



UNIVERSIDAD DE CHILE  
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS  
DEPARTAMENTO DE INGENIERIA MATEMÁTICA

JUEGOS REPETIDOS CON INFORMACIÓN INCOMPLETA EN TIEMPO  
CONTINUO

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERO CIVIL  
MATEMÁTICO

CLAUDIO JAVIER NICOLÁS PAREJA PINEDA

PROFESOR GUÍA:  
NICOLÁS ANDRÉS FIGUEROA GONZÁLEZ

MIEMBROS DE LA COMISIÓN:  
ROBERTO MARIO COMINETTI COTTI-COMETTI  
JUAN FERNANDO ESCOBAR CASTRO

SANTIAGO DE CHILE  
OCTUBRE 2010

RESUMEN DE LA MEMORIA  
PARA OPTAR AL TÍTULO DE  
INGENIERO CIVIL MATEMÁTICO  
POR: CLAUDIO JAVIER NICOLÁS PAREJA  
FECHA: 05/11/2010  
PROF. GUÍA: NICOLÁS FIGUEROA GONZALÉZ

## JUEGOS REPETIDOS A TIEMPO CONTINUO CON INFORMACIÓN INCOMPLETA

El objetivo del presente trabajo de título es estudiar las implicancias de un modelo a tiempo continuo para juegos repetidos con información incompleta, con el objeto de estudiar la versatilidad y verosimilitud de este marco teórico.

Los juegos repetidos comenzaron a ser estudiados desde un contexto teórico discreto llegándose a caracterizar el conjunto de pagos alcanzables en equilibrio, entendiéndose equilibrio como aquella situación desde la cual ninguno de los jugadores quiere cambiarse unilateralmente. En particular, se caracterizaron dichos pagos en el caso en que los jugadores saben exactamente qué jugaron los demás -caso información completa- y en el caso en que los jugadores sólo saben parcialmente cómo sus contrincantes actuaron -caso de información incompleta.

Para el caso en tiempo continuo, aquí, se presentan intuiciones que justifican el modelo, además, se explican algunas de sus características más extensamente que en su presentación original (Sannikov, [5]). El modelo será para dos jugadores y nos permitirá caracterizar completamente el conjunto de pagos de equilibrio. Más aún, podremos encontrar las acciones que los jugadores deben tomar para alcanzar la frontera de dicho conjunto de pagos. Junto a lo anterior, se proponen heurísticas sobre cómo caracterizar la frontera del conjunto de pagos en distintos contextos.

Además, se resuelven varios juegos de bienes complementarios usando el modelo a tiempo continuo. Estos juegos ayudarán a concluir acerca de los efectos que tienen la falta de información, la impaciencia y la incertidumbre sobre el conjunto de pagos alcanzables en equilibrio.

Finalmente, se concluye que el modelo presentado es muy versátil para modelar muchos tipos de juegos repetidos. También, entrega resultados verosímiles e interesantes sobre el comportamiento de los jugadores. Se muestra también que las heurísticas funcionan para obtener resultados numéricos al aplicar el modelo a los juegos presentados aunque se sugiere buscar alguna otra manera para obtener dichos resultados de manera más eficiente y mejor justificada teóricamente.

# Índice general

Índice de figuras	v
Índice de cuadros	vii
<b>1. Introducción</b>	<b>1</b>
<b>2. Tiempo Discreto</b>	<b>3</b>
2.1. Caso de Información Completa. . . . .	3
2.1.1. Juego Repetido Canónico. . . . .	3
2.1.2. Principio de Desviación Simple. . . . .	11
2.1.3. Caracterización de los pagos de equilibrio. . . . .	13
2.1.4. El <i>Folk Theorem</i> . . . . .	20
2.2. Caso de Información Incompleta. . . . .	22
2.2.1. Juego Repetido Canónico. . . . .	22
2.2.2. Desviación Simple y Caracterización de Equilibrios . . . . .	24
<b>3. Tiempo Continuo e Información Incompleta</b>	<b>29</b>
3.1. El Modelo . . . . .	30
3.2. Caracterización de Equilibrios Públicos Perfectos . . . . .	34
3.3. Equilibrios Públicos Perfectos con valores extremos . . . . .	36
3.3.1. El valor de continuar tiene una trayectoria curva . . . . .	36
3.3.2. Ecuación de optimalidad y puntos extremos de $\mathcal{E}$ . . . . .	38
3.3.3. Conclusión sobre PPE con valores extremos . . . . .	40

<b>4. Cálculo de Equilibrios en Tiempo Continuo</b>	<b>42</b>
4.1. Marco Teórico . . . . .	42
4.2. Caso Simétrico . . . . .	43
4.3. Caso Asimétrico . . . . .	44
4.3.1. Conjuntos sin Nash en la frontera de $\mathcal{V}^*$ . . . . .	44
4.3.2. Conjuntos con Nash en la frontera de $\mathcal{V}^*$ . . . . .	46
<b>5. Estudio de Casos</b>	<b>50</b>
5.1. Caso Simétrico . . . . .	52
5.1.1. Efecto del descuento . . . . .	57
5.1.2. Efecto de la incertidumbre . . . . .	57
5.2. Asimetría en la aleatoriedad de las señales . . . . .	58
5.2.1. Comparación con el caso simétrico . . . . .	59
5.3. Asimetría en las señales . . . . .	60
5.3.1. Comparación con caso simétrico y caso asimétrico en la varianza . . . . .	61
5.4. Asimetría en la complementaridad . . . . .	63
<b>6. Conclusiones</b>	<b>67</b>
6.1. Preguntas Abiertas . . . . .	67
<b>Bibliografía</b>	<b>69</b>
<b>Índice alfabético</b>	<b>70</b>

# Índice de figuras

2.1.	$\mathcal{F}$ en la etapa del Dilema del Prisionero. . . . .	7
2.2.	$\mathcal{F}^p$ para la etapa del Dilema del Prisionero . . . . .	7
2.3.	Autómata para la estrategia egoísta pero no tanto . . . . .	10
2.4.	Autómata para la estrategia de Tolerancia Cero . . . . .	15
2.5.	Autómata para la demostración del teorema 2.1.17 . . . . .	21
2.6.	Autómata para la estrategia de Tolerancia Cero, en el caso de información incompleta. . . . .	26
3.1.	$\mathcal{V}^*$ en el juego de bienes complementarios . . . . .	33
3.2.	Pago instantáneo de las acciones en el juego de bienes complementarios	33
3.3.	Análisis de curvatura de $W_t$ . . . . .	37
3.4.	Análisis de curvatura de $\partial\mathcal{E}$ . . . . .	39
3.5.	Análisis de curvatura de $\partial\mathcal{E}$ . . . . .	39
4.1.	Ejemplo de Iteración para el caso simétrico . . . . .	45
4.2.	Ejemplo de Iteración para el caso asimétrico . . . . .	49
4.3.	Otro método para caso asimétrico . . . . .	49
5.1.	Pago instantáneo de las distintas acciones en el juego de referencia . . . . .	52
5.2.	$\mathcal{V}^*$ en el juego de referencia . . . . .	52
5.3.	Conjunto $\mathcal{E}(2)$ para el caso simétrico . . . . .	54
5.4.	Conjunto $\mathcal{E}(2)$ para el caso simétrico y <i>drifts</i> tangenciales . . . . .	55
5.5.	Conjunto $\mathcal{E}(2)$ para el caso simétrico y las volatilidades . . . . .	56
5.6.	Conjunto $\mathcal{E}(r)$ para distintos valores de $r$ . . . . .	57
5.7.	Conjunto $\mathcal{E}(r)$ para distintos valores de incertidumbre . . . . .	58
5.8.	Conjunto $\mathcal{E}(5)$ con incertidumbre no simétrica . . . . .	59

---

5.9. Comparación de $\mathcal{E}(r)$ para dos formas de la varianza . . . . .	60
5.10. Comparación de $\mathcal{E}(r)$ para dos formas de la varianza . . . . .	60
5.11. Conjunto $\mathcal{E}(3)$ cuando uno de los jugadores esconde su precio dentro de su señal . . . . .	61
5.12. Comparación de $\mathcal{E}(r)$ para el caso simétrico y asimétrico en las señales	62
5.13. Comparación de $\mathcal{E}(r)$ para el caso asimétrico en la aleatoriedad de las señales y para el caso asimétrico en las señales . . . . .	62
5.14. Conjunto $\mathcal{E}(3)$ para el caso asimétrico en las señales y las volatilidades	62
5.15. Pagos instantáneos que se pueden alcanzar cuando complementaridad beneficia más al jugador 2 . . . . .	64
5.16. $\mathcal{V}^*$ cuando la complementaridad beneficia más al jugador 2 . . . . .	64
5.17. Pagos instantáneos que se pueden alcanzar cuando complementaridad beneficia más al jugador 1 . . . . .	64
5.18. $\mathcal{V}^*$ cuando la complementaridad beneficia más al jugador 1 . . . . .	64
5.19. $\mathcal{E}(r)$ cuando la complementaridad beneficia más al jugador 2, $\gamma_1 < 1$ .	65
5.20. $\mathcal{E}(r)$ cuando la complementaridad beneficia más al jugador 1 . . . . .	65
5.21. $\mathcal{E}(r)$ cuando la complementaridad beneficia más al jugador 2, $\gamma_1 < 1$ .	66
5.22. $\mathcal{E}(r)$ cuando la complementaridad beneficia más al jugador 1 . . . . .	66

# Índice de cuadros

2.1. Matriz de pagos para el Dilema del Prisionero. . . . .	6
2.2. Matriz de pagos para el juego de los amigos. . . . .	9
2.3. Matriz de pagos para el Dilema del Prisionero. . . . .	17
2.4. Pagos en el estado $w = w_{EE}$ para el juego estático equivalente al Dilema del Prisionero repetido. . . . .	17
2.5. Pagos en el estado $w = w_{AA}$ para el juego estático equivalente al Dilema del Prisionero repetido. . . . .	17
2.6. Pagos ex post y ex antes para el Dilema del Prisionero Repetido con Información Incompleta . . . . .	23
2.7. Pagos en cada estado del juego estático equivalente para el Dilema del Prisionero con Información Incompleta . . . . .	27
3.1. Pagos instantáneos esperados para el juego de bienes complementarios.	33

# Capítulo 1

## Introducción

Al ir por primera, y quizás única, vez a un restaurante ¿quién no se ha sentido impulsado a no dejar propina? Pero, ¿qué ocurre si somos asiduos al mismo restaurante? ¿es acaso más probable que ahora sí dejemos una buena propina? Esta situación la podemos pensar como un juego entre el cliente y quién lo atiende, el primero quiere que lo atiendan bien al menor costo, por su parte, el mesero quiere recibir el mayor beneficio con el menor esfuerzo<sup>1</sup>; la experiencia nos estaría diciendo que sí el juego ocurre una sola vez, el cliente puede aprovecharse del mesero, pero si el juego se repite, ambos tienen incentivos para jugar *limpiamente*.

