THE p-q MODEL BOLTZMANN EQUATION: CONVERGENCE OF THE SOLUTION.
R. FERLAND AND G. GIROUX

1. Introduction
We consider a nonlinear equation for probability measures on $\mathbb{R}_+$:

$$\left( \frac{d}{dt} + 1 \right) \mu(t, B) = \int_0^\infty \mu(t, dx) \int_0^\infty \mu(t, dy) Q(x, y; B)$$

with initial data $\mu_0 \in \mathcal{P}_1$, the set of probability measures on $\mathbb{R}_+$ with finite first moment, $Q$ a Markov kernel and $B \in \mathcal{B}(\mathbb{R}_+)$ the Borel $\sigma$-field of $\mathbb{R}_+$. In [1], the weak convergence of $\mu(t)$ to an equilibrium was proved for a certain class of kernels. The purpose of this paper is to prove that convergence still occurs for a more complex class of kernels introduced by Futcher and Hoare [2], kernels that don’t belong to the class mentioned before.

2. The transition kernel
Let $W_{pq}(x; u)$ be the Beta densities over $[0, u]$:

$$W_{pq}(x; u) = \frac{\Gamma(p+1)\Gamma(q+1)}{\Gamma(p+q)} \frac{x^{p-1}(u-x)^{q-1}}{u^{p+q-1}} 1_{[0,u]}(x)$$

with integers $p, q \geq 1$. The kernels $Q$ of interest here are given by

$$Q(x, y; B) = \int_B K(s; x, y) \, ds$$

where the densities $K$ on $[0, x+y]$ are:

$$K(s; x, y) = \int_0^{x+y} dv W_{pq}(v; x) \int_{x+y}^{x+y} dw W_{qp}(w-x; y) W_{qq}(s-v; w-v)$$

For the limit case $p = 0$, these densities reduce to $W_{qq}(s; x+y)$ and the corresponding kernels fall in the class considered in [1]. We refer to [2] for a detailed account of the origin, significance and properties of the kernels $Q$.

* Research supported in part by NSERC, Canada, under Grant No A-5365 and in part by a grant of FCAR, Gouvernement du Québec.
3. Preliminaries

We need to describe the solution of (1). First define on $\mathcal{P}_1$ a composition operation using a kernel $Q$:

$$
\mu \circ \nu (B) = \int_0^\infty \mu(dx) \int_0^\infty \nu(dy) Q(x,y;B)
$$

For each binary tree $\tau$ with $n$ nodes ($n \geq 1$) associate a $n$-fold product $\varphi(\mu)$ along with a weight $|\tau|$ according to the recursive rule:

$$
|\tau| = \frac{|\tau_1||\tau_2|}{n-1} \quad (n \geq 2), \quad |\tau| = 1 \quad (n = 1)
$$

and

$$
\varphi(\mu) = \varphi_1(\mu) \circ \varphi_2(\mu) \quad (n \geq 2), \quad \varphi(\mu) = \mu \quad (n = 1)
$$

where index 1 and 2 correspond to the left and right sub-tree of the root. E. Wild’s solution of Boltzmann problem for a 3-dimensional Maxwellian gas with cut-off [3,5,8] can be adapted to express $\mu(t)$ as a sum:

$$
\mu(t) = e^{-t} \sum_{n \geq 1} (1 - e^{-t})^{n-1} \sum_{\tau \in T_n} |\tau| \varphi(\mu_0) \quad (2)
$$

with $T_n$ the set of all binary tree with $n$ nodes.

The convergence of $\mu(t)$ will be established with respect to the Kantorovitch metric $\rho$ defined as follows. For $\mu$ and $\nu$ in $\mathcal{P}_1$ let

$$
\rho(\mu, \nu) = \inf \{ \int_0^\infty \int_0^\infty |x-y| \eta(dx,dy) \mid \eta \in \mathcal{C}(\mu, \nu) \},
$$

where $\mathcal{C}(\mu, \nu)$ is the family of all coupling of $\mu$ and $\nu$, namely probability measures $\eta$ on $\mathbb{R}^2_+$ such that $\eta(A \times \mathbb{R}_+) = \mu(A)$ and $\eta(\mathbb{R}_+ \times A) = \nu(A)$ for any $A \in \mathcal{B}(\mathbb{R}_+)$.  

