Kinetic models of Maxwell type. A brief history Part I

Giuseppe Toscani

Department of Mathematics

University of Pavia, Italy

Porto Ercole, June 8-10 2008

Summer School METHODS AND MODELS OF KINETIC THEORY

Outline

Maxwell molecules

- The spatially homogeneous Boltzmann equation for Maxwellian molecules

$$\frac{\partial f}{\partial t} = \int_{\mathbf{R}^n \times S^{n-1}} B\left(\frac{(v-w)\cdot n}{|v-w|}\right)\left(f(v^*)f(w^*) - f(v)f(w)\right) dw\, dn.$$

- Post collision velocities

$$v^* = \frac{1}{2}\left(v + w + |v-w|n\right), \qquad w^* = \frac{1}{2}\left(v + w - |v-w|n\right).$$

- The function $B(\nu)$, when $n = 3$, has a singularity of the form $(1-\nu)^{-5/4}$ as $\nu \to 1$.

Maxwell pseudomolecules

- Maxwell pseudomolecules: The function $\sigma(\nu)$ is summable in $(-1,1)$. Usually obtained by an angular *cut-off*.
- Set

$$\int_{S^{n-1}} B\left(\frac{(v-w)\cdot n}{|v-w|}\right) dw\, dn = 1.$$

- The collision operator simplifies

$$Q(f) = f \circ f(v) - \rho f(v), \quad \rho = \int_{\boldsymbol{R}^n} f(w)\, dw.$$

- $f \circ g(v)$ is known as Wild convolution [E. Wild *Proc. Camb. Phyl. Soc.* (1951)] .

$$f \circ g(v) = \int_{\boldsymbol{R}^n \times S^{n-1}} B\left(\frac{(v-w)\cdot n}{|v-w|}\right) f(v^*) g(w^*)\, dw\, dn.$$

The equilibrium

- The equality

$$\int_{\mathbf{R}^{2n}\times S^{n-1}} f(v)f(w)(\log f(v^*) - \log f(v))\, dv\, dw\, dn = 0.$$

implies $\log f(v)$ belongs to the space of collision invariants.

- The function

$$M_{\rho,m,T}(v) = \rho \left(\frac{1}{2\pi T}\right)^{n/2} \exp\left\{-\frac{(v-m)^2}{2T}\right\}$$

is the equilibrium Maxwellian with mass ρ, momentum m and temperature T.

Wild sum

- Let us consider a density of unit mass. Wild observed that the initial value problem associated with

$$\frac{\partial f}{\partial t} = f \circ f(v) - f(v), \quad f(\cdot, 0) = F(\cdot)$$

can be written as a fixed point problem
[E. Wild *Proc. Camb. Phyl. Soc.* (1951)].

- Define the map $f \mapsto \Phi(f)$

$$\Phi(f)(t) = e^{-t}F + \int_0^t e^{-(t-s)} f \circ f(s)\, ds.$$

- $f(t)$ solves the Boltzmann equation exactly when $\Phi(f) = f$.

Wild sum II

- To find fixed points one iterates. Put $f^{(0)} = 0$, and define

$$f^{(j+1)} = \Phi\left(f^{(j)}\right) \qquad \text{for all } j \geq 1$$

- This yields

$$f^{(1)} = e^{-t}F$$

$$f^{(2)} = e^{-t}F + e^{-t}\left(1 - e^{-t}\right)F \circ F$$

$$f^{(3)} = e^{-t}F + e^{-t}\left(1 - e^{-t}\right)F \circ F$$

$$+e^{-t}\left(1 - e^{-t}\right)^2\left(\frac{1}{2}F \circ (F \circ F) + \frac{1}{2}(F \circ F) \circ F\right),$$

and so on.

Wild sum III

- As one sees,

$$f^{(j+1)} - f^{(j)} \geq 0 \qquad \text{for all } j \geq 0.$$

- Hence

$$\lim_{j \to \infty} f^{(j)}(t) = f(t)$$

exists, and, as Wild showed, is a solution to the Boltzmann equation.