Juegos como el anterior son los que nos llevan a escribir este Trabajo de Título, juegos donde existe una *etapa* que se repite una y otra vez. En estos juegos, se mostrará, los jugadores pueden alcanzar mejores pagos que si sólo jugarán la primera vez, esto, gracias a considerar decisiones de largo plazo en las que todos salgan beneficiados. Claro, alcanzar dichos pagos requerirá incentivos dinámicos, éstos pueden ser: beneficios, o bien castigos, futuros. En nuestro ejemplo del restaurante un beneficio o castigo para el mesero involucra la cantidad de propina que se le dejará en el futuro dependiendo de cómo haga su trabajo hoy. En la otra vereda, al cliente se le ofrecerá un mejor o peor servicio en su próxima visita dependiendo de la propina de hoy.

Si el plato que trae el mesero está frío, ¿es realmente culpa del mesero o puede que sea del cocinero? Supongamos que es más probable que sea del mesero, pero *más probable* no es una certeza: en este caso, la información acerca del trabajo que realiza quién sirve la mesa nos llega *incompleta*. Al considerar este fenómeno, nuestro juego cambia, pues ya no podemos ver completamente las *acciones* de los jugadores y esto traerá consecuencias en los pagos que ellos pueden alcanzar; por ejemplo, puede que para que el acuerdo al que ellos lleguen sea creíble, uno de los jugadores deba ser castigado, a pesar que no se sabe a ciencia cierta si fue él o el azar el verdadero

---

<sup>1</sup>Claro que este es un modelo, probablemente ambos personajes tienen más intereses en juego que sólo los relacionados con el dinero.

culpable del resultado no deseado. Más concreto, volvamos a nuestro contexto de ejemplo: digamos que el mesero y el cliente han convenido que ambos cooperarán pero si alguna vez el mesero trae el plato frío o el cliente deja poca propina, quién se sienta defraudado castigará al otro la siguiente visita; ahora bien, si el plato viene frío, el cliente deberá dejar poca propina, incluso si él sabe que pueden que no sea culpa del mesero, pues si no lo hace, todo el acuerdo, y las amenazas que incluye, se volverían no creíbles.

En el segundo capítulo analizaremos resultados clásicos para juegos que se repiten en tiempo discreto, primero analizaremos el caso cuando las acciones son completamente observables y luego veremos el caso en que las acciones no son completamente públicas. El tercer capítulo trata sobre un modelo distinto para juegos repetidos con información incompleta, en él, los jugadores repiten la etapa en todo momento, es decir, el juego es a tiempo continuo, a pesar que esta característica parece una sofisticación académica, permite observar más claramente cómo son las acciones que los jugadores toman y como ellos pueden alcanzar los máximos pagos posibles. El cuarto capítulo está destinado a explicar cómo es posible aplicar el modelo a tiempo continuo para obtener numéricamente lo que los jugadores harán en una situación de total cooperación. El capítulo siguiente, el quinto, mostrará con ejemplos lo que se obtiene al realizar lo planteado en el capítulo anterior, más precisamente, se analizará un juego de bienes complementarios y se concluirá sobre el efecto de la impaciencia de los jugadores, el de la incertidumbre en la información, entre otros factores. Por último, en el capítulo sexto se explicitan las conclusiones que se alcanzaron durante la realización de este Trabajo de Título.

# Capítulo 2

## Juegos Repetidos a Tiempo Discreto.

En este capítulo sentamos las bases para el análisis posterior, además de fijar notación y repasar los resultados tradicionales para juegos repetidos. Básicamente seguimos a Mailath y Samuelson [3, capítulos 2, 3 y 7].

Comenzaremos con el caso más sencillo donde la información (las acciones de los demás) son completamente observables, esto lo hacemos en 2.1. Después, al pasar a la sección 2.2, se analiza el caso donde no toda la información es observable.

Cada sección esta ordenada de la misma manera, partiendo primero por explicitar cuál es el juego a considerar y cuáles son los equilibrios buscados, esto en 2.1.1 y 2.2.1; luego, buscamos propiedades de éstos y finalmente se caracterizan dichos equilibrios, 2.1.2, 2.1.3 y 2.2.2.

### 2.1. Caso de Información Completa.

#### 2.1.1. Juego Repetido Canónico.

**El juego *a repetir*.**

El juego repetido se compone de un juego *a repetir* infinitas veces, llamaremos a este juego la *etapa* del juego repetido. En la etapa y el juego repetido se tienen  $n$  jugadores.

Hablaremos de *acciones* para las decisiones tomadas en la etapa y de *estrategias* para los comportamientos en el juego repetido, es decir, una estrategia será una regla (o función) que indicará que acción tomar en las distintas repeticiones de la etapa dependiendo de lo ocurrido hasta ese momento. El conjunto de las acciones para el jugador  $i$  lo denotamos por  $A_i$  y a sus elementos por  $a_i$ . El conjunto de los *perfiles*

de acciones puras será  $A := \prod_i A_i$ . Se asumirá que cada  $A_i$  es un conjunto finito en  $\mathbb{R}$ , pero en general muchos de los resultados de este capítulo se pueden generalizar para el caso en que  $A_i$  es un conjunto compacto.

*Nota 2.1.1.* En general se denotará lo relacionado con el jugador  $i$  con un subíndice  $i$ , además usaremos la notación usual del subíndice negativo,  $-i$ , para notar lo relacionado con los demás jugadores. Por ejemplo,  $A_{-i} := \prod_{j \neq i} A_j$ , así,  $A = A_i \times A_{-i}$ .

Con lo anterior en mente, los pagos de la etapa están dados por una función continua:

$$u : \prod_i A_i \rightarrow \mathbb{R}^n$$

Extendemos las acciones posibles considerando ahora las acciones *mixtas*, cuyo conjunto para el jugador  $i$  lo denotamos por  $\Delta(A_i)$ <sup>1</sup>, un elemento en él será  $\alpha_i$ , y el conjunto de perfiles mixtos será denotado por  $\prod_i \Delta(A_i)$ . Los pagos de la etapa son extendidos a funciones mixtas tomando esperanza, esto es,  $u(\alpha_i) = \mathbb{E}_{\alpha_i}(u)$ .

El conjunto de los pagos para las etapas con posibles perfiles en  $A$  lo escribimos como

$$\mathcal{F} := \{v \in \mathbb{R}^n : \exists a \in A, v = u(a)\}$$

Con lo que el conjunto de pagos *alcanzables*,  $\mathcal{F}^\dagger$ , queda dado por la envoltura convexa de  $\mathcal{F}$ , ie,  $\mathcal{F}^\dagger := \text{co}\mathcal{F}$ . A continuación presentamos algunas definiciones que nos permitirán sentar un lenguaje para comparar acciones entre sí, en términos de los pagos que producen.

**Definición 2.1.2** (Pago dominante). Diremos que un pago  $v' \in \mathcal{F}^\dagger$  domina estrictamente (resp. débilmente) a  $v \in \mathcal{F}^\dagger$  si

$$\forall i, v'_i > v_i \text{ (resp } v'_i \geq v_i)$$

**Definición 2.1.3** (Eficiencia de los pagos). Un pago  $v$  será *ineficiente* si existe un pago alcanzable que lo domina, conversamente será *eficiente* (o *Pareto eficiente*) si no es ineficiente.

Se dirá que un pago es *estrictamente eficiente* si no es ineficiente y no es dominado débilmente por ningún otro posible pago.

**Definición 2.1.4** (Mejor Respuesta). Una acción  $a_i$  es una mejor respuesta a  $a_{-i}$  si maximiza, sobre todas las acciones posibles de  $i$ , el pago que obtiene el jugador  $i$  con el perfil dado por la acción de  $i$  y por  $a_{-i}$ , ie,

$$a_i \in \operatorname{argmax}_{a_i \in A_i} u_i(a_i, a_{-i})$$

Con las definiciones anteriores, podemos dejar en claro una primera definición de equilibrio:

<sup>1</sup> $\Delta(C)$  denota el simplex donde a cada acción en  $C$  se le asocia una probabilidad.

**Definición 2.1.5** (Equilibrio de Nash). Un perfil de acciones  $a$  es un equilibrio de Nash del juego a repetir si para todo jugador  $i$  y para toda acción  $\tilde{a}_i$  se tiene que

$$u_i(a) \geq u_i(\tilde{a}_i, a_{-i})$$

Es decir,  $a$  es un equilibrio de Nash si ningún jugador tiene incentivos para desviarse unilateralmente.

Por el teorema de existencia de Nash [4], usando que el número de jugadores es finito y que el conjunto de acciones también lo es, es posible mostrar que existe al menos un equilibrio de Nash en cada etapa (posiblemente un equilibrio en estrategias mixtas).

Es de interés para el análisis posterior definir una cota por abajo para los pagos que el jugador  $i$  busca obtener, una cota interesante se da cuando, dado un jugador  $i$ , los demás jugadores tratan de minimizar el pago de  $i$  mientras ese jugador elige su acción de manera de maximizar su utilidad en esta situación adversa, llamaremos a este pago el *pago minmax*, que queda dado por:

$$\underline{v}_i^p := \min_{a_{-i} \in A_{-i}} \max_{a_i \in A_i} u_i(a_i, a_{-i})$$

La compacidad de  $A$  y la continuidad de  $u_i$  nos aseguran que el pago minmax esta bien definido. Un perfil minmax para el jugador  $i$  es un perfil  $\tilde{a}^i = (\tilde{a}_i^i, \tilde{a}_{-i}^i)$  con la propiedad de que  $\tilde{a}_i^i$  es una mejor respuesta para  $\tilde{a}_{-i}^i$  y  $\underline{v}_i^p = u_i(\tilde{a}_i^i, \tilde{a}_{-i}^i)$ .