**Proposition 1 [4].** $\mu_n \to_{\rho} \mu$ in $\mathcal{P}_1$ if and only if

$$
\mu_n \to_{w} \mu \quad \text{and} \quad \int_0^\infty x\mu_n(dx) \to \int_0^\infty x\mu(dx).
$$

From this proposition we see that the convergence of $\mu(t)$ with respect to $\rho$ is equivalent to the weak convergence since, as we will now show, the first moment $m_1(\mu(t))$ of $\mu(t)$ is constant in $t$. Indeed, simple calculations give

$$
m_1(\mu \circ \nu) = \frac{2p+q}{2(p+q)} m_1(\mu) + \frac{q}{2(p+q)} m_1(\nu),
$$
from which it follows by induction that \( m_1(\varphi(\mu_0)) = m_1(\mu_0) \) for every \( n \)-fold product, and because of formula (2) we get \( m_1(\mu(t)) = m_1(\mu_0) \) for all \( t \geq 0 \).

**Proposition 2** [7]. Let \( \mu \) and \( \nu \) belong to \( \mathcal{P}_1 \) then

\[
\rho(\mu, \nu) = \int_0^\infty |F_\mu(s) - F_\nu(s)| \, ds
\]

where \( F_\mu \) and \( F_\nu \) are the distribution functions of \( \mu \) and \( \nu \).

**Proposition 3** [6]. Let \( (\Omega, \mathcal{F}, \mathbb{P}) \) be an arbitrary probability space and suppose that we are given subfamilies \( \{ \mu^\omega, \omega \in \Omega \} \) and \( \{ \nu^\omega, \omega \in \Omega \} \) of \( \mathcal{P}_1 \) satisfying the following conditions.

(i) For each \( A \in \mathcal{B}(\mathbb{R}_+) \), \( \mu^\omega(A) \) and \( \nu^\omega(A) \) are \( \mathcal{F} \)-measurable in \( \omega \).

(ii) The probability measure \( \mu = \int_\Omega \mu^\omega \, d\mathbb{P}(\omega) \) and \( \nu = \int_\Omega \nu^\omega \, d\mathbb{P}(\omega) \) are in \( \mathcal{P}_1 \).

Then we have

\[
\rho(\mu, \nu) \leq \int_\Omega \rho(\mu^\omega, \nu^\omega) \, d\mathbb{P}(\omega).
\]

**4. A convex type inequality**

The proof of convergence for \( \mu(t) \) is based on a convex-type inequality that makes the connection between the metric \( \rho \) and the composition operation \( \mu \circ \nu \). For \( \mu \) and \( \nu \) in \( \mathcal{P}_1 \), we write \( \mu \star \nu \) and \( \mu \cdot \nu \) for the measures:

\[
\mu \star \nu([0, s]) = \int_0^s \mu(dx) \, F_\nu(s - x) = \int_0^s \nu(dy) \, F_\mu(s - y)
\]

\[
\mu \cdot \nu([0, s]) = \int_0^\infty \mu(dx) \, F_\nu(s/x) = \int_0^\infty \nu(dy) \, F_\mu(s/y)
\]

**Lemma 1.** Let \( \mu_1, \mu_2, \nu_1 \) and \( \nu_2 \) belong to \( \mathcal{P}_1 \). Then we have

\[
\rho(\mu_1 \star \mu_2, \nu_1 \star \nu_2) \leq \rho(\mu_1, \nu_1) + \rho(\mu_2, \nu_2).
\]
The p-q Model Boltzmann Equation

