

Dans une structure quadratique (type Gauss), puisque tous les polynômes SOS sont positifs, une condition nécessaire pour qu'un moment soit réalisable, c'est que l'on ait, pour tout $[a, \mathbf{b}] \neq [0, \mathbf{0}]$

$$[a, \mathbf{b}] \begin{bmatrix} \rho & \mathbf{u} \\ \mathbf{u} & \mathbb{D} \end{bmatrix} \begin{bmatrix} a \\ \mathbf{b} \end{bmatrix} > 0 \quad (24)$$

autrement dit, la matrice de Hankel doit être définie positive.

condition pour la réciproque

La condition sera suffisante lorsque que tous les polynômes positifs seront SOS. C'est une question difficile.

Definition (Maxwellienne de f)

Pour f avec les propriétés d'intégration adéquates on pose sa Maxwellienne

$$\mathcal{M}\left[f\right](\mathbf{v}) := \frac{n_0}{(2\pi k_b T/m)^{d/2}} \exp\left(-\frac{m\left(\mathbf{v}-\mathbf{u}\right)^2}{2k_b T}\right) \tag{25}$$

où l'on a posé

$$\int \left\{1, \mathbf{v}, \mathbf{v}^2\right\} f\left(\mathbf{v}\right) d\mathbf{v} = \left\{n_0, n_0\mathbf{u}, n_0\mathbf{u}^2 + dn_0\frac{k_b T}{m}\right\} \tag{26}$$

On équippe n'importe quel espace polynomial $\mathbb{P}(\mathbf{v})$ avec le produit scalaire $\mathbb{L}^2\left(\mathcal{M}\left[f\right] d\mathbf{v}\right)$.

Dans le contexte de la théorie cinétique, les espaces les plus importants sont:

1. L'espace des quantités conservées: $\mathbb{E}(\mathbf{v}) = \mathrm{span}\left\{1, \mathbf{v}, \mathbf{v}^2\right\}$
2. L'espace total des moments conervés et des moments transportés: $\mathbb{G}(\mathbf{v}) = (1+\mathbf{v}) \vee \mathbb{E}(\mathbf{v})$ qui est exactement l'esapce de Grad:

$$\mathbb{G}(\mathbf{v}) = \mathrm{span}\left\{1, \mathbf{v}, \mathbf{v}\otimes\mathbf{v}, \mathbf{v}^2\mathbf{v}\right\} \tag{27}$$

On utilise la décomposition orthogonale:

$$\mathbb{G}(\mathbf{v}) = \mathbb{E}(\mathbf{v}) \oplus^{\perp} \left[\mathbb{A}(\mathbf{v}-\mathbf{u}) \oplus^{\perp} \mathbf{b}(\mathbf{v}-\mathbf{u})\right], \tag{28}$$

$$\mathbb{A}(\mathbf{v}-\mathbf{u}) = (\mathbf{v}-\mathbf{u}) \otimes (\mathbf{v}-\mathbf{u}) - \frac{1}{d}(\mathbf{v}-\mathbf{u})^2 \mathbb{I}, \tag{29}$$

$$\mathbf{b}(\mathbf{v}-\mathbf{u}) = (\mathbf{v}-\mathbf{u}) \left((\mathbf{v}-\mathbf{u})^2 - (d+2)\frac{k_b T}{m}\right). \tag{30}$$

L'espace

$$\mathbb{D}(\mathbf{v}) := \mathbb{G}(\mathbf{v}) \ominus^{\perp} \mathbb{E}(\mathbf{v}) = \mathbb{A}(\mathbf{v}-\mathbf{u}) \oplus^{\perp} \mathbf{b}(\mathbf{v}-\mathbf{u}) := \mathbb{D}_1 \oplus^{\perp} \mathbb{D}_2 \tag{31}$$

est l'espace de diffusion et chaque sous espace de ce dernier, obtenu comme représentation irréductible pour les rotations, representation définit un moment diffus à part entière

base de diffusion

1. Le tenseur $\mathbb{A}(\mathbf{v}-\mathbf{u})$ est la base polynomiale de diffusion du moment.
2. Le tenseur $\mathbf{b}(\mathbf{v}-\mathbf{u})$ est la base polynomiale de diffusion de l'énergie.