**Definición 2.1.6** (Pago individualmente racional). Un vector de pagos  $v = (v_1, \dots, v_n)$  es *débil* (resp. *estricta*) e *individualmente racional* en acciones puras si  $v_i \geq \underline{v}_i^p$  (resp  $v_i > \underline{v}_i^p$ ) para todo  $i$ .

El nombre *racional* en la definición anterior viene al caso pues a ningún jugador le interesan pagos menores a los que él se puede asegurar en el *peor* de los escenarios. Pensando en lo anterior definimos el conjunto de pagos alcanzables y estrictamente racionales por:

$$\mathcal{F}^{\dagger p} := \{v \in \mathcal{F}^{\dagger} : v_i > \underline{v}_i^p, i = 1, \dots, n\}$$

Similarmente, el conjunto de pagos generados por perfiles en acciones puras esta dado por:

$$\mathcal{F}^p := \{v \in \mathcal{F} : v_i > \underline{v}_i^p, i = 1, \dots, n\}$$

La extensión a acciones mixtas es la natural teniendo en cuenta que *una acción mixta es una mejor respuesta a  $a_{-i}$  si y sólo si cada acción con probabilidad positiva es a su vez una mejor respuesta a  $a_{-i}$* , entonces las definiciones quedan como antes, considerando

$$\underline{v}_i := \min_{\alpha_{-i} \in \prod_{j \neq i} \Delta(A_j)} \max_{a_i \in A_i} u_i(a_i, \alpha_j)$$

en vez de  $\underline{v}_i^p$ , aunque esta vez el conjunto de pagos alcanzables y racionales (con respecto a  $\underline{v}_i$ ) queda denotado por  $\mathcal{F}^*$  y definido por

$$\mathcal{F}^* := \{v \in \mathcal{F}^\dagger : v_i > \underline{v}_i, i = 1, \dots, n\}$$

**Ejemplo 2.1.1.** Consideremos el *Dilema del Prisionero*. Este juego consiste en dos *prisioneros* que son acusados de un crimen, ambos jugadores pueden esforzarse (E) por la liberación de ambos o bien, aprovecharse del otro jugador (A), los pagos recibidos se muestran en el cuadro 2.1.1. La idea tras los pagos es que quién se aprovecha del otro prisionero sale beneficiado a costa del bienestar de él mientras sea uno sólo quién se aprovecha: si son ambos, ambos pierden; en la vereda contraria si ambos cooperan pueden salir bien, aunque no *tan* bien.

		Jugador 2	
		E	A
Jugador 1	E	2,2	-1,3
	A	3,-1	0,0

Cuadro 2.1: Matriz de pagos para el Dilema del Prisionero.

Podemos ver que en este juego, la mejor respuesta frente a cualquier acción del contrario es aprovecharse de él, así, el único equilibrio de Nash, es aquel donde ambos juegan dicha acción, i.e. el equilibrio de Nash es  $(A, A)$ , alcanzando ambos jugadores el pago 0. Es de importancia destacar que el máximo pago alcanzable por los jugadores corresponde al caso en que ambos cooperan  $(E, E)$  y obtienen un pago de 2.

Para ejemplificar las definiciones, calculemos el pago minmax de este juego, para ello notemos que si el jugador 2 juega  $E$ , entonces lo mejor que puede hacer el jugador 1 es jugar  $A$ , similarmente si el jugador 2 juega  $A$ , lo mejor que puede hacer 1 es jugar también  $A$ . Lo anterior nos indica la parte *max* del pago, ahora si el perfil jugado es  $(A, E)$  el jugador 1 recibe un pago de 3, en cambio, si se juega  $(A, A)$ , el jugador 1, recibe 0; luego el jugador 2, elige jugar  $A$  para minimizar el mejor pago posible de 1, i.e.  $\underline{v}_1^p = 0$ . Usando el valor calculado es posible graficar  $\mathcal{F}$  y  $\mathcal{F}^p$  en las figuras 2.1 y 2.2.

### El juego repetido.

El juego repetido consiste en jugar la etapa en cada uno de los periodos  $t \in \{0, 1, \dots\}$ , por eso el apelativo *a tiempo discreto*. El título de *información completa* (o *monitoreo perfecto*) se aplica puesto que al final de cada periodo todos los jugadores pueden ver el perfil escogido, es decir, cada jugador es completamente monitoreado por los demás jugadores; hacemos la salvedad de que cuando un jugador

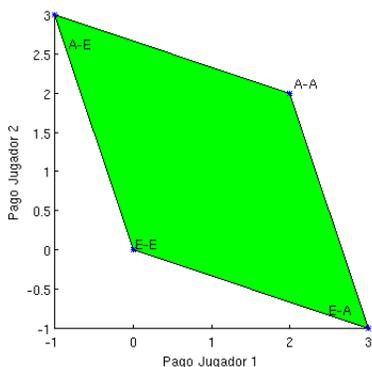


Figura 2.1:  $\mathcal{F}$  en la etapa del Dilema del Prisionero.

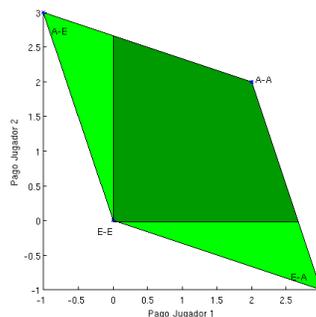


Figura 2.2: En verde oscuro se muestra  $\mathcal{F}^P$  para la etapa del Dilema del Prisionero.

esta aleatorizando (i.e., jugando estrategias mixtas) sólo es observada la realización escogida.

Definamos el conjunto de historias al periodo  $t \geq 0$  por:

$$\mathcal{H}^0 := \phi, \quad \mathcal{H}^t := A^t = \underbrace{A \times \dots \times A}_{t \text{ veces}}$$

Luego, una historia  $h^t \in \mathcal{H}^t$  es una lista de  $t$  perfiles, indicando las acciones jugadas desde el periodo 0 al  $t$ . Si agregamos un nuevo perfil para el periodo  $t + 1$ , se obtiene un elemento  $h^{t+1} \in \mathcal{H}^{t+1} = \mathcal{H}^t \times A$ . El conjunto de todas las posibles historias es

$$\mathcal{H} := \bigcup_{t=0}^{\infty} \mathcal{H}^t$$

Una *estrategia pura*,  $\sigma_i$ , para el jugador  $i$  es una función que dependiendo de en que historia el jugador se encuentra le indica la acción a jugar, ie,

$$\sigma_i : \mathcal{H} \rightarrow A_i$$

Una *estrategia mixta* para el jugador  $i$  es una aleatorización sobre todo el conjunto de estrategias puras. Puede ser más conveniente trabajar, sin pérdida de generalidad, con *estrategias de comportamiento* que consisten en funciones  $\sigma_i$  tal que

$$\sigma_i : \mathcal{H} \rightarrow \Delta(A_i)$$

De ahora en adelante, abusando un poco del lenguaje, nos referiremos simplemente como *estrategias* a las estrategias de comportamiento.

Para cada historia  $h^t \in \mathcal{H}$ , hablamos del *juego continuado* para referirnos al juego repetido que comienza en el periodo  $t$ , siguiendo la historia  $h^t$ . Similarmente,

para cualquier perfil de estrategias  $\sigma = (\sigma_1, \dots, \sigma_n)$ , la *estrategia de continuación inducida por  $h^t$*  para el jugador  $i$ , denotada por  $\sigma_i |_{h^t}$  esta dada por

$$\sigma_i |_{h^t} := \sigma_i(h^t h^\tau), \quad \forall h^\tau \in \mathcal{H}$$

donde  $h^t h^\tau$  es la unión de la historia  $h^t$  seguida de la historia  $h^\tau$ . Es decir, ésta es la estrategia inducida por la estrategia  $\sigma_i$  en el juego continuado posterior a la historia  $h^t$ . Notar que como  $\sigma_i |_{h^t}$  es una estrategia en el juego original, esto es,  $\sigma_i |_{h^t}: \mathcal{H} \rightarrow \Delta(A_i)$ , el juego continuado asociado a cada historia es un subjuego estrategicamente igual al juego original, es decir, los juegos repetidos tienen una estructura recursiva, que será clave en el estudio de ellos. Llevando lo anterior a los demás jugadores, de ahora en adelante escribimos  $\sigma |_{h^t}$  para  $(\sigma_1 |_{h^t}, \dots, \sigma_n |_{h^t})$ .

Un *camino de resultados* (o simplemente un *resultado*) en el juego repetido es una secuencia infinita de perfiles  $\mathbf{a} := (a^0, a^1, \dots) \in A^\infty$ . Notar que a diferencia de una historia, un resultado es una secuencia infinita, aunque si  $\mathbf{a}^t$  denota los primeros  $t$  periodos de un resultado  $\mathbf{a}$ , entonces  $\mathbf{a}^t$  es la historia en  $\mathcal{H}^t$  correspondiente a  $\mathbf{a}$ .

Con un perfil de estrategias puras  $\sigma$  es posible definir el resultado asociado,  $\mathbf{a}(\sigma) := (a^0(\sigma), a^1(\sigma), \dots)$ , como sigue. En el primer periodo, el perfil de acciones

$$a^0(\sigma) := (\sigma_1(\phi), \dots, \sigma_n(\phi))$$

es jugado. En el segundo periodo la historia  $a^0(\sigma)$  lleva al perfil de acciones

$$a^1(\sigma) := (\sigma_1(a^0(\sigma)), \dots, \sigma_n(a^0(\sigma)))$$

luego en el tercero la historia  $(a^0(\sigma), a^1(\sigma))$  lleva a un perfil como el anterior; siguiendo este procedimiento,  $\mathbf{a}(\sigma)$  está completamente definido. Similarmente, una estrategia de comportamiento induce un *camino en el juego*, esto es, una distribución de probabilidad sobre los distintos resultados que pueden ocurrir dado el perfil de acciones mixtas inicial.

Supongamos  $\sigma$  en un perfil de estrategias puras. En el periodo  $t$ , el perfil de acciones puras inducida  $a^t(\sigma)$  entrega el pago  $u_i(a^t(\sigma))$  al jugador  $i$ . Luego, el resultado  $\mathbf{a}(\sigma)$  induce un flujo infinito de pagos para cada jugador  $i$ , dado por

$$(u_i(a^0(\sigma)), u_i(a^1(\sigma)), \dots) \in \mathbb{R}^\infty$$

Cada jugador descuenta estos pagos con un factor  $\delta \in [0, 1)$ , entonces el pago (normalizado) del perfil de estrategias puras para el jugador  $i$  está dado por

$$U_i(\sigma) := (1 - \delta) \sum_{t=0}^{\infty} \delta^t u_i(a^t(\sigma))$$

Como antes, el pago de una estrategia mixta o de comportamiento se define como el valor esperado de los resultados ocurridos.

