UNIVERSIDAD DE BUENOS AIRES
Facultad de Ciencias Exactas y Naturales
Departamento de Matemática

Tesis de Licenciatura

Algunos modelos de formación de opinión

Mariela Rajngewerc

Director: Nicolas Saintier

13 de Abril 2016

II

Índice general

Agradecimientos

A Juan Pablo, por su buena predisposición siempre.

A Nicolas, por su paciencia y por ayudarme en este viaje.

A los jurados.

A los docentes de las teóricas y a los de las prácticas que me dieron oportunidades.

A la UBA.

A Diana y Sergio, por TODO. Por apoyarnos a Chu y a mí en los momentos complicados y potenciar los buenos momentos. Por querer a Chudi como uno más de la manada.

A Lula, Guido y Dan, los adoro.

A Rosita y a Salo por todo el amor .

A Andrea, Sol, Iari, Claudio.

A Carlitos y Marcelo.

A Raquel y a Ernesto.

A Vero Fantini por alimentarme y mimarme, siguen en la lista: Mesple, Lucy (y Pali), Santi, Barbie y Nacho.

A Anita por todo amiga, te quiero!

A Vicu.

A Estefi (me hubiese gustado aprovecharte mas cuando cursábamos juntas), gracias por bancarme y motivarme.

A Diego, mi amigo químico.

A Pau Adaglio, Caro N.(y a Sheima) y Marie Bisso.

A los Computadores: Pigre, Joe , Diego, Tute, Tommy, Jm, por permitirme colarme en tantas tardes de estudio, pomodoros, asados.

A Tamara Goldstein y Zazu, por sentarse arriba de mis apuntes e imponerme recreos.

A Gime, al Gnomo y, en especial, a Silvio y Viktor.

A Flor.

A Adrian Soria, por hacerme reír con su locura.

A Jesi.

A Aye.

A Severius.

A Jujus porque cada vez que nos vemos vuelvo con una sonrisa.
A Wokis por las noches de estudio, por las noches de no estudio, por cuidarme de mi mente.
A Chu por ser el mejor compañero, por tu sonrisa. Porque sin tu apoyo y paciencia no hubiese sido posible aprobar nada. Porque trajiste a Zazu. Porque aunque vivas en algún otro planeta me bajás al mundo y me hacés feliz.

Introducción

Los modelos de dinámica de opinión describen el proceso de formación de opinión en grupos constituidos por individuos. Las personas suelen tener e intercambiar opiniones acerca de distintos tópicos: desde política a salud, desde los productos hasta deportes. Estas opiniones pueden ser el resultado de una seria reflexión, o bien, formadas a través de las distintas interacciones con individuos que dan su punto de vista acerca de determinado tema. Denominamos validación social a la influencia que recibe un individuo por las opiniones expresadas por el resto de la población cuando este debe tomar una decisión. Los modelos de opinión dinámica, intentan comprender cuando la formación de opinión alcanza un consenso, cuando esta polarizada o bien cuando esta fragmentada.

En lo que sigue interpretaremos a los individuos en una sociedad como partículas. Desde el punto de vista de la física, los individuos y sus interacciones constituyen un nivel microscópico de sistema social. La pregunta es si estos comportamientos pueden explicar los fenómenos a escala macroscópica.

En este trabajo presentaremos distintos modelos de formación de opinión. El objetivo de los mismos es conocer la función de distribución de opinión en una sociedad a lo largo del tiempo. Analizaremos el comportamiento asintótico de las ecuaciones diferenciales derivadas de dichos modelos y hallaremos la distribución de opinión en una sociedad para algunos casos particulares.

En el Capítulo 1 presentaremos los modelos de opinión de Sznajd, Ochrombel simplificado y "Dos contra uno". Luego analizaremos los últimos dos sobre un grafo completo (ie: los agentes son todos vecinos entre si) y estudiaremos las ecuaciones de dichos modelos cuando $N \to \infty$.

En el Capítulo 2 mostraremos un modelo cinético de formación opinión considerando el intercambio de opinión entre individuos y la difusión de la información. En este caso tomaremos a la opinión como una variable continua en $[-1, 1]$. Después estudiaremos la distribución de la opinión en algunos casos particulares. Finalmente, al reescalar el tiempo y hacer ciertas consideraciones obtendremos que la distribución verifica una ecuación de tipo Fokker-Planck.

Por último, en el Capítulo 3, estudiaremos un modelo no lineal derivado del Modelo Cinético. Probaremos la existencia de solución y estudiamos su comportamiento asintótico en algunos casos particulares.

Finalmente, compilamos en el apéndice algunos resultados clásicos sobre la ecuación de conservación de masa y algunas herramientas de la teoría de probabilidades que usamos a lo largo de esta tesis.

Capítulo 1

Los modelos de Snazjd y Ochrombel de formación de opinión

1.1. Modelos de opinión

El marco general de los modelos de opinión suele ser el siguiente: consideramos una población formada por N agentes ubicados en una red que modela las relaciones de los agentes entre sí. El estado (u opinión) de cada agente se representa por un parámetro σ que puede ser, por ejemplo, un número entre -1 y 1 donde 1 y -1 representan dos opiniones antagónicas y los números en $(-1, 1)$ representan un continuo de opiniones intermedias. Notemos $\sigma_i(t)$ el estado del agente i a tiempo t y $\sigma(t) = \{\sigma_1(t), ..., \sigma_N(t)\}$ el estado de opiniones de los todos los agentes en el instante t. Cada individuo, como resultado de las interacciones con sus vecinos en la red, al instante siguiente $t + 1$, el agente tendrá un nuevo estado $\sigma_i(t + 1)$. La regla de actualización de los estados $\Sigma(t) \to \Sigma(t+1)$ suele ser simple e involucrar unicamente $\Sigma(t)$ y no otro estado anterior del sistema $\Sigma_{t'}$, $t' < t$. En este caso $(\Sigma_t)_t$ es un procesos de Markov en tiempo discreto, cuya matriz de transición del tiempo t a $t + 1$ difiere según el modelo.

Cuando la cantidad de agentes $N \to +\infty$, uno puede esperar, después de reescalar adecuadamente el tiempo, obtener un modelo continuo que tendrá la forma de una ecuación en derivadas parciales cuyos coeficientes reflejan la dinámica estudiada.

Describimos a continuación algunos modelos de opinión clásicos, el modelo de Sznajd y dos versiones simplificadas, el modelo de Ochrombel y el modelo "Dos contra uno". Cuando la red que modela las interacciones entre los agentes es completa, en el sentido que cualquier agente tiene por vecino directo a cualquier otro agente de la red, es posible obtener el límite continuo de los modeles simplificados siguiendo el paper de F. Slanina y H. Lavicka [9].

1.2. El modelo de Sznajd

En el año 2000 K. Sznajd-Weron and J. Snazjd propusieron en [10] un modelo para describir el fenómeno social global a través de las interacciones locales. Su objetivo era analizar como se esparcen las opiniones en la sociedad. La opinión de una sociedad es, claramente, el resultado de las opiniones individuales. Motivados por el concepto de *validación social*, propusieron el siguiente modelo unidimensional en el que cada miembro de la población modifica su opinión teniendo en cuenta de algún modo la opinión de sus vecinos.

Consideramos una población formada por N agentes de modo tal que el agente i tiene por vecinos a los agentes $i - 1$ y $i + 1$ (salvo los agentes 1 y N que tienen por único vecino al agente 2 y N-1 respectivamente). La preferencia u opinión del agente i esta dada por un número $\sigma_i \in \{-1, +1\}$. En cada paso temporal se actualizan las opiniones de dos agentes de la siguiente forma. Se elige a un par de individuos vecinos, supongamos al individuo i y al individuo $i + 1$, con el objetivo de modificar la opinión de los vecinos de los mismos (o sea la opinión de los agentes $i-1$ e $i+1$) con la siguiente regla:

- Si $\sigma_i \sigma_{i+1} = 1$, es decir si las opiniones de los agentes i e $i+1$ coinciden, entonces se actualizaran las opiniones de los agentes $i-1$ e $i+2$ para que coincidan con las opiniones de los i y $i+1$:

$$\sigma_{i-1} = \sigma_i = \sigma_{i+1} = \sigma_{i+2}.$$

  Esta regla modela una validación social: dos vecinos de acuerdo propagan su opinión a sus vecinos cercanos.

- Si las opiniones de los agentes i e $i+1$ no coinciden, entonces actualizaremos la opinión del agente $i-1$ de forma que coincida con la opion del agente $i+1$ y la del individuo $i+2$ para que coincida con la opinión del agente i:

$$\sigma_{i-1} = \sigma_{i+1}, \qquad \sigma_{i+2} = \sigma_i.$$

  El desacuerdo de dos vecinos se propaga.

Estas dos reglas modelan el adagio "United we stand, divided we fall".

Se puede observar que las configuraciones invariantes para estas reglas son:

- todos los agentes tienen la misma opinión: $\sigma_i = 1$ para todo i, o $\sigma_i = -1$ para todo i,
- cada agente esta en desacuerdo con sus vecinos: $\Sigma = \{-1, 1-1, 1, -1, ...\}$ o $\Sigma = \{1, -1, 1, -1, 1, ...\}$.

K. Sznajd-Weron and J. Snazjd estudiaron este modelo numéricamente. Los experimentos numéricos realizados muestran que cualquiera sea la distribución inicial de opinión en la población, el sistema siempre converge a uno de estos estados estacionarios. También estudiaron numéricamente la influencia de la proporción inicial de opinión $+1$ sobre el estado final obtenido. Notan por ejemplo que hace falta inicialmente una proporción mínima de 0.7 para que esta opinón se propague a toda la población con probabilidad al menos 0.5.

Finalmente analizan la influencia de un ruido externo sobre la dinámica de opiniones. Cada agente cambia de opinión al azar con probabilidad p y sino con probabilidad 1-p, sigue las reglas descriptas anteriormente. Este ruido aleatorio puede pensarse como una "temperatura social" que tiene la capacidad de desorganizar un sistema inicialmente estable. De hecho, si todos los agentes están todos de acuerdo inicialmente, obtienen que si p es bastante chico (del orden de 10^{-6}), la perturbación introducida por este ruido es temporaria y los agentes vuelven a pensar igual que lo hacían inicialmente. En cambio, cuando p es mayor que 10^{-5}, el sistema se desorganiza. Acá el grado de organización se mide gracias a la función de "magnetización" $m = \frac{1}{N}\sum_{i=1}^{N} \sigma_i$. Note que $|m| \leq 1$ con

$|m| = 1$ si y solo si todos los agentes están de acuerdo. Este parámetro nos va a resultar muy útil al momento de estudiar analíticamente los modelos de Ochrombel y "2 contra 1" siguiendo los pasos de [9].

Describimos a continuación dos modelos que pueden verse como versiones simplificadas del modelo de Sznajd.

1.3. El modelo de Ochrombel

1.3.1. Descripción del modelo

En una nota corta [8], S. Ochrombel describe un modelo en el que los agentes están ubicadas en una red cuadrada $\{-L, .., L\} \times \{-L, .., L\}$. A cada paso temporal, se actualiza la opinión de un agente, cuya opinión es un número $\sigma \in \{-1, 1\}$ como en el modelo de Sznajd, de la siguiente forma

- Se elige al azar un agente, digamos el agente i.
- Los cuatro vecinos de i en la red toman la opinión de i.

Según S. Ochrombel, si una mitad (elegida al azar) de los agentes tiene opinión 1 y la otra mitad -1, entonces todos los agentes terminan con la misma opinión. Se obtiene entonces un comportamiento asintótico similar al del modelo de Sznajd con la diferencia ahora que no hace falta un par de agentes con la misma opinión para propagarla sino que un solo agente puede influenciar sus vecinos.

F. Slanina y H. Lavicka proponen en [9] la siguiente modificación del modelo de Ochrombel en el contexto de una red de relaciones cualquiera:

- Se elige al azar un agente, digamos el agente i.
- Se elige al azar otro agente, digamos j, entre los vecinos de i, y se actualiza su opinión de la siguiente manera: $\sigma_j(t+1) = \sigma_i(t)$.

Note que en este modelo, en oposición a los modelos de Sznajd e Ochrombel, se actualiza la opinión de un único agente en cada paso.

1.3.2. Probabilidades de transición

Supongamos a partir de ahora que la red que modela las relaciones entre los agentes es completa por lo que cualquier agente esta conectado con todos los demás. Denotemos con $N^{\pm}(t)$ la cantidad total de agentes con opinión ± 1. En cualquier tiempo t, el estado σ_t del sistema esta por lo tanto caracterizado por la proporción $N^{+}(t)/N$ de agentes con opinión 1. Resulta más conveniente describir el estado del sistema mediante su *magnetización* $m(t)$ definida por

$$m(t) := \frac{1}{N}\sum_{i=1}^{N}\sigma_i(t) = \frac{N^+(t) - N^-(t)}{N} = 1 - 2\frac{N^-(t)}{N} = 2\frac{N^+(t)}{N} - 1. \tag{1.1}$$

Usamos que $N^+ + N^- = N$ en las dos últimas igualdades.

Note que $m(t) \in \{-1, -1+\frac{2}{N}, -1+\frac{4}{N}, ..., 1-\frac{2}{N}, 1\}$ y que en cada paso temporal la magnetización $m(t)$ o se mantiene constante o se incrementa de $\pm\frac{2}{N}$ como consecuencia de las tres configuraciones posibles siguientes:

1. Si $\sigma_i(t) = -\sigma_j(t)$ y $\sigma_i(t) = 1$ entonces
$$m(t+1) = \frac{(N^+(t)+1) - (N^-(t)-1)}{N} = \frac{N^+(t) - N^-(t) + 2}{N} = m(t) + \frac{2}{N}.$$
2. Si $\sigma_i(t) = -\sigma_j(t)$ y $\sigma_i(t) = -1$ entonces
$$m(t+1) = \frac{(N^+(t)-1) - (N^-(t)+1)}{N} = \frac{N^+(t) - N^-(t) - 2}{N} = m(t) - \frac{2}{N}.$$
3. Si $\sigma_i(t) = \sigma_j(t)$, en ese caso m se mantiene constante: $m(t+1) = m(t)$.

La siguente proposición da las probabilidades de ocurrencia de cada una de esas tres transiciones:

Proposición 1.3.1.

$$Prob\left\{m(t) \to m(t) \pm \frac{2}{N}\right\} = \frac{1-m^2}{4}\Big(1 + \frac{1}{N-1}\Big) \tag{1.2}$$

$$Prob\left\{m(t) \to m(t)\right\} = \Big(\frac{1+m^2}{2} - \frac{1}{N}\Big)\Big(1 + \frac{1}{N-1}\Big) \tag{1.3}$$

Demostración. Verifiquemos (1.2). Primero por definición de la dinámica,

$$\begin{aligned}
& Prob\left\{m(t) \to m(t) + \frac{2}{N}\right\} \\
= \; & \text{Probabilidad de elegir un agente con opinión 1 y luego otro con opinión -1} \\
= \; & \frac{N^+(t)}{N}\frac{N^-(t)}{N-1}
\end{aligned}$$

y

$$\begin{aligned}
& Prob\left\{m(t) \to m(t) - \frac{2}{N}\right\} \\
= \; & \text{Probabilidad de elegir un agente con opinión -1 y luego otro con opinión 1} \\
= \; & \frac{N^-(t)}{N}\frac{N^+(t)}{N-1}.
\end{aligned}$$

Reescribimos ahora esta última expresión de la siguiente manera (no escribimos el t):

$$\begin{aligned}
\frac{N^+}{N}\frac{N^-}{N-1} &= \frac{1}{4(N-1)N}(2N^+)(2N^-) \\
&= \frac{1}{4(N-1)N}((N^+ + N^-)^2 - (N^+ - N^-)^2) \\
&= \frac{1}{4(N-1)N}\Big(N^2 - (Nm)^2\Big) \\
&= \frac{1-m^2}{4}\frac{N}{N-1}
\end{aligned}$$

Finalmente la ecuación (1.3) es consecuencia directa de (1.2) ya que

$$\begin{aligned}
& Prob\left\{m(t) \to m(t)\right\} + Prob\left\{m(t) \to m(t) - \frac{2}{N}\right\} \\
& + Prob\left\{m(t) \to m(t) + \frac{2}{N}\right\} = 1
\end{aligned}$$

□

1.3.3. El límite continuo del modelo de Ochrombel

Queremos ahora investigar el límite del modelo de Ochrombel cuando la cantidad de agentes N tiende a infinito. Consideramos para eso la probabilidad $u_N(m,t)$ de hallar el sistema en el estado m a tiempo t cuando hay N agentes. Como $m \in \{-1, -1 + \frac{2}{N}, .., 1 - \frac{2}{N}, 1\}$, la medida de probabilidad $u_N(.,t)$ es de la forma

$$u_N(.,t) = \sum_{k=0}^{N} p_N(k,t)\delta_{-1+\frac{2k}{N}}$$

donde $p_N(k,t)$ es la probabilidad de que la magnetización valga $-1 + \frac{2k}{N}$ a tiempo t. Note que $u_N.,t$ es una medida de probabilidad soportada en el compacto $[-1,1]$. Sabemos por lo tanto que a t fijo una subsucesión converge débilmente a una medida de probabilidad $u(.,t)$ soportada en $[-1,1]$. Nuestro propósito es deducir de un modo informal la ecuación que debería verificar u. Para eso usaremos la ecuación maestra que verifica $u_N.,t$ y su forma débil, una ecuación de tipo Boltzmann que resulta más conveniente. Vamos a ver que u es solución de la ecuación de difusión

$$\frac{\partial u}{\partial \tau} = \frac{\partial^2}{\partial m^2}((1-m^2)u). \tag{1.4}$$

Deducción de la ecuacion maestra

Supongamos que en un intervalo de tiempo de longitud Δt, la cantidad de transiciones X del sistema (o sea la cantidad de interacciones entre agentes) sigue una distribución de Poisson de parámetro Δt es decir que las transiciones ocurren siguiendo un proceso de Poisson de tasa 1. En particular las probabilidades que no ocurra ninguna transición, exactamente 1 o mas de 2 transiciones valen

$$\begin{aligned} P(X=0) &= e^{-\Delta t} = 1 - \Delta t + O((\Delta t)^2), \\ P(X=1) &= e^{-\Delta t}\Delta t = \Delta t + O((\Delta t)^2), \\ P(X \geq 2) &= 1 - P(X=0) - P(X=1) = O((\Delta t)^2) \end{aligned}$$

Por la fórmula de las probabilidades totales, la probabilidad $u(m, t+\Delta t)$ de encontrar el sistema al tiempo $t + \Delta t$ en el estado m es, ignorando la posibilidad de más de una transición, la suma de las probabilidades de que haya una transición desde un estado m' a m por la probabilidad de que el estado sistema estaba en el estado m' al tiempo t mas la probabilidad de que no haya ninguna transición por la probabilidad de que el estado sistema estaba en el estado m al tiempo t:

$$\begin{aligned} & u_N(m, t+\Delta t) \\ = & \sum_{m'} u_N(m', t) P(m' \to m) P(X=1) + P(X=0) P(m \to m) u_N(m,t) + O((\Delta t)^2) \\ = & \Delta t \sum_{m'} u_N(m', t) P(m' \to m) + u_N(m,t) - \Delta t u_N(m,t) + O((\Delta t)^2) \end{aligned}$$

Luego

$$\begin{aligned} & \frac{u_N(m, t+\Delta t) - u_N(m,t)}{\Delta t} \\ = & \sum_{m'} u_N(m', t) P(m' \to m) - u_N(m,t) + O(\Delta t) \\ = & \sum_{m'} u_N(m', t) P(m' \to m) - u_N(m,t) \sum_{m'} P(m \to m') + O(\Delta t). \end{aligned}$$

Pasando al límite $\Delta t \to 0$, obtenemos

$$\frac{\partial u_n}{\partial t}(m,t) = \sum_{m'} u_N(m', t) P(m' \to m) - \sum_{m'} u_N(m,t) P(m \to m'). \tag{1.5}$$

Esta ecuación, llamada ecuación maestra del sistema, afirma que la variación de la densidad de probabilidad de un estado m resulta del balance entre dos términos. El primero expresa una ganancia del estado m debido a las transiciones desde los estados m', mientras que el segunda representa la pérdida de probabilidad debido a las transiciones desde el estado m a los restantes estados m'. Obtuvimos esta ecuación en un marco general sin referencia al modelo de Ochrombel suponiendo únicamente que la dinámica del sistema es markoviana (el estado del sistema en un instante depende únicamente del estado en el instante inmediatamente anterior).

Formulacion débil de la ecuación maestra

Para obtener una formulación débil de la ecuación maestra (1.5), multiplicamos ambos miembros de (1.5) por una función-test $\phi(m)$ y sumamos sobre m. Obtenemos:

$$\begin{aligned}&\frac{d}{dt}\sum_m \phi(m)u_N(m,t)\\&=\sum_{m,m'}\phi(m)P(m'\to m)u_N(m',t)-\sum_{m,m'}\phi(m)P(m\to m')u_N(m,t)\end{aligned}$$

Intercambiamos m y m' en la 1era suma del miembro derecho:

$$\begin{aligned}&\frac{d}{dt}\sum_m \phi(m)u_N(m,t)\\&=\sum_{m,m'}\phi(m')P(m\to m')u_N(m,t)-\sum_{m,m'}\phi(m)P(m\to m')u_N(m,t)\\&=\sum_m\Big\{\sum_{m'}\Big(\phi(m')-\phi(m)\Big)P(m\to m')\Big\}u_N(m,t)\end{aligned}$$

Obtenemos entonces la siguiente formulación débil para la ecuación maestra:

$$\frac{d}{dt}\int\phi(m)\,du_N(m,t)=\int\Big\{\sum_{m'}\Big(\phi(m')-\phi(m)\Big)Prob(m\to m')\Big\}\,du_N(m,t). \tag{1.6}$$

Esta ecuación nos dice que la variación del valor esperado de ϕ resulta de todos los posibles incrementos de ϕ teniendo en cuenta la probabilidad de estar en tal o tal estado y la probabilidad de que el sistema pase de este estado a otro.

Aplicación al modelo de Ochrombel

En el caso del modelo de Ochrombel, ya vimos que la magnetización m puede o incrementarse de $\pm\frac{2}{N}$ o mantenerse constante al cabo de una interacción. La ecuación

(1.6) se escribe entonces

$$
\begin{aligned}
(1.7) \qquad & \frac{d}{dt}\int \phi(m)\, du_N(m,t) \\
&= \int \Big(\phi(m+\frac{2}{N}) - \phi(m)\Big) Prob(m \to m + \frac{2}{N}) \\
&\quad + \Big(\phi(m-\frac{2}{N}) - \phi(m)\Big) Prob(m \to m - \frac{2}{N})\, du_N(m,t)
\end{aligned}
$$

Note que es muy fácil ver usando esta formulación que la masa total y la magnetización promedio se conservan. De hecho tomando $\phi(m) = 1$ obtenemos $\frac{d}{dt}\int du_N(m,t) = 0$ o sea $\int du_N(m,t) = \int du_N(m,0) = 1$ para todo t. Tomando $\phi(m) = m$ e usando (1.2) obtenemos también

$$\frac{d}{dt}\int m\, du_N(m,t) = \frac{2}{N}\int Prob(m \to m + \frac{2}{N}) - Prob(m \to m - \frac{2}{N})\, du_N(m,t) = 0.$$

Haciendo un desarrollo de Taylor de ϕ de orden 2 tenemos

$$\phi(m \pm \frac{2}{N}) - \phi(m) = \pm\frac{2}{N}\phi'(m) + \frac{2}{N^2}\phi''(m) + \frac{4}{3N^3}\phi'''(\tilde{m})$$

para un $\tilde{m}$ entre m y $m \pm \frac{2}{N}$. Como ϕ es suave con soporte compacto obtenemos

$$\phi(m \pm \frac{2}{N}) - \phi(m) = \pm\frac{2}{N}\phi'(m) + \frac{2}{N^2}\phi''(m) + O(N^{-3}).$$

Luego usando (1.2) podemos reescribir (1.7) como

$$\frac{d}{dt}\int \phi(m)\, du_N(m,t) = \int \Big(\frac{4}{N^2}\phi''(m) + O(N^{-3})\Big)\frac{1-m^2}{4}\Big(1 + \frac{1}{N-1}\Big)\, du_N(m,t).$$

Como $u_N(.,t)$ es una medida de probabilidad, obtenemos finalmente

$$\frac{d}{dt}\int \phi(m)\, du_N(m,t) = \frac{1}{N^2}\int \phi''(m)(1-m^2)\, du_N(m,t) + O(N^{-3}).$$

Reescalamos ahora el tiempo (lo aceleramos de hecho) considerando $\tau = \frac{t}{N^2}$. Seguimos notando $u_N(.,\tau) := u_N(.,t)$. La ecuación anterior se reescribe como

$$\frac{d}{d\tau}\int \phi(m)\, du_N(m,\tau) = \int \phi''(m)(1-m^2)\, du_N(m,\tau) + O(\frac{1}{N}).$$

Luego es razonable conjeturar que la medida límite débil $u(.,\tau) := \lim u_N(.,\tau)$ (que existe, a τ dado, al menos por una subsucesión) verifica

$$\frac{d}{d\tau}\int \phi(m)\, du(m,\tau) = \int \phi''(m)(1-m^2)\, du_N(m,\tau)$$

para toda $\phi \in C_c^\infty(\mathbb{R})$. Es la formulación débil de (1.4).

1.4. El modelo «Dos contra uno»

1.4.1. Descripción del modelo

Analicemos ahora el modelo Dos contra uno. Como antes, el conjunto de opiniones posibles es $S = \{-1, 1\}$ y suponemos que todos los agentes son vecinos entre sí. Veamos el algoritmo que propone este modelo para la actualización de las opiniones de los agentes en cada paso temporal:

- Se elige al azar un par de agentes vecinos: el agente i y el algente j.
- Si i y j tienen opiniones distintas (es decir $\sigma_i \sigma_j = -1$) entonces no pasa nada.
- Si i y j tienen la misma opinión (es decir $\sigma_i = \sigma_j$) entonces propagamos esta opinión a uno de sus vecinos: elegimos al azar un agente k (recordar que el grafo de relaciones es completo) y actualizamos la opinión de k con $\sigma_k(t+1) = \sigma_i(t)$.

En casa paso este modelo actualiza la opinión de un único agente.

Como en el modelo de Sznajd, hace falta un grupo de agentes con la misma opinión para propagarla.

1.4.2. Probabilidades de transición

Analicemos la evolución de la variable de $m$agnetización en este modelo. En cada paso temporal la variable m puede modificarse de alguna de las siguientes formas:

1. Si $\sigma_i(t) = \sigma_j(t)$ y $\sigma_i(t) = 1$
   a) Si $\sigma_k(t) = \sigma_i(t)$, en ese caso m se mantiene constante.
   b) Si $\sigma_k(t) = -\sigma_i(t)$, en ese caso $m(t+1) = \frac{N^+ + 1 - (N^- - 1)}{N} = \frac{N^+ - N^- + 2}{N}$
2. Si $\sigma_i(t) = \sigma_j(t)$ y $\sigma_i(t) = -1$
   a) Si $\sigma_k(t) = \sigma_i(t)$, en ese caso m se mantiene constante.
   b) Si $\sigma_k(t) = -\sigma_i(t)$, en ese caso $m(t+1) = \frac{N^+ - 1 - (N^- + 1)}{N} = \frac{N^+ - N^- - 2}{N}$
3. Si $\sigma_i(t) = -\sigma_j(t)$, en ese caso m se mantiene constante.

La siguiente proposición da las probabilidades de ocurrencia de cada una de esas tres transiciones:

Proposición 1.4.1.

$$Prob\left\{m(t) \to m(t) + \frac{2}{N}\right\} = \frac{(1-m^2)}{8} \cdot \left(1 + m + \frac{1+3m}{N}\right) + O(\frac{1}{N^2}) \tag{1.8}$$

$$Prob\left\{m(t) \to m(t) - \frac{2}{N}\right\} = \frac{(1-m^2)}{8} \cdot \left(1 - m + \frac{1-3m}{N}\right) + O(\frac{1}{N^2}) \tag{1.9}$$

$$Prob\left\{m(t) \to m(t)\right\} = 1 - \frac{(1-m^2)}{4} \cdot \left(1 + \frac{1}{N}\right) + O\left(\frac{1}{N^2}\right) \tag{1.10}$$

Note que, contrariamente al modelo de Ochrombel, $P(m \to m + 2/N) \neq P(m \to m + 2/N)$. Esta simetría se va a traducir en la ecuación del modelo continuo por una transición de fase a $m = 0$.

Demostración. Probemos primero (1.8). Por definición del modelo "Dos contra uno", un agente modifica su opinión de -1 a 1 si se eligen sucesivamente dos agentes con opinión 1 y luego un tercer agente con opinión -1. Recordando que la red es completa, obtenemos entonces

$$Prob\left\{m(t) \to m(t) + \frac{2}{N}\right\} = \frac{N^+(t)}{N}\frac{N^+(t)-1}{N-1}\frac{N^-(t)}{N-2}.$$

Como $N^{\pm} = \frac{N}{2}(1 \pm m)$ by (1.1), tenemos $\frac{N^+N^-}{N} = \frac{N}{4}(1-m^2)$ por lo que

$$Prob\left\{m(t) \to m(t) + \frac{2}{N}\right\} = \frac{(1-m^2)}{4} \cdot \frac{N}{N-1} \cdot \frac{N^+-1}{N-2}.$$

Observemos que

$$\begin{aligned}
\frac{N^2}{(N-1)(N-2)} &= \frac{(N-1)(N-2)+3N-2}{(N-1)(N-2)} \\
&= 1 + \frac{3N}{(N-1)(N-2)} + O(1/N^2) \\
&= 1 + 3N(N^{-2} + O(N^{-3})) + O(N^{-2}) \\
&= 1 + \frac{3}{N} + O(N^{-2})
\end{aligned}$$

Luego

$$\frac{N}{(N-1)(N-2)} = \frac{1}{N} + \frac{3}{N^2} + O(N^{-3}).$$

Como $(N^+ - 1) = \frac{N}{2}(m+1) - 1$, obtenemos entonces que

$$Prob\left\{m(t) \to m(t) + \frac{2}{N}\right\} = \frac{(1-m^2)}{8}\Big(N(m+1) - 2\Big)\Big(\frac{1}{N} + \frac{3}{N^2} + O(N^{-3})\Big).$$

Obtenemos (1.8) desarrollando.

La prueba de (1.9) es similar. Recordando que un agente modifica su opinión de 1 a -1 si se eligen sucesivamente dos agentes con opinión -1 y luego un tercer agente con opinión 1, tenemos

$$Prob\left\{m(t) \to m(t) - \frac{2}{N}\right\} = \frac{N^-(N^- - 1)N^+}{N(N-1)(N-2)}.$$

Procediendo como antes podemos escribir

$$Prob\left\{m(t) \to m(t) - \frac{2}{N}\right\} = \frac{1-m^2}{4}(N^- - 1)\frac{N}{(N-1)(N-2)}$$

$$= \frac{1-m^2}{4}\Big(\frac{N}{2}(1-m) - 1\Big)\Big(\frac{1}{N} + \frac{3}{N^2} + O(N^{-3})\Big)$$

Concluimos desarrollando esta expresion.