Pour la suite, on aura donc à considérer la famille pseudo-haar:

$$\begin{aligned}\mathbf{a}(\mathbf{v}-\mathbf{u}) &= \left(1, (\mathbf{v}-\mathbf{u}), (\mathbf{v}-\mathbf{u})^2 - d\frac{k_B T}{m}, \mathbb{A}(\mathbf{v}-\mathbf{u}), \mathbf{b}(\mathbf{v}-\mathbf{u})\right) \\ &:= \left(a_0(\mathbf{v}-\mathbf{u}), \mathbf{a}_1(\mathbf{v}-\mathbf{u}), a_2(\mathbf{v}-\mathbf{u}), \mathbb{A}_3(\mathbf{v}-\mathbf{u}), \mathbf{a}_4(\mathbf{v}-\mathbf{u})\right)\end{aligned}$$

Et on note par $\mathcal{R}_{\mathbf{a}}^{+*}$ l'ensemble des moments réalisables sur cette base. Le cone polaire est donné selon:

$$\mathcal{C}_{\mathbf{a}}^{\circ} = \left\{\boldsymbol{\alpha} \in \mathbb{R}^q, \ \forall \mathbf{v} \in \mathbb{R}^d, \ \boldsymbol{\alpha} \cdot \mathbf{a}(\mathbf{v}-\mathbf{u}) \leq 0\right\} \tag{32}$$

Il est immédiat que $\boldsymbol{\alpha} \in \mathcal{C}_{\mathbf{a}}^{\circ}$ a une composante nulle sur $(\mathbf{v}-\mathbf{u})(\mathbf{v}-\mathbf{u})^2$ et par conséquent sur $\mathbf{b}(\mathbf{v}-\mathbf{u})$. On a donc:

$$\mathcal{C}_{\mathbf{a}}^{\circ} \left\{(\beta, \mathbf{0}), \forall \mathbf{v} \in \mathbb{R}^d, \ \beta \cdot \left(1, (\mathbf{v}-\mathbf{u}), (\mathbf{v}-\mathbf{u})^2 - d\frac{k_B T}{m}, \mathbb{A}_3(\mathbf{v}-\mathbf{u})\right) \leq 0\right\} \tag{33}$$

caractérisation de la réalisabilité de Grad

La réalisabilité sur la base de Grad est indépendante du moment d'ordre 3 sur la diffusion de la chaleur. Elle ne dépend que de la réalisabilité sur la base de Gauss.

Theorem

Soit $\boldsymbol{\rho} := (n, n\mathbf{U}, dne, n\mathbb{A}, n\mathbf{b})$ *associé à la base* $\mathbf{a}(\mathbf{v}-\mathbf{u})$*:*

$$\mathbf{a}(\mathbf{v}-\mathbf{u}) := \left(1, (\mathbf{v}-\mathbf{u}), (\mathbf{v}-\mathbf{u})^2 - d\frac{k_B T}{m}, \mathbb{A}(\mathbf{v}-\mathbf{u}), \mathbf{b}(\mathbf{v}-\mathbf{u})\right) \tag{34}$$

Alors ce moment est réalisable si et seulement si:

$$n > 0, \quad \mathbb{A} + \left(e + \frac{k_B T}{m}\right)\mathbb{I} - \mathbf{U} \otimes \mathbf{U} \ \text{ positive} \tag{35}$$

On en tire le théorème de relaxation suivant

Theorem (relaxation de Grad)

Soit $\boldsymbol{\rho} := (n, n\mathbf{U}, dne, n\mathbb{A}, n\mathbf{b})$ *un moment calculé sur la base* $\mathbf{a}(\mathbf{v}-\mathbf{u})$.