Tal como en la etapa, es posible definir un equilibrio de Nash para el juego repetido:

**Definición 2.1.7** (Equilibrio de Nash). Un perfil de estrategias  $\sigma$  es un *equilibrio de Nash para el juego repetido* si para todo jugador  $i$  y para toda estrategia  $\sigma'_i$  de ese jugador, se tiene que:

$$U_i(\sigma) \geq U_i(\sigma'_i)$$

En juegos con dinámicas complejas el concepto de equilibrio de Nash puede ser muy permisivo (ver Mailath y Samuelson [3, sección 2.1.4] para una discusión más extensa), en el sentido de no cumplir con el requisito de *racionalidad secuencial*, este requisito viene de permitir que los jugadores en cada periodo rehagan su problema de optimización, incluso si en dicho periodo se encuentran fuera de equilibrio.

**Ejemplo 2.1.2.** Consideremos la etapa planteada en el cuadro 2.1.2, en ella dos amigos desean reunirse para pasar un buen rato, el problema es que el amigo/jugador 1 es un fanático del fútbol, mientras el otro es fanático del juego de computador *Starcraft*. Más aún, ninguno de los amigos es bueno en el pasatiempo del otro, por lo que el beneficio conjunto no cambia si se juntan o no. En esta etapa existe un único equilibrio de Nash en acciones puras, a saber  $(F, S)$ . La idea es analizar si al repetir esta etapa, podemos observar que un equilibrio de Nash en el juego repetido, no es lo más natural que uno esperaría.

		Jugador 2	
		S	F
Jugador 1	S	-1,1	0,0
	F	0,0	1,-1

Cuadro 2.2: Matriz de pagos para el juego de los amigos.

Definamos la estrategia *egoísta pero no tanto*, en ella un jugador juega su preferencia en ambas repeticiones, salvo en el caso que un jugador beneficie al otro, en ese caso, el jugador beneficiado en la primera etapa, beneficia al otro en la segunda. Por ejemplo, si en la primera etapa el jugador 1 decide ceder y juega  $S$ , en la segunda el jugador 2 hará lo suyo y elegirá  $F$ . Si ambos jugadores siguen ésta estrategia, cada uno recibe un pago de 0, el perfil para esta situación se representa por el autómata de la figura 2.3 (ver 2.1.3 para una discusión sobre autómatas que representan estrategias), en esa figura los *estados  $w$*  indican cuál es el perfil jugado si se sigue la estrategia, y las flechas indican qué perfil debe ser jugado para que se pase de un estado a otro

Ahora estamos en condiciones de demostrar que la estrategia *egoísta pero no tanto* es un equilibrio de Nash del juego repetido, pero falla al dejar que los jugadores rehagan su problema de optimización en cada etapa. Estudiemos las desviaciones unilaterales que pueden ocurrir:

1. El jugador 1 se desvía sólo en el primer periodo. Es claro que no tiene incentivos para hacerlo, pues juega el Nash de la etapa durante ese periodo.  
Análogamente el jugador 2, tampoco tiene incentivos por la misma razón.

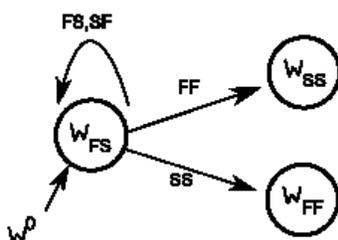


Figura 2.3: Autómata para la estrategia egoísta pero no tanto

2. El jugador 1 se desvía sólo en el segundo periodo. Cómo ambos jugadores siguieron la estrategia en el primer periodo, en el segundo periodo el jugador 2 jugará  $S$ , por lo que nuevamente es óptimo para el jugador 1 seguir la estrategia y jugar  $F$  tal como dice ella para ese periodo. En el caso del jugador 2, nuevamente usando el mismo argumento que para el jugador 1 se concluye que no se puede desviar de esta manera.

3. El jugador se desvía en ambos periodos. Si el jugador 1 elige  $S$  en el primer periodo obtiene una ganancia de  $-1$ . Para el segundo periodo, el otro jugador elegirá  $F$ , luego es óptimo para el jugador 1, jugar  $F$  y obtener una ganancia de 1. Con lo anterior, la utilidad para el jugador 1 es de 0, por lo que tampoco tiene incentivos para jugar esta desviación.

Similarmente, si el jugador 2 elige  $F$  para el primer periodo y obtiene una ganancia de  $-1$ , en el segundo, el otro jugador escogerá  $S$ , por lo que es óptimo para el jugador 2 escoger  $S$  y obtener una ganancia de 1, alcanzado una utilidad de 0. Así, él tampoco tiene incentivos para desviarse.

Luego, la estrategia planteada es un equilibrio de Nash para el juego repetido, pues no existen desviaciones unilaterales que entreguen más utilidad que ella.

Consideremos ahora la situación en la cual *ambos* jugadores rehacen su problema de optimización antes de comenzar la etapa 2. Esto es natural, pues no parece real decir que si el jugador ya analizó sus jugadas antes de comenzar el juego, no lo volverá a hacer una vez que ya se sabe que acciones han ido jugándose, más aún, podríamos decir incluso que si el jugador olvidara las razones por las que jugar cierta estrategia, debiera ser capaz de encontrarlas nuevamente en *cada* etapa y no sólo al principio del juego. En este contexto, supongamos el jugador 1 decide desviarse en la primera repetición, con lo que según la estrategia planteada, el perfil que se dará en la segunda etapa será  $(F, F)$ . Ahora bien, si los jugadores pueden repensar sus decisiones en cada etapa, ¿es óptimo para el jugador 2 jugar  $F$  en la segunda etapa, si es que el jugador 1 también jugará esa acción?, la respuesta es que no. No nos dimos cuenta de esto antes, pues al utilizar la definición de equilibrio de Nash,

sólo nos fijamos que la estrategia sea óptima si *ya* se está jugándola y considerando que sólo un jugador a la vez puede repensar la estrategia durante el juego.

Lo anterior impone un nuevo sabor a los equilibrios llevándonos a la siguiente definición:

**Definición 2.1.8** (Equilibrio perfecto en subjuegos.). Un perfil de estrategias  $\sigma$  es un *equilibrio perfecto en subjuegos del juego repetido* (o equilibrio en subjuegos) si para todas las historias  $h^t \in \mathcal{H}$ ,  $\sigma|_{h^t}$  es un equilibrio de Nash para el juego repetido.

Notar que si la etapa tiene un equilibrio de Nash, entonces un equilibrio perfecto en subjuegos consiste en repetir dicho equilibrio en cada etapa.

### 2.1.2. Principio de Desviación Simple.

Una observación importante es que según la definición de equilibrio en subjuegos, para chequear la optimalidad sería necesario revisar todas las historias posibles en  $\mathcal{H}$  lo que aumenta la complejidad del problema de manera innumerable, las definiciones y propiedades siguientes crean un método simple para caracterizar dichos equilibrios.

**Definición 2.1.9** (Desviación simple rentable). Una *desviación simple* para el jugador  $i$  desde una estrategia  $\sigma_i$  es una estrategia  $\hat{\sigma}_i \neq \sigma_i$  con la propiedad de que existe una única historia  $\hat{h}^t \in \mathcal{H}$  en la cual ambas estrategias difieren, ie,  $\hat{\sigma}_i$  es una desviación simple si

$$\hat{\sigma}_i \neq \sigma_i \quad \wedge \quad \exists! \hat{h}^t \in \mathcal{H} \quad \forall h^\tau \neq \hat{h}^t \quad \hat{\sigma}_i(h^\tau) = \sigma_i(h^\tau)$$

Dado un perfil de estrategias  $\sigma_{-i}$  para los oponentes. Se dirá que una desviación simple  $\hat{\sigma}_i$  desde  $\sigma_i$  es *rentable* si en la historia  $\hat{h}_i$  para la cual  $\hat{\sigma}_i(\hat{h}_i) \neq \sigma_i(\hat{h}_i)$  se tiene que

$$U_i(\hat{\sigma}_i |_{\hat{h}_i}, \sigma_{-i} |_{\hat{h}_i}) > U_i(\sigma |_{\hat{h}_i})$$

**Teorema 2.1.10** (Principio de la desviación simple). *Un perfil de estrategias  $\sigma$  es un equilibrio en subjuegos ssi no hay desviaciones simples rentables.*

*Demostración.* Para simplificar la notación, analizaremos el caso en que las estrategias de equilibrio usan sólo acciones puras.

Si un perfil de estrategias es un equilibrio perfecto en subjuegos, claramente no pueden haber desviaciones rentables.