- Uniqueness proven by D. Morgenstern
[D. Morgenstern *Proc. Nat. Acad. Sci. USA*(1953)].

Convolution iterates

- Consider again the initial value problem

$$\frac{\partial f}{\partial t} = f \circ f(v) - f(v), \quad f(\cdot, 0) = F(\cdot)$$

- Replace the time t and the density f

$$\tau = 1 - e^{-t}, \qquad g(v, \tau) = f(v, t)e^{t}.$$

- g satisfies

$$\frac{\partial g}{\partial \tau} = g \circ g(v), \quad g(\cdot, 0) = F(\cdot)$$

Convolution iterates II

- $\tau \leq 1$. Expand the solution in power series of τ

$$g(v,\tau) = \sum_{k=0}^{\infty} \tau^k f_{(k)}(v), \qquad f_{(0)} = F(v)$$

- The functions $f_{(k)}$ are given by recursion

$$f_{(k+1)}(v) = \frac{1}{k+1} \sum_{i=1}^{k} f_{(i)} \circ f_{(k-i)}, \qquad i \geq 1.$$

- The solution $f(t)$ represented by

$$f(v,t) = e^{-t} \sum_{k=0}^{\infty} (1-e^{-t})^k f_{(k)}(v)$$

Convolution iterates III

- Consider the initial value problem

$$\frac{\partial f}{\partial t} = P_n(f)(v) - f(v), \quad f(\cdot, 0) = F(\cdot)$$

- P is a n-linear operator from the Banach space $B^n \to B$ such that

$$|P(g_1, \dots, g_n)| \leq |g_1| \cdots |g_n|, \qquad g_i \in B.$$

- Then, the unique solution $f(t)$ can be represented (in its interval of existence) by

$$f(v, t) = e^{-t} \sum_{k=0}^{\infty} (1 - e^{-(n-1)t})^k h_{(k)}(v)$$

Convolution iterates III

- for $k \geq 1$

$$h_{(k+1)}(v) = \frac{1}{k+1} \sum_{i_1+\cdots i_n=k} \frac{b_{i_1}\cdots b_{i_n}}{(n-1)b_k} P_n(f_{(i_1)}, \cdots, f_{(i_n)},$$

- The numbers b_k are the coefficients of the Taylor expansion

$$(1-x)^{1/(1-n)} = \sum_{k=0}^{\infty} b_k x^k$$

- Result by Kielek [Z. Kielek, *Acta Jagel.*(1988)]

Convexity and Wild sums

- Suppose for a given convex functional $\Gamma(f)$

$$\Gamma\,(f \circ g) \leq a^2\Gamma(f) + b^2\Gamma(g), \qquad a^2 + b^2 = 1$$

- Then, $\Gamma(\cdot)$ is a Liapunov functional for the Maxwell-Boltzmann equation.
- Jensen's inequality and Wild formula for the solution give the result.
- Functionals super-additive with respect to convolutions are good candidates.

The central limit theorem

- Deep analogies between the large-time behavior of the Boltzmann equation and the central limit theorem.
- The central limit theorem consists of finding the limit distribution M such that, when dealing with the normalized sum

$$S_n = \frac{X_1 + \ldots + X_n}{n^{1/2}}$$

of independent and identically distributed random variables X_i with common distribution function F,

- the distribution F_n of the sum S_n converges to M.

The central limit theorem II

- Let $f(x), x \in \boldsymbol{R}$, the probability density function of the random variables X_i,
- Assume f is normalized (of zero mean and unit variance).
- Denote $f_n(x)$ the density of S_{2^n}, $n \geq 1$.
- Then

$$S_{2^{n+1}} = \frac{1}{2^{1/2}} S_{2^n} + \frac{1}{2^{1/2}} S^*_{2^n},$$

S_{2^n} and $S^*_{2^n}$ are independent and identically distributed.