1. *Alors les moments* $\{\mathbf{U}, \mathbf{b}\}$ *peuvent être mis à* 0 *en gardant la réalizabilité*
2. *Pour tout* $\lambda_u, \lambda_a, \lambda_b$ *tel que* $0 \leq \lambda_u, \lambda_a, \lambda_b \leq 1$ *et* $\lambda_u^2 \leq \lambda_a$ *alors le moment* $(n, n\lambda_u\mathbf{U}, dne, n\lambda_a\mathbb{A}, n\lambda_b\mathbf{b})$ *est réalisable*

En particulier si $(n, \mathbf{0}, dne, n\mathbb{A}, n\mathbf{b})$ *est réalisable, alors pour tout* $\lambda_a, \lambda_b \in [0, 1]$ *le moment relaxé* $(n, \mathbf{0}, dne, n\lambda_a\mathbb{A}, n\lambda_b\mathbf{b})$ *est encore réalisable*

Problem (entropy minimization primal problem)

For the moment, we skip technical details in the presentation of the problem, including the the proofs that all definitions are well posed etc...

- *Consider a prescribed Maxwellian function* $\mathcal{M}(\mathbf{v})$*, as well as a realizable moment* $\boldsymbol{\rho}$ *against the pseudo-haar basis* $\mathbf{a}(\mathbf{v})$ *which contains at least* $(1, \mathbf{v}, \mathbf{v}^2)$*. Let us note by* $\mathcal{R}_{\mathbf{a}}^{+*}$ *the set of realizable moments.*
- *For a suitable convex function* $\phi : \mathbb{R} \mapsto \mathbb{R} \bigcup \{+\infty\}$ *(**this is exactly the*** ϕ***-divergence framework**), with the essential property that* $\mathtt{dom}(\phi) \subset [0, +\infty)$ *(that is* $\phi(x) = +\infty$ *for any negative* x*), define then the following function as an element of the extended reals* $\mathbb{R} \bigcup \{+\infty\}$*:*

$$\forall g,\ g \in \mathbb{L}^1(\mathbf{a}),\ \ \mathcal{H}(g) := \int \phi\left(\frac{g(\mathbf{v})}{\mathcal{M}(\mathbf{v})}\right) \mathcal{M}(\mathbf{v})\, d\lambda(\mathbf{v}) \qquad (36)$$

*the **primal** entropy minimization problem (for the couple* $\mathcal{M}, \boldsymbol{\rho}$*) is then stated as follows: find if possible, both the existence of a real minimum and a minimizer (*$G[\boldsymbol{\rho}]$*) for* $\mathcal{H}(\cdot)$ *under the constraints of the moment* $\boldsymbol{\rho} \in \mathcal{R}_{\mathbf{a}}^{*+}$ *that is:*

$$\int \mathbf{a}(\mathbf{v})\, g(\mathbf{v})\, d\lambda(\mathbf{v}) = \boldsymbol{\rho} \Leftrightarrow g \in D(\boldsymbol{\rho}) \qquad (37)$$

Definition (ϕ-divergence functions)

In this article, we will name as ϕ-diverge function any function ϕ such that:

1. the function $\phi : \mathbb{R} \mapsto \mathbb{R} \bigcup \{+\infty\}$ is strictly convex on its domain $\mathtt{dom}\,(\phi) = [0, +\infty)$ and differentiable on $[0, +\infty)$
2. there holds the following properties:

$$\phi\,(0) = 0, \quad p_0 := \inf_{y>0} \frac{\phi\,(y)}{y} \in \mathbb{R}, \quad \sup_{y>0} \frac{\phi\,(y)}{y} = +\infty \tag{38}$$

3. For any polynomial $\pi\,(\mathbf{v}) = \gamma \cdot \mathbf{a}\,(\mathbf{v})$ then $\phi^*\,(\pi) \in \mathbb{L}^1\,(\mathcal{M} d\lambda\,(\mathbf{v}))$ where ϕ^* is the Legendre transform of ϕ

Theorem

for any N the following family of functions

$$\forall x < 0, \quad \phi_N\,(x) = +\infty, \quad \forall x \geq 0, \quad \phi_N\,(x) = N\left(x^{1+1/N} - x\right) \tag{39}$$

is a ϕ-divergence function. Moreover, there holds respectively

$$p_0\,(N) = -N, \; \forall y \geq p_0\,(N)\,, \; \phi_N^*\,(y) = \left(1 + \frac{y-1}{N+1}\right)^{N+1} \tag{40}$$