Conversamente, supongamos el perfil  $\sigma$  no es un equilibrio perfecto y mostremos que debe existir una desviación simple rentable. Como  $\sigma$  no es un equilibrio debe haber una historia  $\tilde{h}^t$ , un jugador  $i$  y una estrategia  $\tilde{\sigma}_i$  tal que

$$U_i(\sigma_i |_{\tilde{h}^t}, \sigma_{-i} |_{\tilde{h}^t}) < U_i(\tilde{\sigma}_i, \sigma_{-i} |_{\tilde{h}^t})$$

El objetivo de la demostración es crear una desviación simple rentable a partir de  $\tilde{\sigma}$ . Sea  $\varepsilon := U_i(\tilde{\sigma}_i, \sigma_{-i}|\tilde{h}^t) - U_i(\sigma_i|\tilde{h}^t, \sigma_{-i}|\tilde{h}^t)$ . Definamos también  $m := \min_{i,a} u_i(a)$  y  $M = \max_{i,a} u_i(a)$ . Por último sea  $T$  lo suficientemente grande para que  $\delta^T(M - m) < \varepsilon/2$ . Entonces

$$\begin{aligned} & (1 - \delta) \sum_{\tau=0}^{T-1} \delta^\tau u_i(a^\tau(\sigma_i|\tilde{h}^t, \sigma_{-i}|\tilde{h}^t)) + (1 - \delta) \sum_{\tau=T}^{\infty} \delta^\tau u_i(a^\tau(\sigma_i|\tilde{h}^t, \sigma_{-i}|\tilde{h}^t)) = U_i(\sigma_i|\tilde{h}^t, \sigma_{-i}|\tilde{h}^t) \\ &= U_i(\tilde{\sigma}_i, \sigma_{-i}|\tilde{h}^t) + U_i(\sigma_i|\tilde{h}^t, \sigma_{-i}|\tilde{h}^t) - U_i(\tilde{\sigma}_i, \sigma_{-i}|\tilde{h}^t) \\ &= (1 - \delta) \sum_{\tau=0}^{T-1} \delta^\tau u_i(a^\tau(\tilde{\sigma}_i, \sigma_{-i}|\tilde{h}^t)) + (1 - \delta) \sum_{\tau=T}^{\infty} \delta^\tau u_i(a^\tau(\tilde{\sigma}_i, \sigma_{-i}|\tilde{h}^t)) - \varepsilon \end{aligned}$$

Así, usando las cotas para los pagos, además de la primera y última línea de las ecuaciones anteriores, sigue que

$$(1 - \delta) \sum_{\tau=0}^{T-1} \delta^\tau u_i(a^\tau(\sigma_i|\tilde{h}^t, \sigma_{-i}|\tilde{h}^t)) < (1 - \delta) \sum_{\tau=0}^{T-1} \delta^\tau u_i(a^\tau(\tilde{\sigma}_i|\tilde{h}^t, \sigma_{-i}|\tilde{h}^t)) - \frac{\varepsilon}{2} \quad (2.1)$$

Con lo anterior tenemos que la estrategia  $\hat{\sigma}_i$ , definida por

$$\begin{aligned} \hat{\sigma}_i(h^\tau) &= \begin{cases} \tilde{\sigma}_i & si \tau < T \\ \sigma_i|\tilde{h}^t(h^\tau) & si \tau \geq T \end{cases} \\ &= \begin{cases} \tilde{\sigma}_i & si \tau < T \\ \sigma_i(\tilde{h}^t h^\tau) & si \tau \geq T \end{cases} \end{aligned}$$

es una desviación rentable. En particular, al coincidir  $\hat{\sigma}_i$  con  $\tilde{\sigma}_i$  en los primeros  $T$  periodos, le entrega al jugador  $i$  la ganancia  $\varepsilon/2$  prometida en (2.1).

Como  $\hat{\sigma}$  difiere de  $\sigma_i|\tilde{h}^t$  sólo en los primeros  $T$  periodos, hemos demostrado que si una estrategia no es un equilibrio perfecto en sub juegos entonces existe una desviación rentable en  $T$  periodos. Para culminar la demostración, argumentamos recursivamente en el valor de  $T$  hasta hacerlo disminuir a 1.

Sea  $\hat{h}^{T-1} := (\hat{a}^0, \dots, \hat{a}^{T-2})$  la historia de  $T - 1$  periodos inducida por  $(\hat{\sigma}_i, \sigma_{-i}|\tilde{h}^t)$ , entonces hay dos posibilidades

1. Se tiene  $U_i(\sigma_i|\tilde{h}^t \hat{h}^{T-1}, \sigma_{-i}|\tilde{h}^t \hat{h}^{T-1}) < U_i(\hat{\sigma}_i|\hat{h}^{T-1}, \sigma_{-i}|\tilde{h}^t \hat{h}^{T-1})$ . En este caso la estrategia  $\hat{\sigma}_i|\hat{h}^{T-1}$  es una desviación simple rentable después de la historia  $\tilde{h}^t \hat{h}^{T-1}$  pues  $\hat{\sigma}_i|\hat{h}^{T-1}$  coincide con  $\sigma_i$  desde el periodo  $T$  inclusive (note que  $\hat{\sigma}$  comienza en  $t$ , al igual que  $\tilde{\sigma}$  pues el análisis se ha hecho asumiendo que ya ocurrió  $\tilde{h}$ ).
2. Si no se tiene lo anterior, ie, se tiene  $U_i(\sigma_i|\tilde{h}^t \hat{h}^{T-1}, \sigma_{-i}|\tilde{h}^t \hat{h}^{T-1}) \geq U_i(\hat{\sigma}_i|\hat{h}^{T-1}, \sigma_{-i}|\tilde{h}^t \hat{h}^{T-1})$ . Entonces definimos una nueva estrategia  $\bar{\sigma}$  como sigue

$$\begin{aligned} \bar{\sigma}_i(h^\tau) &= \begin{cases} \hat{\sigma}_i & si \tau < T - 1 \\ \sigma_i|\tilde{h}^t(h^\tau) & si \tau \geq T - 1 \end{cases} \\ &= \begin{cases} \tilde{\sigma}_i & si \tau < T - 1 \\ \sigma_i(\tilde{h}^t h^\tau) & si \tau \geq T - 1 \end{cases} \end{aligned}$$

Con lo que,

$$\begin{aligned} U_i(\hat{\sigma}_i|\hat{h}^{T-2}, \sigma_{-i}|\tilde{h}^t\hat{h}^{T-2}) &= (1 - \delta)u_i(\hat{a}^{T-1}) + \delta U_i(\hat{\sigma}_i|\hat{h}^{T-1}, \sigma_{-i}|\tilde{h}^t\hat{h}^{T-1}) \\ &\leq (1 - \delta)u_i(\hat{a}^{T-1}) + \delta U_i(\sigma_i|\tilde{h}^t\hat{h}^{T-1}, \sigma_{-i}|\tilde{h}^t\hat{h}^{T-1}) \end{aligned}$$

Como  $\hat{h}^{T-2}$  es la historia inducida por  $(\hat{\sigma}_i, \sigma_{-i}|\tilde{h}^t)$  hasta el periodo  $T - 2$  y además  $\bar{\sigma}_i$  coincide con  $\hat{\sigma}$  hasta ese periodo se tiene que

$$U_i(\hat{\sigma}_i, \sigma_{-i}|\tilde{h}^t) \leq U_i(\bar{\sigma}_i, \sigma_{-i}|\tilde{h}^t)$$

Recordamos que  $\hat{\sigma}$  era una desviación rentable para concluir que  $\bar{\sigma}$  también es una desviación rentable después de  $\tilde{h}^t$ , más aún,  $\bar{\sigma}$  difiere de  $\sigma_i|\tilde{h}^t$  en sólo  $T - 1$  periodos.

Repetiendo lo anterior, debemos encontrar una desviación simple rentable. □

### 2.1.3. Caracterización de los pagos de equilibrio.

#### Representación a través de autómatas para perfiles de estrategias.

A pesar de que existe diversidad de historias posibles, en general ocurre que muchas de ellas llevan a la misma estrategia de continuación, para poder clasificar y trabajar con dichas clases de equivalencia es de utilidad representar las estrategias como autómatas.

Un *autómata*  $(\mathcal{W}, w^0, f, \tau)$  consiste en un conjunto de estados  $\mathcal{W}$ , un estado inicial  $w^0 \in \mathcal{W}$ , una función de decisión  $f : \mathcal{W} \rightarrow \prod_i \Delta(A_i)$  que asocia a cada estado un perfil de acciones y una función de transición  $\tau : \mathcal{W} \times A \rightarrow \mathcal{W}$ . La función de transición identifica a qué estado se mueve el autómata dado el estado actual y la acción realizada (ver Sipser [8] para una introducción en autómatas).

Escribiremos  $f(w)$  cuando la función especifique un perfil en acciones puras y  $f^w(a)$  para la probabilidad de alcanzar el perfil  $a$  al estar en el estado  $w$  cuando la función especifique un perfil de acciones mixtas; en el primer caso, diremos que el autómata tiene salidas deterministas y en el segundo que las tiene aleatorias.

Cualquier autómata  $(\mathcal{W}, w^0, f, \tau)$  con  $f$  especificando acciones puras induce un resultado  $\{a^0, a^1, \dots\}$  como sigue:

$$a^0 = \sigma(\phi) = f(w^0)$$

$$\begin{aligned}
a^1 &= \sigma(a^0) = f(\tau(w^0, a^0)) \\
a^2 &= \sigma(a^0, a^1) = f(\tau(\tau(w^0, a^0), a^1)) \\
&\vdots
\end{aligned}$$

Es posible extender esto para pasar a trabajar con historias en vez de resultados, para ello es necesario extender la función de transición desde  $\mathcal{W} \times A$  al dominio  $\mathcal{W} \times \mathcal{H} \setminus \{\phi\}$ , se logra esto definiendo recursivamente  $\tau$  de la siguiente manera

$$\tau(w, h^t) = \tau(\tau(w, h^{t-1}), a^{t-1})$$

Con esta definición se tiene la estrategia  $\sigma$  descrita por

$$\sigma(\phi) = f(w^0), \quad \sigma(h^t) = f(\tau(w^0, h^t))$$

Conversamente, es claro que cualquier perfil de estrategias puede ser representado por un autómata. Basta tomar el conjunto de historias  $\mathcal{H}$  como conjunto de estados, la historia nula  $\phi$  como estado inicial y las funciones de decisión y transición como  $f(h^t) = \sigma(h^t)$  y  $\tau(h^t, a) = h^{t+1} := (h^t, a)$ , respectivamente. La extensión de las contrucciones anteriores es la natural para el caso de estrategias de comportamiento.

Diremos que un estado  $w' \in \mathcal{W}$  es *accesible* desde otro estado  $w \in \mathcal{W}$  si existe una secuencia de perfiles de acciones tal que empezando en  $w$  llevan eventualmente al autómata hasta el estado  $w'$ , i.e.  $w'$  es accesible desde  $w$  si  $\exists h^t, w' = \tau(w, h^t)$ .

Hasta el momento la complejidad del problema no ha disminuido, la disminución ocurre, en general, porque es posible agrupar las historias lo que nos deja con un autómata de finita cantidad de estados. Otra ayuda es la siguiente observación: cuando  $(\mathcal{W}, w^0, f, \tau)$  describe la estrategia  $\sigma$ , el perfil de las estrategias de continuación después de la historia  $h^t$  queda completamente descrito por el autómata  $(\mathcal{W}, \tau(w^0, h^t), f, \tau)$ , más aún, si todos los estados en  $\mathcal{W}$  son accesibles, entonces la colección de todas las posibles estrategias continuadas queda descrita por la colección de autómatas  $\{(\mathcal{W}, w, f, \tau) : w \in \mathcal{W}\}$ .