**Proof:**

\[
\rho(\mu_1 \ast \mu_2, \nu_1 \ast \nu_2) = \int_0^\infty |F_{\mu_1 \ast \mu_2}(s) - F_{\nu_1 \ast \nu_2}(s)| \, ds
\]

\[
= \int_0^\infty |F_{\mu_1 \ast \mu_2}(s) - F_{\mu_1 \ast \nu_2}(s) + F_{\nu_2 \ast \mu_1}(s) - F_{\nu_1 \ast \nu_2}(s)| \, ds
\]

\[
\leq \int_0^\infty |F_{\mu_1 \ast \mu_2}(s) - F_{\mu_1 \ast \nu_2}(s)| \, ds
\]

\[
+ \int_0^\infty |F_{\nu_2 \ast \mu_1}(s) - F_{\nu_2 \ast \nu_1}(s)| \, ds
\]

\[
\leq \int_0^\infty ds \int_0^s \mu_1(dx) |F_{\mu_2}(s - y) - F_{\nu_2}(s - y)|
\]

\[
+ \int_0^\infty ds \int_0^s \nu_2(dy) |F_{\mu_1}(s - y) - F_{\nu_1}(s - y)|
\]

\[
= \int_0^\infty \mu_1(dx) \int_0^\infty |F_{\mu_2}(s - x) - F_{\nu_2}(s - x)| \, ds
\]

\[
+ \int_0^\infty \nu_2(dy) \int_0^\infty |F_{\mu_1}(s - y) - F_{\nu_1}(s - y)| \, ds
\]

\[
= \rho(\mu_1, \nu_1) + \rho(\mu_2, \nu_2).
\]

**Lemma 2.** Let \(\lambda, \mu\) and \(\nu\) belong to \(\mathcal{P}_1\). Then we have

\[
\rho(\lambda \cdot \nu, \lambda \cdot \nu) \leq m_1(\lambda) \rho(\mu, \nu)
\]

**Proof:**

\[
\rho(\lambda \cdot \mu, \lambda \cdot \nu) = \int_0^\infty |F_{\lambda \cdot \mu}(s) - F_{\lambda \cdot \nu}(s)| \, ds
\]

\[
= \int_0^\infty ds \left| \int_0^\infty \lambda(dx) \left( F_{\mu}(s/x) - F_{\nu}(s/x) \right) \right|
\]

\[
\leq \int_0^\infty ds \int_0^\infty \lambda(dx) |F_{\mu}(s/x) - F_{\nu}(s/x)|
\]

\[
= \int_0^\infty \lambda(dx) \int_0^\infty |F_{\mu}(s/x) - F_{\nu}(s/x)| \, ds
\]

\[
= \int_0^\infty x \lambda(dx) \int_0^\infty |F_{\mu}(s) - F_{\nu}(s)| \, ds
\]

\[
= m_1(\lambda) \rho(\mu, \nu).
\]
Theorem 1. Let $\mu_1, \mu_2, \nu_1$ and $\nu_2$ belong to $\mathcal{P}_1$. Then we have

$$\rho(\mu_1 \circ \mu_2, \nu_1 \circ \nu_2) \leq \frac{2p + q}{2(p + q)} \rho(\mu_1, \nu_1) + \frac{q}{2(p + q)} \rho(\mu_2, \nu_2).$$

Proof: Let $\Omega$ be the set $\{(\omega_1, \omega_2, \omega_3) \mid 0 \leq \omega_i \leq 1; i = 1, 2, 3\}$ and $\mathbf{P}$ a probability on $\Omega$ with density $W_{qp}(\omega_1; 1)W_{qp}(\omega_2; 1)W_{qq}(\omega_3; 1)$. Define the functions $f(\omega) = \omega_1 \omega_3 + (1 - \omega_1)$ and $g(\omega) = \omega_2 \omega_3$.

Then because of the scaling property $W_{pq}(x; u) = (1/u) W_{pq}(x/u; 1)$ of the Beta densities, one can show (after tedious but simple calculations) that

$$\mu \circ \nu = \int_{\Omega} (\delta_{f(\omega)} \cdot \mu) \ast (\delta_{g(\omega)} \cdot \nu) \, d\mathbf{P}(\omega).$$