Finalmente se obtiene (1.10) haciendo 1 − (1.8) − (1.9). □

1.4.3. El límite continuo del modelo "Dos contra uno".

Como en el estudio del modelo de Ochrombel, notamos $u_N(m,t)$ la probabilidad de hallar el sistema en el estado m a tiempo t cuando hay N agentes. Vamos a ver que, después de acelerar el tiempo adecuadamente, la medida límite $u(m,\tau)$ debería verificar la ecuación de transporte

$$\frac{\partial u(m,\tau)}{\partial \tau} = -\partial_m(m(1-m^2)u(m,\tau)). \tag{1.11}$$

Como antes, obtendremos esta ecuación usando la formulación débil de la ecuación maestra. En el Apéndice, dedujimos la solución de la ecuación (1.11). Llamamos T_t a la solución de:

$$\begin{cases} \partial_t T_t(x) = (1 - T_t^2(x))T_t(x) & t \in \mathbb{R},\ x \in \mathbb{R}^n, \\ T_0(x) = x & x \in \mathbb{R}^n. \end{cases} \tag{1.12}$$

Entonces, dada una medida inicial u_0, la solucion $u(\tau)$ es la medida $u(\tau) = T_t \# u_0$. Si u_0 es una función suave, entonces obtenemos:

$$u(t,m) = \frac{e^{-t}}{(1 + m^2e^{-2t} - x^2)^{\frac{3}{2}}} \cdot u_0\Big(\frac{me^{-t}}{\sqrt{1 + m^2e^{-2t} - m^2}}\Big). \tag{1.13}$$

Observamos que si inicialmente la magnetización es m_0 con probabilidad 1, es decir, $u_0 = \delta_{m_0}$, entonces $u(\tau) = \delta_{T_\tau(m_0)}$. Luego, si $m_0 > 0$, entonces $m(\tau) \to 1$. En el caso $m_0 < 0$, $m(\tau) \to -1$

Deducción de la ecuación límite.

Recordemos la formulación débil de la ecuación maestra que obtuvimos antes:

$$\frac{d}{dt}\int \phi(m)\, du_N(m,t) = \int \Big\{ \sum_{m'} \Big(\phi(m') - \phi(m)\Big) Prob(m \to m') \Big\}\, du_N(m,t) \tag{1.14}$$

Al igual que en el modelo de Ochrombel, en el modelo "Dos contra uno"luego de cada interacción la magnetización $m(t)$ puede incrementarse o disminuirse en $\frac{2}{N}$ o bien, mantenerse constante. Reescribiendo (1.14):

$$\begin{aligned}\frac{d}{dt}\int \phi(m)\, du_N(m,t) \quad = \quad & \int \Big(\phi(m+\frac{2}{N}) - \phi(m)\Big) Prob(m \to m+\frac{2}{N}) \\ & +\Big(\phi(m-\frac{2}{N}) - \phi(m)\Big) Prob(m \to m-\frac{2}{N})\, du_N(m,t)\end{aligned}$$

Haciendo un desarrollo de Taylor de orden 1 de ϕ, obtenemos

$$\begin{aligned}&\frac{d}{dt}\int \phi(m)\, du_N(m,t) \\ &= \frac{2}{N}\int \phi'(m)\Big(Prob(m \to m+\frac{2}{N}) - Prob(m \to m-\frac{2}{N})\Big)\, du_N(m,t) + O(N^{-2}).\end{aligned}$$

En vista de las probabilidades de transición (1.8)-(1.9),

$$\begin{aligned}&\frac{d}{dt}\int \phi(m)\, du_N(m,t) \\ &= \frac{2}{N}\int \phi'(m)\Big(\frac{1}{4}m(1-m^2)(1+\frac{3}{N})\Big)\, du_N(m,t) + O(N^{-2}) \\ &= \frac{1}{2N}\int \phi'(m) m(1-m^2)\, du_N(m,t) + O(N^{-2}).\end{aligned}$$

Aceleramos ahora el tiempo considerando $t = 2N\tau$ y $u_N(m,\tau) := u_N(m,t)$. Luego

$$\frac{d}{d\tau}\int \phi(m)\, du_N(m,\tau) = \int \phi'(m) m(1-m^2)\, du_N(m,\tau) + O(N^{-2}).$$

El límite $u(m,\tau)$ de las $u_N(m,\tau)$ debería por lo tanto verificar

$$\frac{d}{d\tau}\int \phi(m)\, du_N(m,\tau) = \int \phi'(m) m(1-m^2)\, du_(m,\tau)$$

que es la formulación débil de (1.11).

Capítulo 2

Modelos cinéticos de formación de opinión

El estudio de sistemas económicos, sociales y biológicos a través de los típicos métodos de la mecánica estadística, en particular de la teoría cinética, es un campo relativamente nuevo de la matemática aplicada. La idea principal se basa en que el comportamiento macroscópico de un sistema compuesto por un gran número de agentes resulta de interacciones simples al nivel microscópico entre los agentes lo que permite su modelado matemático usando la teoría cinética de gases rarificados.

La teoría cinética de gases rarificados ha proporciónado un exitoso intento para describir las propiedades macroscópicas relevantes de un gas compuesto por una gran cantidad de partículas que obedecen las leyes de la mecánica clásíca. Resolver las ecuaciones de movimiento de cada una de las partículas es una tarea ardua y poco eficiente ya que se debería resolver un sistema acoplado de numerosas ecuaciones no-lineales, una por partícula. El objetivo de la teoría cinética es considerar la evolución, no de una partícula, sino de un sistema de partículas idénticas desde un punto de vista estadístico.

Una sociedad humana puede pensarse, en primera instancia, como un sistema compuesto por un gran cantidad de agentes todos idénticos intercambiando opinión entre sí. Se observa empíricamente que de estas interacciones microscópicas surge una distribución de las opiniones al nivel poblacional. Por ejemplo, en ciertas situaciones, personas inicialmente en desacuerdo pueden llegar al consenso o al revés. El paralelismo con la teoría cinética de gases llevo a proponer modelos de formación de opinión usando herramientas de la mecánica estadística con el propósito de identificar las reglas de interacción microscópicas que permiten obtener comportamientos mas o menos reales al nivel macroscópico.

Investigaciones en ciencias sociales resaltan, entre otros, dos mecanismos importantes en el proceso de formación de opinión en un grupo:

1. compromiso: dos agentes con opiniones distintas tienen tendencia a acercarse uno al otro,

2. difusión: un agente puede modificar levemente su opinión solo (de forma aleatoria o debido al acceso a fuentes de información por ejemplo).

Siguiendo el paper de G. Toscani [11], consideraremos en este capítulo modelos de formación de opinión basados en las interacciones entre dos agentes considerando los aspectos de *compromiso* y *difusión* en el intercambio de opinión entre los individuos. Después de describir las reglas de interacción que modelan los fenómenos de compromiso y difusión, introduciremos la ecuación de tipo Boltzmann que describe la evolución de la densidad de opiniones. Veremos finalmente que con un reescale apropiado del tiempo y de los parámetros del modelo, se puede describir el comportamiento en tiempo grande de la densidad de opinión en la población mediante una ecuación de tipo Fokker-Planck que incluye como casos particulares las ecuaciones obtenidas en el capítulo anterior como lı́mite continuo de los modelos de Ochrombel y «dos contra uno».

2.1. Descripción del modelo

2.1.1. Reglas de interacción microscópicas

Representaremos la opinión de un agente por un número real $w \in [-1, 1]$ por lo que 1 y -1 son las opiniones extremas. Sea $I = [-1, 1]$ el intervalo de las posibles opiniones. Desde un punto de vista microscópico, suponemos que las interacciones son binarias en el sentido que se efectúan entre dos agentes. Las modelamos por las siguientes reglas:

$$\begin{aligned} w' &= w - \gamma P(|w|)(w - w_*) + \eta D(|w|) \\ w'_* &= w_* - \gamma P(|w_*|)(w_* - w) + \eta_* D(|w_*|) \end{aligned} \tag{2.1}$$

donde $(w, w_*) \in I \times I$ son las opiniones de los dos agentes antes de interactuar y (w', w'_*) luego de intercambiar información entre ellos.

En (2.1), $\gamma \in (0, \frac{1}{2})$ es una constante midiendo la tendencia al compromiso y η, η_* son variables aleatorias independientes idénticamente distribuidas con varianza σ^2 y esperanza 0 que representan una difusión intrínseca de la opinión. Veremos mas adelante el tipo de restricción que conviene poner sobre η y η_* para asegurar que $w', w'_* \in I$. Finalmente, las funciones P y D describen la relevancia local del compromiso y la difusión para una opinión dada. Su presencia esta asociada a la hipótesis de que la habilidad de cambiar de opinión esta asociada a la opinión y decrece a medida que se acerca a opiniones extremas modelando así el hecho de que, a priori, las opiniones extremas son mas difíciles de modificar. Se pueden considerar por ejemplo las funciones $1 - w^2$ y $\sqrt{1 - w^2}$. En general supondremos que $P(|w|), D(|w|) \in [0, 1]$ y que son no crecientes respecto a $|w|$.

Dada una distribución inicial de opiniones en la población, la pregunta general de los modelos de formación de opinión es de determinar el comportamiento cuando $t \to \infty$ de la distribución de opiniones. Para eso notamos $f(t, w)$ a la proporción de agentes en la población con opinión w. Conociendo $f(0, .)$, queremos estudiar $f(t, .)$ para t grande.

2.1.2. Ecuación tipo Boltzmann

La función f verifica la siguiente ecuación de tipo Boltzmann: para toda función suave $\phi \in C^\infty([-1, 1])$,

$$\begin{aligned} &\frac{d}{dt} \int_I \phi(w) f(t, w)\, dw \\ &= \int_{I^2} \int_{B^2} \beta_{(w,w_*) \to (w',w'_*)} f(t, w) f(t, w_*)(\phi(w') - \phi(w)) dw_* dw d\eta d\eta_*, \end{aligned} \tag{2.2}$$

donde $\beta_{(w,w_*) \to (w',w'_*)}$ es la función de transición, cuya forma precisaremos mas adelante, que define las transiciones admisibles y su probabilidad de ocurrencia, y B es el soporte de η y η^*.

Esta ecuación describe la variación temporal de la cantidad macroscópica $\int_I \phi(w) f(t,w)\,dw$, que se puede pensar como el valor promedio de ϕ al tiempo t, como resultado de todos los posibles incrementos $\phi(w') - \phi(w)$ que pueden ocurrir al nivel microscópico. De hecho, dicho de modo informal, $f(t,w)dw$ y $f(t,w_*)dw_*$ son las probabilidades de seleccionar un agente con opinión w y otro con opinión w_*, por lo que

$$\beta_{(w,w_*)\to(w',w'_*)} f(t,w) f(t,w_*) dw_* dw d\eta d\eta_*$$

es la probabilidad que ocurra la transición $(w,w_*) \to (w',w'_*)$. Este razonamiento es el mismo que hicimos en el capítulo anterior para deducir los límites de los modelos de Ochrombel y «2 contra 1».

Es fácil ver que una formulación equivalente a (2.2) es

$$\begin{aligned} (2.3) \quad & \frac{d}{dt}\int_I \phi(w) f(t,w) dw \\ &= \frac{1}{2}\int_{I^2}\int_{B^2} \beta_{(w,w_*)\to(w',w'_*)} f(t,w) f(t,w_*)(\phi(w') + \phi(w'_*) - \phi(w) - \phi(w_*)) dw_* dw d\eta d\eta_*. \end{aligned}$$

2.1.3. Marco teórico

Notemos que es razonable suponer que la distribución $f(t,.)$ de opiniones en la población al tiempo t es una medida y no simplemente una función ya que debemos poder contemplar, por ejemplo, el caso en el que todos los agentes opinen lo mismo, en este caso $f(t,.)$ es una masa de Dirac. Por este motivo trabajaremos con funciones $f : [0,+\infty) \to P(I)$ donde $P(I)$ es el conjunto de medidas de probabilidad sobre I con la topología débil de medidas.

De aquí en mas, supondremos que la variable aleatoria η (lo mismo para η^*) es de la forma $\eta = \sigma Y$ para une variable aleatoria Y tal que (i) $E[Y] = 0$, (ii) $V[Y] = 1$, (iii) Y tiene su tercer momento finito en el sentido que $E[|Y|^3] < \infty$, (iv) Y es simétrica. Note que η es también simétrica, tiene su tercer momento finito, tiene esperanza cero y varianza σ^2. Notamos Θ la densidad de η. Notemos finalmente que para $p \leq 3$,

$$\int_{\mathbb{R}} |\eta|^p \; d\Theta(\eta) = E[|\eta|^p] = E[|\sigma Y|^p] = \sigma^p E[|Y|^p].$$

Definición 2.1.1. *Diremos que $f : [0,+\infty) \to P(I)$ es solución de (2.2) con condición inicial $f_0 \in P(I)$ dada si (i) $f \in C^1((0,+\infty), P(I)) \cap C([0,+\infty), P(I))$, (ii) f verifica (2.2) para toda $\phi \in C^\infty(I)$ y (iii) $f(0) = f_0$.*

Observación 2.1.2. *En [11] Toscani considera como clase de función test las funciones $\phi \in C^{2,\delta}(I)$ tales que $\phi(\pm 1) = \phi'(\pm 1) = 0$. Pensamos que esta última condición $\phi(\pm 1) = \phi'(\pm 1) = 0$ no es necesaria ya que, al parecer, se usa unicamente en (2.11) al momento de integrar por partes (las condiciones de borde desaparecen) y también porque*

Toscani mismo usa como funciones test $\phi(v) = 1$ y $\phi(v) = v$ que no cumplen con lo requerido. Por este motivo no pedimos ninguna condición de borde sobre las funciones test.

Admitimos la existencia y unicidad de una solución f en el sentido de la definición anterior. para concentrarnos en analizar el comportamiento asíntótico de $f(t)$ cuando $t \to +\infty$. Antes de eso hacemos unos comentarios sobre el soporte B de Θ, la densidad de η, y luego sobre la evolución del promedio y de la varianza de f.

2.1.4. Algunos comentarios sobre el rango de η

Es primordial que las reglas de interacción produzcan nuevas opiniones w', w'_* que sigan perteneciendo a I. Una forma de asegurar eso consiste en tomar el nucleo β tal que $\beta_{(w,w_*)\to(w',w'_*)} = 1_{w'\in I}1_{w'_*\in I}$. El problema ahora que eso complica el análisis de la ecuación tipo Boltzmann resultante (2.2). Vamos a ver en cambio que en ciertos casos explicitos de D podemos elegir el soporte de Θ de forma que $w', w'_* \in I$.

Primero note que en el caso de ausencia de difusión, los límites laterales no son violados ya que w' y w'_* se escriben como combinaciones convexas de w y w_*:

$$\begin{aligned} w' &= (1-\gamma P(|w|)) + \gamma P(|w|)w_* \\ w'_* &= (1-\gamma P(|w_*|)) + \gamma P(|w_*|)w \end{aligned}$$

Damos ahora un ejemplo explícito de coeficiente de difusión D y de soporte B de Θ para el que $w', w'_* \in I$. Fijemos $D(|w|) = 1 - |w|$. Como

$$\begin{aligned} w' &= (1-\gamma P(|w|))w + \gamma P(|w|)w_* + (1-|w|)\eta \\ &\leq (1-\gamma P(|w|))w + \gamma P(|w|) + (1-|w|)\eta. \end{aligned}$$

Por lo tanto, para garantizar $w' \leq 1$ es suficiente que

$$(1-\gamma P(|w|))w + \gamma P(|w|) + (1-|w|)\eta \leq 1$$

para todo $w \in I$, es decir

$$(1-|w|)\eta \leq (1-\gamma P(|w|))(1-w).$$

Como $0 \leq P(|w|) \leq 1$ basta pedir $\eta \leq 1-\gamma$. De la misma forma, vemos que si $\eta \geq -(1-\gamma)$, entonces $w' \geq -1$. Por lo tanto, si $D(|w|) = 1-|w|$ y $B = (-(1-\gamma), 1-\gamma)$, tendremos que $w', w'_* \in [-1, 1]$.

2.1.5. Evolución del promedio y de la varianza de $f(t)$

Evolución del valor medio de $f(t)$

Notemos $m(t) = \int_I wf(t,w)dw$ el valor medio de la medida $f(t)$, es decir, el valor de la opinón promedio en la población al tiempo t. Para estudiar la evolución de m tomemos $\phi(w) = w$ en la formulación débil (2.2). Obtenemos

$$\begin{aligned} m'(t) &= \int_{I^2}\int_{B^2}(w'-w)f(t,w_*)f(t,w)dw_*dwd\Theta(\eta)d\Theta(\eta_*) \\ &= \int_{I^2}\int_{B}(w'-w)f(t,w_*)f(t,w)dw_*dwd\Theta(\eta). \end{aligned}$$

Reescribimos el miembro derecho usando las reglas de interacción (2.1):

$$\begin{aligned} m'(t) &= \gamma\int_{I^2}\int_{B^2}P(|w|)(w_*-w)f(t,w_*)f(t,w)dw_*dwd\Theta(\eta)d\Theta(\eta_*) \\ &\quad +\int_{I^2}\int_{B^2}\eta D(|w|)f(t,w_*)f(t,w)dw_*dwd\Theta(\eta)d\Theta(\eta_*) \\ &= \gamma\int_{I^2}P(|w|)(w_*-w)f(t,w_*)f(t,w)dw_*dw \\ &\quad +\int_I D(|w|)f(t,w)dw\int_I f(t,w_*)\,dw_*\int_B\eta d\Theta(\eta). \end{aligned}$$

Note que la segunda integral en el miembro derecho vale cero ya que $\int_B \eta d\Theta(\eta) = E[\eta] = 0$. Luego

$$m'(t) = \gamma\int_{I^2}P(|w|)(w_*-w)f(t,w_*)f(t,w)dw_*dw. \tag{2.4}$$

Observe que en el caso $P\equiv 1$, la integral se anula ya que la podemos reescribir como

$$\begin{aligned} &\gamma\int_{I^2}w_*f(t,w_*)f(t,w)dw_*dw-\gamma\int_{I^2}wf(t,w_*)f(t,w)dw_*dw \\ &=\gamma\int_I w_*f(t,w_*)dw_*-\gamma\int_I wf(t,w)dw \\ &=0. \end{aligned}$$

Concluímos entonces que en el caso $P\equiv 1$, la opinión media $m(t)$ se conserva.

Evolución del segundo momento de $f(t)$.

Estudiemos ahora la evolución del segundo momento $\int_I w^2 f(t,w)\,dw$ de $f(t)$. Para esto tomamos $\phi(w) = w^2$ en (2.2). Obtenemos:

$$\frac{d}{dt}\int_I w^2 f(w,t)dw = \int_{I^2}\int_{B^2}({w'}^2-w^2)f(t,w)f(t,w_*)dw_*dwd\Theta(\eta)d\Theta(\eta_*).$$

Usando las reglas de interacción (2.1) vemos que

$$
\begin{aligned}
&\frac{d}{dt}\int_I w^2 f(w,t)dw \\
&= \gamma^2 \int_{I^2}\int_{B^2} (w-w_*)^2 P(|w|)^2 f(t,w)f(t,w_*)dw_* dw d\Theta(\eta) d\Theta(\eta_*) \\
&\quad + \int_{I^2}\int_{B^2} \eta^2 D(|w|)^2 f(t,w)f(t,w_*)dw_* dw d\Theta(\eta) d\Theta(\eta_*) \\
&\quad -2\gamma \int_{I^2}\int_{B^2} wP(|w|)(w-w_*) f(t,w)f(t,w_*)dw_* dw d\Theta(\eta) d\Theta(\eta_*) \\
&\quad +2 \int_{I^2}\int_{B^2} w\eta D(|w|) f(t,w)f(t,w_*)dw_* dw d\Theta(\eta) d\Theta(\eta_*) \\
&\quad -2\gamma \int_{I^2}\int_{B^2} \eta P(|w|)(w-w_*) D(|w|) f(t,w)f(t,w_*)dw_* dw d\Theta(\eta) d\Theta(\eta_*) \\
&= I + II + III + IV + V.
\end{aligned}
$$

Analicemos cada termino. Primero

$$
I = \gamma^2 \int_{I^2} (w-w_*)^2 P(|w|)^2 f(t,w) f(t,w_*) dw_* dw.
$$

De la misma forma,

$$
III = -2\gamma \int_{I^2} wP(|w|)(w-w_*) f(t,w) f(t,w_*) dw_* dw.
$$

Para el segundo término, recordemos primero que $\int \eta^2 \, d\Theta(\eta) = E[\eta^2] = \sigma^2 E[Y^2] = \sigma^2$. Luego

$$
\begin{aligned}
II &= \int_I D(|w|)^2 f(t,w)\, dw \int_B \eta^2 \, d\Theta(\eta) \\
&= \sigma^2 \int_I D(|w|)^2 f(t,w)\, dw.
\end{aligned}
$$

El 4to y el 5to término se anulan ya que $\int_B \eta \, d\Theta(\eta) = E[\eta] = 0$. Concluimos finalmente

$$
\begin{aligned}
\frac{d}{dt}\int_I w^2 f(w,t)dw &= \gamma^2 \int_{I^2} (w-w_*)^2 P(|w|)^2 f(t,w) f(t,w_*) dw_* dw \\
&\quad -2\gamma \int_{I^2} wP(|w|)(w-w_*) f(t,w) f(t,w_*) dw_* dw \\
&\quad +\sigma^2 \int_I D(|w|)^2 f(t,w)\, dw.
\end{aligned}
$$

Supongamos ahora que $P \equiv 1$. En este caso se puede reescribir la ecuación anterior como

$$\begin{aligned}
&\frac{d}{dt}\int_I w^2 f(w,t)dw \\
&= \gamma(\gamma-2)\int_{I^2} w^2 f(t,w)f(t,w_*)dw_*dw + \gamma^2\int_{I^2} w_*^2 f(t,w)f(t,w_*)dw_*dw \\
&\quad +2\gamma(1-\gamma)\int_{I^2} ww_* f(t,w)f(t,w_*)dw_*dw + \sigma^2\int_I D(|w|)^2 f(t,w)\,dw.
\end{aligned}$$

Note que la dos primeras integrales en el miembro derecho son iguales y que la tercera es

$$\int_{I^2} ww_* f(t,w)f(t,w_*)dw_*dw = \int_I wf(t,w)dw \int_I w_* f(t,w_*)dw_* = m(t)^2$$

donde $m(t) = \int_I wf(t;w)\,dw$ es el valor promedio de $f(t)$. Obtenemos entonces

$$\frac{d}{dt}\int_I w^2 f(w,t)dw \;\; = -2\gamma(1-\gamma)\Big[\textstyle\int_I w^2 f(w,t)dw - m(t)^2\Big] + \sigma^2\int_I D(|w|)^2 f(w,t)dw.$$

Vimos en la sección anterior que si $P \equiv 1$ entonces $m(t)$ es constante. Obtenemos entonces

$$\begin{aligned}
&\frac{d}{dt}\int_I w^2 f(w,t)dw \\
&= -2\gamma(1-\gamma)\left[\int_I w^2 f(w,t)dw - m(0)^2\right] + \sigma^2\int_I D(|w|)^2 f(w,t)dw.
\end{aligned} \tag{2.5}$$

En la ausencia de difusión, es decir, si $\sigma = 0$ obtenemos que

$$\frac{d}{dt}\int_I w^2 f(w,t)dw \;\; = -2\gamma(1-\gamma)\left[\textstyle\int_I w^2 f(w,t)dw - m(0)^2\right]$$

Notemos que $Var(f(t)) = \int_I (w-m(t))^2 f(w,t)dw = \int_I w^2 f(w,t)dw - m(0)^2$ la varianza de $f(t)$. La ecuación anterior se puede reescribir como

$$\frac{d}{dt}Var(f(t)) = -2\gamma(1-\gamma)Var(f(t))$$

Integrando esta EDO deducimos

$$Var(f(t)) = Var(f(0))e^{-2\gamma(1-\gamma)t}.$$

En particular, la varianza de $f(t)$ va a 0 cuando $t \to +\infty$ ya que $\gamma \in (0,1/2)$. El comportamiento asíntótico de $f(t)$ cuando $t \to +\infty$ es en este caso particularmente simple:

Proposición 2.1.3. *En el caso $P \equiv 1$ y $\eta = 0$, tenemos $f(t) \rightharpoonup \delta_{m(0)}$ cuando $t \to +\infty$, donde $m(0)$ es la opinión promedio inicial. Entonces todos los agentes de la población tienden a tener la misma opinón que es la opinión promedio inicial.*

Note que se puede intuir este resultado fácilmente a partir de las reglas de interacción (2.1). En el caso $P \equiv 1$ y $\eta = 0$, estas reglas se escriben

$$\begin{aligned} w' &= w - \gamma(w - w_*) \\ w'_* &= w_* - \gamma(w_* - w) \end{aligned}$$

En particular

$$|w' - w'_*| = (1 - 2\gamma)|w - w_*|.$$

Entonces después de una interacción entre dos agentes cualquiera, la distancia entre las opiniones de los agentes que interactúan decrece estrictamente ya que $\gamma \in (0, 1/2)$. De ahi es razonable conjeturar que todos los agentes tienden a pensar los mismo asintóticamente.

Demostración. Como I es compacto, $P(I)$ es compacto por lo que existe una subsucesión $t_n \to +\infty$ y una medida $f_\infty \in P(I)$ tales que $f(t_n) \to f_\infty$ en el sentido que

$$\int_I \phi(w) f(t_n, w) dw \to \int_I \phi \, df_\infty$$

para toda $\phi \in C(I)$. Tomando $\phi(w) = w$ y luego $\phi(w) = w^2$, obtenemos la convergencia del promedio y de la varianza de las $f(t_n)$ hacia el promedio y la varianza de f_∞. Deducimos que f_∞ tiene promedio $m(0)$ y varianza cero por lo que $f_\infty = \delta_{m(0)}$. Luego cualquier subsucesión convergente de $f(t)$ converge a $\delta_{m(0)}$. La proposición sigue. □

El hecho de que la varianza $Var(f(t))$ de $f(t)$ converge a cero exponencialmente sugiere que $f(t)$ debería converger a $\delta_{m(0)}$ exponencialmente también. Eso se puede ver muy fácilmente usando la distancia de Monge-Kantorovich W_2 (ver Apéndice B.2)

Proposición 2.1.4. *Se tiene que $W_2(f(t), \delta_{m(0)}) \leq \sqrt{Var(f(0))} e^{-\gamma(1-\gamma)t}$.*

Demostración. Por definición de W_2, tenemos

$$W_2(f(t), \delta_{m(0)})^2 = \text{ínf} \int_{I^2} |w - w'|^2 \, d\pi(w, w')$$

donde el ínfimo se toma sobre todas las $\pi \in P(I \times I)$ con marginal $f(t)$ y $\delta_{m(0)}$. Tomamos $\pi = f(t) \otimes \delta_{m(0)}$. Entonces

$$\begin{aligned} W_2(f(t), \delta_{m(0)})^2 &\leq \int_I |w - m(0)|^2 \, f(t, w) dw = \int_I |w - m(t)|^2 \, f(t, w) dw \\ &= Var(f(t)) \leq Var(f(0)) e^{-2\gamma(1-\gamma)t}. \end{aligned}$$

□

Aclaramos que las dos propocisiones anteriores no aparecen en el paper [11] de Toscani.

El propósito de la siguiente sección es obtener informaciones sobre el comportamiento asíntótico de $f(t)$ cuando $t \to +\infty$ en el caso general.

2.2. Ecuación de Fokker-Planck como "grazing limit" de la ecuación de Boltzmann.

2.2.1. Introducción

Asumiremos de ahora en adelante que $P \equiv 1$. Tenemos entonces conservación de masa y del promedio. Consideremos ahora el caso en el que las interacciones entre los individuos producen un pequeño cambio en la opinión de los mismos, o sea $\gamma \to 0$ y $\sigma \to 0$. Este procedimiento, en la literatura sobre la ecuación de Boltzmann, se llama "grazing limit" .

Para intuir el tipo de relación entre σ y γ que conviene mantener cuando $\gamma, \sigma \to 0$, consideramos la varianza de f. Sabemos por (2.5) que la varianza $Var(f(t))$ de $f(t)$ satisface

$$\frac{d}{dt}Var(f(t)) = -2\gamma(1-\gamma)Var(f(t)) + \sigma^2 \int_I D(|w|)^2 f(w,t)dw.$$

Para seguir observando el decaimiento a cero de $Var(f(t))$ en ausencia de difusión (es decir $\eta = 0$) cuando $\gamma \sim 0$ conviene reescribir esta expresión como

$$\frac{1}{\gamma}\frac{d}{dt}Var(f(t)) = -2(1-\gamma)Var(f(t)) + \frac{\sigma^2}{\gamma} \int_I D(|w|)^2 f(w,t)dw.$$

Cambiamos ahora de escala temporal considerando $\tau := \gamma t$ y $g(\tau) := f(t)$. Como $\frac{d}{d\tau} = \frac{1}{\gamma}\frac{d}{dt}$, podemos reescribir la ecuación anterior como

$$\frac{d}{d\tau}Var(g(\tau)) = -2(1-\gamma)Var(g(\tau)) + \frac{\sigma^2}{\gamma} \int_I D(|w|)^2 g(w,\tau)dw.$$

Se ve ahora a partir de esta expresión que el cociente σ^2/γ es fundamental para estudar el comportamiento asíntótico de g cuando $\gamma, \sigma \to 0$.

Note que la nueva escala temporal con τ es mayor a t ya que $\frac{\tau}{t} = \gamma \to 0$ o sea $t \gg \tau$. Tal cambio de escala temporal era de esperar: como los cambios después de cada interacción son infinitesimalmente chicos, es necesario esperar mucho tiempo para ver un resultado al nivel macroscópico.

Observe también que por este motivo, el comportamiento de g cuando $\tau \to +\infty$ da el comportamiento de f cuando $t \to +\infty$. La ventaja del procedimiento de grazing limit es que g verifica ya no una ecuación tipo Boltzmann como f sino una ecuación de Fokker-Planck. Es el objeto de las siguientes secciones.