**Ejemplo 2.1.3.** En el Dilema del Prisionero, definimos la estrategia *Tolerancia Cero* como aquella estrategia en que el jugador coopera mientras el otro lo haga, si en algún momento el otro jugador deja de cooperar, entonces el primero nunca volverá a hacerlo, por eso, la tolerancia es *cero*.

Es posible representar esta estrategia en un autómata de 2 estados. Para ello consideremos el conjunto de estados  $\mathcal{W} = (w_{EE}, w_{AA})$ , la función de decisión  $f$  tal que  $f(w_{II}) = II$  ( $I = E, A$ ) y la función de transición  $\tau$  como:

$$\tau(a, w) = \begin{cases} w_{EE} & , \text{ si } w = w_{EE} \wedge a = EE \\ w_{AA} & , \text{ si no} \end{cases}$$

Finalmente consideramos  $w^0 = w_{EE}$ . Es posible ver gráficamente este autómata en la figura 2.4.

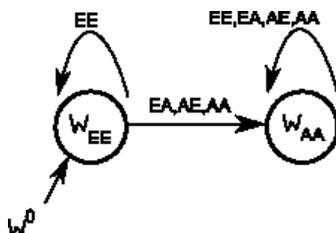


Figura 2.4: Autómata para la estrategia de Tolerancia Cero

### Promesas creíbles.

Fijemos un autómata  $(\mathcal{W}, w^0, f, \tau)$  donde todos los  $w \in \mathcal{W}$  son accesibles desde  $w^0$  y definamos  $V_i(w)$  como la utilidad descontada del juego que comienza en  $w$ <sup>2</sup>. Dada la recursividad propia del autómata es posible calcular el valor de  $V_i$  recursivamente. En efecto, teniendo en cuenta que el pago instantáneo de seguir la estrategia viene dado por

$$\sum_a u_i(a) f^w(a)$$

y que los pagos continuados corresponden a la esperanza de los pagos de los distintos estados a los que se podría llegar

$$\sum_a V_i(\tau(w, a)) f^w(a).$$

Es posible darse cuenta que  $V_i$  satisface el sistema de ecuaciones lineales,

$$V_i(w) = (1 - \delta) \sum_a u_i(a) f^w(a) + \delta \sum_a V_i(\tau(w, a)) f^w(a), \quad \forall w \in \mathcal{W}$$

De lo anterior es posible interpretar  $V_i(\tau(w, a))$  (o su esperanza) como una *promesa* para el jugador  $i$  al jugar  $a_i$ , más aún, si  $a = (a_i, f_{-i}(w))$ , entonces  $V_i$  es la promesa (o castigo) que obtiene  $i$  al desviarse de la estrategia dada por el autómata. Esta intuición es la raíz del siguiente teorema.

**Teorema 2.1.11** (Equivalencia con un juego estático). *Suponga que la estrategia  $\sigma$  esta descrita por un autómata  $(\mathcal{W}, w^0, f, \tau)$  y que  $V(w) = (V_1(w), \dots, V_n(w))$ . Entonces el perfil de estrategias  $\sigma$  es un equilibrio perfecto en subjuegos ssi para todo*

<sup>2</sup> $V_i(w^0) = U_i(\sigma)$  donde  $\sigma$  es la estrategia definida por el autómata.

$w \in \mathcal{W}$  accesible desde  $w^0$ , la acción  $f(w)$  es un equilibrio de Nash del juego en forma normal descrito por los pagos  $g^w : A \rightarrow \mathbb{R}^n$ , donde

$$g^w(a) = (1 - \delta)u(a) + \delta V(\tau(w, a))$$

*Demostración.* Nuevamente estudiaremos sólo el caso en que las acciones son puras, es decir, cuando el autómatata tiene una función de decisión determinista, hacemos esto para simplificar la notación.

Sea  $\sigma$  el perfil de estrategias inducido por  $(\mathcal{W}, w^0, f, \tau)$ . Primero demostraremos que si  $f(w)$  es un equilibrio de Nash para el juego en forma normal  $g^w$ , para todo  $w \in \mathcal{W}$ , entonces no hay desviaciones simples rentables desde  $\sigma$ . Esto, como vimos en el teorema anterior, basta para demostrar que  $\sigma$  es un equilibrio de Nash perfecto en subjuegos. Sean  $\hat{\sigma}_i$  una desviación simple y  $\hat{h}^t$  la historia tal que  $\sigma_i(\hat{h}^t) \neq \hat{\sigma}_i(\hat{h}^t)$ , respectivamente, además sea  $\hat{w}$  el estado alcanzado por la historia  $\hat{h}^t$ , es decir,  $\hat{w} := \tau(w^0, \hat{h}^t)$ . Finalmente, sea  $a_i$  la acción jugada en la desviación, es decir,  $a_i := \hat{\sigma}_i(\hat{h}^t)$ . Con lo anterior, se tiene:

$$U_i(\sigma_i|_{\hat{h}^t}, \sigma_{-i}|_{\hat{h}^t}) = V_i(\hat{w})$$

Además, como  $\hat{\sigma}_i$  es una desviación simple

$$\begin{aligned} U_i(\hat{\sigma}_i|_{\hat{h}^t}, \sigma_{-i}|_{\hat{h}^t}) &= (1 - \delta)u_i(a_i, \sigma_{-i}(\emptyset)) + \delta V_i(\tau(\hat{w}, (a_i, \sigma_{-i}|_{\hat{h}^t}(\emptyset)))) \\ &= u_i(a_i, f_{-i}(\hat{w})) + \delta V_i(\tau(\hat{w}, (a_i, f_{-i}(\hat{w})))) \\ &\leq u_i(f_i(\hat{w}), f_{-i}(\hat{w})) + \delta V_i(\tau(\hat{w}, (f_i(\hat{w}), f_{-i}(\hat{w})))) \end{aligned}$$

La última desigualdad se tiene por la hipótesis:  $\hat{w} \in \mathcal{W}$  luego  $f(\hat{w}) = \sigma(\hat{h}^t)$  es un equilibrio de Nash para el juego  $g^{\hat{w}}$ . Con dicha desigualdad no es posible que hayan desviaciones simples rentables.

Demostraremos la conversa por contrarecíproca, es decir, demostraremos que si existe  $\hat{w} \in \mathcal{W}$  tal que  $f(\hat{w})$  no es un equilibrio de Nash para el juego  $g^{\hat{w}}$ , entonces  $\sigma$  no es un equilibrio perfecto en subjuegos, nuevamente por el teorema anterior, basta demostrar que hay una desviación simple rentable desde  $\sigma$  para concluir que  $\sigma$  no es un equilibrio perfecto en subjuegos. Con lo anterior en mente, supongamos existe  $\hat{w} \in \mathcal{W}$ , accesible desde  $w^0$ , y  $a_i$  tal que

$$(1 - \delta)u_i(a, f_{-i}(\hat{w})) + \delta V_i(\tau(\hat{w}, (a_i, f_{-i}(\hat{w})))) > V_i(\hat{w}) \quad (2.2)$$

Es decir,  $f(\hat{w})$  no es un equilibrio de Nash del juego  $g^{\hat{w}}$ . Sea  $\hat{h}^t$  la historia tal que  $\hat{w} = \tau(w^0, \hat{h}^t)$  y definamos

$$\hat{\sigma}_i(\mathcal{H}^\tau) = \begin{cases} a_i, & h^\tau = \hat{h}^t \\ \sigma_i(h^\tau), & h^\tau \neq \hat{h}^t \end{cases}$$

Luego,  $\hat{\sigma}_i$  es una desviación simple desde  $\sigma_i$ , y debido a la desigualdad (2.2) es rentable.  $\square$

**Ejemplo 2.1.4.** Los pagos para la etapa del Dilema del Prisionero considerado anteriormente se encuentran nuevamente en el cuadro 2.3. El objetivo es mostrar que la estrategia tolerancia cero es un equilibrio de Nash en subjuegos, para ello usaremos el teorema anterior y la construcción del autómata hecha en la figura 2.4. También necesitaremos la utilidad descontada  $V$ , por ellos hagamos explícito que toma los siguientes valores:  $V(w_{EE}) = (2, 2)$  y  $V(w_{AA}) = (0, 0)$ , calculados a partir del cuadro y figura recién mencionados.

		Jugador 2	
		E	A
Jugador 1	E	2,2	-1,3
	A	-1,3	0,0

Cuadro 2.3: Matriz de pagos para el Dilema del Prisionero.

Así las cosas, según el teorema 2.1.11 debemos considerar dos juegos estáticos: uno para  $w_{EE}$  y otro para  $w_{AA}$ . Para  $w_{EE}$ , los pagos quedan dados por  $(1 - \delta)u_i(a) + \delta V_i(\tau(w_{EE}, a))$ , dichos pagos están representados en el cuadro 2.1.4. Sabiendo dichos pagos es fácil ver que  $f(W_{EE}) = EE$  será un equilibrio de Nash ssi  $\delta > 1/3$ . Ahora bien, revisando los pagos asociados a  $f(w_{AA}) = w_{AA}$  en el cuadro 2.1.4, vemos que cualquier valor para  $\delta$  hace de  $AA$  un equilibrio. Tomando ambas consideraciones, se concluye que la tolerancia cero es un equilibrio perfecto en subjuegos ssi  $\delta > 1/3$ .

		E	A
		2,2	$-(1-\delta), 3(1-\delta)$
E	$3(1-\delta), -(1-\delta)$	0,0	
A			

Cuadro 2.4: Pagos en el estado  $w = w_{EE}$  para el juego estático equivalente al Dilema del Prisionero repetido.

		E	A
		$2(1-\delta), 2(1-\delta)$	$-(1-\delta), 3(1-\delta)$
E	$3(1-\delta), -(1-\delta)$	0,0	
A			

Cuadro 2.5: Pagos en el estado  $w = w_{AA}$  para el juego estático equivalente al Dilema del Prisionero repetido.

### Construcción.

La construcción que se presenta a continuación es debida a Abreu, Pearce y Stacchetti [1].