So applying successively Proposition 3, Lemma 1 and 2 we obtain

$$\rho(\mu_1 \circ \mu_2, \nu_1 \circ \nu_2) \leq \int_{\Omega} \rho((\delta_{f(\omega)} \cdot \mu_1) \ast (\delta_{g(\omega)} \cdot \nu_1), (\delta_{f(\omega)} \cdot \mu_2) \ast (\delta_{g(\omega)} \cdot \nu_2)) \, d\mathbf{P}(\omega)$$

$$\leq \int_{\Omega} \rho(\delta_{f(\omega)} \cdot \mu_1, \delta_{f(\omega)} \cdot \nu_1) \, d\mathbf{P}(\omega) + \int_{\Omega} \rho(\delta_{g(\omega)} \cdot \mu_2, \delta_{g(\omega)} \cdot \nu_2) \, d\mathbf{P}(\omega)$$

$$\leq \int_{\Omega} (f(\omega) \rho(\mu_1, \nu_1) + g(\omega) \rho(\mu_2, \nu_2)) \, d\mathbf{P}(\omega)$$

$$\leq \frac{2p + q}{2(p + q)} \rho(\mu_1, \nu_1) + \frac{q}{2(p + q)} \rho(\mu_2, \nu_2).$$

A measure $\gamma$ is called an equilibrium if $\gamma \circ \gamma = \gamma$. The Gamma laws with parameters $p + q$ and $\lambda > 0$ are equilibrium measures. The next result shows that in $\mathcal{P}_1$, there is no other equilibrium measures.

Corollary. Let $\gamma$ be a Gamma law with parameters $p + q$ and $\lambda > 0$ and suppose $\mu \in \mathcal{P}_1$ is such that $m_1(\mu) = m_1(\gamma)$. Then

$$\rho(\mu \circ \mu, \gamma) < \rho(\mu, \gamma)$$

if $\mu \neq \gamma$.

Proof: Since $\gamma \circ \gamma = \gamma$ the preceding theorem already gives

$$\rho(\mu \circ \mu, \gamma) \leq \rho(\mu, \gamma).$$
Assuming that equality holds, we prove that \( \mu = \gamma \). Indeed, looking at the proof of Theorem 1, we see that if there is equality then
\[
\rho((\delta_f(\omega) \cdot \mu) \ast (\delta_g(\omega) \cdot \mu), (\delta_f(\omega) \cdot \gamma) \ast (\delta_g(\omega) \cdot \gamma)) = \rho(\delta_f(\omega) \cdot \mu, \delta_f(\omega) \cdot \gamma) + \rho(\delta_g(\omega) \cdot \mu, \delta_g(\omega) \cdot \gamma),
\]
and this almost surely with respect to \( \mathbf{P} \). Pick \( \omega \in \Omega \) for which both \( \delta_f \) and \( \delta_g \) are zero and such that equality above holds. Now looking at the proof of Lemma 1, we get
\[
\left| \int_0^{s} (\delta_g(\omega) \cdot \gamma)(dx) (F_{\delta_f(\omega)} \mu(s-x) - F_{\delta_f(\omega)} \gamma(s-x)) \right|
= \int_0^{s} (\delta_g(\omega) \cdot \gamma)(dx) \left| F_{\delta_f(\omega)} \mu(s-x) - F_{\delta_f(\omega)} \gamma(s-x) \right|
\]
almost everywhere in \( s \). This implies that the sign of the integrand \( F_{\delta_f(\omega)} \mu(x) - F_{\delta_f(\omega)} \gamma(x) \) is the same (say positive) almost everywhere for \( x \geq 0 \). Thus
\[
\rho(\delta_f(\omega) \cdot \mu, \delta_f(\omega) \cdot \gamma) = \int_0^{\infty} \left| F_{\delta_f(\omega)} \mu(s) - F_{\delta_f(\omega)} \gamma(s) \right| ds
= \int_0^{\infty} \left[ F_{\delta_f(\omega)} \mu(s) - F_{\delta_f(\omega)} \gamma(s) \right] ds
= \int_0^{\infty} f(\omega) [F_{\mu}(s) - F_{\gamma}(s)] ds
= f(\omega) \int_0^{\infty} ([1 - F_{\gamma}(s)] - [1 - F_{\mu}(s)]) ds
= f(\omega) (m_1(\gamma) - m_1(\mu)) = 0.
\]
But, by Lemma 2 we have \( \rho(\delta_f(\omega) \cdot \mu, \delta_f(\omega) \cdot \gamma) = f(\omega) \rho(\mu, \gamma) \). Hence \( \rho(\mu, \gamma) = 0 \) and \( \mu = \gamma \).

5. Convergence to equilibrium

In this section we make use of the results of Section 4 to prove the convergence to equilibrium assuming only the existence of the first moment of the initial law. To prove the convergence theorem we use a compactness argument and this can be done because of the following lemma.