The central limit theorem III

- Then

$$f_{n+1}(x) = \int_{\boldsymbol{R}} 2^{1/2} f_n\left(2^{1/2}(x-y)\right) 2^{1/2} f_n\left(2^{1/2} y\right)\, dy\,.$$

- Change variable into the integral $x - y = \frac{1}{2}(x+z)$ (which implies $y = \frac{1}{2}(x-z)$).
- Obtain

$$f_{n+1}(x) = \int_{\boldsymbol{R}} f_n\left(\frac{x+z}{2^{1/2}}\right) f_n\left(\frac{x-z}{2^{1/2}}\right)\, dz.$$

The central limit theorem IV

- $f_n(x)$ has unit mass,

$$f_{n+1}(x) = f_n(x) + \int_{R} \left\{ f_n\left(\frac{x+z}{2^{1/2}}\right) f_n\left(\frac{x-z}{2^{1/2}}\right) - f_n(x)f_n(z) \right\} dz$$

- Explicit Euler scheme (at discrete times $\Delta t = 1$) of the kinetic equation

$$\frac{\partial f}{\partial t} = Q_s^*(f, f)$$

- The post-collision velocities $x^* = \frac{1}{\sqrt{2}}(x+z)$, $z^* = \frac{1}{\sqrt{2}}(x-z)$ satisfy

$$(x^*)^2 + (z^*)^2 = x^2 + z^2.$$

Kac equation

- Kac's caricature of a Maxwellian gas is a model for the motion of a molecule in a chaotic bath of molecules moving on a line
[M. Kac *Probability and related topics in Physical sciences* (1959)]

$$\frac{\partial f}{\partial t} = \int_{\mathbf{R}\times[0,2\pi]} \frac{d\theta}{2\pi} \left(f(v^*)f(w^*) - f(v)f(w)\right)\, dw\,.$$

- The post collisional velocities (v^*, w^*) are defined by a *rotation*

$$v^* = v\cos\theta - w\sin\theta, \qquad w^* = v\sin\theta + w\cos\theta.$$

- Wild convolution for Kac model

$$f \circ g(v) = \frac{1}{2\pi}\int_{\mathbf{R}\times[0,2\pi]} f(v\cos\theta - w\sin\theta)g(v\sin\theta + w\cos\theta)\, dw\, d\theta.$$

Other models

- Introduced by Kac to justify the propagation of chaos [M. Kac *Probability and related topics in Physical sciences* (1959)].
- Kac model with a singular collision kernel was introduced by Desvillettes [Desvillettes *Comm. Math. Phys.* (1995)]

$$\frac{1}{2\pi} \to \beta(\theta), \qquad \beta(\theta) \cong \theta^{-1+\mu}, \quad \theta \ll 1, \quad 0 < \mu < 2$$

- The model is the analogous of the true Maxwell molecules.

$$\frac{\partial f}{\partial t} = \int_{\mathbf{R}\times[0,2\pi]} \beta(\theta)\left(f(v^*)f(w^*) - f(v)f(w)\right) dw\, d\theta.$$

- A singular kernel induces regularity!

Dissipative Kac model

- A dissipative Kac model was introduced in [A. Pulvirenti, G. Toscani, *J. Stat. Phys* (2003)]

$$\frac{\partial f}{\partial t} = \int_{\mathbf{R}\times[0,2\pi]} \frac{d\theta}{2\pi}\, dw \left[\chi^{-1} f(v_p^{**}) f(w_p^{**}) - f(v)f(w)\right] .$$

- Kac equation contains the pre-collisional velocities (v_p^{**}, w_p^{**}), and the Jacobian χ.
- The post-collisional velocities (v_p^*, w_p^*) are defined generalizing the Kac rule

$$v_p^* = v\cos\theta|\cos\theta|^p - w\sin\theta|\sin\theta|^p, \quad w_p^* = v\sin\theta|\sin\theta|^p + v\cos\theta|\cos\theta|^p.$$

- $0 < p < +\infty$ measures the degree of inelasticity.