2.2.2. Deducción informal de la ecuación de Fokker-Plack.

Sabemos que $g(v,\tau) = f(v,t)$ satisface

$$\begin{aligned}&\frac{d}{d\tau}\int_I \phi(w)g(w,\tau)dw \\ &= \frac{1}{\gamma}\int_{I^2}\int_{B^2}(\phi(w') - \phi(w))g(w,\tau)g(w_*,\tau)dw_*dwd\Theta(\eta)d\Theta(\eta_*).\end{aligned} \tag{2.6}$$

Seguimos suponiendo por el momento que $P \equiv 1$. Las reglas de interacción (2.1) son entonces

$$\begin{aligned} w' &= w - \gamma(w - w_*) + \eta D(|w|) \\ w'_* &= w_* - \gamma(w_* - w) + \eta D(|w_*|) \end{aligned}$$

Recordamos que $\gamma, \sigma \sim 0$ por lo que w' y w son muy parecidos. Tiene sentido entonces aproximar el incremento $\phi(w') - \phi(w)$ que aparece en la ecuación de Boltzmann (2.6) por su desarrollo de Taylor de orden 2. Olvidándonos por el momento del resto tenemos entonces

$$\begin{aligned}&\phi(w') - \phi(w) \simeq \phi'(w)(w'-w) + \frac{1}{2}\phi''(w)(w'-w)^2 \\ &= -\gamma\phi'(w)(w-w_*) + \phi'(w)\eta D(|w|) + \frac{1}{2}\phi''(w)(\gamma(w_*-w) + \eta D(|w|))^2\end{aligned} \tag{2.7}$$

Reemplazando en (2.6) obtenemos

$$\begin{aligned}&\frac{d}{d\tau}\int_I \phi(w)g(\tau,w)dw \simeq \\ &\int_{I^2}\int_{B^2}(w_*-w)\phi'(w)g(\tau,w)g(\tau,w_*)dw_*dwd\Theta(\eta)d\Theta(\eta_*) \\ &+ \frac{1}{\gamma}\int_{I^2}\int_{B^2}\eta D(|w|)\phi'(w)g(\tau,w)g(\tau,w_*)dw_*dwd\Theta(\eta)d\Theta(\eta_*) \\ &+ \frac{1}{2\gamma}\int_{I^2}\int_{B^2}\left(\gamma(w_*-w) + \eta D(|w|)\right)^2\phi''(w)g(\tau,w)g(\tau,w_*)dw_*dwd\Theta(\eta)d\Theta(\eta_*) \\ &= I + II + III.\end{aligned} \tag{2.8}$$

Analicemos ahora cada término. Primero, si notamos $m(\tau) = \int_I w_*g(\tau,w_*)dw_*$ el valor promedio de $g(\tau)$, tenemos

$$\begin{aligned} I &= \int_I w_*g(\tau,w_*)dw_*\int_I\phi'(w)g(\tau,w)dw - \int_I w\phi'(w)g(\tau,w)dw \\ &= m(\tau)\int_I\phi'(w)g(\tau,w)dw - \int_I w\phi'(w)g(\tau,w)dw \\ &= \int_I(m(\tau)-w)\phi'(w)g(\tau,w)dw. \end{aligned}$$

Luego como $\int_B \eta \, d\Theta(\eta) = E[\eta] = 0$ tenemos $II = 0$. Finalmente,

$$\begin{aligned} III &= \frac{\gamma}{2}\int_{I^2}(w_* - w)^2\phi''(w)g(\tau,w)g(\tau,w_*)dw_*dw \\ &+ \left(\int_B \eta\Theta(\eta)d\eta\right)\left(\int_{I^2}(w_* - w)D(|w|)\phi''(w)g(\tau,w)g(\tau,w_*)dw_*dw\right) \\ &+ \frac{1}{2\gamma}\int_B \eta^2 d\Theta(\eta)\int_{I^2} D^2(|w|)\phi''(w)g(\tau,w)dw \end{aligned}$$

Recordando que Θ tiene experanza cero y varianza es σ^2, obtenemos

$$III = \frac{\gamma}{2}\int_{I^2}(w_* - w)^2\phi''(w)g(\tau,w)g(\tau,w_*)dw_*dw + \frac{\sigma^2}{2\gamma}\int_{I^2} D(|w|)^2\phi''(w)g(\tau,w)dw$$

Volviendo a (2.8), obtenemos finalmente que

$$\begin{aligned} &\frac{d}{d\tau}\int_I \phi(w)g(\tau,w)dw \\ &\simeq \int_I (m(\tau) - w)\phi'(w)g(\tau,w)dw + \frac{\sigma^2}{2\gamma}\int_{I^2} D(|w|)^2\phi''(w)g(\tau,w)dw \\ &\quad + \frac{\gamma}{2}\int_{I^2}(w_* - w)^2\phi''(w)g(\tau,w)g(\tau,w_*)dw_*dw. \end{aligned} \tag{2.9}$$

Luego, si $\gamma, \sigma \to 0$ de manera tal que σ^2/γ converge a algún $\lambda \geq 0$ y si $g(\tau)$, que depende de σ y γ, tiene un límite cuando $\gamma, \sigma \to 0$, este límite debería verificar

$$\begin{aligned} &\frac{d}{d\tau}\int_I \phi(w)g(\tau,w)dw \\ &= \int_I (m(\tau) - w)\phi'(w)g(\tau,w)dw + \frac{\lambda}{2}\int_{I^2} D(|w|)^2\phi''(w)g(\tau,w)dw \end{aligned} \tag{2.10}$$

para toda $\phi \in C^\infty(I)$. Si suponemos por un instante que g es suave e integramos por partes el miembro derecho obtenemos

$$\begin{aligned} &\int_I (m(\tau) - w)\phi'(w)g(\tau,w)dw + \frac{\lambda}{2}\int_{I^2} D(|w|)^2\phi''(w)g(\tau,w)dw \\ &= \Big[(m(\tau) - w)\phi(w)g(\tau,w)\Big]_{-1}^{1} - \int_I \phi(w)\partial_w\left((m(\tau) - w)g(\tau,w)\right)dw \\ &\quad + \frac{\lambda}{2}\Big[D(|w|)^2\phi'(w)g(\tau,w)\Big]_{-1}^{1} - \frac{\lambda}{2}\Big[\phi(w)\partial_w(D(|w|)^2 g(\tau,w))\Big]_{-1}^{1} \\ &\quad + \frac{\lambda}{2}\int_{I^2}\phi(w)\partial_{ww}\left(D(|w|)^2 g(\tau,w)\right)dw \end{aligned}$$

Reordenando,

$$
\begin{aligned}
&\int_I (m(\tau)-w)\phi'(w)g(\tau,w)dw + \frac{\lambda}{2}\int_{I^2} D(|w|)^2\phi''(w)g(\tau,w)dw \\
&= \int_{I^2}\phi(w)\Big(\frac{\lambda}{2}\partial_{ww}\left(D(|w|)^2 g(\tau,w)\right) - \partial_w\left((m(\tau)-w)g(\tau,w)\right)\Big)dw \\
&\quad + \Big[\phi(w)\Big((m(\tau)-w)g(\tau,w) - \partial_w(D(|w|)^2 g(\tau,w))\Big)\Big]_{-1}^{1} \\
&\quad + \frac{\lambda}{2}\Big[D(|w|)^2\phi'(w)g(\tau,w)\Big]_{-1}^{1}
\end{aligned}
$$

Vemos entonces que (2.10) es la formulación débil de la ecuación de tipo Fokker-Planck

$$
\frac{\partial g}{\partial \tau} = \partial_w\left((w-m(\tau))g\right) + \frac{\lambda}{2}\partial_{ww}\left(D(|w|)^2 g\right) \tag{2.11}
$$

con condición de borde

$$
(m(\tau)-w)g(\tau,w) - \partial_w(D(|w|)^2 g(\tau,w)) = 0 \qquad \text{en } w=\pm 1, \tag{2.12}
$$
$$
D(|w|)^2 g(\tau,w) = 0 \qquad \text{en } w=\pm 1. \tag{2.13}
$$

Observación 2.2.1. *El 2ndo miembro de de la ecuación* (2.11) *se compone de dos términos cuyos efectos son opuestos. El 1er término* $\partial_w\left((w-m(\tau))g\right)$ *es un término de transporte que lleva* g *hacia el promedio* $m(\tau)=m(0)$ *(recuerde que el promedio se conserva en el caso* $P=1$*). Traduce la tendencia al compromiso del modelo. El 2ndo término* $\frac{\lambda}{2}\partial_{ww}\left(D(|w|)^2 g\right)$ *es un término de difusión. Traduce el hecho de cada agente a nivel microscópico puede cambiar de opinión por si mismo. La evolución de* g *es entonces el resultado de dos mecanismos opuestos: uno de concentración al promedio y el otro de difusión. Se puede esperar por lo tanto que exista en general estados estacionarios límites. Volveremos sobre este punto después.*

Observación 2.2.2. *Si consideramos* $\gamma \to 0$ *y* $\sigma \to 0$ *de forma tal que* $\frac{\sigma^2}{\gamma} = \lambda$*, podremos recuperar en el límite las contribuciones tanto de la tendencia al compromiso como de la difusión de información. Otros límites pueden resultar de interés como difusión dominante (*$\frac{\sigma^2}{\gamma}\to\infty$*) o compromiso dominante (*$\frac{\sigma^2}{\gamma}\to 0$*).*

Observación 2.2.3. *Toscani en [11] considera funciones-test* ϕ *tales que* $\phi(\pm 1) = \phi'(\pm 1) = 0$*. En este caso, los términos de borde en la integración por partes desaparecen. Como lo explicamos antes, nos parece preferible no considerar esta restricción sobre las* ϕ *por lo que aparecen las condiciones de bordes* (2.12)*. Esas son análogas a las condiciones obtenidas en [4].*

La deducción que hicimos de (2.11) no es para nada rigurosa por dos motivos. En primer lugar nos olvidamos del resto en el desarrollo de Taylor (2.7). En segundo lugar las $g(\tau)$ que consideramos dependen de γ y σ ya que las reglas de interacción dependen de γ y σ. Debemos por lo tanto ver que alguna subsucesión de las funciones $g_{\gamma,\sigma} : t \to g_{\gamma,\sigma}(t) \in P(I)$ converge en un sentido tal que nos permita pasar al límite en la ecuación de Boltzmann (2.6). Cabe precisar que Toscani en [11] no dice nada acerca de este último problema.

2.2.3. Deducción rigurosa de la ecuación de Fokker-Planck

Empezamos haciendo rigurosa la deducción de la sección anterior tomando en cuenta el resto en el desarrollo de Taylor (2.7). Escribimos

$$\phi(w') - \phi(w) = \phi'(w)(w' - w) + \frac{1}{2}\phi''(\tilde{w})(w' - w)^2$$

donde $\tilde{w} = \theta w' + (1 - \theta)w$ para algún $\theta \in (0, 1)$. Reemplazando en (2.6),

$$\begin{aligned} &\frac{d}{d\tau}\int_I \phi(w)g(w)dw = \\ &\frac{1}{\gamma}\int_{I^2}\int_{B^2}\Big[(\gamma(w_* - w) + \eta D(|w|))\phi'(w) \\ &+ \frac{1}{2}(\gamma(w_* - w) + \eta D(|w|))^2\phi''(w)\Big]g(w)g(w_*)dw_*dwd\Theta(\eta)d\Theta(\eta_*) + R(\gamma, \sigma) \end{aligned} \tag{2.14}$$

donde

$$\begin{aligned} &R(\gamma, \sigma) = \\ &\frac{1}{2\gamma}\int_{I^2}\int_{B^2}(\gamma(w_* - w) + \eta D(|w|))^2(\phi''(\tilde{w}) - \phi''(w))g(w)g(w_*)dw_*dwd\Theta(\eta)d\Theta(\eta_*) \end{aligned} \tag{2.15}$$

Como $|\tilde{w} - w| = \theta\,|w - w'|$, tenemos:

$$|\phi''(\tilde{w}) - \phi''(w)| \leq \|\phi'''\|_\infty\,|\tilde{w} - w| \leq \|\phi'''\|_\infty\,|w' - w|$$

Entonces,

$$|R(\gamma, \sigma)| \leq \frac{\|\phi'''\|_\infty}{2\gamma}\int_{I^2}\int_{B^2}|\gamma(w_* - w) + \eta D(|w|)|^3\, g(w_*)g(w)dw_*dwd\Theta(\eta)d\Theta(\eta_*)$$

Aplicando la desigualdad de convexidad $(a + b)^3 = 8(a/2 + b/2)^3 \leq 4(a^3 + b^3)$ y recordando que $D(|w|) \in [0, 1]$, $w_*, w \in [-1, 1]$ y $\gamma \in [0, 1]$,

$$\begin{aligned} |\gamma(w_* - w) + \eta D(|w|)|^3 &\leq 4\left(|\gamma(w_* - w)|^3 + |\eta D(|w|)|^3\right) \\ &\leq 32\gamma^3 + 4[\eta|^3. \end{aligned}$$

Luego

$$|R(\gamma,\sigma)| \leq \|\phi'''\|_\infty \left(16\gamma^2 + \frac{2}{\gamma}\int_B |\eta|^3 d\Theta(\eta)\right).$$

Como Θ es la densidad de σY, tenemos

$$\int_B |\eta|^3 d\Theta(\eta) = E[(\sigma Y)^3] = \sigma^3 E[|Y|^3].$$

Luego

$$|R(\gamma,\sigma)| \leq \|\phi'''\|_\infty \left(16\gamma^2 + 2\frac{\sigma^2}{\gamma}\sigma E[|Y|^3]\right). \tag{2.16}$$

Como $\sigma^2/\gamma \to \lambda$, $\gamma \to 0$ y $E[|Y|^3| < \infty$, deducimos que $R(\gamma,\sigma) \to 0$. De esta estimación y de las cuentas de la sección anterior, Toscani [11] deduce inmediatamente el siguiente teorema:

Teorema 2.2.4. *Sea $f_0 \in P(I)$ una densidad de probabilidad. Entonces, cuando $\gamma \to 0$ y $\sigma \to 0$ de manera tal que $\sigma^2 = \lambda\gamma$, la solución débil de la ecuación de Boltzman para la densidad reescalada $g_\gamma(v,\tau) = f(v,t)$, con $\tau = \gamma t$, converge, a extracción de una subsucesión, a una densidad de probabilidad $g(w,\tau)$. Dicha densidad es una solución débil de la ecuación de Fokker-Planck* (2.11)*. Además, la opinión promedio se conserva.*

Como ya dijimos Toscani no explica porque se puede extraer una subsucesión convergente de $g(\tau)$ ni tampoco detalla el sentido de la convergencia. Para aclarar este punto proponemos el siguiente argumento. A partir de ahora suponemos que $\sigma = \sigma(\gamma)$ y notamos $g_\gamma(\tau)$ la densidad reescalada. Sea $\phi \in C^3(I)$. Sabemos que $g_\gamma(\tau)$ verifica (2.6) por lo que

$$\begin{aligned}
&\int_I \phi(w) g_\gamma(\tau,w)dw \\
&= \int_I \phi(w) g_\gamma(0,w)dw \\
&\quad + \frac{1}{\gamma}\int_0^\tau \int_{I^2}\int_{B^2} (\phi(w') - \phi(w)) g_\gamma(s,w) g_\gamma(s,w_*) dw_* dw d\Theta(\eta) d\Theta(\eta_*).
\end{aligned}$$

Luego para $\tau < \tau'$,

$$\begin{aligned}
&\int_I \phi(w)(g_\gamma(\tau,w)dw - g_\gamma(\tau',w)dw) \\
&= \frac{1}{\gamma}\int_\tau^{\tau'} \int_{I^2}\int_{B^2} (\phi(w') - \phi(w)) g_\gamma(s,w) g_\gamma(s,w_*) dw_* dw d\Theta(\eta) d\Theta(\eta_*).
\end{aligned}$$

Usando el desarrollo de Taylor de $\phi(w') - \phi(w)$ y las cuentas que llevaron a (2.9), podemos escribir

$$\begin{aligned}
&\int_I \phi(w)(g_\gamma(\tau, w)dw - g_\gamma(\tau', w)dw) \\
&= \int_\tau^{\tau'} \int_I (m_\gamma(s) - w)\phi'(w)g_\gamma(s,w)dwds + \frac{\sigma^2}{2\gamma}\int_\tau^{\tau'}\int_{I^2} D(|w|)^2\phi''(w)g_\gamma(s,w)dwds \\
&\quad + \frac{\gamma}{2}\int_\tau^{\tau'}\int_{I^2}(w_* - w)^2\phi''(w)g_\gamma(s,w)g_\gamma(s,w_*)dwdw_*ds + \int_\tau^{\tau'} R(\gamma,\sigma,\phi,s)ds
\end{aligned}$$

donde $m_\gamma(s)$ es el promedio de $g_\gamma(s)$ y

$$\begin{aligned}
&R(\gamma,\sigma,s,\phi) = \\
&\frac{1}{2\gamma}\int_{I^2}\int_{B^2}\left(\gamma(w_* - w) + \eta D(|w|)\right)^2(\phi''(\tilde{w}) - \phi''(w))g_\gamma(s,w)g_\gamma(s,w_*)dw_*dwd\Theta(\eta)d\Theta(\eta_*).
\end{aligned}$$

En vista de (2.16),

$$|R(\gamma,\sigma,s,\phi)| \leq \|\phi'''\|_\infty\left(16\gamma^2 + 2\frac{\sigma^2}{\gamma}\sigma E[|Y|^3]\right). \tag{2.17}$$

Normamos $C^3(I)$ con la norma usual $\|\phi\|_{C^3(I)} = \sum_{k=0}^3 \|\phi^{(k)}\|_\infty$. Como $m_\gamma(s) \in I$, $D(|w|) \in [0,1]$ y $g(s) \in P(I)$, podemos acotar el valor absoluto de $\int_I \phi(w)(g_\gamma(\tau,w)dw - g_\gamma(\tau',w)dw)$ por

$$\begin{aligned}
& 2\|\phi'\|_\infty(\tau' - \tau) + \frac{\sigma^2}{2\gamma}\|\phi''\|_\infty(\tau'-\tau) + 2\gamma\|\phi''\|_\infty(\tau'-\tau) \\
& + \|\phi'''\|_\infty\left(16\gamma^2 + 2\frac{\sigma^2}{\gamma}\sigma E[|Y|^3]\right)(\tau'-\tau) \\
\leq \; & \left(2 + \frac{\sigma^2}{2\gamma} + 2\gamma + 16\gamma^2 + 2\frac{\sigma^2}{\gamma}\sigma E[|Y|^3]\right)\|\phi\|_{C^3(I)}(\tau'-\tau).
\end{aligned}$$

Como γ y σ^2/γ son acotados obtenemos finalmente que

$$\sup_{\phi\in C^3(I),\, \|\phi\|_{C^3(I)}\leq 1}\left|\int_I \phi(w)(g_\gamma(\tau,w)dw - g_\gamma(\tau',w)dw)\right| \leq C(\tau'-\tau)$$

con una constante C independiente de $\tau, \tau', \phi, \gamma$. La expresión

$$\|\mu\| := \sup_{\phi\in C^3(I),\, \|\phi\|_{C^3(I)}\leq 1}\int_I \phi\, d\mu$$

define una norma sobre $P(I)$ que metriza la convergencia débil (ver Apéndice B.3). Por lo tanto se puede reescribir la desigualdad anterior como

$$\|g_\gamma(\tau) - g_\gamma(\tau')\| \leq C(\tau' - \tau), \qquad 0 \leq \tau < \tau' < \infty.$$

Esto prueba que la sucesión de funciones continuas $g_\gamma : \tau \in [0, +\infty) \to g_\gamma(\tau) \in P(I)$ es uniformemente equicontinua. Como $P(I)$ es compacto, aplicando el teorema de Arzela-Ascoli (ver Apéndice B.1.2) y un argumento diagonal, obtenemos la existencia de $g \in C([0, +\infty), P(I))$ y de una subsucesión $\gamma_n \to 0$ tales que g_{γ_n} converge a g en $C([0, T], P(I))$ para todo $T > 0$ o sea

$$\lim_{n\to+\infty} \max_{0\leq\tau\leq T} \|g_{\gamma_n}(\tau) - g(\tau)\| = 0$$

para todo $T > 0$. En vista de los resultados enunciados en el Apéndice B.2.21, obtenemos en consecuencia que

$$\lim_{n\to+\infty} \max_{0\leq\tau\leq T} W_1(g_{\gamma_n}(\tau), g(\tau)) = 0. \tag{2.18}$$

Admitamos por un momento que de ahí se deduce que para toda $\phi \in C(I)$,

$$\int_I \phi(w) g_{\gamma_n}(t, w) dw \to \int_I \phi(w) g(t, w) dw \tag{2.19}$$

uniformemente en $t \in [0, T]$. Podemos ahora pasar al límite $n \to +\infty$ en

$$\begin{aligned}
&\int_I \phi(w) g_{\gamma_n}(\tau, w) dw - \int_I \phi(w) f_0(w) dw) \\
&= \int_0^\tau \int_I (m_{\gamma_n}(s) - w)\phi'(w) g_{\gamma_n}(s, w) dw ds + \frac{\sigma^2}{2\gamma} \int_0^\tau \int_{I^2} D(|w|)^2 \phi''(w) g_{\gamma_n}(s, w) dw ds \\
&\quad + \frac{\gamma_n}{2} \int_0^\tau \int_{I^2} (w_* - w)^2 \phi''(w) g_{\gamma_n}(s, w) g_{\gamma_n}(s, w_*) dw dw_* ds + \int_0^\tau R(\gamma_n, \sigma, \phi, s) ds
\end{aligned}$$

para obtener

$$\begin{aligned}
&\int_I \phi(w) g(\tau, w) dw - \int_I \phi(w) f_0(w) dw) \\
&= \int_0^\tau \int_I (m(s) - w)\phi'(w) g(s, w) dw ds + \frac{\lambda}{2} \int_0^\tau \int_{I^2} D(|w|)^2 \phi''(w) g(s, w) dw ds.
\end{aligned}$$

donde $m(s)$ es el promedio de $g(s)$. Aca usamos que el valor absoluto de $\int_0^\tau R(\gamma_n, \sigma, \phi, s) ds$ se puede acotar por $o(1)\|\phi'''\|_\infty \tau$ por (2.17) y que

$$m_{\gamma_n}(s) - m(s) = \int_I w(g_{\gamma_n}(s, w) dw - g(s, w) dw) \to 0$$

cuando $n \to +\infty$ uniformemente para s en un compacto de $[0, +\infty]$.

Para terminar nos falta justificar que vale (2.19) uniformemnte en $t \in [0, T]$. Adaptamos la prueba (B.10)-(B.11). Sea $\phi \in C(I)$ y $(a_m)_m$ (resp. $(b_m)_m$) una sucesión creciente (resp. decreciente) de funciones Lipschitz que converge puntualmente a ϕ. De (2.18) deducimos que para todo m fijo,

$$\limsup_{n\to+\infty} \int_I \phi(w) g_{\gamma_n}(t,w)dw \leq \int_I b_m(w)g(t,w)dw$$

uniformemnte en $t \in [0, T]$. Como $(b_m)_m$ es una sucesión decreciente de funciones continuas sobre el compacto I que converge punutualmente a $\phi \in C(I)$, el teorema de Dini asegura que la convergencia es uniforme sobre I. Entonces

$$\Big| \int_I b_m(w)g(t,w)dw - \int_I \phi(w)g(t,w)dw \Big| \leq \|b_m - \phi\|_\infty \int_I g(t,w)dw = \|b_m - \phi\|_\infty$$

converge a 0 uniformemente en $t \in [0, T]$. Luego

$$\limsup_{n\to+\infty} \int_I \phi(w) g_{\gamma_n}(t,w)dw \leq \int_I \phi(w)g(t,w)dw$$

uniformemente en $t \in [0, T]$. Obtenemos la desigualdad opuesta considerando $(a_m)_m$ al lugar de $(b_m)_m$.

Hemos probado entonces el siguiente teorema:

Teorema 2.2.5. *Sea $f_0 \in P(I)$. Suponemos que $E[|Y|^3] < \infty$ y que $\sigma = \sigma(\gamma) \to 0$ de manera tal que $\sigma^2/\gamma \to \lambda$. Entonces existe una subsucesión $\gamma_n \to 0$ tal que la solución débil de la ecuación de Boltzman para la densidad reescalada $g_\gamma(v, \tau) = f(v, t)$, con $\tau = \gamma t$, converge en $C([0, T], P(I))$ para todo $T > 0$ a un $g \in C([0, +\infty, P(I))$ que es solución débil de la ecuación de Fokker-Planck* (2.11).

2.2.4. Otros modelos de formación de opinión del tipo Fokker-Planck

Función de propensión al compromiso

El Teorema (2.2.4) puede ser generalizado. Seguiremos considerando que el núcleo es de la forma $\beta(\eta, \eta_*) = \Theta(\eta)\Theta(\eta_*)$. Consideremos ahora una función $P(|w|)$ mas general en (2.1). Así, la demostración del teorema se modificará en el primer término de (2.8).

En este caso, recordando que la esperanza de Θ es cero, tendremos:

$$\begin{aligned}
I &= \frac{1}{\gamma}\int_{I^2}\int_B \left(\gamma(w_*-w)P(|w|)+\eta D(|w|)\right)\phi'(w)g(\tau,w)g(\tau,w_*)dw_*dwd\Theta(\eta) \\
&= \int_{I^2} P(|w|)(w_*-w)\phi'(w)g(\tau,w)g(\tau,w_*)dw_*dw \\
&= \int_I w_*g(\tau,w_*)dw_*\int_I P(|w|)\phi'(w)g(\tau,w)dw - \int_{I^2} P(|w|)w\phi'(w)g(\tau,w)dw \\
&= \int_I P(|w|)(m(\tau)-w)\phi'(w)g(\tau,w)g(\tau,w_*)dw_*dw
\end{aligned}$$

donde $m(\tau)=\int_I wg(w,\tau)dw$ es el promedio de opinión a tiempo $\tau\geq 0$. Siguiendo como antes, obtenemos finalmente, tomando el límite $\gamma\to 0$, la ecuación de Fokker-Planck

$$\frac{\partial g}{\partial \tau}=\frac{\lambda}{2}\frac{\partial^2}{\partial w^2}\left(D^2(|w|)g(w)\right)+\frac{\partial}{\partial w}(P(|w|)(w-m(t))g).$$

Note que al considerar una funcion no constante $P(|w|)$, estamos introduciendo una dificultad ya que ahora el promedio $m(\tau)$ no será constante y para hallar una solución de la ecuación de Fokker-Planck debemos resolver

$$\frac{dm(\tau)}{d\tau}=m(\tau)\int_I P(|w|)g(w,\tau)dw-\int_I w\ P(|w|)g(w,\tau)dw$$

en vista de (2.4).

Modelo con difusión dominante

Analicemos el caso donde la difusión sea dominante, es decir, cuando el único término que sobreviva en la ecuación (2.11) sea el relacionado a la derivada segunda. Este será el caso $\frac{\sigma^2}{\gamma}\to\infty$. Para ves eso, supongamos que

$$\frac{\sigma^2}{\gamma^a}\to\lambda>0 \qquad \text{con } a<1.$$

Note que en este caso $\sigma^2/\gamma=\gamma^{-(1-a)}\sigma^2/\gamma^a\to+\infty$. Consideremos el cambio de tiempo

$$\tau=\gamma^a t,\ \ g(w,\tau)=f(w,t).$$

Entonces la ecuación que satisface g análoga a (2.8) es

$$\begin{aligned}
&\frac{d}{d\tau}\int_I g(w)\phi(w)dw= \\
&\frac{1}{\gamma^a}\int_{I^2}\int_{B^2}\Big[\left(\gamma(w_*-w)+\eta D(|w|)\right)\phi'(w) \\
&+\frac{1}{2}(\gamma(w_*-w)+\eta D(|w|))^2\phi''(w)\Big]\,g(w)g(w_*)dw_*dwd\Theta(\eta)d\Theta(\eta_*)+\gamma^{1-a}R(\gamma,\sigma)
\end{aligned}$$

con $R(\gamma,\sigma)$ dado por (2.15). Procediendo como antes usando que Θ tiene esperanza cero y varianza σ^2, vemos que

$$\begin{aligned}
&\frac{d}{d\tau}\int_I g(w)\phi(w)dw = \\
&\int_{I^2}\Big(\gamma^{1-a}(w_*-w)\phi'(w)+\frac{\gamma^{2-a}}{2}(w_*-w)^2\phi''(w)\Big)g(w)g(w_*)dw_*dw \\
&+\frac{\sigma^2}{2\gamma^a}\int_I D(|w|)^2\phi''(w)g(w)dw)+\gamma^{1-a}R(\gamma,\sigma).
\end{aligned}$$

Tomando $\gamma\to 0$ y recordando que $a<1$ (la justificación rigurosa es análoga a la que hicimos recién) obtenemos

$$\frac{d}{d\tau}\int_I g(w)\phi(w)dw = \frac{\lambda}{2}\int_I D(|w|)^2\phi''(w)g(w)dw).$$

Es la formualción débil de

$$\frac{\partial g}{\partial \tau} = \frac{\lambda}{2}\frac{\partial^2}{\partial w^2}\left(D^2(|w|)g(w)\right).$$

En particular, la elección

$$D(|w|)=\sqrt{1-w^2}, \qquad \lambda = 2$$

nos conduce a la ecuación de difusión

$$\frac{\partial g}{\partial \tau} = \frac{\partial^2}{\partial w^2}\left((1-w^2)g(w)\right).$$

Esta ecuación es la que obtuvimos en la sección anterior al describir el modelo de Sznajd y la simplificación de Ochrombel sobre un grafo completo de N nodos cuando $N\to+\infty$.

Modelo donde el compromiso es dominante

Analicemos ahora el caso donde el compromiso es dominante, es decir, el caso donde $\frac{\sigma^2}{\gamma}\to 0$. Hacemos el mismo reescale de tiempo que en el caso $\sigma^2\sim\gamma$ es decir

$$\tau=\gamma t, \qquad g(w,\tau)=f(w,t).$$

Por lo tanto $g(w,\tau)$ satisface (2.8):

$$\begin{aligned}
\frac{d}{d\tau}\int_I g(w)\phi(w)dw = \frac{1}{\gamma}\int_{I^2}\int_{B^2}&\left[\left(\gamma(w_*-w)+\eta D(|w|)\right)\phi'(w)\right. \\
&\left.+\frac{1}{2}\left(\gamma(w_*-w)+\eta D(|w|)\right)^2\phi''(w)\right]g(w)g(w_*)dw_*dwd\Theta(\eta)d\Theta(\eta_*) \\
&+R(\gamma,\sigma)
\end{aligned}$$

Procediendo como antes llegamos a (2.9) con la diferencia que ahora los dos últimos términos desaparecen al tomar el límite $\gamma \to 0$. Sobrevive únicamente el término de 1er orden:

$$\frac{d}{d\tau}\int_I \phi(w)g(w)dw = \int_I P(|w|)(w-m)g(w)\phi'(w)dw.$$

Es la formulación débil de

$$\frac{\partial g}{\partial \tau} = \frac{\partial}{\partial w}\Big(P(|w|)(m(t)-w)g(w)\Big).$$

Si consideramos por ejemplo $P(|w|) = 1 - w^2$, obtenemos

$$\frac{\partial g}{\partial \tau} = \frac{\partial}{\partial w}\Big((1-w^2)(m(t)-w)g(w)\Big).$$

Estudiaremos esta ecuación en el próximo capítulo siguiendo el paper [1].

Capítulo 3

Modelo continuo del 1er orden para la formación de opinión

3.1. Introducción

El objetivo de este capítulo es estudiar la ecuación

$$\partial_t f = \gamma \partial_x \Big((1-x^2)(x-m(t))f \Big), \qquad x \in (-1,1),\ t > 0 \tag{3.1}$$

donde $f(t,x)$ es una densidad de probabilidad en $[-1,1]$ de valor promedio $m(t) = \int_{-1}^{1} x f(t,x) dx$ y $\gamma = \pm 1$. Completamos esta ecuación con la condición inicial

$$f(0) = f_0 \tag{3.2}$$

donde f_0 es una densidad de probabildad en $[-1,1]$. Queremos probar la existencia y unicidad de una solución (en un sentido que explicitaremos más adelante) y estudiar su comportamiento asintótico cuando $t \to +\infty$. El contenido de este capítulo se basa en el paper [1].

Recordemos que obtuvimos esta ecuación con $\gamma = 1$ en el capítulo anterior como "grazing limit" de una ecuación de Boltzmann cuando las interacciones entre los agentes no tienen difusión propia. Vimos también en el 1er capítulo que el límite continuo del modelo "2 contra 1" es una ecuación parecida a (3.1) pero sin el término $m(t)$. Podemos pensar entonces que (3.1) modela la evolución en el tiempo de la distribucion de opiniones en una población. Acá una opinión se representa por un numero $x \in [-1,1]$, y $f(t,x)$ es la proporción de individuos en la población con opnión x en el instante t.

Vamos a ver que el valor de γ tiene un impacto fundamental sobre el comportamiento de f cuando $t \to +\infty$. Para convencerse de eso intuitivamente, basta observar que (3.1) es una ecuación de conservación de masa con campo de velocidades asociado $v(t,x) = \gamma(1-x^2)(m(t)-x)$. Notemos que este campo depende de la f misma a través de m. Notemos que $v(t,\pm 1) = 0$ por lo que esperamos que las eventuales masas de Dirac inicialmente presente se mantengan. Es la consecuencia macroscópica de las reglas microscópicas de interacción del capítulo anterior tomando $P(w) = 1 - w^2$ y $\eta = 0$. Como $x \in [-1,1]$, el signo de v es el signo de γ por el signo de $m(t) - x$. Luego si $\gamma = -1$ entonces $v(t,.)$ es positivo si $x > m(t)$ y negativo si $x < m(t)$. Luego v "empuja" f hacia los extremos de $[-1,1]$. Es razonable esperar entonces que $f(t)$ converja cuando $t \to +\infty$ a alguna combinación convexa de masas de Dirac localizadas en -1 y 1. En este caso, la población se divide asintóticamente en dos grupos de individuos: los de opinión 1 y los de opinion -1. En cambio si $\gamma = 1$ entonces $v(t,.)$ es positivo si y solo si $x < m(t)$. Luego f tiene tendencia a concentrarse por lo que podemos esperar que f converja a alguna masa de Dirac localizada en algun punto de $[-1,1]$. En este caso todos los individuos de la población tienen tenden asintoticamente a tener todos la misma opinión.