Lemma 3. The second moments \( m_2(\mu(t)) \), \( t \geq 0 \), are bounded as soon as \( m_2(\mu_0) \) exist.

Proof: Since
\[
m_2(\mu(t)) = \sum_{n \geq 1} e^{-t} (1 - e^{-t})^{n-1} \sum_{\tau \in T_n} |\tau| m_2(\varphi(\mu_0))
\]
it is enough to show that the second moments \( m_2(\varphi(\mu_0)) \) are bounded.
It is shown in [2] that
\[
m_2(\mu \circ \nu) = a_{20} m_2(\mu) + 2a_{21} m_1(\mu)m_1(\nu) + a_{22} m_2(\nu)
\] (3)
where the coefficients \(a_{20}, a_{21}\) and \(a_{22}\) may be explicitly computed. Let \(\gamma\) be an equilibrium measure with its first moment equal to the one of \(\mu_0\). It is surely possible to find a constant \(C \geq 1\) such that \(m_2(\mu_0) \leq C m_2(\gamma)\).

Now suppose \(m_2(\varphi(\mu_0)) \leq C m_2(\gamma)\), for every \(k\)-fold product with \(k \leq n\). Then for a \((n + 1)\)-fold product
\[
m_2(\varphi(\mu_0)) = a_{20} m_2(\varphi_1(\mu_0)) + 2a_{21} m_1(\varphi_1(\mu_0))m_1(\varphi_2(\mu_0)) \\
+ a_{22} m_2(\varphi_2(\mu_0)) \\
\leq C(a_{20} m_2(\gamma) + 2a_{21} m_1(\gamma)m_1(\gamma) + a_{22} m_2(\gamma)) \\
= C m_2(\gamma).
\]
The last equality follows from (3) and the fact that \(\gamma \circ \gamma = \gamma\).

**Theorem 2.** Let \(\mu_0 \in \mathcal{P}_1\) and suppose \(\gamma\) is an equilibrium measure such that \(m_1(\gamma) = m_1(\mu_0)\). Then, \(\rho(\mu(t), \gamma)\) decreases to 0 as \(t \uparrow \infty\).

**Proof:** An approximation argument shows that we can restricted ourselves to the case \(m_2(\mu_0) < \infty\). As a consequence of Theorem 1 it is possible [6] to prove that
\[
\rho(\mu(t), \gamma) \leq \rho(\mu_0, \gamma) - \int_0^t \bar{\rho}(\mu(s), \gamma) \, ds,
\] (4)
where \(\bar{\rho}(\mu(s), \gamma) = \rho(\mu(s), \gamma) - \rho(\mu(s) \circ \mu(s), \gamma)\). This shows that \(\rho(\mu(t), \gamma)\) is decreasing. Now define \(\mathcal{P}(\epsilon, M)\) as the set of all \(\lambda\) in \(\mathcal{P}_1\) such that
\[
m_1(\lambda) = m_1(\gamma) \quad m_2(\lambda) \leq M \quad \rho(\lambda, \gamma) \geq \epsilon,
\]
then \(\mathcal{P}(\epsilon, M)\) is compact with respect to \(\rho\). Suppose \(\rho(\mu(t), \gamma)\) doesn’t decrease to zero. By Lemma 3, \(\mu(t) \in \mathcal{P}(\epsilon, M)\) for some \(\epsilon\) and \(M\). But since \(\bar{\rho}(\cdot, \gamma)\) is continuous on \(\mathcal{P}(\epsilon, M)\), the infimum of \(\bar{\rho}\) is achieved for \(\mu^* \in \mathcal{P}(\epsilon, M)\) and
\[
\bar{\rho}(\mu(t), \gamma) \geq \bar{\rho}(\mu^*, \gamma) = \delta > 0,
\]
because of the Corollary. Then by (4), \(\rho(\mu(t), \gamma) \leq \rho(\mu_0, \gamma) - \delta t\) which leads to a contradiction for \(t\) large enough.
REFERENCES


Département de mathématiques et d’informatique
Université de Sherbrooke
Sherbrooke, Québec
Canada J1K 2R1