Approach to equilibrium

- In 1965, McKean investigated the speed of approach to equilibrium for Kac model
  [H.P. McKean *Arch. Rat. Mech. Anal.* (1965)]
- He introduced interesting and new ideas.
- Let $f_\delta = f * M_\delta$ denote the convolution of f with a normalized Maxwellian of variance δ

$$M_\delta(v) = \left(\frac{1}{2\pi\delta}\right)^{1/2} e^{-\frac{v^2}{2\delta}}$$

- The convolution $f \to f_\delta$ distributes across the Wild convolution

$$[f \circ f]_\delta = f_\delta \circ f_\delta.$$

The strategy

- Let $f(t)$ be the solution to Kac equation corresponding to the initial value f_0
- Mollify the initial value by convolution with a Maxwellian $f_0 \longrightarrow f_{0,\delta}$.
- The solution to Kac equation is now $f_\delta(t) = f(t) * M_\delta$.
- Prove the H-theorem for the regular solution.
- Relax δ to zero and prove the convergence to equilibrium for the original solution.
- Requires a uniform control of the difference between H-functionals

$$H(f) - H(f_\delta).$$

Fisher information

- Let f a probability density on $\mathbf{R}^n$, $n \geq 1$. Fisher's quantity of information associated to f is

$$I(f) = \int_{\mathbf{R}^n} \frac{|\nabla f|^2}{f} \, dv = 4 \int_{\mathbf{R}^n} |\nabla \sqrt{f}|^2 \, dv.$$

- This formula defines a convex, isotropic functional I, first used by Fisher [R.A. Fisher *Proc. Cambrige Philos. Soc.* (1925)]
- In 1959 Linnik [Yu.V. Linnik *Th. Prob. Appl.* (1959)] used the functional to prove the central limit theorem.
- Connection between I and H through the heat equation

$$\lim_{\delta \to 0} \frac{2}{\delta} (H(f) - H(f_\delta)) = I(f).$$

Fisher information decreases

- Let $H^*(f)$ denote the entropy production of Kac equation

$$H^*(f) = -\int_{\mathbf{R}\times[0,2\pi]} \frac{d\theta}{2\pi} \log f(v)\left(f(v^*)f(w^*) - f(v)f(w)\right)\, dw \geq 0$$

- Prove that

$$H^*(f) \geq H(f_\delta)$$

- This implies $H(f)(t) - H(f_\delta(t)$ is a decreasing function of time.
- $I(f)(t)$ is decreasing along the solution of Kac equation
- McKean proof of monotonicity of I not based on properties of the functional.

Other functionals

- In addition to the H-functional

$$I_0(f) = H(f) = \int f \log f$$

- and Fisher information

$$I_1(f) = \int (f')^2 f^{-1}$$

- McKean considered higher order coefficients of Maclaurin expansion of $H(f_\delta)$. The second

$$I_2(f) = 2\int (f'')^2 f^{-1} - \int f''(f')^2 f^{-2}$$

Other functionals III

- Let σ be the second moment. Consider that

$$0 \leq 2 \int \left[\frac{f''}{f} - \left(\frac{f'}{f} \right)^2 + \sigma^{-2} \right]^2$$

- Compute the square into the integral

$$2 \int (f'')^2 f^{-1} - \int f''(f')^2 f^{-2} - 4\sigma^{-2} I(f) + 2\sigma^{-4}$$

- Use

$$I(f) \geq I(M) = \sigma^{-4},$$

- and

$$0 \leq 2 \int (f'')^2 f^{-1} - \int f''(f')^2 f^{-2} - 2\sigma^{-4} = I_2(f) - I_2(M).$$

New tools for Kac model

- McKean analysis of Kac caricature of Maxwellian molecules introduced into kinetic theory Fisher information
- Relationships between Boltzmann H-functional and Fisher treated by Stam [A. Stam *Inform. Control* (1959)] and Blachman [N.M. Blachman *IEEE Trans. Info.* (1965)]
- Results linked to Shannon's entropy power inequality

$$e^{-2H(f*g)} \geq e^{-2H(f)} + e^{-2H(g)}$$

- Functional I_2 used in [P.L. Lions, G. Toscani *J. Funct. Anal.* (1995)] in connection with the central limit theorem and in [U. Gianazza, G. Savaré, G. Toscani (2006)] in problems related to higher order diffusions.