En esta sección nos restringiremos a acciones puras para hacer la exposición más clara. Denotemos por  $\mathcal{E}^p \subset \mathbb{R}^n$  el conjunto de los pagos que son alcanzados en un equilibrio perfecto en subjuegos utilizando acciones puras. Para cada valor  $v \in \mathcal{E}^p$ , sea  $\sigma^v$  una estrategia pura que es un equilibrio y que otorga  $v$ . De la discusión en la sección anterior, sabemos existe una función  $\gamma : A \rightarrow \mathcal{E}^p$  (a saber,  $V(\tau(w, \cdot))$ ) tal

que si  $a^*$  es parte del equilibrio entonces, para todo jugador  $i$  y para toda acción  $a_i \in A_i$ ,

$$(1 - \delta)u_i(a^*) + \delta\gamma_i(a^*) \geq (1 - \delta)u_i(a_i, a_{-i}^*) + \delta\gamma_i(a_i, a_{-i}^*).$$

La idea ahora es ver si una acción para la cual existe una función similar a cierto conjunto  $\mathcal{W}$  (en vez de a  $\mathcal{E}^p$ ), hace que dicho conjunto sea alcanzable en equilibrio. La respuesta será que sí, más aún, todo conjunto de pagos alcanzables en equilibrio se podrá caracterizar de esta forma.

**Definición 2.1.12** (Acción Vinculable). Un perfil de acciones  $a^*$  es *vinculable en*  $\mathcal{W}$  si existe una función (promesa)  $\gamma : A \rightarrow \mathcal{W}$  tal que para todo jugador  $i$ , y todo  $a_i \in A_i$ , se tiene que:

$$(1 - \delta)u_i(a^*) + \delta\gamma_i(a^*) \geq (1 - \delta)u_i(a_i, a_{-i}^*) + \delta\gamma_i(a_i, a_{-i}^*).$$

Para cada par  $(a^*, a_i)$  donde se cumple lo anterior es posible decir que  $\gamma$  vincula a  $a^*$  frente a  $a_i$ .

Es decir, una acción será vinculable si existe alguna función que hace dicha acción *atractiva* en términos de la función y de la utilidad.

**Definición 2.1.13** (Descomposición de un pago). Un pago  $v \in \mathcal{F}^\dagger$  es *descomponible con acciones puras en*  $\mathcal{W}$  si existe un perfil de acciones puras  $a^* \in \mathcal{W}$  vinculable en  $\mathcal{W}$  tal que

$$v_i = (1 - \delta)u_i(a^*) + \delta\gamma_i(a^*)$$

donde  $\gamma$  es la función que hace vinculable a  $a^*$ .

**Definición 2.1.14** (Conjunto Auto-Generante). Un conjunto  $\mathcal{W}$  es *auto-generante en acciones puras* si cada pago  $v \in \mathcal{W}$  es descomponible con acciones puras en  $\mathcal{W}$ .

La intuición nos sugiere que los conjuntos auto-generantes deberían ser alcanzables en equilibrio pues cada pago dentro de él nos lleva a un pago y una promesa dentro de dicho conjunto, digamos  $(1 - \delta)u_i(a^*)$  y  $v' := \delta\gamma_i(a^*)$ , respectivamente. Pero, la promesa  $v'$  es a su vez vinculable (*llevable a cabo*), y la segunda promesa (la asociada a  $v'$ ) también lo será, y así sucesivamente, por lo que un pago en un conjunto auto-generante sería como un pago creíble para todos los periodos. El siguiente resultado hace formal dicha intuición.

**Teorema 2.1.15** (Conjuntos Auto-Generantes son Alcanzables en Equilibrios). *Cualquier elemento de un conjunto  $\mathcal{W} \subset \mathcal{F}^\dagger$  autogenerante en acciones puras es un pago alcanzable a través de un equilibrio perfecto en subjuegos con estrategias puras.*

*Demostración.* Para cada pago  $v \in \mathcal{W}$ , sea  $\tilde{a}(v)$  el perfil de acciones que descompone a  $v$  y  $\gamma^v : A \rightarrow \mathcal{W}$  la función que vincula dicho perfil. Consideremos la colección de autómatas  $\{(\mathcal{W}, v, f, \tau) : v \in \mathcal{W}\}$ , donde el conjunto de estados es el mismo para todos ellos y viene dado por  $\mathcal{W}$ , la función de decisión también es común, siendo dada por

$$f(v) = \tilde{a}(v), \quad \forall v \in \mathcal{W}$$

por último la función de transición también es igual para todos los autómatas, ella viene dada por

$$\tau(v, a) = \gamma^v(a), \quad \forall a \in A$$

Es decir, el autómata siempre en cada estado/pago juega la acción que lo descompone y se mueve al siguiente estado/pago según la promesa que vincula a dicha acción frente a la acción jugada. Notemos que los autómatas sólo difieren en su estado inicial  $v \in \mathcal{W}$ .

Necesitamos demostrar que para cada  $v \in \mathcal{W}$ , el autómata  $(\mathcal{W}, v, f, \tau)$  describe un equilibrio perfecto en subjuegos con pago  $v$ . Esto será una implicancia del teorema 2.1.11 y la descomponibilidad del pago  $v$ , una vez que mostremos que

$$v_i = V_i(v)$$

siendo  $V_i(v)$  el valor para el jugador  $i$  de estar en el estado (pago)  $v$ .

Como cada pago  $v \in \mathcal{W}$  es descomponible en un perfil de acciones puras, podemos definir una secuencia de pares pagos-acciones  $\{(v^k, a^k)\}_{k=0}^{\infty}$  como sigue:

- $v^0 = v$ .  
 $a^0 = \tilde{a}(v^0)$ .
- $v^k = \gamma^{v^{k-1}}(a^{k-1}) = \tau(v^{k-1}, a^{k-1})$ , es decir, el estado al que se llega según lo que se jugó en el estado (pago) anterior. Otra forma de verlo:  $v^k$  es el pago involucrado en la descomposición de  $v^{k-1}$ .  
 $a^k = \tilde{a}(v^k) = f(v)$ , la acción que se juega en el nuevo estado, o bien, la acción que descompone al pago  $v^k$ .

Con las definiciones anteriores y la definición de descomposición, se tiene

$$\begin{aligned} v_i &= (1 - \delta)u_i(a^0) + \delta v_i^1 \\ &= (1 - \delta)u_i(a^0) + \delta[(1 - \delta)u_i(a^1) + \delta v_i^2] \\ &= (1 - \delta) \sum_{s=0}^{t-1} \delta^s u_i(a^s) + \delta^t v_i^t \end{aligned}$$

Como  $v_i^t \in \mathcal{F}^\dagger$ ,  $\{v_i^t\}$  es una sucesión acotada, por lo cual tomando  $t \rightarrow \infty$ , sigue que

$$v_i = (1 - \delta) \sum_{s=0}^{\infty} \delta^s u_i(a^s)$$

Con ello  $v_i = V_i(v)$ , pues la suma en la relación anterior es exactamente los pagos esperados al seguir  $\sigma$  partiendo en  $v$ . Ahora que ya se tiene la relación buscada el resultado sigue inmediatamente: recordemos que los pagos del juego estático (ver teorema 2.1.11) son

$$\begin{aligned} g^v(a) &= (1 - \delta)u(a) + \delta V(\tau(v, a)) \\ &= (1 - \delta)u(a) + \delta V(\gamma^v(a)) \\ &= (1 - \delta)u(a) + \delta \gamma^v(a) \end{aligned}$$

Claramente de la definición de vinculabilidad sigue que  $f(w) = \tilde{a}(v)$  es un equilibrio de Nash del juego  $g^v$ . En efecto, sea  $a_i \neq \tilde{a}_i(v)$ , luego tenemos

$$\begin{aligned} g^v(\tilde{a}(v)) &= (1 - \delta)u(\tilde{a}(v)) + \delta \gamma^v(\tilde{a}(v)) \\ &\geq (1 - \delta)u(a_i, \tilde{a}_{-i}(v)) + \delta \gamma^v(a_i, \tilde{a}_{-i}(v)) \\ &= (1 - \delta)u(a_i, \tilde{a}_{-i}(v)) + \delta V(\tau(v, (a_i, \tilde{a}_{-i}(v)))) \end{aligned}$$

Con lo anterior, para cada  $v$ , la estrategia definida por el autómata asociado es un equilibrio perfecto en subjuegos con pago  $v$ .  $\square$

**Corolario 2.1.16** (Caracterización de  $\mathcal{E}^p$ ). *El conjunto de pagos alcanzables con equilibrios perfectos en subjuegos a través de estrategias puras,  $\mathcal{E}^p$ , es el mayor conjunto auto-generante en acciones puras.*

Dado una pago  $v \in \mathcal{E}^p$  es posible explicitar con que equilibrio es posible alcanzar dicho pago, para una discusión al respecto ver Mailath y Samuelson [3, sección 2.6].

#### 2.1.4. El Folk Theorem

Dentro de la literatura de los juegos, comunes e importantes son los teoremas *Folk* (o *del pueblo*), estos teoremas se caracterizan por demostrar que bajo ciertas condiciones, el conjunto de pagos racionales  $\mathcal{F}^*$  es completamente alcanzable a través de equilibrios perfectos en subjuegos. En general basta con la condición de que los jugadores sean lo suficientemente paciente ( $\delta \approx 1$ ), o sea, se demuestra que si los jugadores están dispuestos a esperar *mucho* por promesas futuras, entonces para cualquier pago mayor que los pagos minmax de ambos jugadores existe un equilibrio perfecto en subjuegos que alcanza dicho pago.

En esta subsección sólo trataremos con el teorema más simple: el teorema del pueblo para dos jugadores con información completa, en él demostraremos que los pagos en  $\mathcal{F}^p$  son alcanzables si los jugadores son suficientemente pacientes (para una discusión sobre cómo alcanzar los pagos en  $\mathcal{F}^*$  o sobre el caso con más jugadores ver Mailath y Samuelson [3, capítulo 3]).