3.2. Noción de solución

Aclaramos ahora en que sentido decimos que una f es solución de (3.1). Primero notemos que por la idea intuitiva que esbozamos recién sobre el comportamiento asintótico cuando $t \to +\infty$ de f, es de esperar la aparición de masas de Dirac por lo que es razonable pedir que a cada instante, $f(t)$ pertenece al conjunto K de las medidas de probabilidad sobre $[-1, 1]$. De esta forma se pierde toda esperanza de tomar el ∂_x en sentido clásico.

Para hallar la formulación débil adecuada, nos inspiramos de la intuición usada para llegar a la formulación débil de una ecuación cinética. Consideramos una variable aleatoria X_t a valor en $[-1, 1]$ de distribución $f(t)$. Para conocer la evolución temporal de X_t, estudiamos la evolución del valor promedio de $h(X_t)$ donde h es una función suave cualquiera (una observable, la cantidad que uno ve al nivel macroscópico). Esto lleva a considerar $\frac{d}{dt} \int_{\mathbb{R}} h(x)f(t, x)dx$. Si suponemos que f es una función suave, tenemos

$$
\begin{aligned}
\frac{d}{dt} \int_{\mathbb{R}} h(x)f(t, x)dx &= \int_{\mathbb{R}} h(x)\partial_t f(t, x)dx \\
&= \gamma \int_{-1}^{1} h(x)\partial_x\Big((1 - x^2)(x - m(t))f\Big)dx \\
&= -\gamma \int_{-1}^{1} h'(x)(1 - x^2)(x - m(t))f(t, x)\, dx
\end{aligned}
$$

donde usamos (3.1) e hicimos integración por partes. Ahora, olvidando las etapas intermedias de esta cuenta que necesitan la suavidad de f, podemos quedarnos con la primera y con la última etapa para dar sentido (3.1).

Pediremos también que f sea continua de $[0, +\infty)$ en K donde se considera la topología débil en K. Note que K es compacto (ya que $[-1, 1]$ lo es) y que la topologiá débil puede ser metrizada de distintas maneras, por ejemplo usando las distancias de Monge-Kantorovich W_p, $1 \leq p < \infty$, asociadas a la distancia eulcideana en $[-1, 1]$ (ver Apéndice B.2). Note que estas distancias son todas equivalentes.

Finalmente aclaramos que, siguiendo las notaciones de [1], a veces notaremos una medida de probabilidad $f \in K$ como si fuese una función $f(x)$ (aunque f no sea absolutamente continua con respecto a la medida de Lebesgue).

Estas consideraciones llevan a la siguiente noción de solución de (3.1).

Definición 3.2.1. *Decimos que f es solucioń de* (3.1) *con condición inicial $f_0 \in K$ si $f : t \in [0, +\infty) \to f(t) \in K$ es continua con $f(0) = f_0$ y si para toda función $h \in C_c^1(\mathbb{R})$ vale que $\int_{\mathbb{R}} h(x)f(t, x)dx$ es derivable con respecto a $t > 0$ y vale*

$$
\frac{d}{dt} \int_{\mathbb{R}} h(x)f(t, x)dx = -\gamma \int_{\mathbb{R}} h'(x)(1 - x^2)(x - m(t)f(t, x)dx. \tag{3.3}
$$

Ejemplo 3.2.2. *Sea $h \in C_c^1(\mathbb{R})$ tal que $h(x) = x$ en $[-1,1]$. Entonces $\int h(x)f(t,x)\,dx = m(t)$ es derivable en $(0,+\infty)$ con*

$$m'(t) = -\gamma \int_{-1}^{1} (1-x^2)(x - m(t)f(t,x)dx.$$

3.3. Una formulación alternativa de (3.1)

Reescribimos ahora (3.1) en término de la inversa generalizada de $F(t,.)$, la función de distribución acumulada de $f(t)$.

3.3.1. Sobre la inversa generalizada de la función de distribución acumulada

Recordemos primero que, en general, la función de distribución acumulada de una medida de probabilidad $f \in P(\mathbb{R})$ es la función $F : \mathbb{R} \to [0,1]$ definida por $F(x) = f((-\infty, x])$. Note que F es no-decreciente, continua a derecha y tiene por límite 0 y 1 en $-\infty$ y $+\infty$ respectivamente. Se define la inversa generalizada de F como la función $X : (0,1] \to \mathbb{R}$ tal que:

$$X(r) = \text{ínf}\,\{x \in \mathbb{R} :\ F(x) \geq r\},$$

es decir, $X(r)$ es el 1er momento en el que se acumula por lo menos una fracción r de la masa total. Notemos que X es no-decreciente, continua a izquierda, que $F(X(r)) \geq r$ y que para todo $\rho \in [0,1]$,

$$F(x) \geq \rho \Longleftrightarrow X(\rho) \leq x.$$

Además conocer X es equivalente a conocer F y luego a conocer f.

La reescritura de (3.1) se basa en el siguiente lema:

Lema 3.3.1. *Con las notaciones anteriores. Sea m la medida de Lebesgue sobre $\mathbb{R}$ restringida al intervalo $[0,1]$. Se tiene que*

$$f = X\#m \tag{3.4}$$

donde $X\#m$ es la medida push-forward de m por X (ver def. en Apéndice A.3.1). En particular para cualquier función medible $h : \mathbb{R} \to \mathbb{R}$ no-negativa o acotada,

$$\int_{\mathbb{R}} h\,df = \int_0^1 h(X(x))\,dx. \tag{3.5}$$

Ejemplo 3.3.2. *Con $h(x) = x$ y f se soporte compacto, se tiene que*

$$m := \int x\,df(x) = \int_0^1 X(x)\,dx. \tag{3.6}$$

Demostración. Como la σ-algebra de los borelianos es generada por los intervalos de la forma $(-\infty, a]$, $a \in \mathbb{R}$, basta demostrar que (3.4) para estos conjuntos. Fijamos entonces un $a \in \mathbb{R}$. Tenemos

$$X\#m((-\infty, a]) = m(\{r \in [0,1] : X(r) \leq a\}) = F(a) = f((-\infty, a])$$

donde usamos en la segunda igualdad que $\{r \in [0,1] : X(r) \leq a\}$ es $[0, F(a))$ o $[0, F(a)]$. Finalmente (3.5) es consecuencia de la definición del push-forward de una medida (ver apéndice A). $\square$

3.3.2. Formulación de (3.1) en término de la inversa generalizada

Consideramos ahora $f \in C([0, +\infty), K)$. Notamos $F(t, .)$ la función de distribución acumulada de $f(t)$ y $X(., t)$ su inversa generalizada. Una forma eficiente de tratar la ecuación (3.1) es reescribirla en términos de $X(., t)$. Veremos en el teorema (3.3.4) abajo que la ecuacion de evolución para $X(., t)$ viene dada por:

$$\partial_t X(\rho, t) = -\gamma(X(t, \rho) - m(t))(1 - X(t, \rho)^2) \qquad \text{para todo } \rho \in (0,1),\ t > 0, \tag{3.7}$$

donde $m(t) = \int_0^1 X(\rho, t) d\rho$ por (3.6), y con condición inicial

$$X(., 0) = X_0, \text{ la inversa generalizada de } f_0 \tag{3.8}$$

Aclaramos en que sentido buscamos una solución de (3.7)-(3.8):

Definición 3.3.3. *Decimos que $X(t, \rho)$ es solución de la ecuación (3.7)-* (3.8) *si $X_\rho := X(., \rho)$ es derivale para todo $\rho \in (0,1)$ y $X_\rho(t)$ verifica (3.7)-*(3.8).

Teorema 3.3.4. *Si existe una solución $X(t, \rho)$ de* (3.7)-(3.8) *entonces existe una solución f de* (3.1)-(3.2) *en el sentido de la definición (3.2.1) Además, $X(t, .)$ es la inversa generalizada de la función de distribución acumulada de $f(t)$.*

Demostración. Supongamos que existe una solución $X(\rho, t)$ de (3.7)-(3.8). Como X_0 es la inversa generalizada de f_0, X_0 es no-decreciente, continua a izquierda y a valor en $[-1, 1]$. En vista del Lema (3.3.5) abajo, $X(., t)$ verifica esas mismas propiedades para todo $t \geq 0$. Deducimos que existe una única $f(t) \in K$ tal que $X(., t)$ es la inversa generalizada de la función de distribución acumulada de $f(t)$. Probemos que f es solución de (3.1)-(3.2). Sea $h \in C_c^1(\mathbb{R})$. Verificamos que se cumple (3.3). Aplicando (3.5) con $f(t)$ al lugar de f, se tiene que

$$\int_{-1}^{1} h(x) f(t, x)\, dx = \int_0^1 h(X(t, \rho))\, d\rho.$$

Derivamos ambos miembros con respecto a t aplicando el teorema de convergencia dominada a la derecha (notemos que h' es continua con soporte compacto y por lo tanto acotada). Obtenemos

$$\begin{aligned}\frac{\partial}{\partial t}\int_{\mathbb{R}} h(x)f(x,t)dx &= \int_0^1 h'(X(\rho,t))\frac{\partial}{\partial t}X(\rho,t)\,d\rho \\ &= -\gamma\int_0^1 h'(X(\rho,t))((X(\rho,t)-m(t))(1-X^2(\rho,t)))\,d\rho.\end{aligned}$$

Aplicamos nuevamente (3.5) al miembro derecho y deducimos

$$\frac{\partial}{\partial t}\int_{\mathbb{R}} h(x)f(x,t)dx = -\gamma\int_0^1 h'(x)(1-x^2)(x-m(t))f(t,x)dx.$$

Luego f verifica (3.3).

Falta probar que $f(t)$ es continua en t. Fijamos $t_0 \in [0,+\infty)$. Debemos probar que para toda $\phi \in C_b(\mathbb{R})$,

$$\lim_{t\to t_0}\int \phi\; df_t = \int \phi\; df_{t_0}.$$

Sabemos que $\int \phi\; df_t = \int \phi\,(X_t(\rho))\,d\rho$. Como $X_t(\rho)$ es continua en t para todo ρ, tenemos por convergencia mayorada que

$$\int_{-1}^1 \phi\; df_t = \int_0^1 \phi(X_t(\rho))d\rho \overset{t\to t_0}{\to} \int_0^1 \phi\,(X_{t_0}(\rho))\,d\rho = \int_{-1}^1 \phi\; df_{t_0}.$$

□

Falta probar el siguiente lema que usamos en la prueba del teorema 3.3.4.

Lema 3.3.5. *Consideremos el siguiente sistema:*

$$\begin{cases}\frac{\partial}{\partial t}X(t,\rho) = -\gamma(X(t,\rho)-m(t))(1-X^2(t,\rho)) \\ X(0,\rho) = X_0(\rho) \qquad \forall\rho\in(0,1)\end{cases}$$

Si el dato inicial $X_0 : [0,1] \to \mathbb{R}$ es no-decreciente, continuos a izquierda y a valor en $[-1,1]$, entonces esas mismas propiedades se preservan para $X(t,.)$, $t \geq 0$.

Este lema es una consecuencia directa del siguiente lema:

Lema 3.3.6. *Dada una función continua $m : \mathbb{R} \to [-1,1]$, la solución de la EDO*

$$\begin{cases} y'(t) = \phi(y(t),t) := -\gamma(y(t)-m(t))(1-y(t)^2) \\ y(0) = y_0 \in [-1,1]\end{cases}$$

esta bien definida de $\mathbb{R}$ en $[-1,1]$, es creciente y continua con respecto a la condicíon inicial y_0. Además, vale $|y(t)-\tilde{y}(t)| \leq |y(t)-\tilde{y}(t)|\,e^{5t}$

Demostración. Como $\phi(\pm 1, t) = 0$, la funciones constantes ± 1 son soluciones. Luego si $|y_0| \leq 1$ entonces la solución y esta a valor en el compacto $[-1, 1]$ lo que implica que esta definida para todo $t \in \mathbb{R}$. La continuidad y la montónía con respecto a la condición inicial son resultados clásicos (ya que $|\phi(y,t) - \phi(\tilde{y},t)| \leq 5\,|y - \tilde{y}|$, ver Lema (3.4.3)).

□

3.4. Existencia y unicidad de una solución

Teorema 3.4.1. *Para toda $f_0 \in K$, existe una única solución f de* (3.1)-(3.2) *en el sentido de la definición (3.2.1). Mas aún, la solución depende continuamente de la condición inicial. Por lo tanto el problema* (3.1)-(3.2) *esta bien definido en $C(\mathbb{R}, K)$.*

Observación 3.4.2. *En el capítulo anterior obtuvimos la existencia de una solución de la ecuación mas general* (2.11)*. la prueba de [1] que detallamos a continuación tiene la ventaja ser constructiva y por lo tanto permite obtener un algoritmo para realizar una aproximación num'erica.*

En la prueba de este teorema, necesitaremos el siguiente lema:

Lema 3.4.3. *Sea $\phi(x,y) = -\gamma(1-x^2)(x-y)$ con $\gamma = \pm 1$. Entonces*

$$|\phi(x_1,y_1) - \phi(x_2,y_2)| \leq 5(|x_1 - x_2| + |y_1 - y_2|), \qquad (x_1,y_1),(x_2,y_2) \in [-1,1]\times[-1,1] \tag{3.9}$$

Además, sea $g : [a,c] \to [-1,1]$ continua en $[a,b]$ y $[b,c]]$. Dado $f_a \in [-1,1]$, definimos $f : [a,c] \to [-1,1]$ como la solución en $[a,b]$ de $f'(t) = \phi(f(t), g(t))$ con $f(a) = f_a$, y luego como la solución en $[b,c]$ de la misma EDO pero con condición inicial $f(b)$. Entonces f es continua en $[a,c]$, derivable en $[a,b]$ y $[b,c]$, a valor en $[-1,1]$ y

$$|f(s) - f(t)| \leq 2|s-t| \qquad s,t \in [a,c].$$

Obviamente el resultado de la 2nda parte del Lema vale si g tiene una cantidad finita cualquiera de discontinuidades.

Demostración. Para la primera parte del lema basta con aplicar el teorema del valor intermedio:

$$|\phi(x_1,y_1) - \phi(x_2,y_2)| \leq \|\nabla\phi\|_\infty(|x_1 - x_2| + |y_1 - y_2|)$$

con $|\nabla\phi(x,y)| \leq |\partial_x\phi| + |\partial_y\phi| = |1 - 3x^2 + 2xy| + |x^2 - 1| \leq 4 + 1 = 5$.

Para la segunda parte del lema, la existencia de f en $[a,b]$ y luego en $[b,c]$ esta asegurada por el teorema de Cauchy-Lipschitz ya que cada vez resolvemos $f'(t) = \phi(f(t), g(t)) =: F(t, f(t))$ con F continua y localmente lipschitz en f. Como las funciones constantes 1 y -1 son sur y subsolución y $|f(a)| \leq 1$, tenemos que $|f(t)| \leq 1$, $a \leq t \leq b$. En particular $|f(b)| \leq 1$ por lo que $|f(t)| \leq 1$ también para $t \in [b,c]$ por el mismo motivo. Finalmente $|f'(t)| \leq |1 - f(t)^2||f(t) - g(t))| \leq |f(t) - g(t))|$. Como $\|g\|_\infty \leq 1$ deducimos que $|f'(t)| \leq 2$. Luego para $s, t \in [a, c]$, por ejemplo $s \in [a, b]$ y $t \in [b, c]$,

$$|f(s) - f(t)| \leq \int_s^b |f'(r)|\, dr + \int_b^t |f'(r)|\, dr \leq 2|s - t|.$$

□

Podemos ahora probar el teorema 3.4.1.

Demostración. **Existencia.** Sabemos por el Teorema 3.3.4 que para probar la existencia de una solución de (3.1)-(3.2), basta probar la existencia de una solución de

$$(3.10) \qquad \begin{aligned} \partial_t X(\rho, t) &= -\gamma(X(\rho, t) - m(t))(1 - X^2(\rho, t)) \qquad \rho \in (0, 1), \\ X(\rho, 0) &= X_0(\rho) \qquad \forall \rho \in (0, 1) \end{aligned}$$

donde X_0 es la inversa generalizada de la acumulada de f_0 y $m(t) = \int_0^1 X(t, \rho)\, d\rho$. Notemos que la existencia de X es un problema no-trivial debido a la presencia del término $m(t)$ en el miembro derecho. Probaremos la existencia de una solución de este último problema de forma construtiva en tres etapas. La idea general consiste en subdividir el intervalo de tiempo $[-T, T]$ en subintervalos de longitud 2^{-N} con N grande en los que resolvemos problemas como (3.10) pero reemplazando m por una constante. En este caso, es trivial resolver (3.10) ya que, para cada ρ, se trata de una EDO para $t \to X(t, \rho)$.

Etapa 1 *Construimos una sucesion $\{X_n, n \in \mathbb{N}\}$ que tiende a la solución buscada.*

Sea $[-T, T]$ un intervalo fijo. Dado $N \in \mathbb{N}$, subdividimos al intervalo $[-T, T]$ en intervalos de longitud $\frac{1}{2^N}$. Construimos la sucesión $X^{(N)}(., t)$, $t \in [-T, T]$ por inducción de la siguiente forma:

1.a Calculamos $m_0^{(N)} = \int_0^1 X_0(\rho) d\rho$

1.b Resolvemos (3.10) en el intervalo $\left[\frac{-T}{2^N}, \frac{T}{2^N}\right]$ con $m := m_0^{(N)}$ conocido. Notamos $X^{(n)}(., t)$, $t \in \left[\frac{-T}{2^N}, \frac{T}{2^N}\right]$, la solución obtenida.

1.c Para todo $k = 1, ..., 2^n - 1$, definimos recursivamente $X^{(n)}(., t)$ para $t \in \left(\frac{kT}{2^N}, \frac{(k+1)T}{2^N}\right]$ y $t \in \left(-\frac{(k+1)T}{2^N}, -\frac{kT}{2^N}\right]$ de la siguiente forma:

- Calculamos

$$m_{\pm k}^{(N)} = \int_0^1 X^{(N)}\left(\rho, \pm k \frac{T}{2^N}\right) d\rho.$$

- Resolvemos (3.10) en el intervalo $\left(\frac{kT}{2^N}, \frac{(k+1)T}{2^N}\right]$ con $m := m_k^{(N)}$ y dato inicial $X^{(N)}\left(., k\frac{T}{2^N}\right)$. Notamos $X^{(N)}(., t)$, $t \in \left(\frac{kT}{2^N}, \frac{(k+1)T}{2^N}\right]$ la solución obtenida.
- Resolvemos (3.10) en el intervalo $\left(-\frac{(k+1)T}{2^N}, -\frac{kT}{2^N}\right]$ con $m := m_{-k}^{(N)}$ y dato inicial $X^{(N)}\left(., -k\frac{T}{2^N}\right)$. Notamos $X^{(N)}(., t)$, $t \in \left(-\frac{(k+1)T}{2^N}, -\frac{kT}{2^N}\right]$ la solución obtenido.

Se va facilmente por inducción sobre k usando el lema 3.3.5 que para cada $t \in [-T, T]$, $X^{(N)}(., t)$ es no-decreciente, continua a izquierda y a valor en $[-1, 1]$. Luego existe $f^{(N)} : t \in [-T, T] \to K$ tal que para todo t, $X^{(N)}(., t)$ es la inversa generalizada de la acumulada de $f^{(N)}(t)$.

También si definimos $m^{(N)}(t) = m_k^{(N)}(t)$ si $t \in \left(\frac{kT}{2^N}, \frac{(k+1)T}{2^N}\right]$, $k = -(2^N - 1), .., 2^N - 1$, vemos por inducción que $m^{(N)}(t) \in [-1, 1]$ para todo $t \in [-T, T]$.

Etapa 2 *A través de argumentos de compacidad, hallamos una subsucesión convergente* $\{X_{n_l}\}_{l \in \mathbb{N}}$.

Aclaro que por comodidad de notacion, seguiré notando cualquier subsucesión extraída como la sucesión original.

Se deduce del Lema 3.4.3 que $X^{(N)}(\rho, t) \in [-1, 1]$ para todo $t \in [-T, T]$ y todo $N \in \mathbb{N}$. Entonces $m^{(N)}(t) \in [-1, 1]$, $t \in [-T, T]$. Luego, obtenemos del Lema (3.4.3) que para todo $N \in \mathbb{N}$, $s, t \in [-T, T]$, $\rho \in [0, 1]$,

$$\left|X^{(N)}(\rho, s) - X^{(N)}(\rho, t)\right| \leq 2\,|s - t|\,. \tag{3.11}$$

En particular, para todo $\rho \in [0, 1]$ fijo, las funciones $X^{(N)}(\rho, .) : [-T, T] \to [-1, 1]$, $N \in \mathbb{N}$, forman una sucesión de funciones uniformemente equicontinuas y acotadas. Luego, por el Teorema de Arzela-Ascoli (ver Apéndice B.1.2), sabemos que una subsucesión converge uniformemente. Usando un argumento diagonal, podemos extraer una subsucesión tal que $X^{(N)}(\rho, .)$ convergen uniformemente para cualquier $\rho \in (0, 1) \cap \mathbb{Q}$ (ya que $(0, 1) \cap \mathbb{Q}$ es numerable). Notamos $X(\rho, t)$ la función límite. Está definida para $\rho \in (0, 1) \cap \mathbb{Q}$ y $t \in [-T, T]$.

Como cada $X^{(N)}(., t)$ es no-decreciente, $X(., t)$ es no-decreciente por lo que podemos extender la definición de $X(., t)$ a todo $(0, 1)$ a una función continua a izquierda definiendo por $\rho \notin \mathbb{Q}$ que $X(\rho, t) = \sup_{\rho' < \rho, \rho' \in \mathbb{Q}} X(\rho', t))$. Veamos que para todo $t \in [-T; T]$,

$$X^{(N)}(\rho, t) \overset{N \to \infty}{\longrightarrow} X(\rho, t) \text{ para casi todo } \rho \in [0, 1]. \tag{3.12}$$

Sea $t \in [-T, T]$ fijo y sea $\rho_0 \in [0, 1] - \mathbb{Q}$ un punto de continuidad de $X(., t)$. Probemos que $X^{(N)}(\rho_0, t) \overset{N \to \infty}{\longrightarrow} X(\rho_0, t)$. Dado un $\varepsilon > 0$, existe $\gamma > 0$ tal que

$$|X(\rho, t) - X(\rho_0, t)| < \epsilon \qquad \text{si } |\rho - \rho_0| < \gamma.$$

Sean $\rho_1, \rho_2 \in \mathbb{Q} \cap [0,1]$ tales que

$$|\rho_1 - \rho_0| < \gamma, \quad |\rho_2 - \rho_0| < \gamma, \quad \rho_1 < \rho_0 < \rho_2.$$

Recordemos que para cada $N \in \mathbb{N}$, $X^N(.,t)$ es no decreciente, por lo que

$$X^N(\rho_1, t) \leq X^N(\rho_0, t) \leq X^N(\rho_2, t).$$

Como $\rho_1, \rho_2 \in \mathbb{Q}$, podemos pasar al límite:

$$X(\rho_1, t) \leq \liminf_{N\to\infty} X^N(\rho_0, t) \leq \limsup_{N\to\infty} X(\rho_0, t) \leq X(\rho_2, t).$$

Obtenemos finalmente

$$X(\rho_0, t) - \epsilon \leq \liminf_{N\to\infty} X^N(\rho_0, t) \leq \limsup_{N\to\infty} X(\rho_0, t) \leq X(\rho_0, t) + \epsilon$$

para todo $\varepsilon > 0$. Deducimos que $X^{(N)}(\rho_0, t) \overset{N\to\infty}{\longrightarrow} X(\rho_0, t)$ Para finalizar la prueba de (3.12), observemos que, por ser monótona, la función $X(.,t)$ tiene a lo sumo numerables puntos de discontinuidad.

Volviendo a la ecuación (3.11), al integrarla tenemos

$$W_1(f^{(N)}(t), f^{(N)}(s)) = \int_0^1 \left| X^{(N)}(\rho, s) - X^{(N)}(\rho, t) \right| \, d\rho \leq 2\,|s-t|\,, \tag{3.13}$$

donde W_1 es la distancia de Monge-Kantorovich (ver Apéndice B.2). Luego las funciones $f^{(N)} : t \in [-T,T] \to f^{(N)}(t) \in K$ son uniformemente equicontinuas para la distancia d_1. Como (K, W_1) es compacto (ver Apéndice B.2), existe $f : [-T,T] \to K$ y una subsucesión tal que $\max_{-T\leq t\leq T} W_1(f^{(N)}(t), f(t)) \to 0$ o sea

$$\max_{-T\leq t\leq T} \int_0^1 |X^{(N)}(\rho, t) - \tilde{X}(\rho, t)|\, d\rho \to 0$$

si notamos $\tilde{X}(.,t)$ la inversa generalizada de la acumulada de $f(t)$. Como para todo $t \in [-T,T]$, $X^{(N)}(\rho, t) \to X(\rho, t)$ para casi todo $\rho \in [0,1]$, podemos aplicar convergencia dominada (ya que $|X, \tilde{X}| \leq 1$) para obtener que para todo t, $\int_0^1 X(\rho,t) - \tilde{X}(\rho,t)|\, d\rho = 0$. Luego $X = \tilde{X}$.

Etapa 3 *Probamos que para todo $t \in [-T,T]$,*

$$m(t) := \int_0^1 X(\rho, t)\, d\rho = \lim_{N\to+\infty} m^{(N)}(t). \tag{3.14}$$

Fijamos un $t \in [0,T]$ (el caso $t \in [-T,0]$ es análogo). Sea $k = k(t) \in \{0,..,2^N-1\}$ tal que $kT2^{-N} < t \leq (k+1)T2^{-N}$. Entonces $m^{(N)}(t) = m_k^{(N)} = \int_0^1 X^{(N)}(\rho, kT2^{-N})\, d\rho$. Aplicando (3.11) con $|t - kT2^{-N}| \leq T2^{-N}$ obtenemos

$$\left| m^{(N)}(t) - \int_0^1 X^{(N)}(\rho,t)\, d\rho \right| \leq 2|kT2^{-N} - t| \leq 2^{1-N}T. \tag{3.15}$$

Por otro lado, por el Teorema de Convergencia Mayorada,

$$m(t) = \int_0^1 X(\rho,t) d\rho = \lim_{N\to\infty} \int_0^1 X^{(N)}(\rho,t) d\rho.$$

Deducimos (3.14).

Etapa 4 *Probamos que para todo $\rho \in (0,1)$, $X(\rho,.)$ es diferenciable y que es solución de (3.10).*

Sea $\rho \in (0,1) \cap \mathbb{Q}$, fijo. Definimos

$$\begin{aligned} H(N,\rho,t) &:= \partial_t X^{(N)}(\rho,t) = -\gamma(X^{(N)}(\rho,t) - m^{(N)}(t))(1 - X^{(N)}(\rho,t)^2) \\ &= \phi(X^{(N)}(\rho,t), m^{(N)}(t)). \end{aligned}$$

donde ϕ esta definida en el lema 3.4.3. Como $X^{(N)}(\rho,t) \to X(\rho,t)$ y $m^{(N)}(t) \to m(t)$ por la etapa anterior, tenemos

$$H(\rho,t) := \lim_{N\to\infty} H(N,\rho,t) = -\gamma(X(\rho,t) - m(t))(1 - X(\rho,t)^2).$$

Sabemos por la Etapa 2 que $X^{(N)}(\rho,t) \to X(\rho,t)$ uniformemente en $t \in (-T,T)$. Luego para probar que $X(\rho,t)$ es derivable en $t \in (-T,T)$ con $\partial_t X(\rho,t) = H(\rho,t)$, basta probar que

$$H(N,\rho,t) \to H(\rho,t) \qquad \text{uniformemente en } t \in (-T,T). \tag{3.16}$$

Para eso, vamos a verificar que $(H(N,\rho,.))_N$ es de Cauchy en $C([-T,T])$ o sea

$$\max_{-T\leq t\leq T} |H(N,\rho,t) - H(N',\rho,t)| \overset{N,N'\to+\infty}{\longrightarrow} 0. \tag{3.17}$$

Observamos primero que, en consecuencia de (3.15) y (3.13), vale para todo $s,t \in [-T,T]$ que

$$\begin{aligned} &\left| m^{(N)}(s) - m^{(N)}(t) \right| \\ &\leq \left| m^{(N)}(t) - \int_0^1 X^{(N)}(\rho,t)\, d\rho \right| + \left| \int_0^1 X^{(N)}(\rho,t)\, d\rho - \int_0^1 X^{(N)}(\rho,s)\, d\rho \right| \\ &\qquad + \left| \int_0^1 X^{(N)}(\rho,s)\, d\rho - m^{(N)}(s) - \right| \\ &\leq 2^{2-N}T + 2|s-t| \end{aligned}$$

Deducimos aplicando (3.9) y (3.11) que

$$\begin{aligned}
|H(N,\rho,t)-H(N,\rho,s)| &= |\phi(X^{(N)}(\rho,t),m^{(N)}(t))-\phi(X^{(N)}(\rho,s),m^{(N)}(s))| \\
&\leq 5|X^{(N)}(\rho,t)-X^{(N)}(\rho,s)|+5|m^{(N)}(t)-m^{(N)}(s)| \\
&\leq 10|s-t|+5(2^{2-N}T+2|s-t|) \\
&\leq 20|s-t|+52^{2-N}T \qquad (3.18)
\end{aligned}$$

Escribimos ahora para $N'>N$ que

$$\begin{aligned}
&|H(N,\rho,t)-H(N',\rho,t)| \\
&\leq \left|H(N',\rho,t)-H(N',\rho,kT2^{-N})\right| + \left|H(N',\rho,kT2^{-N})-H(N,\rho,kT2^{-N})\right| \\
&\quad + \left|H(N,\rho,kT2^{-N})-H(N,\rho,t)\right|
\end{aligned}$$

donde k tiene el mismo significado que en la prueba de la Etapa 3. En particular $|t-kT2^{-N}|\leq T2^{-N}$. En vista de (3.18) tenemos entonces

$$\begin{aligned}
&|H(N,\rho,t)-H(N',\rho,t)| \\
&\leq 52^{2-N}T+52^{2-N'}T+20T2^{-N}+20T2^{-N'}+\left|H(N',\rho,kT2^{-N})-H(N,\rho,kT2^{-N})\right|
\end{aligned}$$

Finalmente, por el Lema 3.4.3 sabemos que:

$$\begin{aligned}
&\left|H(N',\rho,kT2^{-N})-H(N,\rho,kT2^{-N})\right| \\
\leq\ & 5\left|X^N(\rho,kT2^{-N})-X^{N'}(\rho,kT2^{-N})\right|+5\left|m^N(kT2^{-N})-m^{N'}(kT2^{-N})\right| \\
\leq\ & 5\left|X^N(\rho,kT2^{-N})-X^{N'}(\rho,kT2^{-N})\right|+5\int_0^1\left|X^N(\rho,kT2^{-N})-X^{N'}(\rho,kT2^{-N})\,d\rho\right.
\end{aligned}$$

que va a 0 cuando $N\to+\infty$ uniformemente en $t\in[-T,T]$ ya que, por la etapa 2, $X^N(\rho,t)\to X(\rho,t)$ y $W_1(X^N(.,t),X(.,t))\to 0$ uniformemente en $t\in[-T,T]$. Se deduce (3.17) y por lo tanto, tenemos (3.16). Luego $X(\rho,.)$ es derivable con derivada $H(\rho,.)$ para todo ρ racional.