Approach to equilibrium

- In 1973, Tanaka investigated the evolution to equilibrium for Kac model [H. Tanaka *Z. Wahrsch. Verw. Gebiete* (1973)]
- He independently introduced in kinetic theory the distance nowadays known as Wasserstein metric.
- Let M denote the normalized Maxwellian. He defined

$$\mathfrak{e}(f, M) = \inf_{(X,V)\in\tilde{\Gamma}} \mathbb{E}\left[|X - V|^2\right]$$

- $\tilde{\Gamma}$ is the set of all possible couples of random variables (X, V) with f and M as respective laws.
- The main idea behind the introduction of $\mathfrak{e}$ functional was the discovery of its monotonicity in time along the solution to Kac equation.

Approach to equilibrium II

- Tanaka's proof [H. Tanaka *Z. Wahrsch. Verw. Gebiete* (1973)] was very intricated.
- At the end of the paper he added a note of the Referee, who suggested to simplify proofs by means of Hoeffding theorem
- This remarks opened a bridge between Tanaka's work, and the previous history of minimal metrics
- The extension of these ideas to higher dimensions is contained in [H. Tanaka *Z. Wahrsch. Verw. Gebiete* (1978)].
- In this paper Tanaka recovered most of properties of his functional, and the monotonicity in time of $\mathfrak{e}(f(t), g(t))$.

Tanaka functional in one dimension

- Discuss the Referee remark and the key result of Hoeffding [W. Höffding *Schriften des Mathematisches* . . . (1940)]
- Let Γ denote the set of (transference plans) joint probability measures on $\boldsymbol{R} \times \boldsymbol{R}$ with marginals f and $g \in \mathcal{P}(\boldsymbol{R})$. Denote by $F(v)$ the distribution function of f,

$$F(v) = \int_{-\infty}^{v} df,$$

- Within $\Gamma(F, G)$ there are cumulative probability distribution functions Π^* and Π_* discovered by Hoeffding and Fréchet [M. Fréchet *Ann. Univ. Lyon Ser. A* (1957)] which have maximum and minimum correlation.
- Let $x^+ = \max\{0, x\}$ and $x \wedge y = \min\{x, y\}$.

Tanaka functional in one dimension

- Easy to conclude that in $\Gamma(F, G)$ for all $(v, x) \in \mathbb{R}^2$

$$\Pi^*(v,x) = F(v) \wedge G(x) \quad \text{and} \quad \Pi_*(v,x) = [F(v) + G(x) - 1]^+.$$

- The extremal distributions can also be characterized in another way, based on certain familiar properties of uniform distributions.
- Given any $\rho \in (0, 1)$, the pseudo inverse function of the distribution function $F(v)$

$$F^{-1}(\rho) = \inf\{v : F(v) > \rho\}$$

- If X is a real–valued random variable with distribution function F, and U is a random variable uniformly distributed on $[0, 1]$, $F^{-1}(U)$ has distribution function F

Tanaka functional in one dimension II

- For any pair F, G with finite variances, the pair of random variables $[F^{-1}(U), G^{-1}(U)]$ has cumulative distribution function $\Pi^* = \min(F(v), G(x))$.
- Consequently

$$\mathfrak{e}(f,g) = \inf_{\Pi \in \Gamma} \iint_{\boldsymbol{R}\times\boldsymbol{R}} |v-x|^2 \, d\Pi(v,x) = \iint_{\boldsymbol{R}\times\boldsymbol{R}} |v-x|^2 \, d\Pi^*(v,x)$$