It is really worth noting the following point wise convergence aspect:

pointwise limit

$$\forall x \geq 0, \quad \lim_N \phi_N(x) = x \ln(x), \quad \forall y \in \mathbb{R}, \quad \lim_N \phi_N^*(y) = \exp(y-1) \tag{41}$$

that is, the entropy ϕ_N is a point wise approximation of the Boltzmann entropy density $x \ln(x)$ and ϕ_N^* is an approximation of $\exp(y-1)$ which is the Legendre transform of $x \ln(x)$.

Moreover we have the important limits:

derivatives pointwise limit

$$\forall x \in \mathbb{R}^{*+}, \quad \partial_x \phi_N(x) = (N+1)\, x^{1/N} - N, \tag{42}$$

$$\forall y \geq -N, \quad \partial_y \phi_N^*(y) = \left(1 + \frac{y-1}{N+1}\right)^N \tag{43}$$

$$\forall x \in \mathbb{R}^{*+}, \quad \lim_N \partial_x \phi_N(x) = \ln(x) + 1, \tag{44}$$

$$\forall y \in \mathbb{R} \quad \lim_N \partial_y \phi_N^*(y) = \exp(y-1) \tag{45}$$

Theorem (functional entropy definition)

We can define the function $\mathcal{H} : \mathbb{L}^1(\mathbf{a}) \mapsto \mathbb{R} \bigcup \{+\infty\}$ *by the formula:*

$$\forall g \in \mathbb{L}^1(\mathbf{a}), \quad \mathcal{H}(g) := \int \phi\left(\frac{g}{\mathcal{M}}\right) \mathcal{M} d\lambda(\mathbf{v}) \tag{46}$$

More over this function is strictly convex.

Note that the functional entropy is still not the moment entropy. To define properly the moment entropy, for any set D consider the so called indicator of D, noted as $i_D(\cdot) \mapsto \{0, +\infty\}$ such that

1. there is $i_D(X) = 0$ if $X \in D$
2. there is $i_D(X) = +\infty$ if $X \notin D$

it is well known that i_D is convex if and only if D is convex and i_D is semi-lower continuous if and only if D is (weakly) closed.

Definition (moment entropy definition)

Let $\boldsymbol{\rho} \in \mathbb{R}^q$ and consider the closed convex domain $D(\boldsymbol{\rho}) = \left\{g \in \mathbb{L}^1(\mathbf{a}),\ \int \mathbf{a}g = \boldsymbol{\rho}\right\}$. Then we define the real extended (that is $\overline{\mathbb{R}}$) valued function for any $\boldsymbol{\rho} \in \mathbb{R}^q$ by:

$$h_{\mathbf{a}}(\boldsymbol{\rho}) := \inf_{g \in \mathbb{L}^1(\mathbf{a})} \left[\mathcal{H}(g) + i_{D(\boldsymbol{\rho})}(g)\right] = \inf_{g \in \mathbb{L}^1(\mathbf{a})} \left[\mathcal{H}(g) + i_{\{\mathbf{0}\}}\left(\int \mathbf{a}g - \boldsymbol{\rho}\right)\right] \quad (47)$$

One can preice the convex attributes of the moment etnropy function $h_{\mathbf{a}}$

Theorem (moment entropy theorem)

The following properties are satisfied

1. *the moment entropy function $h_{\mathbf{a}}(\boldsymbol{\rho})$ has returned values in $\mathbb{R} \bigcup \{+\infty\}$*
2. *its domain is exactly $\mathtt{dom}(h_{\mathbf{a}}) = \mathcal{R}_{\mathbf{a}}^{+} = \left\{\int \mathbf{a}g, g \in \mathbb{L}^{1+}(\mathbf{a})\right\}$.*
3. *the interior of the domain is given by $\mathtt{int}(\mathtt{dom}(h_{\mathbf{a}})) = \mathcal{R}_{\mathbf{a}}^{+*} = \mathtt{dom}(h_{\mathbf{a}}) \setminus \{0\}$*
4. *the function $h_{\mathbf{a}}$ is convex.*