Extendemos ahora esta demostración a cualquier $\rho\in(0,1)$: sea $\rho\in(0,1)-\mathbb{Q}$ y sea $\rho'\in(0,1)\cap\mathbb{Q}$, $\rho'<\rho$. Tenemos:

$$\begin{aligned}
&|H(N,\rho,t)-H(N',\rho,t)| \\
&\leq |H(N,\rho,t)-H(N,\rho',t)|+|H(N,\rho',t)-H(N',\rho',t)| \\
&\quad +|H(N',\rho',t)-H(N',\rho,t)| \\
&= I+II+III
\end{aligned}$$

Como ρ' es racional sabemos que $II\to 0$ cuando $N,N'\to+\infty$ uniformemente en $t\in[-T,T]$. Veamos que ocurre lo mismo para I y III. Por el lema 3.4.3 y el lema 3.3.6,

$$I\leq 5|X^N(\rho,t)-X^N(\rho',t)|\leq 5|X(\rho,0)-X(\rho',0)|e^{5T}.$$

Como $X(.,0)$ es continua a izquierda, obtenemos el resultado tomando $\rho' \to \rho$.

Unicidad y continuidad con respecto a la condición inicial.
Sean $Y(\rho,t), X(\rho,t)$ dos soluciones de (3.10) con condición inicial $X(\rho,0)$, $Y(\rho,0)$. Notamos $f(t)$ y $g(t)$ la soluciones de (3.1)-(3.2) correspondientes y $m_Y(t) = \int_0^1 Y(\rho,t)\,d\rho$, $m_X(t) = \int_0^1 X(\rho,t)\,d\rho$ los promedios. Estudiamos la evolución temporal de la distancia $W_2(f(t),g(t))$ que notamos por abuso $W_2(X(.,t),Y(.,t))$. Tenemos

$$\begin{aligned}
&\frac{d}{dt}\frac{1}{2}W_2(X(.,t),Y(.,t))^2 = \frac{d}{dt}\frac{1}{2}\int_0^1 (X(\rho,t)-Y(\rho,t))^2\,d\rho \\
&= \int_0^1 (X(\rho,t)-Y(\rho,t))\,(\partial_t X(\rho,t)-\partial_t Y(\rho,t))\,d\rho \\
&= \int_0^1 (X(\rho,t)-Y(\rho,t))\,(\phi(X(\rho,t),m_X(t))-\phi(Y(\rho,t),m_Y(t)))\,d\rho.
\end{aligned}$$

Usamos el teorema de convergencia mayorada para derivar bajo la integral observando que el integrando en la última igualdad es acotado por una constante independiente de $\rho \in [0,1]$ y $t \in \mathbb{R}$ (ya que $|X|,|Y|| \le 1$ y ϕ es acotada en $[-1,1]\times[-1,1]$.)

En vista de (3.9), obtenemos

$$\begin{aligned}
&\frac{d}{dt}\frac{1}{2}W_2(X(.,t),Y(.,t))^2 \\
&\le 5\int_0^1 (X(\rho,t)-Y(\rho,t))\Big(|X(\rho,t)-Y(\rho,t)|+|m_X(t)-m_Y(t)|\Big)\,d\rho \\
&\le 5W_2(X(.,t),Y(.,t))^2 + 5\Big(\int_0^1 |X(\rho,t)-Y(\rho,t)|\,d\rho\Big)^2.
\end{aligned}$$

Usando la desigualdad de Cauchy-Schwarz, vemos que la integral en el miembro derecho es menor que $W_2(X(.,t),Y(.,t))^2$. Luego

$$\frac{d}{dt}\frac{1}{2}W_2(X(.,t),Y(.,t))^2 \le 10W_2(X(.,t),Y(.,t))^2.$$

Por el Lema de Gronwall deducimos que

$$W_2(X(.,t),Y(.,t)) \le W_2(X(\rho,0),Y(\rho,0))\;e^{10t}$$

lo que prueba en particular la continuidad de la solución de (3.10) con respecto a la condición inicial. Tomando $X(\rho,0)=Y(\rho,0)$, conluimos $X(\rho,t)=Y(\rho,t)$ para todo $\rho\in(0,1)$ y todo t lo que prueba la unicidad de la solución. $\square$

3.5. Comportamiento asintótico cuando $t \to +\infty$

En esta sección analizaremos el comportamiento asintótico $t \to +\infty$ de la solución f de la ecuación (3.1)-(3.2) que obtuvimos en la sección anterior en el Teorema 3.4.1. Por lo explicado de forma intuitiva en la introducción a este capítulo, esperamos el siguiente comportamiento asintótica de $f(t)$:

- $\gamma = 1$: concentración de $f(t)$ en el sentido que $f(t)$ debería converger a una combinación convexa de 3 masas de Dirac, dos correspondientes a las masas de Dirac en $f(0)$ (si las hay).

- $\gamma = -1$: esparcimiento de $f(t)$ hacia los extremos. Esperamos que $f(t)$ converja a una combinación convexa de δ_1 y δ_{-1}.

Antes de comenzar el análisis veamos unos resultados simples.

3.5.1. Resultados preliminares

Permanencia de las masas de Dirac

El siguiente Lema nos dice que si para algún tiempo t, f(t) tiene una masa de Dirac en x_0, lo que se traduce del punto de vista de X_t por una igualdad de la forma $X_t(\rho) = X_t(\tilde{\rho}) = x_0$ con $\rho \neq \tilde{\rho}$, entonces f siempre tendrá una masa de Dirac con el mismo peso (cuya localización puede cambiar con el tiempo).

Lema 3.5.1. *Para todo $t \neq s \in \mathbb{R}$ y $\rho \neq \tilde{\rho} \in (0,1)$, vale:*

$$X(\rho, t) = X(\tilde{\rho}, t) \Longleftrightarrow X(\rho, s) = X(\tilde{\rho}, s).$$

Demostración. Es una consecuencia de la unidad de solución del Lema (3.3.5). □

Como consecuencia de este lema, si hay inicialmente proporciones p_1 y p_{-1} de opiniones ± 1, es decir

$$f_0(\{-1\}) = p_{-1} \qquad \text{y} \qquad f_0(\{1\}) = p_1,$$

entonces esas masas permanecerán con esas opiniones para todo $t > 0$:

$$f_t(\{-1\}) = p_{-1} \qquad \text{y} \qquad f_t(\{1\}) = p_1.$$

Lo mismo ocurre con la masa $1 - p_{-1} - p_1$ que a tiempo inicial se encuentra en el intervalo $(-1, 1)$.

Simetría de $f(t)$

Supongamos que la distribución inicial f_0 de opiniones es simétrica con respecto a 0 es decir $X_0(/\rho) = -X(1-\rho)$, $\rho \in [0,1]$. Entonces $f(t)$ es simétrica para todo t. Es consecuencia de que para todo ρ, las funciones $X(t,\rho)$ y $-X(t,1-\rho)$ verifican la misma ecuación:

$$\begin{aligned}
\partial_t \tilde{X}(t,\rho) &= -\partial_t X(t,1-\rho) = \gamma((1-X(t,1-\rho)^2)(X(t,1-\rho)-m(t))) \\
&= -\gamma((1-\tilde{X}(t,\rho)^2)(\tilde{X}(t,\rho)+m(t)))
\end{aligned}$$

con $m(t) = \int_0^1 X(t,\rho)\,d\rho = -\tilde{m}(t) =: \int_0^1 \tilde{X}(t,\rho)\,d\rho$. Además X y $\tilde{X}$ son iguales inicialmente si f_0 es simétrica. Luego $X(t,\rho) = \tilde{X}(t,\rho)$ por la unicidad en el teorema 3.4.1.

En particular $m(t) = \int_{-1}^1 xf(t,x)dx = 0$ para todo t. Entonces la ecuación para X es

$$\partial_t X(t,\rho) = -\gamma(X(t,\rho)(1-X(t,\rho)^2))).$$

Notemos que el miembro derecha se anula en 0. Luego si f_0 tiene una masa de Dirac en 0 es decir $X(0,\rho) = X(0,\rho+\delta) = 0$ para ciertos ρ, δ, entonces $X(t,\rho) = X(t,\rho+\delta) = 0$ para todo t: $f(t)$ tendrá siempre una masa de Dirac en 0 con el mismo peso.

Evolución del promedio y de la varianza

Observamos que la varianza de $f(t)$ viene dada por:

$$V(t) = \int_{-1}^1 (x-m(t))^2\, f(t,x)dx = \int_0^1 (X(\rho,t)-m(t))^2\, d\rho \tag{3.19}$$

donde usamos (3.5).

Lema 3.5.2. *La función varianza $V(t)$ definida anteriormente es una funciń montona, diferenciable y $V(t) \in [0,1]$.*

Demostración. Como $X(t,\rho)$ es diferenciable en t para todo ρ con $|\partial_t X(t,\rho)| \leq Cste$, podemos derivar V aplicado el teorema de convergencia mayorada:

$$\begin{aligned}
V'(t) &= \frac{d}{dt}\int_0^1 (X(\rho,t)-m(t))^2\, d\rho \\
&= 2\int_0^1 \partial_t(X(t,\rho)(X(t,\rho)-m(t))\,d\rho - m'(t)\int_0^1 (X(t,\rho)-m(t))\,d\rho \\
&= -2\gamma\int_0^1 \left(1-X^2(\rho,t)\right)(X(\rho,t)-m(t))^2 d\rho.
\end{aligned}$$

Entonces, $V(t)$ es monótona (creciente si $\gamma = -1$, decreciente sino). Además esta acotada por 1:

$$V(t) = \int_{-1}^{1} x^2 f(t,x)dx - m(t)^2 \leq \int_{-1}^{1} x^2 f(t,x)dx \leq 1.$$

□

De

$$V(t) = \int_{-1}^{1} x^2 f(t,x)dx - m(t)^2 \leq 1 - m(t)^2,$$

deducimos que

Lema 3.5.3. *Para todo $t > 0$, $m(t) \in \left[-\sqrt{1-V(t)}, \sqrt{1-V(t)}\right]$.*

Lema 3.5.4. *La funcion $m' : \mathbb{R} \to [-1,1]$ esta en $L^2(\mathbb{R})$. Mas aun, existe $\lim_{t\to\infty} m'(t) = 0$.*

Demostración. Derivando $m(t) = \int_0^1 X(t,\rho)\, d\rho$, tenemos

$$m'(t) = -\gamma \int_0^1 (1 - X(t,\rho)^2)(X(t,\rho) - m(t))\, d\rho. \tag{3.20}$$

Elevamos al cuadrado aplicando Cauchy-Schwarz:

$$\begin{aligned} (m'(t))^2 &\leq \int \left(1 - X(\rho,t)^2\right) \left(1 - X(\rho,t)^2\right) \left(X(\rho,t) - m(t)\right)^2 d\rho \\ &\leq \int \left(1 - X(\rho,t)^2\right) \left(X(\rho,t) - m(t)\right)^2 d\rho \\ &= -\frac{1}{2\gamma} V'(t). \end{aligned}$$

Como V es acotada, concluimos que $m' \in L^2(\mathbb{R})$. Además m' es diferenciable y podemos derivar bajo la integral en (3.20). Se ve fácilmente que $|m''(t)| \leq Cte$. De ahi la conclusión. □

3.5.2. Esparcimiento de opiniones ($\gamma = -1$)

En esta sección estudiamos el comportamiento asintótico $t \to +\infty$ de $f(t)$ cuando $\gamma = -1$. Por lo visto anteriormente, si por ejemplo f_0 es simétrica entonces $f(t)$ es simétrica para todo t y si hay una masa de Dirac inicialmente en 0, esta masa de Dirac permanecerá siempre. Para evitar eso, supondremos que las unicas masas de Dirac posibles en f_0 estan en los extremos ± 1:

$$X(\rho_1, 0) = X(\rho_2, 0) \qquad \text{si y solo si } \rho_1 = \rho_2 \text{ o } X(\rho_1, 0) = \pm 1. \tag{3.21}$$

Teorema 3.5.5. *Supongamos que se cumple* (3.21). *Entonces cuando* $t \to +\infty$, $f(t)$ *converge a una medida soportada en* 1 *y* -1.

Demostración. Dividimos la prueba en varias etapas.

Etapa 1 *Identificamos los pesos de las masas de Dirac límite en* ± 1.

Recordemos que $X(., \rho)$ verifica la ecuación

$$\partial_t X(\rho, t) = (1 - X(\rho, t)^2)(X(\rho, t) - m(t)).$$

Luego el comportamiento de $X(., \rho)$ esta determinado por el signo de $X(t, \rho) - m(t)$. Sabemos por el lema 3.5.2 que $V(t)$ es creciente. Luego obtenemos del lema 3.5.3 que

$$-\sqrt{1 - V(0)} \leq m(t) \leq \sqrt{1 - V(0)}.$$

Deducimos que si para algún t_0 y algún ρ se tiene que

- $X(\rho, t_0) > \sqrt{1 - V(0)}$ entonces claramente $X(\rho, t) > \sqrt{1 - V(0)}$ para todo $t \geq t_0$ y $\lim_{t \to +\infty} X(\rho, t) = 1$ en forma creciente;
- $X(\rho, t_0) < -\sqrt{1 - V(0)}$ entonces claramente $X(\rho, t) < -\sqrt{1 - V(0)}$ para todo $t \geq t_0$ y $\lim_{t \to +\infty} X(\rho, t) = -1$ en forma decreciente.

Esta alternativa sale de estudiar una EDO en una dimension de la forma $x' = (1 - x^2)(x - a)$ con $|a|, |x(0)| < 1$. Si $x_0 > a$ entonces $x(t)$ es monótona creciente a 1 y si $x_0 < a$ entonces $x(t)$ es monótona decreciente a -1.

En consecuencia vemos también que las funciones

$$p_{-1}^h(t) = \sup \left\{ \rho \in (0, 1) : X(\rho, t) < -\sqrt{1 - V(0)} \right\}$$

$$p_1^h(t) = \inf \left\{ \rho \in (0, 1) : X(\rho, t) > \sqrt{1 - V(0)} \right\}$$

son no-decreciente y no-creciente respectivamente. Por ejemplo, dado $\varepsilon > 0$, sea ρ tal que $\rho \geq p_{-1}^h(t) - \varepsilon$ y $X(\rho, t) < -\sqrt{1 - V(0)}$. Entonces para $t' > t$, tenemos $X(\rho, t') \leq X(\rho, t) < -\sqrt{1 - V(0)}$ por lo que $p_{-1}^h(t') \geq \rho \geq p_{-1}^h(t) - \varepsilon$. Entonces $p_{-1}^h(t') \geq p_{-1}^h(t)$.

Podemos entonces definir

$$p_{-1}^h := \lim_{t \to \infty} p_{-1}^h(t) \qquad \text{y} \qquad p_1^h := \lim_{t \to \infty} p_1^h(t).$$

Tenemos finalmente

$$\lim_{t \to \infty} X(\rho, t) = -1 \qquad \text{para todo } \rho \in [0, p_{-1}^h),$$

$$\lim_{t \to \infty} X(\rho, t) = 1 \qquad \text{para todo } \rho \in (p_1^h, 1].$$

De hecho si por ejemplo $\rho \in [0, p_{-1}^h)$ entonces $\rho < p_{-1}^h(t_0)$ para algún t_0 por lo que $X(\rho, t_0) < -\sqrt{1 - V(0)}$. Entonces $\lim_{t\to+\infty} X(\rho, t) \to -1$.

Etapa 2 *Probamos que $p_{-1}^h = p_1^h$.*

Supongamos por el absurdo que $p_{-1}^h \neq p_1^h$. Entonces

$$X(\rho, t) \in [-\sqrt{1 - V(0)}, \sqrt{1 - V(0)}] \qquad \text{para todo } \rho \in (p_{-1}^h, p_1^h) \text{ y todo } t.$$

es decir

$$1 - X(\rho, t)^2 \geq V(0) \qquad \text{para todo } \rho \in (p_{-1}^h, p_1^h) \text{ y todo } t. \tag{3.22}$$

Por el Lema 3.5.4 sabemos que existe $T > 0$ tal que $|m'(t)| \leq \frac{V(0)^2}{4}$ para todo $t \leq T$. Veamos que

$$|X(\rho, t) - m(t)| \leq \frac{V(0)}{4} \qquad \text{para todo } t > T \text{ y todo } \rho \in (p_{-1}^h, p_1^h). \tag{3.23}$$

Supongamos por el absurdo que existe $t_0 \geq T$ y $\rho \in (p_{-1}^h, p_1^h)$ tal que $|X(\rho, t_0) - m(t_0)| > \frac{V(0)}{4}$, digamos $f(t_0) := X(\rho, t_0) - m(t_0) > \frac{V(0)}{4}$. Notemos que, para todo $t \geq T$, si $f(t) > \frac{V(0)}{4}$ entonces $f'(t) > 0$. De hecho usando (3.22) y la definición de T,

$$f'(t) = \partial_t X(\rho, t) - m'(t) = (1 - X(\rho, t)^2) f(t) - m'(t) > 0.$$

De ahi concluimos que $f(t) > \frac{V(0)}{4}$ para todo $t \geq t_0$. Entonces para $t \geq t_0$,

$$\partial_t X(\rho, t) = (1 - X(\rho, t)^2) f(t) \geq \frac{V(0)^2}{4}$$

lo que es absurdo ya que $|X(\rho, t)| \leq 1$. Entonces vale (3.23).

Consideramos ahora la función diferenciable $F(x, y) = (1 - x^2)(x - y)$. Sean $x_1 \geq x_2$. Por el Teorema de Lagrange existe $\psi \in (x_1, x_2)$ tal que:

$$F(x_1, y) - F(x_2, y) = (x_1 - x_2) \frac{\partial}{\partial x} F(x, y) \mid_{x=\psi}$$

Note que si $1 - x^2 \geq V(0)$ y $|x - y| \leq \frac{V(0)}{4}$ entonces

$$\frac{\partial}{\partial t} F(x, y) = 1 - 3x^2 + 2xy \geq V(0) + 2x(y - x) \geq \frac{V(0)}{2}.$$

Luego

$$F(x_1, y) - F(x_2, y) \geq (x_1 - x_2) \frac{V(0)}{2}$$

si $x_1^2, x_2^2 \leq 1 - V(0)$ y $|x_1 - y|, |x_2 - y| \leq \frac{V(0)}{4}$.

Sean ρ_1, ρ_2 tales que $p_{-1}^h < \rho_2 \leq \rho_1 < p_1^h$. Luego para todo $t > T$, como valen (3.22) y (3.23), la función $x(t) := X(\rho_1, t) - X(\rho_2, t)$ verifica

$$\begin{aligned} x'(t) &= \frac{\partial}{\partial t}\left(X(\rho_1, t) - X(\rho_2, t)\right) = F(X(\rho_1, t), m(t)) - F(X(\rho_2, t), m(t)) \\ &\geq \left(X(\rho_1, t) - X(\rho_2, t)\right)\frac{V(0)}{2} \\ &= \frac{V(0)}{2} x(t). \end{aligned}$$

Como $V(0) > 0$, si $x(t') > 0$ para algún $t' > T$, entonces $x(t) \geq x(t')e^{\frac{V(0)}{2}(t-t')}$ para $t > t'$ por lo que $x(t) = X(\rho_1, t) - X(\rho_2, t) \to +\infty$ cuando $t \to +\infty$ lo cual es absurdo pues $x(t)$ es acotada. Entonces

$$X(\rho_1, t) = X(\rho_2, t) \qquad \text{para todo } t \geq T \text{ y } \rho_1, \rho_2 \in (p_{-1}^h, Zp_1^h.$$

Por el Lema (3.5.1), esta propiedad debe valer también al tiempo $t = 0$ lo cual es imposible por la hipótesis (3.21). Concluimos que $p_{-1}^h = p_1^h$

Etapa 3 *Conclusión de la prueba.*

Notemos $f_\infty := p_{-1}^h\ \delta_{-1} + (1 - p_{-1}^h)\ \delta_1$. La inversa generalizada de la acumulada de f_∞ es la función X definida por $X(\rho) = -1$ si $\rho \in [0, p_1^h]$ y $X(\rho) = 1$ si $\rho \in (p_1^h, 1]$.

Por las etapas anteriores sabemos que $\lim_{t\to+\infty} X(\rho, t) = X(\rho)$ para todo $\rho \in [0, 1] - \{p_{-1}^h\}$. Luego por el teorema de convergencia mayorada

$$W_1(f(t), f_\infty) = \int_0^1 |X(\rho, t) - X(\rho)|\, d\rho \overset{t\to+\infty}{\longrightarrow} 0.$$

Hemos entonces probado que

$$\lim_{t\to+\infty} f(t) = p_{-1}^h\ \delta_{-1} + (1 - p_{-1}^h)\ \delta_1. \tag{3.24}$$

□

El siguiente corolario identifica el límite del promedio $m(t)$:

Corolario 3.5.6. *Supongamos que se cumple* (3.21)*. Entonces, existe* $\lim_{t\to\infty} m(t) = 1 - 2p_{-1}^h$.

Demostración. De (3.24), tenemos

$$\begin{aligned} m(t) &= \int x f(t, x) dx \to \int x(p_{-1}^h\ \delta_{-1} + (1 - p_{-1}^h)\ \delta_1) = -p_{-1}^h + (1 - p_{-1}^h) \\ &= 1 - 2p_{-1}^h. \end{aligned}$$

□

3.5.3. Concentración de opiniones ($\gamma = 1$)

Analizamos ahora el comportamiento asintótico cuando $t \to +\infty$ de la solución $f(t)$ de la ecuación (3.1) en el casi $\gamma = -1$. Vamos a suponer en toda esta sección que existe el límite del promedio $m_\infty := \lim_{t\to+\infty} m(t)$. Probaremos que existe el límite $f_\infty := \lim_{t\to+\infty} f(t)$ y que en general es una combinación convexa de masas de Dirac centrada en ± 1 y en m_∞. Proporcionaremos además fórmulas para determinar m_∞ en varios casos.

Empezamos por un caso muy fácil: cuando $m_\infty = \pm 1$.

Proposición 3.5.7. *Si* $m_\infty = \pm 1$ *entonces* $\lim_{t\to+\infty} f(t) = f_\infty := \delta_{m_\infty}$.

Demostración. Supongamos que $m_\infty = -1$ (el caso $m_\infty = 1$ es completamente análogo). Como I es compacto, sabemos que $P(I)$ es compacto por lo que existe $t_n \to +\infty$ y $f_\infty \in P(I)$ tal que $f(t_n) \to f_\infty$. Luego,

$$-1 = m_\infty = \int_{-1}^{1} w f_\infty(w) dw = \lim_{n\to+\infty} \int_{-1}^{1} w f(t_n, w) dw \geq -1.$$

Luego $\int_{-1}^{1} w f_\infty(w) dw = -1$ por lo que f_∞ es una medida de probabilidad soportada en $\{-1\}$ o sea $f_\infty = \delta_{-1}$. Como $P(I)$ es compacto deducimos que $\lim_{t\to+\infty} f(t) = \delta_{-1}$. $\square$

Examinamos ahora el caso $m_\infty \in (-1, 1)$. Enunciamos primero un resultado preliminar que será útil al momento de identificar m_∞.

Lema 3.5.8. *Supongamos que* $\log\left(\frac{1+X(\rho,0)}{1-X(\rho,0)}\right) \in L^1(0,1)$. *Luego, para todo* $t \in \mathbb{R}$, *la función*

$$T(t) := \int_0^1 \log\left(\frac{1+X(\rho,t)}{1-X(\rho,t)}\right)$$

esta bien definida. Además, T *es una función diferenciable y* $T'(t) = 0$.

Demostración. Como $\log\left(\frac{1+X(\rho,0)}{1-X(\rho,0)}\right) \in L^1(0,1)$ concluimos que $X(\rho, 0) \in (-1, 1)$ para casi todo $\rho \in (0, 1)$. Fijamos ρ tal que $X(\rho, 0) \in (-1, 1)$. Entonces $X(\rho, t) \in (-1, 1)$ para todo t ya que las funciones constantes ± 1 son soluciones de (3.7). Luego la función

$$G(\rho, t) := \ln\left(\frac{1 + X(\rho, t)}{1 - X(\rho, t)}\right)$$

está bien definida. Observamos que existe $\partial_t G(\rho, t)$ para todo $t \in [-T, T]$ y que es una función acotada ya que un cálculo directo usando la ecuación (3.7) da

$$|\partial_t G(\rho, t)| = | - 2(X(\rho, t) - m(t)| \leq 2.$$

Luego
$$|G(\rho,t)| \leq |G(\rho,t) - G(\rho,0)| + |G(\rho,0)| \leq |G(\rho,0)| + 2T.$$
Entonces $G(\rho,t) \in L^1((0,1))$ y $T(t)$ esta bien definido. Además como $|\partial_t G| \leq 2 \in L^1(0,1)$, podemos derivar T derivando el integrando:

$$T'(t) = \int_0^1 \partial_t G(\rho,t)\, d\rho = -2\int_0^1 X(\rho,t) - m(t) d\rho = 0.$$

□

Definimos
$$p_{-1} := f_0(\{-1\}) \qquad \text{y} \qquad p_1 := f_0(\{1\})$$
la proporción iniciales de opinión ± 1. Recordando el Lema (3.5.1), vale que

$$p_{-1} := f_t(\{-1\}) \qquad \text{y} \qquad p_1 := f_t(\{1\}) \tag{3.25}$$

par todo $t > 0$. Entonces, para todo t, queda un total de $1 - p_{-1} - p_1$ masa distribuida entre las opiniones -1 y 1.

Enunciamos ahora el resultado acerca del comportamiento asintótico de $f(t)$ cuando $m_\infty \in (1,1)$:

Teorema 3.5.9. *Supongamos que $m_\infty \in (-1,1)$. Se tiene que*

$$\lim_{t\to+\infty} f(t) = f_\infty$$

donde la medida límite f_∞ viene dada por

$$f_\infty = p_{-1}\delta_{-1} + p_1\delta_{-1} + (1 - p_{-1} - p_1)\delta_{m_\infty}.$$

Demostración. Notamos X_∞ la inversa generalizada de la acumulada de f_∞. Tenemos

$$X_\infty(\rho) = \begin{cases} -1 & \text{si } 0 \leq \rho \leq p_{-1} \\ m_\infty & \text{si } p_{-1} < \rho \leq 1 - p_1 \\ 1 & \text{si } 1 - p_1 < \rho \leq 1 \end{cases}$$

Debemos probar (ver Apéndice B.2.3) que

$$W_1(f(t), f_\infty) = \int_{-1}^1 |X(\rho,t) - X_\infty(\rho)|\, d\rho \overset{t\to+\infty}{\longrightarrow} 0.$$

Por el teorema de convergencia mayorada, basta ver que

$$X(\rho,t) \to X_\infty(\rho) \qquad \text{para casi todo } \rho \in [0,1].$$

Por (3.25), ya tenemos esta convergencia para $\rho \in [0, p_{-1}]$ y $\rho \in (1-p_1, 1]$. Sea entonces $\rho \in (p_{-1}, 1-p_1)$. Consideramos $z(t) = (X(\rho, t) - m(t))^2$, entonces

$$\begin{aligned} z'(t) &= 2(X(\rho,t) - m(t))(\partial_t X(\rho,t) - m'(t)) \\ &= 2(X(\rho,t) - m(t))\partial_t X(\rho,t) + \varepsilon(t) \end{aligned}$$

con $\varepsilon(t) := -(X(\rho,t) - m(t))m'(t)$. Por el Lema 3.5.4, sabemos que $\lim_{t\to+\infty} m'(t) = 0$ y como tanto $X(\rho,t)$ como $m(t)$ son acotadas, tenemos $\lim_{t\to+\infty} \epsilon(t) = 0$. Como $X(\rho, .)$ verifica (3.7), tenemos

$$\begin{aligned} z'(t) &= -2(X(\rho,t) - m(t))^2(1 - X^2(\rho,t)) + \varepsilon(t) \\ &= -2z(t)(1 - X^2(\rho,t)) + \varepsilon(t). \end{aligned}$$

Supongamos que existe un $\alpha > 0$ y $t_0 > 0$ tal que :

$$1 - X^2(\rho,t) \geq \alpha > 0 \qquad \text{par todo } t \geq t_0. \tag{3.26}$$

Entonces para todo $t \geq t_0$, $z'(t) \leq -2z(t) + \varepsilon(t)$ por lo que la función $y(t) := e^{2\alpha(t-t_0)}z(t)$ verifica $y'(t) \leq e^{2\alpha(t-t_0)}\varepsilon(t)$. Integrando esta relación de t' a t con $t_0 < t' < t$, obtenemos

$$z(t) \leq z(t')e^{-2\alpha(t-t')} + \int_{t'}^{t} e^{2\alpha(s-t)}\epsilon(s)ds.$$

Como $\int_{t'}^{t} e^{2\alpha(s-t)}ds \leq 1/2\alpha$, obtenemos

$$z(t) \leq z(t')e^{-2\alpha(t-t')} + \sup_{s\geq t'} |\varepsilon(s)|/2\alpha.$$

Eligiendo $t' \gg 1$ deducimos que $\lim_{t\to+\infty} z(t) \to_{t\to\infty} 0$. Finalmente, recordando que $m(t) \to m_\infty$, obtenemos que $X(\rho,t) \to m_\infty = X_\infty(\rho)$.

Finalmente probamos (3.26). Como $m(t) \to m_\infty \in (-1, 1)$, existen $\xi > 0$ y $t_0 > 0$ tales que $-1 + \xi < m(t) < 1 - \xi$ para todo $t \geq t_0$. Como $\rho \in (p_{-1}, p_1)$ y vale (3.25), $X(\rho,t) \in (-1,1)$ para todo $t > 0$. Por lo tanto existe $\delta > 0$ tal que $-1+\delta < X(\rho, t_0) < 1 - \delta$. Sea $\epsilon = \min\{\xi, \delta\}$. Veamos que $-1 + \epsilon < X(\rho,t) < 1 - \epsilon$ para todo $t \geq t_0$. Para eso consideremos

$$T = \sup\left\{t \geq t_0 :\, -1 + \epsilon < X(\rho, s) < 1 - \epsilon \text{ para todo } s \in [t_0, t]\right\}.$$

Si por el absurdo $T < \infty$ entonces $X(\rho,T) = -1 + \epsilon$ o $X(\rho,T) = 1 - \epsilon$. Analicemos el caso $X(\rho,T) = -1 + \epsilon$ (se procede analogamente si $X(\rho,T) = 1 - \epsilon$).

- Si $X(\rho,T) \geq m(T)$ entonces $X(\rho,T) > -1 + \xi \geq -1 + \epsilon$, absurdo.
- Si $X(\rho,T) < m(T)$ entonces $\partial_t X(\rho,T) = -(1 - X^2(\rho,T))(X(\rho,T) - m(T)) > 0$. Por lo tanto, existe $h > 0$ chico tal que $X(\rho, T - h) < X(\rho,T) = -1 + \epsilon$, lo cual es absurdo por la definicion de T.

Por lo tanto, $T = \infty$ y vale (3.26). $\square$

El siguiente resultado da formulas para calcular m_∞ en término de p_1 y p_{-1} en varios casos:

Teorema 3.5.10. *En el marco del Teorema 3.5.9, se tiene que*

$$m_\infty = \begin{cases} \dfrac{p_1 - p_{-1}}{p_1 + p_{-1}} & \text{si } p_1 \text{ o } p_{-1} > 0, \\ \dfrac{e^{T(0)} - 1}{e^{T(0)} + 1} & \text{si } p_1 = p_{-1} = 0 \end{cases}$$

En el caso $p_1 = p_{-1} = 0$ asumimos además que f_0 esta soportada en $(-1+\delta, 1-\delta)$.

Demostración. Por el Teorema(3.5.9) sabemos que

$$f(t) \overset{t\to\infty}{\longrightarrow} f_\infty := p_1\delta_1 + p_{-1}\delta_{-1} + (1 - p_{-1} - p_1)\delta_{m_\infty}.$$

Luego

$$\begin{aligned} m_\infty &= \lim m(t) = \lim \int_{-1}^{1} w f(w,t)dw = \int_{-1}^{1} w f_\infty(w)dw \\ &= p_1 - p_{-1} + (1 - p_{-1} - p_1)m_\infty. \end{aligned}$$

Si $p_1 + p_{-1} > 0$, podemos despejar $m\infty$ para obtener $m_\infty = \frac{p_1 - p_{-1}}{p_1 + p_{-1}}$.