- Clear from Hoeffding result

$$\begin{aligned}\mathbb{E}(VX) - \mathbb{E}(V)\mathbb{E}(X) &= \iint_{\boldsymbol{R}\times\boldsymbol{R}} [H(v,x) - F(v)G(x)] \, dv\, dx \\ &\le \iint_{\boldsymbol{R}\times\boldsymbol{R}} [\Pi^*(v,x) - F(v)G(x)] \, dv\, dx,\end{aligned}$$

Tanaka functional in one dimension III

- $[F^{-1}(U), G^{-1}(U)]$ has cumulative distribution function Π^*
- Tanaka (Wasserstein) distance between F and G can be rewritten as the L^2-distance of the pseudo inverse functions

$$W_2(f,g) = \mathfrak{e}(f,g)^{1/2} = \left(\int_0^1 [F^{-1}(\rho) - G^{-1}(\rho)]^2 \, d\rho \right)^{1/2}.$$

- For all $1 \le p < \infty$

$$W_p(f,g) = \left(\int_0^1 |F^{-1}(\rho) - G^{-1}(\rho)|^p \, d\rho \right)^{1/p}.$$

- and

$$W_\infty(f,g) := \lim_{p \nearrow \infty} W_p(f,g) = \|F^{-1} - G^{-1}\|_{L^\infty(0,1)}.$$

Properties of Wasserstein metric

- **Relation to Temperature:** If f belongs to $\mathcal{P}_2(\boldsymbol{R}^n)$ and δ_a is the Dirac mass at a in $\boldsymbol{R}^n$, then

$$W_2^2(f, \delta_a) = \int_{\boldsymbol{R}^n} |v - a|^2 df(v).$$

- **Scaling:** Given f in $\mathcal{P}_2(\boldsymbol{R}^n)$ and $\theta > 0$, let us define

$$\mathcal{S}_\theta[f] = \theta^{N/2} f(\theta^{1/2} v)$$

- Then for any f and g in $\mathcal{P}_2(\boldsymbol{R}^n)$, we have

$$W_2^2(\mathcal{S}_\theta[f], \mathcal{S}_\theta[g]) = \frac{1}{\theta} W_2^2(f, g).$$

Properties of Wasserstein metric II

- **Convexity:** Given f_1, f_2, g_1 and g_2 in $\mathcal{P}_s(\boldsymbol{R}^n)$ such that d_s is finite and α in $[0,1]$

$$W_2^2(\alpha f_1+(1-\alpha)f_2, \alpha g_1+(1-\alpha)g_2) \le \alpha W_2^2(f_1,g_1)+(1-\alpha)W_2^2(f_2,g_2).$$

- **Additivity with respect to convolution:** Given f_1, f_2, g_1 and g_2 in $\mathcal{P}_s(\boldsymbol{R}^n)$

$$W_2^2(f_1 * f_2, g_1 * g_2) \le W_2^2(f_1,g_1) + W_2^2(f_2,g_2).$$

- The main property derived by Tanaka is called Superadditivity with respect to convolution.

Superadditivity

- **Superadditivity with respect to convolution:** Coupling Convolution property with the Scaling property, shows that, for any constant λ such that $0 < \lambda < 1$

$$W_2^2(\mathcal{S}_{1/\lambda}[f_1] * \mathcal{S}_{1/(1-\lambda)}[f_2], \mathcal{S}_{1/\lambda}[g_1] * \mathcal{S}_{1/(1-\lambda)}[g_2]) \leq \lambda W_2^2(f_1, g_1) + (1-\lambda) W_2^2(f_2, g_2).$$

- This property is at the basis of most of the applications of Wasserstein metric to Maxwell models.
- It implies contractivity of Wasserstein metric of two solutions to Boltzmann equation for Maxwell molecules.