Problem (entropy minimization dual problem)

The dual problem consists in define the Legendre dual function h^* from $\mathbb{R}^q$ to the extended reals $\overline{\mathbb{R}}$ - that is $\mathbb{R} \bigcup \{+\infty, -\infty\}$- as follows:

$$\forall \boldsymbol{\alpha} \in \mathbb{R}^q, \quad h^*(\boldsymbol{\alpha}) = \sup_{\boldsymbol{\rho} \in \mathbb{R}^q} \left(\boldsymbol{\rho} \cdot \boldsymbol{\alpha} - h_{\mathbf{a}}(\boldsymbol{\rho})\right) \tag{48}$$

Theorem

the following properties are satisfied for the dual function:

1. *For any $\boldsymbol{\alpha} \in \mathbb{R}^q$ we have*

$$h_{\mathbf{a}}^*(\boldsymbol{\alpha}) = \sup_{g \in \mathbb{L}^1(\mathbf{a})} \int \left[\boldsymbol{\alpha} \cdot \mathbf{a}(\mathbf{v})\, g(\mathbf{v}) - \phi\left(\frac{g}{\mathcal{M}}\right) \mathcal{M}(\mathbf{v})\right] d\mathbf{v} \in \mathbb{R} \tag{49}$$

2. *Moreover one can also compute for any $\boldsymbol{\alpha} \in \mathbb{R}^q$*

$$h_{\mathbf{a}}^*(\boldsymbol{\alpha}) = \int \phi^*(\boldsymbol{\alpha} \cdot \mathbf{a}(\mathbf{v}))\, \mathcal{M}(\mathbf{v})\, d\mathbf{v} \tag{50}$$

3. *The function $h_{\mathbf{a}}^*$ is continuously differentiable on $\mathbb{R}^q$ and there holds*

$$h_{\mathbf{a}}^{*'}(\boldsymbol{\alpha}) = \int \phi^{*'}(\boldsymbol{\alpha} \cdot \mathbf{a}(\mathbf{v}))\, \mathbf{a}(\mathbf{v})\, \mathcal{M}(\mathbf{v})\, d\mathbf{v} \tag{51}$$

Theorem (well posedness of the primal problem)

The following asseertions are true:

1. *For any* $\boldsymbol{\rho} \in \mathcal{R}_{\mathbf{a}}^{+*}$ *there exists a unique* $\boldsymbol{\alpha} \in \mathbb{R}^q$ *such that*

$$\boldsymbol{\rho} = \int \phi^{*'} \left(\boldsymbol{\alpha} \cdot \mathbf{a}\left(\mathbf{v}\right)\right) \mathbf{a}\left(\mathbf{v}\right) \mathcal{M} d\lambda \left(\mathbf{v}\right) \tag{52}$$

   and the moments $\boldsymbol{\rho}$ *and its conjugate moment* $\boldsymbol{\alpha}$ *are linked thanks to the sub-differential equation:*

$$h_{\mathbf{a}}\left(\boldsymbol{\rho}\right) + h_{\mathbf{a}}^{*}\left(\boldsymbol{\alpha}\right) = \boldsymbol{\alpha} \cdot \boldsymbol{\rho} \tag{53}$$

2. *Moreover, for any* $\boldsymbol{\alpha}$, *the function* $G = \mathcal{M} \phi^{*'} \left(\boldsymbol{\alpha} \cdot \mathbf{a}\left(\mathbf{v}\right)\right)$ *such that*

$$\int \mathbf{a} G d\lambda \left(\mathbf{v}\right) = \boldsymbol{\rho} \tag{54}$$

   is the unique minimizer of the primal problem and satisfies

$$h_{\mathbf{a}}\left(\boldsymbol{\rho}\right) = \int \phi \left(\frac{G}{\mathcal{M}}\right) \mathcal{M} d\lambda \left(\mathbf{v}\right) \tag{55}$$