Supongamos ahora que $p_1 = p_{-1} = 0$ y que f_0 esta soportada en $(-1+\delta, 1-\delta)$. Entonces existe $\delta' > 0$ tal que $f(t)$ esta soportada en $(-1+\delta', 1-\delta')$ para todo t. Se puede ver eso exactamente como en la última parte de la prueba del teorema anterior. Por el lema 3.5.8, la funcion $T(t)$ es constante. Como $f_\infty = \delta_{m_\infty}$ obtenemos entonces

$$\begin{aligned} T(0) = T(t) &= \int_0^1 \ln\left(\frac{1 + X(\rho,t)}{1 - X(\rho,t)}\right) d\rho = \int_{-1}^{1} \ln\left(\frac{1+x}{1-x}\right) f(x,t)dx \\ &\to \int_{-1}^{1} \ln\left(\frac{1+x}{1-x}\right) f_\infty(x)dx = \ln\left(\frac{1+m_\infty}{1-m_\infty}\right) \end{aligned}$$

Pasamos el límite usando el teorema de convergencia mayorada ya que $f(t)$ esta soportada en $(-1+\delta', 1-\delta')$ para todo t. Concluímos despejando m_∞. $\square$

Apéndice A

La ecuación de conservación de masa

La ecuación de conservación de masa es la ecuación en derivadads parciales del 1er orden dada por

$$\partial_t \rho + \operatorname{div}(v\rho) = 0 \tag{A.1}$$

donde $v = \mathbb{R}^n \to \mathbb{R}^n$ es un campo de vectores dado y $\rho : \mathbb{R}^n \to \mathbb{R}$. Esta ecuación aparece por ejemplo en mecánica de los fluidos. En este caso ρ es la densidad de masa de un fluido y v es el campo de velocidades. En la ausencia de creación o de destrucción de masa, esta ecuación expresa matemáticamente el hecho que mientras fluye, la masa se conserva.

Presentamos dos formas de deducir (A.1), la primera desde el punto de vista lagrangiano y la segunda desde una perspectiva euleriana. Presentamos luego, en el caso $n = 1$, un método elemental que permite resolver facilmente (A.1) una vez conocido el flujo de v. Aplicamos este método a la resolución de la ecuación obtenida como límite continuo del modelo “dos contra uno”. Aplicamos finalmente el método de las características para resolver (A.1) en el caso que la condición inicial $\rho_{|t=0}$ es una medida.

A.1. Deducción de la ecuación de conservación de masa.

El movimiento de los fluídos puede describirse desde dos puntos de vista: desde la perspectiva euleriana o desde la perspectiva lagrangiana. En la perspectiva Lagrangiana, se describe el fluido por el conocimiento de la trayectoria de cada partícula o sea mediante las funciones $t \to \phi_t(x)$ si la partícula esta inicialmente localizada en el punto $x \in \mathbb{R}^n$. Acá $\phi_t(x)$ es el flujo de v. En la perspectiva Euleriana, se pone el acento en el campo de velocidades $v(t,x)$. Las ecuaciones de la mecánica de los fluidos usan principalmente la descripción euleriana. Obviamente se puede pasar de la perspectiva Euleriana a la Lagrangina resolviendo para cada $x \in \mathbb{R}^n$ la ecuación $\partial_t\phi_t(x) = v(t,\phi_t(x))$ con la condición inicial $\phi_0(x) = x$.

Mostramos dos maneras de deducir (A.3) usando las perspectivas antes descriptas. Asumiremos por simplicidad que v no depende del tiempo o sea que el fluido está en estado estacionario.

A.1.1. Perspectiva Lagrangiana

Consideramos un fluido en movimiento en $\mathbb{R}^n$ de densidad de masa $\rho(t,x)$ que suponemos en estado estacionario en el sentido que el campo de velocidad v del fluido no depende del tiempo t por lo que $v : \mathbb{R}^n \to \mathbb{R}^n$. La trayectoria de una partícula del fluido inicialmente en $x \in \mathbb{R}^n$ es la función $t \to \phi_t(x)$ solución de

$$\begin{cases} \partial_t\phi_t(x) = v(\phi_t(x)) \quad t \in \mathbb{R},\ x \in \mathbb{R}^n, \\ \phi_0(x) = x \quad x \in \mathbb{R}^n. \end{cases} \tag{A.2}$$

Suponemos además que v es globalmente Lipschitz. Entonces $\phi : (t,x) \in \mathbb{R}\times\mathbb{R}^n \to \mathbb{R}^n$ está bien definida, verifica la propiedad de semi-grupo $\phi_t(\phi_s(x)) = \phi_{t+s}(x)$ para todo $s,t \in \mathbb{R}$, $x \in \mathbb{R}^n$, y, finalmente, para todo $t \in \mathbb{R}$, $\phi_t : \mathbb{R}^n \to \mathbb{R}^n$ es un difeomorfismo con inversa ϕ_{-t}.

Consideramos un volumen de fluido $U \subset \mathbb{R}^n$ al instante inicial. En un tiempo t cualquiera, el fluido inicialmente en U ocupa el volumen $U_t := \phi_t(U)$. La masa de fluido en U_t es

$$m(t) = \int_{U_t} \rho(t,x)\,dx.$$

Como supusimos que no hay creación ni tampoco destrucción de masa, el principio físico de conservación de masa afirma que la función m debe ser constante. Calculando la derivada de m, vamos a ver que este principio se traduce por el hecho que ρ debe ser solución de la ecuación del 1er orden

$$\partial_t\rho(t,x) + div(\rho(t,x)v) = 0. \tag{A.3}$$

Teorema A.1.1. *Supongamos que $v : \mathbb{R}^n \to \mathbb{R}^n$ es de clase C^2 con $v, \partial_i v, \partial_{ij} v$ acotadas en $\mathbb{R}^n$ para todo i,j. Supongamos además que ρ es de clase C^2 en (t,x) con $\partial_i\rho, \partial_{ij}\rho$ acotadas en $\mathbb{R}^n$ para todo i,j. Entonces*

$$\frac{d}{dt}\int_{U_t} \rho(t,x)\,dx = \int_{U_t} \partial_t\rho(t,y) + div(v(y)\rho(t,y))dy.$$

Demostración. Vamos a hacer un desarrollo de Taylor de la función $m(t) = \int_{U_t} \rho(t,x)\,dx$ donde $U_t = \phi_t(U)$. Como el flujo $\phi_t(x)$ tiene la propiedad de semi-grupo, tenemos $U_{t+h} = \phi_{t+h}(U) = \phi_h(\phi_t(U)) = \phi_h(U_t)$ por lo que

$$m(t+h) = \int_{U_{t+h}} \rho(t+h,x)dx = \int_{\phi_h(U_t)} \rho(t+h,x)dx.$$

Recordando que $\phi_h : \mathbb{R}^n \to \mathbb{R}^n$ es un difeomorfismo, podemos realizar el cambio de variables $x = \phi_h(y)$ para obtener

$$m(t+h) = \int_{U_t} \rho(t+h,\phi_h(y))|Det(D\phi_h(y))|\,dy. \tag{A.4}$$

En lo siguiente, C denotará una constante que puede cambiar de un renglón al siguiente pero que será siempre independiente de $x \in \mathbb{R}^n$ y h cerca de 0. Además los restos $O(h^2)$ serán siempre uniformes en $x \in \mathbb{R}^n$.

Aplicando la fórmula de Taylor con resto integral, recordando que $\phi_0(x) = x$, tenemos

$$\begin{aligned} \phi_h(x) &= \phi(h,x) = \phi(0,x) + h\partial_t\phi(0,x) + R(h,x) \\ &= x + hv(x) + R(h,x) \end{aligned} \tag{A.5}$$

con

$$R(h,x) = \int_0^h \tilde{R}(s,x)(h-s)\,ds, \qquad \tilde{R}(s,x) = \partial_{tt}\phi(s,x).$$

Derivando (A.2) con respecto a t, tenemos

$$\tilde{R}(s,x) = \partial_{tt}\phi(t,x) = \partial_t(v(\phi(t,x)) = Dv(\phi(t,x))\partial_t\phi(t,x) = Dv(\phi(t,x))v(t,x).$$

Luego $|\tilde{R}(s,x)| \leq C$ por lo que $|R(h,x)| \leq Ch^2$. Entonces

$$\phi_h(x) = x + hv(x) + O(h^2). \tag{A.6}$$

Como ρ es de clase C^2 con derivadas primeras y segundas uniformemente acotadas, obtenemos

$$\begin{aligned} \rho(t+h,\phi_h(y)) &= \rho(t+h, y + hv(y) + O(h^2)) \\ &= \rho(t,y) + h(\partial_t\rho(t,y) + \nabla\rho(t,y)v(y)) + O(h^2). \end{aligned} \tag{A.7}$$

Además derivando (A.5) con respecto a x obtenemos que

$$D\phi_h(x) = Id + hDv(x) + DR(h,x).$$

Justifiquemos ahora que $|DR(h,x)| \leq O(h^2)$. Derivando $\tilde{R}(s,x) = Dv(\phi(t,x))v(t,x)$ con respecto a x y recordando que las derivadas de v hasta el orden dos estan acotodas, obtenemos

$$|D\tilde{R}(s,x)| \leq C|D\phi_s(x)|.$$

Para acotar $A(s) := D\phi_s(x)$, derivamos (A.2) con respecto a x para obtener

$$A'(s) = Dv(\phi_s(x))A(s)$$

o sea para todo i,

$$A_i'(s) = Dv(\phi_s(x))A_i(s)$$

donde A_i es la i-ésima columna de A. Sea $x(s) := \frac{1}{2}\|A_i(s)\|^2$. Como Dv es acotada obtenemos

$$x'(s) = (A_i'(s), A_i(s)) = (Dv(\phi_s(x))A_i(s), A_i(s)) \leq Cx(s).$$

Deducimos que $x(s) \leq x(0)e^{Cs}$. Como $\phi_0 = Id$ tenemos $A(0) = Id$ y luego $x(s) \leq e^{Cs}$. Entonces $|D\phi_s(x)| \leq C$ y $|D\tilde{R}(s,x)| \leq C$. Deducimos que $|DR(h,x)| \leq O(h^2)$. Finalmente

$$D\phi_h(x) = Id + hDv(x) + O(h^2). \tag{A.8}$$

Recordamos que $Det(Id + tH) = 1 + tTr(H) + O(t^2)$ donde el $O(t^2)$ es uniforme en H para H es un conjunto acotado de matrices. Como Dv es acotada, deducimos de (A.8), observando que $Tr(Dv(x)) = \mathrm{div}v$, que

$$Det\,(D\phi_h(x)) = 1 + h\ div(v(x)) + O(h^2). \tag{A.9}$$

Volvamos ahora a (A.4). En vista de (A.7) y (A.9), tenemos

$$m(t+h) = \int_{U_t} (\rho(t,x) + h(\partial_t\rho(t,x) + \nabla\rho(t,x)v(x)) + O(h^2))(1 + h\mathrm{div}(v(y)) + O(h^2))dy.$$

Luego

$$\begin{aligned} m(t+h) - m(t) &= h\int_{U_t} \partial_t\rho(t,y) + v(y)\nabla\rho(t,y) + \rho(t,y)div(v(y))dy + O(h^2)|U_t| \\ &= h\int_{U_t} \partial_t\rho(t,y) + div(v(y)\rho(t,y))dy + O(h^2). \end{aligned}$$

Deducimos el resultado. $\square$

Obtenemos en consecuencia que

Teorema A.1.2. *Bajo las mismas hipótesis que en el teorema A.1.1, tenemos conservación de la masa si y solo si*

$$\partial_t \rho(t,y) + \rho(t,y) div(v(y)\rho(t,y)) = 0.$$

Demostración. En vista del teorema anterior, decir que hay conservación de la masa significa que

$$\int_{U_t} \partial_t \rho(t,y) + \rho(t,y) div(v(y)\rho(t,y)) dy = 0$$

para todo tiempo t y cualquier abierto U. Como $\phi_t : \mathbb{R}^n \to \mathbb{R}^n$ es un difeomorfismo, obtenemos

$$\int_{V} \partial_t \rho(t,y) + \rho(t,y) div(v(y)\rho(t,y)) dy = 0$$

para cualquier abierto V y cualquier tiempo t. Obtenemos el resultado. $\square$

A.1.2. Perspectiva euleriana

La segunda forma que proponemos ahora para obtener la ecuación de conservación de masa (A.3) es mas bien formal. Consideramos un abierto acotado suave $U \subset \mathbb{R}^n$. Como hicimos anteriormente, estamos interesados en evaluar la variación de la masa de fluidos en U es decir la derivada de $m(t) = \int_U \rho(t,x)\, dx$.

Primero, suponiendo que se puede derivar bajo la integral, tenemos

$$m't) = \int_U \partial_t \rho(t,x)\, dx.$$

Por otro lado, como no hay ni creación ni tampoco destrucción de masa, la variaión de masa en U se debe al balance total entre la masa que sale y entra en U o sea $m'(t) = -\int_{\partial U} \rho v n\, d\sigma$ donde n es el vector normal exterior unitario. Igualando las dos expresiones obtenemos

$$\int_U \partial_t \rho(t,x)\, dx = m'(t) = -\int_{\partial U} \rho v n\, d\sigma.$$

Aplicando el teorema de la divergencia, deducimos

$$\int_U \partial_t \rho(t,x)\, dx = -\int_U \operatorname{div}(\rho v)\, dx$$

o sea

$$\int_U \partial_t \rho(t,x) + \operatorname{div}(\rho v)\, dx = 0.$$

Como esta relación vale para todo $U \subset \mathbb{R}^n$, obtenemos (A.3).

A.2. Resolución de la ecuación de conservación de masa en $n = 1$.

El principal objetivo de esta sección consiste en obtener la solución de la ecuación

$$\frac{\partial u(x,t)}{\partial t} + \partial_x(v(x)u(x,t)), \qquad v(x) = x(1-x^2), \tag{A.10}$$

que obtuvimos en el primer capítulo como límite continuo del modelo «2 contra 1». Para eso recordaremos las consecuencias del método de las caracteristicas para la ecuación de conservación de masa (A.1). Luego una simple observación nos permitirá obtener facilmente la solución de (A.1) en el caso $n = 1$, en particular de (A.10). Para concluir, daremos la solución de (A.1) cuando el dato inicial es una medida.

A.2.1. El método de las características en el caso $n = 1$.

Consideramos la ecuación (A.1) con campo de velocidades $v = v(x)$ que suponemos C^1 y globalmente Lipschitz. Seguimos notando $\phi : (t,x) \to \phi_t(x)$ su flujo.

Notando $\tilde{x} := (t,x)$, $\tilde{v} = (1,v)$, $Du = \nabla_{t,x}u$ y $c = \text{div } v$, podemos reescribir (A.1) como

$$\tilde{v}Du + cu = 0.$$

Según [6][3.2.2.a], una aplicación directa del método de las caracteristicas da que la función $z(t) := u(t,\phi_t(x))$ verifica

$$z'(t) = -c(\phi_t(x))z(t).$$

Obviamente se puede obtener esta relación directamente sin usar toda la teoría desarrollada en [6] derivando $u(t,\phi_t(x))$ y usando la ecuación (A.1).

En el caso $n = 1$ que nos interesa para resolver (A.10), obtenemos entonces

$$\frac{d}{dt}u(t,\phi_t(x)) = -v'(\phi_t(x))u(t,\phi_t(x)).$$

El siguiente resultado es una consecuencia directa de esta igualdad:

Proposición A.2.1. *La funcion $t \to v(\phi_t(x))u(t,\phi_t(x))$ es constante y*

$$v(x)u(t,x) = v(\phi_t^{-1}(x))u(0,\phi_t^{-1}(x)). \tag{A.11}$$

Demostración. Vimos que para un $x \in \mathbb{R}$ dado, la función $\widehat{u}(t) = u(t,\phi_t(x))$ verifica

$$\widehat{u}'(t) + v'(\phi_t(x))\widehat{u}(t) = 0 \qquad t \in \mathbb{R}.$$

Multiplicando por $v(\phi_t(x)) = \frac{d}{dt}\phi_t(x)$ esta ecuación, obtenemos

$$\begin{aligned} 0 &= v(\phi_t(x))\widehat{u}'(t) + v(\phi_t(x))v'(\phi_t(x))\widehat{u}(t) \\ &= \frac{d}{dt}\Big(v(\phi_t(x) \cdot \widehat{u}(t))\Big) \end{aligned}$$

por lo que la función $v(\phi_t(x) \cdot \widehat{u}(t)) =$ es constante en t. Luego

$$v(\phi_t(x))u(t, \phi_t(x)) = v(x)u(0, x)$$

para todo $(t, x) \in \mathbb{R} \times \mathbb{R}$. Deducimos finalmente (A.11) recordando que $\phi_t : \mathbb{R} \to \mathbb{R}$ es un difeomorfismo para todo $t \in \mathbb{R}$. $\square$

Como ejemplo de aplicación de la relación (A.11), resolvemos ahora (A.10). La ecuación (A.10) es un caso particular de ecuación de masa (A.1) con $v(x) = (1 - x^2)x$. Sabemos por la prop. A.2.1 que su solución verifica (A.11). Empecemos por hallar el flujo $\phi_t(x)$ asociado con v con $x \in (-1, 1)$. Consideramos entonces la ecuación diferencial

$$x'(t) = v(x(t)) = x(t)(1 - x(t)^2) \tag{A.12}$$

con $x(0) = x_0 \in (-1, 1)$ dado. Note que las funciones constantes iguales a 0,± 1 son soluciones. Luego si $x_0 \in (-1, 1)$, $x_0 \neq 0$, entonces para todo $t \in \mathbb{R}$, $0 < x(t) < 1$ si $0 < x_0 < 1$ y $-1 < x(t) < 0$ si $-1 < x_0 < 0$. Ademas, como v es impar, si x y $\tilde{x}$ son dos soluciones tales que $x(0) = -\tilde{x}(0)$ entonces $x(t) = -\tilde{x}(t)$ para todo t. Podemos entonces suponer que $x_0 \in (0, 1)$ (por lo que $x(t) \in (0, 1)$ para todo t).

Note que (A.12) es una ecuación con variables separadas:

$$1 = \frac{x'}{x(1 - x^2)} = \frac{x'}{x(1 - x)(1 + x)} = \frac{x'}{x} + \frac{1}{2}\frac{x'}{1 - x} - \frac{1}{2}\frac{x'}{1 + x}.$$

Integrando, obtenemos

$$t + c = \ln(x) - \frac{1}{2}\ln(1 - x) - \frac{1}{2}\ln(1 + x) = \ln\Big(\frac{x}{\sqrt{1 - x^2}}\Big)$$

y luego

$$\frac{x(t)}{\sqrt{1 - x(t)^2}} = \tilde{c}e^t.$$

Obtenemos el valor de la constante $\tilde{c}$ evaluando a $t = 0$. Finalmente

$$\frac{x(t)}{\sqrt{1 - x(t)^2}} = \frac{x_0}{\sqrt{1 - x_0^2}}e^t.$$

Elevando ambos miembros al cuadrado, podemos despejar $x(t)$ obteniendo

$$x(t) = \frac{x_0 e^t}{\sqrt{1 + x_0^2 e^{2t} - x_0^2}}.$$

Luego, recordando que ϕ_t es impar,

$$\phi_t(x) = \frac{xe^t}{\sqrt{1 + x^2e^{2t} - x^2}} \qquad x \in \mathbb{R},\, t \in \mathbb{R}.$$

Reemplazamos ahora en (A.11). Primero

$$\phi_t^{-1}(x) = \phi_{-t}(x) = \frac{xe^{-t}}{\sqrt{1 + x^2e^{-2t} - x^2}},$$

y luego

$$\begin{aligned}
v(\phi_t^{-1}(x)) &= \phi_{-t}(x)(1 - \phi_{-t}(x)^2) = \frac{xe^{-t}}{\sqrt{1 + x^2e^{-2t} - x^2}}\Big(1 - \frac{x^2e^{-2t}}{1 + x^2e^{-2t} - x^2}\Big) \\
&= \frac{xe^-t}{\sqrt{1 + x^2e^{-2t} - x^2}} - \frac{x^3e^{-3t}}{(1 + x^2e^{-2t} - x^2)^{\frac{3}{2}}} = \frac{xe^{-t}(1 - x^2)}{(1 + x^2e^{-2t} - x^2)^{\frac{3}{2}}} \\
&= \frac{v(x)e^{-t}}{(1 + x^2e^{-2t} - x^2)^{\frac{3}{2}}}.
\end{aligned}$$

Reemplazamos en (A.11) obtenemos que la solución de (A.10) es

$$u(t,x) = \frac{e^{-t}}{(1 + x^2e^{-2t} - x^2)^{\frac{3}{2}}} \cdot u_0\Big(\frac{xe^{-t}}{\sqrt{1 + x^2e^{-2t} - x^2}}\Big). \tag{A.13}$$

A.2.2. Usando la inversa generalizada.

Supongamos que u_0 es una densidad de probabilidad suave. por lo que $u(t,.)$, la solución de (A.10), es también una densidad de probabilidad suave para todo t tratandose de una ecuación de conservación de masa. En [1][Section 7] los autores resuelven (A.10) usando la inversa generalizada de la acumulada de $u(t,.)$. Vamos a ver que este método es equivalente a (A.11) en el caso general de la ecuación de conservación de masa (A.11).

Notamos X_t la inversa generalizada de la acumulada F_t de $u(t,.)$. Por lo visto en el capítulo anterior, para todo $\rho \in (0,1)$, $X_t(\rho)$ satisface la ecuación diferencial ordinaria

$$\partial_t X_t(\rho) = v(X_t(\rho))$$

ya que para toda función suave ψ, derivando con respecto a t la igualdad

$$\int \psi(x)u(t,x)dx = \int_0^1 \psi(X_t(\rho))\, d\rho$$

obtenemos

$$\begin{aligned}\frac{d}{dt}\int \psi(x)u(t,x)dx &= \int_0^1 \psi'(X_t(\rho))v(X_t(\rho))\,d\rho = \int \psi'(x)v(x)u(t,x)dx \\ &= -\int \psi(x)\partial_x(u(t,x)v(x))dx.\end{aligned}$$

Luego $X_t(\rho) = \phi_t(X_0(\rho))$. Tomando $\rho = F_0(x)$, obtenemos $X_t(F_0(x)) = \phi_t(x)$ o sea

$$X_t(F_0(\phi_{-t}(x))) = x.$$

Deducimos que

$$F_t(x) = F_0(\phi_{-t}(x)).$$

Derivando con respecto a x, obtenemos

$$u(t,x) = u_0(\phi_{-t}(x))\frac{d}{dx}\phi_{-t}(x).$$

Notemos $\psi_t(x) := \frac{d}{dx}\phi_t(x)$. Entonces $\partial_t\psi_t(x) = v'(\phi_t(x))\psi_t(x)$ con $\psi_0(x) = 1$. Es fácil ver que $\psi_t(x) = v(\phi_t(x))/v(x)$. Luego

$$u(t,x) = u_0(\phi_{-t}(x))v(\phi_{-t}(x))/v(x)$$

que coincide con (A.11).

A.3. Resolución en el marco de funciones a valor medida.

Esta sección esta dedicada a establecer una extensión de la fórmula de las caracteristicas para resolver una ecuación de conservación de masa en $\mathbb{R}^n$ de la forma

$$\begin{cases} \frac{\partial \rho}{\partial t} + div(\rho v) = 0 \\ \rho_0 = \mu \end{cases}$$

donde la condición inicial es una medida $\mu \in P(\mathbb{R}^n)$. Informalmente, vamos a ver que se obtiene la solución "llevando hacia adelante" la medida μ siguiendo el flujo del campo de velocidad v.

Introducimos primero la noción de pushforward de una medida por una función:

Definición A.3.1. *Sean $(X, \mathcal{B}, \mu)$ y $(Y, \mathcal{B}')$ dos espacios medibles y $T : X \to Y$ una función medible. Definimos una medida sobre $(Y, \mathcal{B}')$ llamada push-forward de μ a través de T, que notamos $T\#\mu$, por*

$$(T\#\mu)[A] = \mu[T^{-1}(A)], \qquad A \in \mathcal{B}',$$

donde $T^{-1}(A) = \{x \in X \ \ t.q. \ T(x) \in A\}$.

El siguiente resultado explica como se integra con respecto a $T\#\mu$:

Proposición A.3.2. *En el marco de la definición A.3.1 se tiene que*

$$\int_Y f \, d(T\#\mu) = \int_X f \circ T \, d\mu \tag{A.14}$$

para toda $f : Y \to [0, +\infty)$ medible.

Demostración. El esquema de prueba es estándar: verificamos (A.14) primero para funciones caracteristicas, luego para combinaciones lineales finitas de funciones caracteristicas, lo que permite finalmente probar (A.14) para funciones no-negativas por aproximación.

Si f es la función característica de un conjunto medible $A \in \mathcal{B}'$, o sea $f = 1_A$, entonces:

$$\int_Y f \, d(T\#\mu) = \int_Y 1_A \, d(T\#\mu) = (T\#\mu)(A) = \int_{T^{-1}(A)} d\mu = \int_X 1_A \circ T \, d\mu.$$

Por lo tanto (A.14) vale para funciones características. Los miembros izquierda y derecha de (A.14) siendo lineales con respecto a f, la fórmula (A.14) vale también cuando f es simple es decir cuando f es una combinación lineal finita de funciones características de conjuntos medibles.

Supongamos ahora que f es una función medible no-negativa. Sabemos que existe una sucesión creciente $(\varphi_n)_{n\in\mathbb{N}}$ de funciones simples que converge puntualmente a f. Por el teorema de convergencia monótona,

$$\int_Y f\, d(T\#\mu) = \int_Y \lim_{n\to\infty} \varphi_n\, d(T\#\mu) = \lim_{n\to\infty} \int_Y \varphi_n\, d(T\#\mu).$$

Como (A.14) vale para funciones simples, vale en particular para cada φ_n por lo que

$$\int_Y f\, d(T\#\mu) = \lim_{n\to\infty} \int_X \varphi_n \circ T\, d\mu.$$

Como $(\varphi_n \circ T)_{n\in\mathbb{N}}$ es una sucesión creciente de funciones medibles que converge puntualmente a $f \circ T$, podemos aplicar nuevamente el Teorema de Beppo-Levi para obtener

$$\int_Y f\, d(T\#\mu) = \int_X (\lim_{n\to\infty} \varphi_n \circ T)\, d\mu = \int_X f \circ T\, d\mu.$$

Concluimos entonces que (A.14) vale para funciones medibles no-negativas. □

Volvemos ahora a la ecuación de transporte (A.3). Precisemos primero el sentido en el que una función $\rho(t, x)$ es solución de (A.3).

Definición A.3.3. *Decimos que ρ es solución de (A.3) en $\mathbb{R} \times \mathbb{R}^n$ si*

1. $\rho : t \in \mathbb{R} \to \rho_t \in P(\mathbb{R}^n)$ *es continua para la topología débil en $P(\mathbb{R}^n)$ con $\rho_0 = \mu$,*
2. *para toda función test $\phi \in C_c^\infty(\mathbb{R}^n)$ la función $t \in \mathbb{R} \to \int_{\mathbb{R}^n} \phi\, d\rho_t$ es C^1 con derivada*

$$\frac{d}{dt} \int_{\mathbb{R}^n} \phi(x)\, d\rho_t(x) = \int_{\mathbb{R}^n} \nabla\phi(x) v(t, x)\, d\rho_t(x). \tag{A.15}$$

Observación A.3.4. *Observe que el 2do miembro de* (A.15) *corresponde a la formulación débil de $-div(\rho v)$ ya que si ρ_t y v fuesen funciones C^1 entonces tendríamos, integrando por partes, que para cualquier $\phi \in C_c^\infty(\mathbb{R}^n)$,*

$$-\int_{\mathbb{R}^n} div(\rho v)\phi(x)\, dx = -\sum_{i=1}^n \int_{\mathbb{R}^n} \frac{\partial(\rho v_i)}{\partial x_i}\phi(x)\, dx = \sum_{i=1}^n \int_{\mathbb{R}^n} \rho v_i \frac{\partial \phi}{\partial x_i}\, dx = \int_{\mathbb{R}^n} \rho(v\nabla\phi)\, dx.$$

Teorema A.3.5. *Supongamos que $v : \mathbb{R}^n \to \mathbb{R}^n$ es C^1 globalmente Lipschitz y notamos $T_t(x)$ su flujo. Dado $\mu \in P(\mathbb{R}^n)$, definimos $\rho : \mathbb{R}^n \to P(\mathbb{R}^n)$ por $\rho_t = T_t\#\mu$. Entonces ρ es la única solución en $\mathbb{R} \times \mathbb{R}^n$ de la ecuación lineal de transporte*

$$\begin{cases} \frac{\partial\rho}{\partial t} + div(\rho v) = 0 \\ \rho_0 = \mu \end{cases}$$

en el sentido de la definición A.3.3.

Demostración. Veamos primero que $\rho_t = T_t\#\mu$ es una solucion de (A.3) en el sentido de la definición A.3.3. Como $T_0 = Id$ tenemos $\rho_0 = Id\#\mu = \mu$. Verifiquemos ahora que ρ es continua de $\mathbb{R}$ en $P(\mathbb{R}^n)$ con la topología débil. Dado $t \in [0, T^*)$, debemos probar que $\rho_s \overset{s\to t}{\rightharpoonup} \rho_t$ o sea $\lim_{s\to t} \int_{\mathbb{R}^n} \phi\, d\rho_s = \int_{\mathbb{R}^n} \phi\, d\rho_t$ para toda $\phi \in C_b(\mathbb{R}^n)$. Por definición de ρ_s, ρ_t y del push-forward de una medida, eso es equivalente a probar que

$$\text{(A.16)} \qquad \lim_{s\to t} \int_{R^n} \phi(T_s(x)) - \phi(T_t(x))\, d\mu = 0$$

para toda $\phi \in C_b(\mathbb{R}^n)$. Es una consecuencia directa del teorema de convergencia dominada ya que (i) la función $s \to T_s(x)$ es continua en t por lo que $\phi(T_s(x)) - \phi(T_t(x)) \xrightarrow{s\to t} 0$, y (ii) se puede acotar el valor absoluto del integrando por $2\|\phi\|_\infty$ que es integrable con respecto a μ ya que μ es una medida de probabilidad.

Fijamos ahora una función-test $\phi \in C_c^\infty(\mathbb{R}^n)$ y verifiquemos que función $F(t) := \int_{\mathbb{R}^n} \phi\, d\rho_t$ es C^1 con derivada dada por (A.15). De hecho por definición de ρ_t, podemos reescribir F como

$$F(t) = \int_{\mathbb{R}^n} \phi(T_t(x))\, d\mu(x).$$

El integrando $\psi(t,x) := \phi(T_t(x))$ es una función C^1 en (t,x) (porque $T_t(x)$ lo es) con derivada en t dada por

$$\partial_t \psi(t,x) = \nabla\phi(T_t(x))\partial_t T_t(x) = \nabla\phi(T_t(x))v(T_t(x)).$$

Como ϕ tiene soporte compacto podemos aplicar el teorema de convergencia mayorada para obtener que F es C^1 con

$$F'(t) = \int_{\mathbb{R}^n} \partial_t\psi(t,x)\, d\mu(x) = \int_{\mathbb{R}^n} \nabla\phi(T_t(x))v(T_t(x))\, d\mu(x).$$

Recordando la definición del push-forward $\rho_t = T_t\sharp\mu$, obtenemos finalmente

$$F'(t) = \int_{\mathbb{R}^n} \nabla\phi(x)v(x)\, d\rho_t(x).$$

Veamos ahora que $\rho_t = T_t\#\mu$ es la única solución de (A.3). Supongamos que existen dos soluciones distintas ρ_1 y ρ_2 del problema (A.3). Entonces por linealidad $\widetilde{\rho} := \rho_1 - \rho_2$ es solucion de:

$$\begin{cases} \frac{\partial\rho}{\partial t} + \text{div } (\rho v) = 0 \\ \rho_0 = 0 \end{cases}$$

Luego para probar que $\rho_t = T_t\#\mu$ es la única solución de (A.3), es suficiente probar que si ρ es una solución de (A.3) entonces $\rho_t = 0$ para todo $t \in \mathbb{R}$. Para eso usaremos el método de dualidad. Dado $T > 0$ y ϕ_T una funcion arbitraria en $C_c^\infty(\mathbb{R}^n)$, admitimos por el momento que existe $\phi : [0,T] \times \mathbb{R}^n \to \mathbb{R}$ Lipschitz con soporte compacto que verifique

$$\begin{cases} \frac{\partial \phi}{\partial t}(t,x) = -v(x)\nabla\phi(t,x) \\ \phi_{t=T} = \phi_T. \end{cases} \tag{A.17}$$

El mismo razonamiento que en la 1ra parte de esta demostra-
-ción permite probar que la función $G(t) := \int_{R^n} \phi(t,x) d\rho_t(x)$ es C^1 con derivada

$$G'(t) = \frac{d}{dt}\int_{R^n} \phi(t, T_t(x))\, d\mu = \int_{\mathbb{R}^n} \partial_t\phi(t, T_t(x)) + \nabla\phi(t, T_t(x)) v(t, T_t(x))\, d\mu.$$

Usando la definición de ϕ, vemos que el integrando es cero. Entonces G es constante por lo que

$$\int_{\mathbb{R}^n} \phi_T d\rho_T = \int_{\mathbb{R}^n} \phi_0\, d\rho_0 = 0$$

ya que supusimos que $\rho_0 = 0$. Como ϕ_T era una función-test arbitraria, concluimos que $\rho_T = 0$ para todo $T > 0$.