Theorem (entropy majoration)

Let $\boldsymbol{\rho} := (\rho_1, \mathbf{U}, \rho_2, \mathbb{A}, \mathbf{b})$ *be a realizable moment against the basis* $\mathbf{a}(\mathbf{v}-\mathbf{u})$*. Then we have:*

$$h_{\mathbf{a}}(\rho_0, \mathbf{0}, \rho_2, \mathbb{A}, \mathbf{0}) \leq h_{\mathbf{a}}(\boldsymbol{\rho}) \tag{56}$$

$$h_{\mathbf{a}}(\rho_0, \mathbf{0}, \rho_2, \mathbb{O}, \mathbf{0}) \leq h_{\mathbf{a}}(\boldsymbol{\rho}) \tag{57}$$

In particular, for any $\lambda_a \in [0,1]$, $\lambda_b \in [0,1]$*, in the special case when* $\mathbf{U} = \mathbf{0}$*, we have*

$$h_{\mathbf{a}}(\rho_0, \mathbf{0}, \rho_2, \lambda_a\mathbb{A}, \lambda_b\mathbf{b}) \leq h_{\mathbf{a}}(\rho_0, \mathbf{0}, \rho_2, \mathbb{A}, \mathbf{b}) \tag{58}$$

Finally, assume that $\lambda_a \in [0,1)$ *and* $\lambda_b \in [0,1)$*. Then the following relation holds*

$$[h_{\mathbf{a}}(\rho_0, \mathbf{0}, \rho_2, \lambda_a\mathbb{A}, \lambda_b\mathbf{b}) = h_{\mathbf{a}}(\rho_0, \mathbf{0}, \rho_2, \mathbb{A}, \mathbf{b})] \Rightarrow [\mathbb{A} = \mathbb{O} \;\; \text{and} \;\; \mathbf{b} = \mathbf{0}] \tag{59}$$

Now we are going to construct a BGK operator using the former results that is:

- The results on moment relaxation
- The results on ϕ-divergence solution constrained by realizable moments

We are going to give the receipt to build the BGK operator. Let consider $f \geq 0$. Then

BGK construction method

- Compute its associated Maxwellian $\mathcal{M}_f$. Note by n the density and by $\mathbf{u}$ its velocity and T its temperature and form the basis $\mathbf{a}(\mathbf{v}-\mathbf{u})$.
- Compute for f its moment against the basis $\mathbf{a}(\mathbf{v}-\mathbf{u})$, that is:

$$\boldsymbol{\rho}_f = (n, \mathbf{0}, 0, n\mathbb{A}, n\mathbf{b}) = \int \left(1, (\mathbf{v}-\mathbf{u}), (\mathbf{v}-\mathbf{u})^2 - d\frac{k_B T}{m}, \mathbb{A}(\mathbf{v}-\mathbf{u}), \mathbf{b}(\mathbf{v}-\mathbf{u})\right) f(\mathbf{v})\, d\lambda(\mathbf{v}) \tag{60}$$

BGK construction method (continued)

- Choose the following relaxation times

$$\nu_A = \frac{nk_B T}{\mu}, \nu_b = \frac{5}{2}\frac{nk_B T}{m\kappa} \tag{61}$$

where μ is the viscosity and κ the Fourrier conductance. Take any ν such that $\nu > \nu_A, \nu_B$

- Consider the relaxed moment:

$$\mathtt{L}\left(\boldsymbol{\rho}_f\right) = \left(n, \mathbf{0}, 0, \left(1-\alpha_a\right) n\mathbb{A}, n\left(1-\alpha_b\right)\mathbf{b}\right), \quad \alpha_a = \frac{\nu_a}{\nu}, \quad \alpha_b = \frac{\nu_b}{\nu} \tag{62}$$

we know from the grad relaxation theorem that it is still realizable.

- Consider now the solution of the ϕ-divergence optimization program (using ϕ_N and $\mathcal{M}_f$) under the constraint $\mathtt{L}\left(\boldsymbol{\rho}_f\right)$. This function is G_f
- Your BGK operator reads as $K\left(f\right) = \nu\left(G_f - f\right)$

Remarque

The modelling function G_f reads as

$$G_f = \mathcal{M}_f \phi^{*'} (\boldsymbol{\alpha} \cdot \mathbf{a}(\mathbf{v})) \tag{63}$$

$$\forall y \geq -N,\ \phi_N^*(y) = \left(1 + \frac{y-1}{N+1}\right)^{N+1} \tag{64}$$

$$\forall y < -N,\ \phi_N^*(y) = 0 \tag{65}$$

Theorem (entropy)

There holds the following inequality

$$\forall f \geq 0 \in \mathbb{L}^{1,+,*}(\mathbf{a}),\ \left\langle K_N(f)\, \partial_x \phi_N \left(\frac{f}{\mathcal{M}_f}\right)\right\rangle \leq 0 \tag{66}$$

Theorem (Extended H-theorem)

The following equivalences hold

$$K(f) = 0 \Leftrightarrow \left\langle K_N(f)\, \phi_N \left(\frac{f}{\mathcal{M}_f}\right)\right\rangle = 0 \Leftrightarrow f \sim \mathcal{M}_f \tag{67}$$

Theorem (linearized operator)

for any Maxwellian $\mathcal{M}$, the linearized operator:

$$\mathcal{L}_K(g) := \lim_{\epsilon\to 0} \frac{1}{\epsilon\mathcal{M}} K(\mathcal{M}(1+\epsilon g)) \tag{68}$$

which is defined for $g \in \mathcal{L}^2(\mathcal{M} d\lambda(\mathbf{v}))$ reads as:

$$\mathcal{L}_K(g) = \nu\left(\mathcal{P}_{\mathbb{E}} + (1-\alpha_a)\mathcal{P}_{\mathbb{A}} + (1-\alpha_b)\mathcal{P}_{\mathbf{b}} - \mathcal{I}\right) \tag{69}$$

where for any $\mathbb{D}$ the operator $\mathcal{P}_{\mathbb{D}}$ is the orthogonal projection. In particular

1. *The kernel of the operator $\mathcal{L}_{\mathcal{M}}$ is exactly $\mathbb{E} = span(1, \mathbf{v}, \mathbf{v}^2)$ and as a consequence we have also:*

$$\forall f, \; \left[\int K(f)\phi = 0\right] \Leftrightarrow \phi \in span\left(1, \mathbf{v}, \mathbf{v}^2\right) \tag{70}$$

2. *The operator is Fredholm, self-adjoin and negative on $\mathbb{E}^{\perp}$*
3. *there holds for the transport coefficients:*

$$\mu_K = \frac{2}{(d+2)(d-1)} \int \mathbb{A}(\mathbf{v}-\mathbf{u}) : \mathcal{L}_K^{-1}(\mathbb{A}(\mathbf{v}-\mathbf{u}))\mathcal{M} d\lambda(\mathbf{v}) = \mu \tag{71}$$

$$\kappa_K = \frac{1}{d} \int \mathbf{b}(\mathbf{v}-\mathbf{u}) : \mathcal{L}_K^{-1}(\mathbf{b}(\mathbf{v}-\mathbf{u}))\mathcal{M}(\mathbf{v}) d\lambda(\mathbf{v}) = \kappa \tag{72}$$

Conclusions

1. Study of partial troncated multivariate moment problem for diffusive phenomena enables to derive relaxation map which are able to preserve realizability.
2. Thanks to the ϕ-divergence framework, the primal optimization problem has always solution and the dual problem enables to gets a priori expression of the minimizer.
3. Using minimizer enables to preserve positivity of the Modelling function (hence BGK are al to preserve positivity of the pdf f and to have entropy structure for the corresponding BGK.
4. Using relaxation enables to control the linearized BGK and then to maintain the correctness of diffusive transport (viscosity, Fourrier)
5. Hope the method is able to work in the context of gas mixture that is: getting entropic BGK, preserving positivity of pdf while having the correct transport coefficients (Fick, Soret, Newton, Duffour, Fourrier)