Por último nos falta probar la existencia de ϕ. Podemos reescribir la ecuación en (A.17) como $\frac{d}{dt}\phi(t, T_t(x)) = 0$. Luego esta ϕ debe cumplir $\phi(t, T_t(x)) = \phi(T, T_T(x)) = \phi_T(T_T(x))$. Tiene por lo tanto sentido proponer

$$\phi_t := \phi_T \circ T_T \circ T_t^{-1}, \qquad \text{es decir} \qquad \phi(t,x) = \phi_T(T_T(T_t^{-1}(x))).$$

Es fácil ver que esta ϕ conviene. $\square$

Obviamente cuando u_0 es suave recuperamos bien (A.11): para $\phi \in C^\infty(\mathbb{R}^n)$ se tiene que

$$\begin{aligned} \int \phi(x) u(t,x) dx &= \int \phi(T_t(x)) u_0(x) dx = \int \phi(y) u_0(T_{-t}(y)) \frac{d}{dt} T_{-t}(x) dx \\ &= \int \phi(y) u_0(T_{-t}(y)) \frac{v(T_{-t}(x))}{v(x)} dx \end{aligned}$$

Concluimos que $u(t,x) = u_0(T_{-t}(y))\frac{v(T_{-t}(x))}{v(x)}$ como obtuvimos en la sección anterior.

Además de su elegancia, el teorema anterior permite resolver la ecuación de conservación cuando el dato inicial es una medida. En particular si u_0 es una masa de dirac en x_0 entonces $u(t) = T_t\sharp\delta_{x_0} = \delta_{T_t(x_0)}$. Se puede apreciar sobre este ejemplo que el campo v transporta la masa.

Apéndice B

Algunas herramientas de probabilidad

Dado un espacio métrico (X, d), notaremos $\mathcal{B}(X)$ a la $\sigma-$álgebra de Borel en X.

Definición B.0.1. *Sea (X, d) un espacio métrico. Se dice que $\mu : \mathcal{B}(X) \to [0, \infty]$ es una medida de probabilidad de Borel en X si:*

1. $\mu(X) = 1$

2. $\mu(\emptyset) = 0$

3. *Si $A_1, A_2, , \ldots \in \mathcal{B}(X)$, disjuntos dos a dos entonces* $\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu\left(A_i\right)$

Notamos $P(X)$ el espacio de las medidas de probabilidad borelianas sobre (X, d). Observamos que si μ es una medida de probabilidad de Borel en X entonces, μ es regular y tiene soporte σ-compacto.

Dado un espacio métrico (X, d) notaremos

$$C_b(X) := \{f : X \to \mathbb{R} : f escontinuayacotada\}$$

Si $f \in C_b(X)$ entonces, $f \in L^1(\mu)$ $\forall \mu$ medida de probabilidad de Borel.

Definición B.0.2. *Decimos que una sucesión $(\mu_k)_k \subset P(X)$ converge a $\mu \in P(X)$, y notamos $\mu_k \rightharpoonup \mu$. si $\int_X \phi \, d\mu_k \to \int_X \phi \, d\mu$ para toda $\phi \in C_b(X)$.*

La convergencia débil define una topología Hausdorff separable en $P(X)$, llamada *topología débil*. En particular $\mu = \nu$ si y solo si $\forall \varphi \in C_b(X), \int \varphi d\mu = \int \varphi d\nu$.

B.1. Resultados de compacidad

B.1.1. El teorema de Prokhorov

Definición B.1.1. *Sea (X, d) un espacio métrico y sea Γ una familia de medidas de probabilidad en $(X, \mathcal{B}(X))$. Decimos que Γ es relativamente compacto si para toda sucesión de Γ existe una subsucesión débilmente convergente.*

Definición B.1.2. *Sea (X, d) un espacio métrico y sea Γ una familia de medidas de probabilidad en $(X, \mathcal{B}(X))$. Decimos que Γ es tight si $\forall \epsilon > 0$, $\exists K_\epsilon$ compacto, tal que $\mu(K_\epsilon) \geq 1 - \epsilon \ \forall \mu \in \Gamma$*

Teorema B.1.3. *(Teorema de Prokhorov) Sea (X, d) un espacio métrico y sea Γ una familia de medidas de probabilidad en $(X, \mathcal{B}(X))$. Son equivalentes:*

1. *Γ es precompacto.*
2. *Γ es tight.*

Demostración. Ver el capitulo 5 de [[3]]. □

B.1.2. El teorema de Arzela-Ascoli

Teorema B.1.4. *Sea X un espacio métrico compacto, Y un espacio métrico completo y sea H un subconjunto de $C(X, Y)$. Entonces, H es relativamente compacto si y solo si H es equicontinuo y $\forall x \in X$, el conjunto $H(x) = \{f(x) \ : \ f \in H\}$ es relativamente compacto en Y.*

Demostración. Ver □

Observación B.1.5. *Si tomamos $Y = P(I)$ y $X = [0, T], T > 0$. Y es compacto, por lo tanto completo. Sea $H = \{g \in C([0, T], P(I))\}$. Entonces H es relativamente compacto si y solo si H es equicontinua y $\forall \ \tau \in [0, T]$ toda sucesion $g_n(t) \in H$ (t fijo) tiene una subsucesión convergente en $P(I)$.*

B.2. Distancia de Monge-Kantorovich

El propósito de esta sección es estudiar una familia de distancias W_p, llamadas distancias de Monge-Kantorovich (M-K), sobre el espacio de las medidas de probabilidad de un espacio métrico (X, d). Estas distancias tienen la propiedad de metrizar la convergencia débil y de tener una interpretación intuitiva clara en término de transporte óptimo de masa. Además la distancia W_1 tiene una caracterización dual muy útil (teorema de Kantorovich-Rubinstein) que se obtiene como consecuencia de una teorema de dualidad mas general (teorema de Kantorovich).

Después de definir estas distancias, probaremos que son efectivamente distancias genuinas que metrizan la convergencia débil. Cabe destacar que las pruebas resultas ser no-triviales. Luego enunciaremos el teorema de dualidad de Kantorovich que aplicaremos para obtener el teorema de Kantorovich-Rubinstein. Esta sección esta basada en el libro de C. Villani [12].

En toda esta capitulo, (X, d) es un espacio métrico que suponemos Polaco, es decir completo y separable.

B.2.1. Definición de las distancias de M-K.

Definición B.2.1. *Dadas $\mu, \nu \in P(X)$, decimos que una medida de probabilidad π sobre el espacio producto $X \times X$ tiene por marginal a μ y ν si para todo $A \subset X$ Borel,*

$$\pi(A \times X) = \mu(A) \qquad y \qquad \pi(X \times A) = \nu(A),$$

o si de manera equivalente,

$$\int_{X\times X} \phi(x) + \psi(y)\, d\pi(x, y) = \int_X \phi\, d\mu + \int_X \psi\, d\nu \tag{B.1}$$

para toda $\phi, \psi \in C_b(X)$ (se puede tomar también $\phi \in L^1(\mu)$ y $\psi \in L^1(\mu)$ - ver [12][1.1.1]) Una tercera manera equivalente consiste en decir que μ (resp. ν) es la proyeccion de π sobre el 1er (resp. 2ndo) factor del producto $X \times X$ es decir $pr_1\sharp\pi = \mu$ y $pr_2\sharp\pi = \nu$ donde $pr_1 : (x, y) \in X \times X \to x \in X$ y $pr_2 : (x, y) \in X \times X \to y \in X$ son las proyecciones.

Notamos $\Pi(\mu, \nu)$ el conjunto de las $\pi \in P(X \times X)$ que tienen por marginal a μ y ν.

Ejemplo B.2.2. *Dado dos puntos $x_0, y_0 \in X$, $\Pi(\delta_{x_0}, \delta_{y_0}) = \{\delta_{x_0} \otimes \delta_{y_0}\}$.*

Por definición, $\delta_{x_0} \otimes \delta_{y_0}(A \times B) = \delta_{x_0}(A)\delta_{y_0}(B)$ por lo tanto $\delta_{x_0} \otimes \delta_{y_0} \in \Pi(\delta_{x_0}, \delta_{y_0})$.

Veamos ahora que si $\pi \in \Pi(\delta_{x_0}, \delta_{y_0})$ entonces $\pi = \delta_{x_0} \otimes \delta_{y_0}$: sean $A, B \in X$, queremos ver que $\pi(A \times B) = \delta_{x_0}(A)\delta_{y_0}(B)$. Supongamos que $y_0 \notin B$,

$$\pi(A \times B) \leq \pi(X \times B) = \delta_{y_0}(B) = 0$$

Por lo tanto, $\pi(A \times B) = \delta_{x_0} \otimes \delta_{y_0}(A \times B)$.

Veamos que ocurre si $y_0 \in B$:

$$\delta_{x_0}(A) = \pi(A \times X) \geq \pi(A \times B)$$

Análogamente,

$$\delta_{x_0}(X - A) = \pi(X - A \times X) \geq \pi(X - A \times B)$$

Entonces,

$$\delta_{x_0}(A) = 1 - \delta_{x_0}(X - A) \leq$$
$$1 - \pi(X - A \times B) = \pi(X - A \times X - B) + \pi(A \times B) \leq$$
$$\pi(X \times X - B) + \pi(A \times B) = \pi(A \times B)$$

Por los cálculos anteriores y recordando que $y_0 \in B$*, tenemos*

$$\delta_{x_0}(A)\delta_{y_0}(B) \leq \pi(A \times B) \leq \delta_{x_0}(A)\delta_{y_0}(B)$$

Concluimos,

$$\pi(A \times B) = \delta_{x_0}(A)\delta_{y_0}(B)$$

Observación B.2.3. $\Pi(\mu, \nu)$ *es no vacio dado que siempre contiene a la medida* $\mu \otimes \nu$.

Observación B.2.4. $\Pi(\mu, \nu)$ *es un conjunto convexo: si* $\pi, \tilde{\pi} \in \Pi(\mu, \nu)$ *y* $t \in [0, 1]$ *entonces* $t\pi + (1-t)\tilde{\pi} \in \Pi(\mu, \nu)$ *ya que para todo boreliano* $A \subset X$,

$$(t\pi + (1-t)\tilde{\pi})(A \times X) = t\pi(A \times X) + (1-t)\tilde{\pi}(A \times X) = t\mu(A) + (1-t)\mu(A) = \mu(A).$$

Analogamente, $(t\pi + (1-t)\tilde{\pi})(X \times A) = \nu(A)$.

Definición B.2.5. *Dado* $p \in [1, \infty)$*, notamos* $P_p(X)$ *el subconjunto de* $P(X)$ *de las medidas de probabilidad cuyo p-ésimo momento es finito:*

$$P_p(X) = \left\{ \mu \in P(X) \ tq \ \int_X d(x_0, x)^p d\mu(x) < +\infty \ \textit{para algún} \ x_0 \in X \right\}.$$

Observación B.2.6. *La elección del punto* x_0 *es irrelevante: si* $\int_X d(x_0, x)^p d\mu(x) < +\infty$ *entonces* $\int_X d(y_0, x)^p d\mu(x) < +\infty$ *para cualquier* $y_0 \in X$ *ya que*

$$d(y_0, x)^p \leq (d(y_0, x_0) + d(x_0, x))^p \leq 2^{p-1}(d(y_0, x_0)^p + d(x_0, x)^p).$$

Observación B.2.7. *Si* d *es acotada entonces* $P_p(X) = P(X)$.

Teorema B.2.8. ***(Distancias Monge-Kantorovich)*** *Dado* $p \geq 1$*, la función* $W_p : P_p(X) \times P_p(X) \to \mathbb{R}_+$ *definida por*

$$W_p(\mu, \nu) = \left\{ \inf_{\pi \in \Pi(\mu, \nu)} \int_{XxX} d(x, y)^p d\pi(x, y) \right\}^{1/p}$$

define una distancia en $P_p(X)$.

Ejemplo B.2.9. *Usando el ejemplo B.2.2 vemos que* $W_p(\delta_{x_0}, \delta_{y_0}) = d(x_0, y_0)$ *ya que*

$$W_p(\delta_{x_0}, \delta_{y_0})^p = \int_{XxX} d(x, y)^p d(\delta_{x_0} \otimes \delta_{y_0})(x, y) = d(x_0, y_0)^p.$$

Observación B.2.10. *Si d es acotada, entonces como corolario del teorema anterior, W_p define una distancia en $P(X)$ en vista de la observación B.2.7. En el caso que d no sea acotada, muchas veces es conveniente reemplazarla por la métrica acotada $\tilde{d} = \text{ínf}(d, 1)$, la cual induce la misma topología que d.*

Observación B.2.11. *Se puede interpretar $W_p(\mu, \nu)^p$ desde un punto de vista de transporte óptimo de masa entre μ y ν en el que μ representa una pila de material que se debe llevar a un hueco modelado por ν, el costo que se paga para transportar una unidad de masa de x a y viene dado por $d(x, y)^p$ y $\pi \in \Pi(\mu, \nu)$ es un plan de transferencia de masa. $W_p(\mu, \nu)^p$ es entonces el costo mínimo que se puede pagar para transportar el material al hueco.*

Para probar que W_p es una distancia, necesitaremos los dos resultados siguientes. El primero afirma que siempre existe una π optima para $W_p(\mu, \nu)$.

Teorema B.2.12. *Para todas $\mu, \nu \in P(X)$, el ínfimo en la definición de $W_p(\mu, \nu)$ se alcanza: existe $\pi \in \Pi(\mu, \nu)$ tal que $W_p(\mu, \nu)^p = \int_{X \times X} d(x, y)^p \, d\pi(x, y)$.*

La prueba usa el teorema de Prohorov que da una condición necesaria y suficiente para que una parte de $P(X)$ sea precompacto para la topología débil (la convergencia se testea contra funciónes continuas acotadas).

Teorema B.2.13. *Sea (X, d) un espacio métrico polaco. Un subconjunto $A \subset P(X)$ es precompacto en $P(X)$ para la topología débil si y solo si es tight en el sentido que para todo $\varepsilon > 0$ existe un compacto $K \subset X$ tal que $\mu(X \backslash K) < \varepsilon$ para todo $\mu \in A$.*

Podemos ahora demostrar el teorema B.2.12.

Demostración. Probemos primero que $\Pi(\mu, \nu)$ es compacto para la topología débil. Verifiquemos que es un conjunto cerrado de $P(X \times X)$. Consideramos una sucesión $(\pi_n)_n \subset \Pi(\mu, \nu)$ que converge débil a $\pi \in P(X \times X)$. Sabemos que para todo n,

$$\int_X \phi(x) + \psi(y) \, d\pi_n(x, y) = \int_X \phi \, d\mu + \int_X \psi \, d\nu$$

para cualquier $\phi, \psi \in C_b(X)$. Como la función $(x, y) \to \phi(x) + \psi(y)$ pertenece a $C_b(X \times X)$, podemos pasar al limite para obtener

$$\int_X \phi(x) + \psi(y) \, d\pi(x, y) = \int_X \phi \, d\mu + \int_X \psi \, d\nu$$

para cualquier $\phi, \psi \in C_b(X)$. Eso prueba que μ y ν son las marginales de π por lo que $\pi \in \Pi(\mu, \nu)$. Veamos ahora que $\Pi(\mu, \nu)$ es precompacto en $P(X \times X)$. Por el teorema de Prohorov, es equivalente a mostrar que es tight. Fijamos un $\varepsilon > 0$. Como los conjuntos

$\{\mu\}$ y $\{\nu\}$ son precompactos, son tight, por lo que existe un compacto $K \subset X$ tal que $\mu(X\backslash K) < \varepsilon$ y $\nu(X\backslash K) < \varepsilon$. Luego para cualquier $\pi \in \Pi(\mu, \nu)$ se tiene que

$$\begin{aligned} \pi((X \times X)\backslash(K \times K)) &\leq \pi((X\backslash K) \times X) + \pi(X \times (X\backslash K)) \\ &= \mu(X\backslash K) + \nu(X\backslash K) \\ &< 2\varepsilon. \end{aligned}$$

Eso prueba que $\Pi(\mu, \nu)$ es tight.

Verifiquemos ahora que la función $F : \pi \in P(X \times X) \to \int_{X\times X} c\, d\pi \in \mathbb{R}_+$, donde $c(x, y) := d(x, y)^p$, es semi-continua inferiormente para la topología débil. Consideramos entonces una sucesión $(\pi_n)_n \subset P(X \times X)$ cualquiera que converge a un $\pi \in P(X \times X)$. Debemos probar que $\liminf \int_{X\times X} c\, d\pi_n \geq \int_{X\times X} c\, d\pi$. Si c fuese acotada, F seria continua y listo. Como c no es acotada, la truncamos introduciendo $c_k := \min\{c, k\}$, $k \in \mathbb{N}$. Note que la sucesión $(c_k)_k$ es no-decreciente y converge puntualmente a c. El teorema de convergencia monótona da

$$\int_{X\times X} c\, d\pi = \lim_k \int_{X\times X} c_k\, d\pi.$$

Como $c_k \in C_b(X \times X)$ para todo k y $\pi_n \to \pi$,

$$\int_{X\times X} c\, d\pi = \lim_k \lim_n \int_{X\times X} c_k\, d\pi_n.$$

Finalmente, $c_k \leq c$ para todo k por lo que

$$\int_{X\times X} c\, d\pi \leq \liminf_n \int_{X\times X} c\, d\pi_n.$$

Podemos ahora probar fácilmente que existe una $\pi \in \Pi(\mu, \nu)$ optima. Tomamos una sucesión minimizante $(\pi_n)_n \subset \Pi(\mu, \nu)$. Luego $F(\pi_n) \to W_p(\mu, \nu)^p$. Como $\Pi(\mu, \nu)$ es compacto, podemos extraer una subsucesión, que seguimos notando (π_n), que converge a un $\pi \in \Pi(\mu, \nu)$. Note que $F(\pi) \geq W_p(\mu, \nu)^p$. La semi-continuidad inferior de F da entonces

$$W_p(\mu, \nu)^p = \lim_n F(\pi_n) \geq F(\pi) \geq W_p(\mu, \nu)^p.$$

Deducimos que π es óptima.

$\square$

Observación B.2.14. *Bajo ciertas hipótesis, existe una única π óptima y esta proviene de una función $T : X \to X$ tal que $T\sharp\mu = \nu$ en el sentido que $\pi = (Id \times T)\sharp\mu$. Esto significa simplemente que, en la interpretación de la observación B.2.11, conviene transportar la unidad de masa en x al lugar $y = T(x)$. Estas hipotesis se verifican si por ejemplo $X = \mathbb{R}^n$ con la distancia euclideana, $p > 1$, y μ es absolutamente continua con respecto a la medida de Lebesgue (ver [12][Thm 2.44]).*

El segundo resultado, que admitimos, será útil al momento de probar que W_p verifica la desigualdad triangular.

Lema B.2.15. ***(Lema de encolado)*** *Sean μ_1, μ_2, μ_3 tres medidas de probabilidad soportadas en los espacios Polacos X_1, X_2, X_3 respectivamente. Sea $\pi_{12} \in \Pi(\mu_1, \mu_2)$, $\pi_{23} \in \Pi(\mu_2, \mu_3)$. Entonces, existe una medida de probabilidad $\pi \in P(X_1 \times X_2 \times X_3)$ con marginales π_{12} en $X_1 \times X_2$ y π_{23} en $X_2 \times X_3$.*

Podemos ahora probar el teorema B.2.8.

Demostración. Verifiquemos primero que $W_p(\mu, \nu)$ esta bien definido si $\mu, \nu \in P_p(X)$. Tomemos $x_0 \in X$ tales que

$$\int_X d(x_0, x)^p d\mu(x) < +\infty \qquad \text{y} \qquad \int_X d(x_0, x)^p d\nu(x) < +\infty.$$

Dada una $\pi \in \Pi(\mu, \nu)$ cualquiera, tenemos usando la desigualdad triangular en $L^p(X \times X, \pi)$ que

$$\begin{aligned} W_p(\mu, \nu) &= \left(\int d(x, y)^p d\pi(x, y) \right)^{1/p} \\ &\leq \left(\int \left[d(x, x_0) + d(x_0, y) \right]^p d\pi(x, y) \right)^{1/p} \\ &\leq \left(\int d(x, x_0)^p d\pi(x, y) \right)^{1/p} + \left(\int d(x_0, y)^p d\pi(x, y) \right)^{1/p}. \end{aligned}$$

Por hipótesis, las funciónes $\phi(x) := d(x, x_0)^p$ y $\psi(y) := d(y, x_0)^p$ son integrables con respecto a μ y ν respectivamente. Recordando que π tiene por marginal μ y ν, obtenemos entonces que

$$W_p(\mu, \nu) \leq \left(\int d(x, x_0)^p d\mu(x) \right)^{1/p} + \left(\int d(x_0, y)^p d\nu(y) \right)^{1/p} < +\infty.$$

Verifiquemos ahora que W_p es una distancia sobre $P_p(X)$. Claramente W_p es simétrica y no negativa. Veamos que $W_p(\mu, \mu) = 0$. Consideramos el homeomorfismo $f : X \to \Delta$, donde $\Delta = \{(x, x) / \ x \in X\}$, definido por $f(x) = (x, x)$. Sea $\nu = f \# \mu \in P(\Delta)$ y $\pi \in P(X \times X)$ definida por $\pi(A) := \nu(A \times \Delta)$, $A \subseteq X \times X$ Boreliano. Luego para todo $A_x \subset X$ Borel, se tiene que

$$\pi[A_x \times X] = \nu[(A_x \times X) \cap \Delta] = \mu[f^{-1}((A_x \times X) \cap \Delta)] = \mu[A_x].$$

Analogamente, $\pi[X \times A_y] = \mu[A_y]$ para todo $A_y \subset X$ Borel. Entonces $\pi \in \Pi(\mu, \mu)$. Usamos esta π para estimar $W_p(\mu, \mu)$. Observando que π esta concentrada en la diagonal Δ, vemos que $d(x, y) = 0$ π-ctp por lo que

$$W_p(\mu, \mu)^p \leq \int_{X \times X} d(x, y)^p d\pi(x, y) = 0.$$

Probamos ahora que si $W_p(\mu, \nu) = 0$ entonces $\mu = \nu$, o sea $\int_X \phi \, d\mu = \int_X \phi \, d\nu$ para toda $\phi \in C_b(X)$. Por la prop. B.2.12 sabemos que existe una $\pi \in \Pi(\mu, \nu)$ óptima para $W_p(\mu, \nu)$:

$$\int_{X \times X} d(x, y)^p \, d\pi(x, y) = W_p(\mu, \nu)^p = 0.$$

Entonces $d(x, y) = 0$ π-ctp por lo que π esta soportada en la diagonal Δ de $X \times X$. Para toda $\phi \in C_b(X)$ obtenemos por lo tanto que

$$\int_X \phi(x) \, d\mu(x) = \int_{X \times X} \phi(x) \, d\pi(x, y) = \int_{X \times X} \phi(y) \, d\pi(x, y) = \int_X \phi(y) \, d\nu(y).$$

Ahi usamos que π tiene por marginal μ y nu en la 1era y 3era igualdad, y que esta soportada en Δ en la 2nda igualdad.

Veamos para terminar que vale la desigualdad triangular. Sean μ_1, μ_2, μ_3 en $P_p(X)$ y X_1, X_2, X_3 sus respectivos soportes. Sean $\pi_{12} \in \Pi(\mu_1, \mu_2)$ y $\pi_{23} \in \Pi(\mu_2, \mu_3)$ óptimos. Sea $\pi \in P(X_1 \times X_2 \times X_3)$ como el Lema de encolado, y π_{13} la medida marginal de π en $X_1 \times X_3$. Verifiquemos que $\pi_{13} \in \Pi(\mu_1, \mu_3)$. Para toda $\phi \in C_b(X_1)$,

$$\begin{aligned}
\int_{X_1 \times X_3} \phi(x_1) \, d\pi_{13}(x_1, x_3) &= \int_{X_1 \times X_2 \times X_3} \phi(x_1) \, d\pi(x_1, x_2, x_3) \\
&= \int_{X_1 \times X_2} \phi(x_1) \, d\pi_{12}(x_1, x_2) \\
&= \int_{X_1} \phi(x_1) \, d\mu_1(x_1)
\end{aligned}$$

Luego μ_1 es la marginal de π_{13} sobre X_1. Analogamente, μ_3 es la marginal de π_{13} sobre X_3. Podemos entonces usar π_{13} para estimar $W_p(\mu_1, \mu_3)$:

$$\begin{aligned}
W_p(\mu_1, \mu_3) &\leq \left(\int_{X_1 \times X_3} d(x_1, x_3)^p \; d\pi_{13}(x_1, x_3) \right)^{1/p} \\
&= \left(\int_{X_1 \times X_2 \times X_3} d(x_1, x_3)^p \; d\pi(x_1, x_2, x_3) \right)^{1/p} \\
&\leq \left(\int_{X_1 \times X_2 \times X_3} [d(x_1, x_2) + d(x_2, x_3)]^p \; d\pi(x_1, x_2, x_3) \right)^{1/p}.
\end{aligned}$$

Aplicando la desigualdad triangular en $L^p(X_1 \times X_2 \times X_3, \pi)$, obtenemos:

$$\begin{aligned}
& W_p(\mu_1, \mu_3) \\
\leq \; & \left(\int_{X_1 \times X_2 \times X_3} d(x_1, x_2)^p \; d\pi(x_1, x_2, x_3) \right)^{1/p} + \left(\int_{X_1 \times X_2 \times X_3} d(x_2, x_3)^p \; d\pi(x_1, x_2, x_3) \right)^{1/p} \\
\leq \; & \left(\int_{X_1 \times X_2} d(x_1, x_2)^p \; d\pi_{12}(x_1, x_2) \right)^{1/p} + \left(\int_{X_2 \times X_3} d(x_2, x_3)^p \; d\pi_{23}(x_2, x_3) \right)^{1/p} \\
= \; & W_p(\mu_1, \mu_2) + W_p(\mu_2, \mu_3).
\end{aligned}$$

□

Proposición B.2.16. *Si $1 \leq p \leq q < \infty$ entonces $W_p \leq W_q$ en $P_q(X)$. Si suponemos ademas que d es acotada entonces*

$$W_q \leq W_p^{\frac{p}{q}} \, diam(X)^{1-\frac{p}{q}}$$

donde $diam(X) = \sup_{x \neq y} d(x,y)$. Entonces si d es acotada las distancias W_p, $p \geq 1$, son equivalentes sobre $P(X)$.

Demostración. Observe que $P_q(X) \subset P_p(X)$. Sea $\mu, \nu \in P_q(X)$ y $\pi \in \Pi(\mu,\nu)$. Aplicando la desigualdad de Holder recordando que $\pi(X \times X) = 1$, tenemos

$$W_p(\mu,\nu)^p \leq \int_{X\times X} d(x,y)^p \, d\pi(x,y) \leq \Big(\int_{X\times X} d(x,y)^q \, d\pi(x,y) \Big)^{\frac{p}{q}}.$$

Tomando el ínfimo sobre todas las $\pi \in \Pi(\mu,\nu)$ en el miembro derecho obtenemos $W_p(\mu,\nu) \leq W_q(\mu,\nu)$.

Supongamos ahora que d es acotada. Entonce para todo $x \neq y \in X$,

$$d(x,y)^q = d(x,y)^p d(x,y)^{q-p} \leq d(x,y)^p \mathrm{diam}(X)^{q-p}.$$

Integrando con respecto a un $\pi \in \Pi(\mu,\nu)$ obtenemos

$$W_q(\mu,\nu)^q \leq \int_{X\times X} d(x,y)^q \, d\pi(x,y) \leq \int_{X\times X} d(x,y)^p \, d\pi(x,y) \mathrm{diam}(X)^{q-p}.$$

Tomando el ínfimo sobre todas las $\pi \in \Pi(\mu,\nu)$ en el miembro derecho obtenemos $W_q(\mu,\nu)^q \leq W_p(\mu,\nu)^p \mathrm{diam}(X)^{q-p}$. □

B.2.2. Expresión dual de la distancia W_1.

Observe que en la definicion de $W_p(\mu,\nu)^p$, estamos minimizando el funciónal $I(\pi) := \int_{X\times X} d(x,y)^p \, d\pi(x,y)$ que es lineal, y por lo tanto convexo, sobre el conjunto convexo $\Pi(\mu,\nu)$ (ver obs.B.2.4). Por le tanto es de esperar que se puede reescribir el problema de minimización definiendo $W_p(\mu,\nu)^p$ como un problema de maximización. Este problema fue resuelto por Kantorovich en un marco mas general que el considerado aca.

La motivacion proviene de la interpretación en termino de transporte optimo de masa evocado en la observacion B.2.11. En este contexto tiene sentido considerar que μ y ν son medidas de probabilidad sobre dos espacios X e Y a priori distintos y que el costo de transporte de una unidad de masa desde el punto $x \in X$ hasta el punto $y \in Y$ es una función $c : X \times Y \to [0, +\infty]$. Como antes una medida $\pi \in P(X \times Y)$ con marginal μ y ν representa un plan de transporte de masa entre μ y ν de costo total $\int_{X\times Y} c(c,y)\,d\pi(x,y)$. Llegamos de esta forma a estudiar el problema de encontrar el costo mínimo de transporte

$$\inf_{\pi\in\Pi(\mu,\nu)} \int_{X\times Y} c(x,y)\,d\pi(x,y). \tag{B.2}$$

Mencionamos que bajo hipótesis bastante leves sobre c, se puede probar que el ínf se alcanza (ver citeV2[Chap.4]). La demostración es casi idéntica a la demostración del teorema B.2.12.

Podemos tener una idea intuitiva del problema dual asociado reescribiendo (B.2) como

$$\inf_{\pi\in M^+(\mu,\nu)} \int_{X\times Y} c(x,y)\,d\pi(x,y) + \chi_{\Pi(\mu,\nu)}(\pi) \tag{B.3}$$

donde $M^+(\mu,\nu)$ es el conjunto de medidas borelianas no-negativas sobre $X\times Y$ y $\chi_{\Pi(\mu,\nu)}(\pi)=0$ si $\pi\in\Pi(\mu,\nu)$ y $+\infty$ sino. Recuerde que $\pi\in\Pi(\mu,\nu)$ si se cumple (B.1) para todo ϕ,ψ. Es facil ver entonces que

$$\chi_{\Pi(\mu,\nu)}(\pi) = \sup_{\phi\in C_b(X),\psi\in C_b(Y)} \int_X \phi\,d\mu + \int_Y \psi\,d\nu - \int_{X\times Y} \phi(x)+\psi(y)\,d\pi(x,y).$$

Luego podemos reescribir (B.3) como

$$\inf_{\pi\in\Pi(\mu,\nu)} \int_{X\times Y} c(x,y)\,d\pi(x,y) + \sup \int_X \phi\,d\mu + \int_Y \psi\,d\nu.$$

Intercambiando el ínf y el sup, lo que necesitaría una justificación, obtenemos

$$\sup_{\phi\in C_b(X),\psi\in C_b(Y)} \int_X \phi\,d\mu + \int_Y \psi\,d\nu + \inf_{\pi\in M^+(\mu,\nu)} \int_{X\times Y} c(x,y) - (\phi(x)+\psi(y))\,d\pi(x,y)$$

Es fácil ver que el ínf es $-\chi_{\{\phi+\psi\le c\}}$. Luego

$$\inf_{\pi\in\Pi(\mu,\nu)} \int_{X\times Y} c(x,y)\,d\pi(x,y) = \sup_{\phi+\psi\le c} \int_X \phi\,d\mu + \int_Y \psi\,d\nu.$$

Claramente el razonamiento que hicimos no es riguroso ya que no en general no se puede intercambiar sup e ínf. Pero se puede justificar y llegar al siguiente teorema cuya demostración se puede encontrar en [12]:

Teorema B.2.17. *Dados dos espacios polacos X e Y, dos medidas de probabilidad $\mu\in P(X)$, $\nu\in P(Y)$ y una función semi-continua inferiormente $c: X\times Y\to[0,+\infty]$, se tiene que*

$$\inf_{\pi\in\Pi(\mu,\nu)} \int_{X\times Y} c(x,y)\,d\pi(x,y) = \sup \int_X \phi\,d\mu + \int_Y \psi\,d\nu \tag{B.4}$$

donde y el sup *se toma sobre todos los pares $(\phi,\psi)\in L^1(X,\mu)\times L^1(Y,\nu)$ tales que $\phi(x)+\psi(y)\le c(x,y)$ para μ-casi todo $x\in X$ y ν-casi todo $y\in Y$.*

Sea (ϕ,ψ) un par de funciones admisibles. De $\phi(x) \leq c(x,y) - \psi(y)$ para todo x, y, obtenemos que

$$\phi(x) \leq \psi^c(x) := \inf_{y\in Y} c(x,y) - \psi(y)$$

y luego $J(\phi,\psi) \leq J(\psi^c,\psi)$ donde $J(\phi,\psi) = \int_X \phi\, d\mu + \int_Y \psi\, d\nu$. Note que la restricción $\psi^c(x) + \psi(y) \leq c(x,y)$ se satisface automáticamente por definición misma de ψ^c. Como se busca maximizar J, conviene entonces restringirse a pares (ψ^c,ψ), $\psi \in L^1(\nu)$. Repitiendo esta idea llegamos a $J(\psi^c,\psi) \leq J(\psi^c,\psi^{cc})$ donde $\psi^{cc}(y) = \inf_x c(x,y) - \psi^c(x)$. Podriamos seguir obteniendo $J(\psi^c,\psi^{cc}) \leq J(\psi^{ccc},\psi^{cc})$ con $\psi^{ccc}(x) = \inf_y c(x,y) - \psi^{cc}(y)$ pero ocurre que $\psi^{ccc} = \psi^c$:

Lema B.2.18. *$\psi^{ccc} = \psi^c$.*

Demostración. Por construcción, $\psi^c(x) \leq \psi^{ccc}(x) \forall x \ \in X$.

Sea $x \in X$, veamos que $\psi^{ccc}(x) \leq \psi^c(x)$:

$$\begin{aligned}\psi^{ccc}(x) = \inf_{y\in Y} c(x,y) - \psi^{cc}(y) &= \inf_{y\in Y}\inf_{z\in X} c(x,y) - c(z,y) + \psi^c(z)\\ &\leq \psi^c(x)\end{aligned}$$

□

Finalmente se puede probar (ver [?][Thm 5.10], [12][Ejercicio 2.36]) que en el teorema de dualidad de Kantorovich podemos restringirnos a pares (ψ^c,ψ) o (ψ^c,ψ^{cc}):

$$\inf_{\pi\in\Pi(\mu,\nu)} \int_{X\times Y} c(x,y)\, d\pi(x,y) = \sup_{\psi\in L^1(\mu)} \int_X \psi^c\, d\mu + \int_Y \psi\, d\nu = \sup_{\psi\in L^1(\nu)} \int_X \psi^c\, d\mu + \int_Y \psi^{cc}\, d\nu.$$

Si en el último sup, nos restringimos a las funciones ψ tales que $\psi^{cc} = \psi$, obtenemos, recocrdando que $\psi^{ccc} = \psi^c$ por el lema anterior, que

$$\inf_{\pi\in\Pi(\mu,\nu)} \int_{X\times Y} c(x,y)\, d\pi(x,y) = \sup_{\psi\in L^1(\mu),\phi^{cc}=\phi} \int_X \psi^c\, d\mu + \int_Y \psi\, d\nu. \tag{B.5}$$

Supongamos a partir de ahora que $c(x,y) = d(x,y)$ con $X = Y$. Note que el miembro izquierdo de (B.5) es ahora la distancia $W_1(\mu,\nu)$. El siguiente lema identifica las funciónes ϕ tales que $\phi = \phi^{cc}$ que aparecen en el miembro derecho:

Lema B.2.19. *Una función $\phi : X \to \mathbb{R}$ verifica que $\phi = \phi^{cc}$ si y solo si es 1-Lipschitz en el sentido que $|\phi(x) - \phi(y)| \leq d(x,y)$ para todo $x, y \in X$, $x \neq y$. En este caso, $\phi^c = -\phi$.*

Obtenemos entonces:

Teorema B.2.20 (Teorema de Kantorovich-Rubinstein)**.** *Sea (X,d) un espacio métrico polaco. Entonces para toda $\mu,\nu \in P_1(X)$,*

$$W_1(\mu,\nu) = \sup\left\{\int_X \phi\, d(\mu-\nu) : \ \phi \in L^1(|\mu-\nu|),\ \phi \text{ 1-Lipschitz}\right\}.$$

B.2.3. Propiedades topologicas de las distancias de M-K.

En esta sección consideraremos algunas propiedades topológicas de la distancia W_p. El primer resultado afirma en particular que cualquier distancia W_p, $p \geq 1$, metriza la topología débil.

Teorema B.2.21. *Sea $p \geq 1$, $(\mu_k)_{k\in\mathbb{N}}$ una sucesión de medidas de probabilidad en $P_p(X)$, y $\mu \in P_p(X)$. Entonces son equivalentes:*

(i) $\lim_{k\to\infty} W_p(\mu_k, \mu) = 0$.

(ii) *$\mu_k \rightharpoonup \mu$, y $(\mu_k)_{k\in\mathbb{N}}$ satisface la siguiente condición: para algún $x_0 \in X$,*

$$\lim_{R\to\infty} \limsup_{k\to\infty} \int_{d(x_0,x)\geq R} d(x_0,x)^p d\mu_k(x) = 0$$

(iii) *$\mu_k \rightharpoonup \mu$, y hay convergencia del momento de orden p: para algún $x_0 \in X$,*

$$\lim_{k\to\infty} \int d(x_0,x)^p d\mu_k(x) = \int d(x_0,x)^p d\mu(x)$$

.

(iv) *Sea $\varphi \in C(X)$. Si φ satisface $|\varphi(x)| \leq C[1 + d(x_0,x)^p]$ para algún $x_0 \in X, C \in \mathbb{R}$, entonces:*

$$\lim_{k\to\infty} \int \varphi d\mu_k(x) = \int \varphi d\mu(x)$$

Observación B.2.22. a) *En los puntos (i) y (ii) es fácil ver usando la desigualdad triangular que si la afirmación se satisface para algún $x_0 \in X$, entonces se satisface para cualquier $x_0 \in X$.*

b) *Como consecuencia del teorema, W_p metriza la topología débil de las medidas en cualquier subconjunto de $P_p(X)$ que satisface la propiedad de ser tight uniforme, descripta en ii).*

c) *En particular si d es acotada, W_p metriza la topología débil de $P(X)$*

Demostración. Probaremos primero que las afirmaciones $(ii), (iii)$ y (iv) son equivalentes.
Claramente $(iv) \Rightarrow (iii)$. Verifiquemos que $(ii) \Rightarrow (iv)$. Sea $\varphi \in C(X)$ tal que $|\varphi(x)| \leq C[1 + d(x_0,x)^p]$ para algún $x_0 \in X$ e algún $C > 0$. Dado $R > 1$, reescribimos:

$$\varphi(x) = \varphi_R(x) + \psi_R(x)$$

donde $\varphi_R(x) = \text{mín}\{(\varphi(x), C(1+R^p))\}$ y $\psi_R(x) = \varphi(x) - \varphi_R(x)$. Observamos que: $|\psi_R(x)| \leq Cd(x_0,x)^p 1_{\{d(x_0,x)\geq R\}}(x)$. Entonces,

$$\left|\int \varphi d\mu_k - \int \varphi d\mu\right| \leq \left|\int \varphi_R d(\mu_k - \mu)\right| + C\int_{\{d(x_0,x)\geq R\}} d(x_0,x)^p d(\mu_k+\mu)(x).$$

Como $\phi_R \in C_b(X)$, la 1ra integral en el miembro derecho tiende a 0 cuando $k \to +\infty$. Luego,

$$\limsup_{k\to\infty}\left|\int \varphi d\mu_k - \int \varphi d\mu\right| \leq \limsup_{k\to\infty} C\int_{d(x_0,x)\geq R} d(x_0,x)^p d[\mu_k+\mu](x).$$

Tomamos ahora limite $R \to \infty$ observando que

$$\lim_{R\to\infty}\limsup_{k\to\infty}\int_{d(x_0,x)\geq R} d(x_0,x)^p d[\mu_k+\mu](x) = 0$$

donde usamos (ii) para la integral con μ_k y el teorema de convergencia dominada para la integral con μ (recuerde que $\mu \in P_p(X)$). Concluímos entonces que $\lim_{k\to\infty}\int \varphi d\mu_k = \int \varphi d\mu$.

Probamos ahora que $(iii) \Rightarrow (ii)$. Escribimos primero que para todo $R > 0$ y todo k,

$$\begin{aligned} 0 &\leq \int d(x_0,x)^p - [d(x_0,x)\wedge R]^p \; d\mu_k(x) \\ &= \int d(x_0,x)^p\, d(\mu_k-\mu)(x) + \int d(x_0,x)^p - (d(x_0,x)\wedge R)^p \; d\mu(x) \\ &\quad + \int (d(x_0,x)\wedge R)^p \; d(\mu-\mu_k)(x) \end{aligned}$$

Observamos ahora que la 1ra integral tiende a 0 cuando $k \to +\infty$ en vista de (iii); la 2nda tiende a 0 cuando $R \to +\infty$ por convergencia dominada ya que $\mu \in P_p(X)$; la 3ra tiende a 0 cuando $k \to +\infty$ con R fijo ya que $\mu_k \rightharpoonup \mu$. Obtenemos por lo tanto

$$\lim_{R\to\infty}\lim_{k\to\infty}\int d(x_0,x)^p - [d(x_0,x)\wedge R]^p\, \mu_k(x) = 0.$$

En particular, como el integrand es nonegativo,

$$\lim_{R\to\infty}\lim_{k\to\infty}\int_{\{d(x_0,x)\geq 2R\}} d(x_0,x)^p - [d(x_0,x)\wedge R]^p\, \mu_k(x) = 0.$$

Deducimos (ii) observando que $d(x_0,x)^p - R^p \geq (1-2^{-p})d(x_0,x)^p$ si $d(x_0,x) \geq 2R$.

Hasta ahora hemos probado que (ii), (iii) y (iv) son equivalentes. Para terminar de probar el teorema vamos a mostrar que (i) y (iii) son equivalentes.

Supongamos que $(\mu_k)_k \subset P_p(X)$ y $\mu \in P_p(X)$ verifican $W_p(\mu_k, \mu) \to 0$. Por el teorema B.2.12 existe $\pi_k \in \Pi(\mu_k, \mu)$ óptimo:

$$\int d(y,x)^p d\pi_k(x,y) = W_p(\mu_k,\mu)^p \stackrel{k\to\infty}{\longrightarrow} 0. \tag{B.6}$$

Probamos primero que

$$\limsup_{k\to\infty} \int d(x_0,x)^p d\mu_k(x) \le \int d(x_0,x)^p d\mu(x). \tag{B.7}$$

Recordemos la siguiente desigualdad elemental: para todo $\epsilon > 0$, existe $C_\epsilon > 0$ tal que para todo $a, b \in \mathbb{R}_{\ge 0}$,

$$(a+b)^p \le (1+\epsilon)a^p + C_\epsilon b^p.$$

Combinando esta desigualdad junto con la desigualdad triangular, obtenemos

$$d(x_0,x)^p \le (1+\epsilon)d(x_0,y)^p + C_\epsilon\, d(y,x)^p.$$

Integrando con respecto a π_k recordando que π_k tiene por marginal μ_k y μ, deducimos

$$\int d(x_0,x)^p d\mu_k(x) \le (1+\epsilon)\int d(x_0,y)^p d\mu(x) + C_\epsilon \int d(y,x)^p d\pi_k(x,y).$$

Usando (B.6) concluimos

$$\limsup_{k\to\infty} \int d(x_0,x)^p d\mu_k(x) \le (1+\epsilon)\int d(x_0,y)^p d\mu(x).$$

Obtenemos (B.7 tomamdo $\epsilon \to 0$.

Admitimos por un momento que $\mu_k \rightharpoonup \mu$ y probemos que

$$\lim_{k\to\infty} \int d(x_0,x)^p d\mu \le \liminf_{k\to\infty} \int d(x_0,x)^p d\mu_k(x). \tag{B.8}$$

Junto con (B.7), obtenemos la propiedad de convergencia del momento de orden p de (iii). Para probar (B.8) escribimos, usando $\mu_k \rightharpoonup \mu$, que para todo $R > 0$,

$$\lim_{k\to\infty} \int [d(x_0,x) \wedge R]^p\, d\mu_k = \int [d(x_0,x)\wedge R]^p\, d\mu.$$

Por convergencia dominada deducimos

$$\int d(x_0,x)^p\, d\mu(x) = \lim_{R\to\infty} \lim_{k\to\infty} \int [d(x_0,x)\wedge R]^p\, d\mu_k.$$

Finalmente como $d(x_0,x)\wedge R \le d(x_0,x)$, deducimos (B.8).

Para terminar la prueba de la equivalencia $(i) \Leftrightarrow (iii)$, falta probar que si $W_p(\mu_k, \mu) \to 0$ entonces $\mu_k \rightharpoonup \mu$ y que (iii) implica (i). Veamos que es suficiente probar estas afirmaciones suponiendo d acotada. En efecto, sea $\tilde{d} = \text{ínf}\,(d, 1)$ y $\tilde{W}_p$ la distancia de Wasserstein asociada a $\tilde{d}$. Claramente, $W_p \geq \tilde{W}_p$ por lo que si queremos probar que la convergencia en W_p implica convergencia en sentido débil, basta probar que la convergencia en $\tilde{W}_p$ implica la convergencia en sentido débil. Por otro lado supongamos que vale (iii) y que $\tilde{W}_p(\mu_k, \mu) \to 0$. Veamos que $W_p(\mu_k, \mu) \to 0$. Para esto usaremos la siguiente desigualdad:

$$d(x, y) \leq d(x, y) \wedge R + 2\ d(x, x_0)\chi_{d(x,x_0)\geq R/2}(x) + 2\ d(x_0, y)\chi_{d(x_0,y)\geq R/2}(y)$$

y por lo tanto, para algúna constante $C_p > 0$ que depende solo de p, tenemos:

$$d(x, y)^p \leq C_p \left([d(x, y) \wedge R]^p + d(x, x_0)^p\chi_{d(x,x_0)\geq R/2}(x) + d(x_0, y)^p\chi_{d(x_0,y)\geq R/2}(y)\right)$$

Sea π_k un plan de transferencia optimo entre μ_k y μ para la función costo $c(x, y) = d^p(x, y)$. Por la desigualdad anterior, si $R \geq 1$, tenemos:

$$\begin{aligned} W_p(\mu_k, \mu) = \int d^p(x, y)d\pi_k &\leq C_p \int [d(x, y) \wedge R]^p\, d\pi_k + C_p \int_{\left\{d(x_0,x)\geq \frac{R}{2}\right\}} d(x, x_0)^p d\pi_k \\ &+ C_p \int_{\left\{d(x_0,y)\geq \frac{R}{2}\right\}} d(y, x_0)^p d\pi_k \\ &\leq C_p R^p \tilde{W}_p(\mu_k, \mu) + C_p \int_{\left\{d(x_0,x)\geq \frac{R}{2}\right\}} d(x, x_0)^p d\mu_k \\ &+ C_p \int_{\left\{d(x_0,y)\geq \frac{R}{2}\right\}} d(y, x_0)^p d\mu \end{aligned}$$

Tomando limite $k \to \infty$, obtenemos:

$$W_p(\mu_k, \mu) \leq C_p \int_{\left\{d(x_0,x)\geq \frac{R}{2}\right\}} d(x, x_0)^p d\mu + C_p \int_{\left\{d(x_0,y)\geq \frac{R}{2}\right\}} d(y, x_0)^p d\mu$$

De ahora en mas asumiremos entonces que la función distancia d es acotada, spg: $d \leq 1$. Por la prop. B.2.16, todas las distancias W_p son equivalentes y por lo tanto podemos assumir que $p = 1$.

Supongamos que $W_1(\mu_k, \mu) \to 0$ y probemos que $\mu_k \rightharpoonup \mu$. Por lo dicho anteriormente, esto termina la prueba de la implicacion $(i) \implies (iii)$. Recuerde que por el teorema B.2.20 de Kantorovich-Rubinstein,

$$W_1(\mu_k, \nu) = \sup_{\|\varphi\|_{Lip}\leq 1} \int \varphi d(\mu_k - \mu) \overset{k\to\infty}{\longrightarrow} 0. \tag{B.9}$$

Note que si $\phi : X \to \mathbb{R}$ es una función lipschitz, entonces $\phi/\mathrm{Lip}(\phi)$ es 1-Lipschitz por lo que

$$\int \varphi d(\mu_k - \mu) \overset{k\to\infty}{\longrightarrow} 0.$$

Tomemos ahora una función $\phi \in C_b(X)$ cualquiera y probemos que $\int \phi\, d(\mu_k - \mu) \to 0$. En un espacio métrico, cualquier función continua y acotada puede ser aproximada tanto superior como inferiormente por sucesiones de funciones Lipschitz. Existen entonces dos sucesiones $(a_n)_{n\in\mathbb{N}}$ y $(b_n)_{n\in\mathbb{N}}$, respectivamente creciente y decreciente, de funciones Lipchitz e uniformemente acotadas convergiendo puntualmente a ϕ.

Basta considerar $a_n(x) = \inf_{y\in X} \{\phi(y) + nd(x,y)\}$ y $b_n(x) = \sup_{y\in X} \{\phi(y) - nd(x,y)\}$.

Como cada es b_n es Lipschitz,

$$\lim_{k\to\infty} \int b_n\, d\mu_k = \int b_n\, d\mu.$$

para todo n. Luego, usando que $b_n \geq \phi$,

$$\limsup_{k\to\infty} \int \varphi d\mu_k \leq \liminf_{n\to\infty} \int b_n d\mu. \tag{B.10}$$

Finalmente, aplicando el teorema de convergencia monotona obtenemos:

$$\limsup_{k\to\infty} \int \varphi d\mu_k \leq \int \varphi d\mu. \tag{B.11}$$

Aplicando las mismas ideas con la sucesión $(a_n)_{n\in\mathbb{N}}$, tenemos que

$$\liminf_{k\to\infty} \int \varphi d\mu_k \geq \int \varphi d\mu. \tag{B.12}$$

Concluimos $\int \phi\, d(\mu_k - \mu) \to 0$.

Probemos finalmente que $(iii) \implies (i)$. Supongamos que $\mu_k \rightharpoonup \mu$ con convergencia del 1er momento (recuerde que $p = 1$) y mostremos que (B.9) vale. Dado $x_0 \in X$, definimos $Lip_{1;x_0}$ como el espacio de funciónes $\varphi : X \to \mathbb{R}$ 1-Lipschitz tales que $\varphi(x_0) = 0$. Note que para probar (B.9), basta probar que

$$\sup_{\varphi\in Lip_{1;x_0}} \int \varphi\; d(\mu_k - \mu) \overset{k\to\infty}{\longrightarrow} 0. \tag{B.13}$$

En efecto, dado φ 1-Lipschitz, definimos $\tilde{\varphi(x)} = \varphi(x) - \varphi(x_0)$. Entonces $\tilde{\varphi} \in Lip_{1;x_0}$. Por lo tanto, tenemos que:

$$\int \varphi(x) d(\mu_k - \mu) - \int \varphi(x_0)\; d(\mu_k - \mu) = \int \tilde{\varphi}\; d(\mu_k - \mu) \overset{k\to\infty}{\longrightarrow} 0.$$

Como $\mu_k \rightharpoonup \mu$, la 2nda integral en el miembro izquierda tiende a 0 por lo que obtenemos (B.9).

Probemos ahora (B.13). Por definicion del supremo existe una sucesión $(\varphi_k)_k \in Lip_{1;x_0}(X)$ tal que

$$\sup_{\varphi \in Lip_{1;x_0}} \int \varphi \; d(\mu_k - \mu) \leq \int \varphi_k \; d(\mu_k - \mu) + \frac{1}{k}.$$

Debemos entonces probar que $\int \varphi_k \; d(\mu_k - \mu) \to 0$. Si supieramos que las φ_k convergen en $C(X)$ hacia algúna $\varphi_\infty \in C_b(X)$ entonces podriamos escribir

$$\int \varphi_k \; d(\mu_k - \mu) = \int (\varphi_k - \varphi_\infty \; d(\mu_k - \mu) + \int \varphi_\infty \; d(\mu_k - \mu)$$

y cada integral tiende a 0, la primera gracias a la convergencia uniforme de las φ_k y la segunda por $\mu_k \rightharpoonup \mu$. La idea de la prueba consiste en obtener la convergencia de las φ_k uniformemente sobre un compacto K que contiene casi toda la masa de las μ_k y de μ. Obtendremos la convergencia uniforme de las φ_k aplicando el teorema de Arzela-Ascoli. De hecho es facil ver que

(B.14) $\qquad Lip_{1;x_0}(X)$ es uniformemente equicontinuo y acotado en $C(X)$.

ya que cada función $\varphi \in Lip_{1;x_0}(X)$ es 1-Lipschitz por lo que $Lip_{1;x_0}(X)$ es uniformemente equicontinuo, y para todo $x \in X$, $|\varphi(x)| = |\varphi(x) - \varphi(x_0) \leq d(x, x_0) \leq 1$. Esta propiedad justifica la introduccion del espacio $Lip_{1;x_0}(X)$. Elegimos ahora el compacto K. Como la sucesión $(\mu_k)_k$ es débilmente convergente, sabemos por el teorema B.2.13 de Prohorov que existe una sucesión creciente de conjuntos compactos $(K_n)_{n\in\mathbb{N}}$ tales que $\sup_{k\in\mathbb{N}} \mu_k \left(K_n^c\right) \leq \frac{1}{n}$ y $\mu\left(K_n^c\right) \leq \frac{1}{n}$ para todo $n \in \mathbb{N}$. Sin perder generalidad, podemos suponer que $x_0 \in K_1$. Cada conjunto

$$H_n = \left\{\varphi_{|K_n}, \; \varphi \in Lip_{1;x_0}(X)\right\}, \; n \in \mathbb{N},$$

es, por (B.14), uniformemente equicontinuo y acotado en $C(K_n)$ y por lo tanto compacto en $C(K_n)$ por el teorema de Arzela-Ascoli. Aplicando el argumento de la diagonal, podemos entonces extraer de la sucesión $(\varphi_k)_k$ una subsucesión, que seguimos notando $(\varphi_k)_{k\in\mathbb{N}}$, que converge, uniformemente en cada K_n, a una función φ_∞ definida y continua sobre $S := \cup_{n\in\mathbb{N}} K_n$. Como $\varphi_k \in Lip_{1;x_0}(S)$ para todo k, tenemos ademas que $\varphi_\infty \in Lip_{1;x_0}(S)$ - en particular $\|\varphi_\infty\|_\infty \leq 1$ por (B.14). Podemos extender φ_∞ en una función $\tilde{\varphi}_\infty$ 1-Lipschitz en todo X considerando $\tilde{\varphi}_\infty(x) = \inf_{y\in S} [\varphi_\infty(y) + d(x,y)]$. Por comodidad seguimos notando φ_∞ por $\tilde{\varphi}_\infty$. Entonces $\varphi_\infty \in Lip_{1;x_0}(X)$. Para concluir la prueba, basta probar que $\int \varphi_k d(\mu_k - \mu) \to 0$. Para esto escribimos que para todo n,

$$\begin{aligned} &\left|\int \varphi_k d(\mu_k - \mu)\right| \\ \leq\; &\left|\int_{K_n} \varphi_k - \varphi_\infty d(\mu_k - \mu)\right| + \left|\int_{K_n^c} \varphi_k - \varphi_\infty d(\mu_k - \mu)\right| + \left|\int_X \varphi_\infty d(\mu_k - \mu)\right| \end{aligned}$$

Recordando que ϕ_k y φ_∞ están acotadas por 1, podemos acotar los dos primeros términos por

$$\begin{aligned} & (\mu_k(K_n)+\mu(K_n))\|\varphi_k-\varphi_\infty\|_{L^\infty(K_n)} + (\|\varphi_k\|_{L^\infty(X)}+\|\varphi_\infty\|_{L^\infty(X)})(\mu_k(K_n^c)+\mu(K_n^c)) \\ \leq\ & 2\|\varphi_k-\varphi_\infty\|_{L^\infty(K_n)}+\frac{4}{n}. \end{aligned}$$

Obtenemos entonces

$$\left|\int \varphi_k d(\mu_k-\mu)\right| \leq 2\|\varphi_k-\varphi_\infty\|_{L^\infty(K_n)}+\frac{4}{n}+\left|\int_X \varphi_\infty d(\mu_k-\mu)\right|.$$

Note que la integral en el miembro derecha tiende a cero cuando k tiende a infinito porque $\mu_k \rightharpoonup \mu$ y $\varphi_\infty \in C_b(X)$. Para finalizar, fijamos un n tal que $4/n$ sea lo suficientemente chico y luego, con este n fijo, hacemos $k \to \infty$. □

La recta real

El siguiente teorema muestra la relación entre la medida que realiza al $\inf_{\{\pi\in\Pi(\mu,\nu)\}}\int_{XxX} d(x,y)^p d\pi(x,y)$ y las funciones de distribución acumulada de las medidas μ y ν.

Teorema B.2.23. *Sean μ y ν dos medidas de probabilidad en $\mathbb{R}$, sean F y G sus respectivas funciones de distribucion acumulada. Sea π la medida de probabilidad en $\mathbb{R}^2$ con funcion de distribucion acumulada dada por:*

$$H(x,y)=\min\{F(x),G(y)\}$$

Luego, $\pi\in\Pi(\mu,\nu)$ y $W_1(\mu_k,\mu)=\int|x-y|\,d\pi$. Mas aun,

$$W_1(\mu_k,\mu)=\int_0^1\left|F^{-1}(t)-G^{-1}(t)\right|dt=\int_{\mathbb{R}}|F(x)-G(x)|\,dx$$

donde F^{-1} y G^{-1} son las inversas generalizadas de F y G.

Demostración. ver [12]. □

B.3. Otras distancias que metrizan la convergencia débil

Para $x\in X$ y $A\subset X$ sea

$$d(x,A)=\inf\{d(x,y): y\in A\}$$

y para $\delta \geq 0$
$$A^{\delta} = \{x \in X : d(x, A) < \delta\}$$

Dadas las medidas $R, Q \in P(X)$ definimos

$$\sigma(R, Q) = \text{ínf}\,\{\epsilon > 0 : \; R(A) \leq Q(A^{\epsilon}) + \epsilon \; \text{ para todo cerrado } A \subset X\}$$

donde $R(A) = \int_A dR$. Sea

$$\rho(R, Q) = \text{máx}\,\{\sigma(R, Q), \sigma(Q, R)\}$$

ρ es conocida como la distancia de Prokohorov.

Proposición B.3.1. *La distancia de Prokohorov metriza la topología débil en $P(X)$.*

Demostración. Ver pagina 28 de [5]. □

Definición B.3.2. *Dado $n \geq 1$,*

$$C^n(\mathbb{R}^d) = \left\{ \phi : \mathbb{R}^d \to \mathbb{R} \; / \; \sup_{v \in \mathbb{R}^d} \sum_{k=1}^{n} \left| D^k(v) \right| < \infty \right\}$$

Definición B.3.3. *Sea $F \in P(\mathbb{R}^d)$ definimos*

$$\|F\|_*^n = \sup \left\{ \int_{\mathbb{R}^d} \phi dF : \phi \in C^n, \|\phi\| \leq 1 \right\}$$

Teorema B.3.4. *Sean $F, G \in P(\mathbb{R}^d)$. Para cada $n > 0$ existe una constante c_n que solo depende de n y de d tal que:*

$$\rho(F, G) \leq \text{máx} \left\{ c_n \, [\; \|F - G\|_n^* \;]^{\frac{1}{n+1}}, \|F - G\|_n^* \right\}$$

Demostración. Ver [7]. □

Corolario B.3.5. *La distancia inducida por $\|.\|_n^*$ metriza la topología débil en $P(X)$.*

Demostración. Se sigue de la Proposicion (B.3.1) y del Teorema (B.3.4). □

Observación B.3.6. *Los mismos teoremas y proposiciones valen si en vez de considerar el espacio $\mathbb{R}^d$, consideramos $B_R = \left\{ v \in \mathbb{R}^d \; / \; |v| < R \right\}$.*

Bibliografía

[1] G. Aletti, G. Naldi, G. Toscani, First-order continuous models of opinion formation, *SIAM J. Appl. Math.*, 67 (3), 837-853.

[2] L. Boudin, F. Salvarani, Modelling opinion formation by means of kinetic equations, chapter in *Mathematical Modeling of Collective Behavior in Socio-Economic and Life Sciences*, 245-270.

[3] Billingsley, P. Convergence of probability measures, second ed. John Wiley & SonsInc., New York, 1999. A Wiley-Interscience Publication.

[4] B. During, P. Markowich, J-F Pietschmann, M-T Wolfram, Boltzmann and Fokker-Planck equation modelling opinion formation in the presence of strong leaders, *Proc. R. Soc. A*, 465, 2009, 3687-3708.

[5] Huber, P. J. (1981) The Weak Topology and its Metrization, in Robust Statistics, John Wiley & Sons, Inc., Hoboken, NJ, USA. doi: 10.1002/0471725250.ch2

[6] L.C. Evans, Partial differential equations, GSM 19, American Mathematical Society.

[7] G. Gabetta, G. Toscani, B. Wennberg, Metrics for probability distributions and the trend to equilibrium for solutions of the Boltzmann equation, *Journal of statistical physics*, 881 (5-6), 1995, 901-934.

[8] R. Ochrombel, Simulations of Sznajd sociophysics model with convincing single opinions, *Int. J. Mod. Phys.C*, 12 (07), 2001, 1091-1091.

[9] F. Slanina, H. Lavicka, Analytical results for the Sznajd model of opinion formation,*Eur. Phys. J. B*, 35, 2003, 279-288.

[10] K. Sznajd-Weron, J. Snazjd, Opinion evolution in closed community, *Int. J. Mod. Phys.C*, 11, 2000, 1157-1165.

[11] G. Toscani, Kinetic models of opinion formation, *Comm. Math. Sci.*, 4 (3), 2006, 481-496.

[12] C. Villani, Topics in optimal transportation, Graduate Studies in Mathematics vol 58, Americam Mathematical Society, 2003.