**Teorema 2.1.17** (Teorema Folk para  $\mathcal{F}^p$ ). *Para cualquier perfil de acciones puras,  $\tilde{a}$ , estrictamente racional, es decir,  $u_i(\tilde{a}) > \underline{v}_i^p$ ,  $i = 1, 2$ , existe  $\underline{\delta} \in [0, 1)$  tal que para todo  $\delta \in (\underline{\delta}, 1)$  hay un equilibrio perfecto en subjugos con descuento  $\delta$  en el cual  $\tilde{a}$  se juega en cada periodo.*

*Demostración.* Fijemos  $\tilde{a}$  tal que  $u_i(\tilde{a}) > \underline{v}_i^p$  para todo  $i$ . Definamos  $M$  como el mayor pago posible para cualquier jugador dentro de la etapa, ie,

$$M := \max_{i \in \{1,2\}, a \in A} u_i(a).$$

El perfil de estrategias buscado es tal que si algún jugador deja de jugar  $\tilde{a}$ , ambos jugadores se castigan el uno al otro jugando el perfil minmax mutuo durante  $L$  periodos (donde  $L$  lo determinaremos en lo que sigue). Más precisamente, si definimos

$$\hat{a}_j = \operatorname{argmín} \left\{ \min_{a_j \in A_j} \max_{a_i \in A_i} u_i(a_i, a_j) \right\}, \quad j = 1, 2$$

el perfil de estrategias de minmax mutuo, entonces la estrategia viene dada por el autómata con estados  $\mathcal{W} := \{w(l) : l = 0, \dots, L\}$ , estado inicial  $w^0 := w(0)$ , función de decisión

$$f(w(l)) = \begin{cases} \tilde{a}, & l = 0 \\ \hat{a}, & l = 1, \dots, L \end{cases}$$

y función de transición

$$\tau(w(l), a) = \begin{cases} w(0), & (l = 0 \wedge a = \tilde{a}) \text{ ó } (l = L \wedge a = \hat{a}) \\ w(l+1), & 0 < l < L \wedge a = \hat{a} \\ w(1), & \sim \end{cases}$$

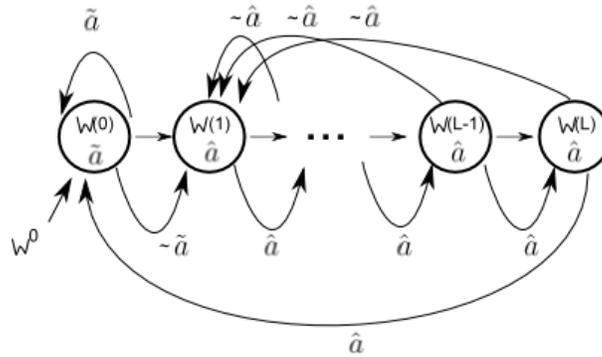


Figura 2.5: Autómata para la demostración del teorema 2.1.17. Dentro de cada estado  $w(l)$ , su etiqueta y  $f(w(l))$ . Al mostrar las transiciones, denotamos por  $\sim a$  al conjunto  $A \setminus \{a\}$ .

Como para todo  $i$  se tiene  $u_i(\hat{a}) \leq \underline{v}_i^p < u_i(\tilde{a})$ , existe  $L$  tal que

$$L \min_i (u_i(\tilde{a}) - u_i(\hat{a})) > M - \min_i u_i(\tilde{a}), \quad \forall i \quad (2.3)$$

Con lo anterior, cualquier jugador lo *suficientemente* paciente, preferirá, estrictamente,  $L + 1$  periodos de  $u_i(\tilde{a})$  en vez de desviarse, obteniendo a lo más  $M$  en el periodo actual y luego sufriendo  $L$  periodos de  $u_i(\hat{a})$ . Ahora determinaremos que significa *suficientemente* paciente, para ello, definamos primero

$$v_i^* = (1 - \delta^L)u_i(\hat{a}) + \delta^L u_i(\tilde{a})$$

Este pago corresponde a la fase de *castigo* del juego, donde por  $L$  periodos se recibe, a lo más, el pago minmax y luego se recibe el pago dado por  $\tilde{a}$ . Notemos que existe  $\delta^*$  tal que para todo  $i$  y  $\delta > \delta^*$ , se tiene  $v_i^* > \underline{v}_i^p$ .

La idea es demostrar que existe  $\delta$  suficientemente grande, para que

$$u_i(\tilde{a}) \geq (1 - \delta)M + \delta v_i^* \quad (2.4)$$

se cumpla. Es decir, buscamos  $\delta$  de manera que la utilidad de seguir la acción  $\tilde{a}$  sea mayor que la utilidad de desviarse en el periodo actual y luego caer en la fase de castigo. Reemplazando la definición de  $v_i^*$  en la ecuación anterior y reordenando, llegamos a:

$$\begin{aligned} (1 - \delta^{L+1})u_i(\tilde{a}) &\geq (1 - \delta)M + \delta(1 - \delta^L)u_i(\hat{a}) \\ \sum_{t=0}^L \delta^t u_i(\tilde{a}) &\geq M + \delta \sum_{t=0}^{L-1} \delta^t u_i(\hat{a}) \end{aligned}$$

Entonces, usando  $L$  como en (2.3) sigue que existe  $\underline{\delta} \geq \delta^*$  tal que la inecuación anterior se cumple para todo  $\delta \in (\underline{\delta}, 1)$  y todo  $i$ . Es decir, hemos demostrado que mientras los jugadores jueguen  $\tilde{a}$  no tienen incentivos para desviarse *simplemente*.

Falta revisar que no existen desviaciones simples durante la fase de castigo cuando  $\delta \in (\underline{\delta}, 1)$ , esto no necesariamente es verdad pues puede ser costoso aplicar el minmax al oponente. Reparemos en que si existe una desviación rentable durante la fase de castigo, entonces es rentable desviarse en  $w(1)$ , pues el menor pago durante la fase de castigo se alcanza en el primer periodo y cualquier desviación induce el mismo camino en lo que sigue del juego. Ahora bien, seguir el perfil en  $w(1)$  entrega un pago de  $v_i^*$ , mientras una desviación entrega un pago inmediato de  $\underline{v}_i^p < v_i^*$  (pues a cada jugador se le está aplicando el minmax) con un pago subsecuente de  $v_i^*$ , luego desviarse es estrictamente peor que seguir la estrategia propuesta.  $\square$

## 2.2. Caso de Información Incompleta.

### 2.2.1. Juego Repetido Canónico.

El juego *a repetir*.

Al igual que en el caso de información completa, el juego repetido se compone de una etapa repetida en cada periodo. En este caso se siguen considerando  $n$  jugadores, cada uno con un conjunto de posibles acciones  $A_i$ , subconjunto finito de  $\mathbb{R}$ .

La diferencia radica en que ahora la información acerca de las acciones tomadas en cada periodo no están disponibles para todos los jugadores: los jugadores sólo ven una señal pública  $y$  (además de la acción personal que eligieron). La señal  $y$  es obtenida desde un conjunto  $Y$ , finito. La probabilidad de que cada señal  $y$  ocurra depende del perfil de acciones  $a$  de la etapa, denotamos esta probabilidad como  $\rho(y|a)$ . Consideraremos que la función  $\rho : Y \times A \rightarrow [0, 1]$  es continua en las acciones, la extensión a un perfil de acciones mixtas  $\alpha$  es la natural.

Denotaremos los pagos del jugador  $i$  al elegir la acción  $a_i$  y después de ocurrir la señal  $y$  como  $u_i^*(y, a_i)$ . Con ese pago (el pago ex post), definimos el pago ex ante de la etapa como:

$$u_i(a) = \sum_{y \in Y} u_i^*(y, a_i) \rho(y|a)$$

Es decir, el pago ex ante corresponde a la esperanza, sobre las señales, de los pagos ex post.

**Ejemplo 2.2.1.** Consideremos nuevamente el *Dilema del Prisionero*, pero esta vez con monitoreo imperfecto. La imperfección en la información es capturada por dos señales  $\bar{y}$  y  $\underline{y}$ , cuya distribución de probabilidades esta dada por

$$\rho(\bar{y}|a) = \begin{cases} p, & a = EE \\ q, & a = AE, EA \\ r, & a = AA \end{cases}$$

con  $0 < q < p < 1$  y  $0 < r < p$ , entonces las acciones de los jugadores determinan si la señal es alta ( $\bar{y}$ ) o baja ( $\underline{y}$ ). Al interpretar el dilema del prisionero como un juego de cooperación, se tiene que una salida alta, una señal *buena*, tiene más probabilidad de ocurrir si es que ambos jugadores deciden cooperar. Luego la señal entrega información con ruido acerca de si los jugadores cooperaron o no. Los pagos ex post y ex ante se muestran en el cuadro 2.2.1.

	$\bar{y}$	$\underline{y}$	
E	$\frac{3-p-2q}{p-q}$	$-\frac{p+2q}{p-q}$	
A	$\frac{3(1-r)}{q-r}$	$-\frac{3r}{q-r}$	

	E	A
E	2,2	-1,3
A	3,-1	0,0

Cuadro 2.6: En la matriz de la izquierda se encuentran los pagos ex post para el dilema del prisionero repetido con información incompleta. A la derecha, la matriz de pagos ex ante implicada.

### El juego repetido.

A diferencia del caso con información completa ahora una historia *pública* corresponde sólo al vector de señales públicas, i.e.  $h^t := (y^0, y^1, \dots, y^{t-1})$ . El conjunto de

historias públicas queda entonces dado por:

$$\mathcal{H} := \bigcup_{t=0}^{\infty} Y^t, \quad Y^0 = \phi$$

Sin embargo, para un jugador  $i$  una historia incluye las señales realizadas y sus propias acciones, esto es,  $h_i^t := (a_i^0, y^0; a_i^1, y^1; \dots; a_i^{t-1}, y^{t-1})$ . Así, el conjunto de historias para el jugador  $i$  está definido como

$$\mathcal{H}_i := \bigcup_{t=0}^{\infty} (A_i \times Y)^t$$

Ahora ya es posible definir una estrategia pura para el jugador  $i$ :

$$\sigma_i : \mathcal{H}_i \rightarrow A_i$$

Luego una estrategia mixta, será una aleatorización sobre las estrategias puras, y una estrategia de comportamiento será una función

$$\sigma_i : \mathcal{H}_i \rightarrow \Delta(A_i).$$

### 2.2.2. Principio de Desviación Simple y Caracterización de equilibrios.

Antes de establecer el principio de desviación simple para el caso de información incompleta es importante hacer un alcance sobre la estructura recursiva del problema. Notemos que dado que cada jugador tiene información privada (sus acciones pasadas), entonces una historia pública  $h^t$  no induce un juego continuado estratégicamente idéntico al original, ya que una misma historia pública puede haber provenido de distintos perfiles de acciones, y dichos perfiles afectan el juego pues determinan la decisiones que cada jugador toma. El problema anterior nos lleva definir un espacio restringido de estrategias posibles para los jugadores, afortunadamente se demuestra que estratégicamente no se pierde generalidad al considerar este nuevo conjunto (Mailath y Samuelson [3, Lema 7.1.2]).

#### Estrategias y Equilibrios Públicos.

**Definición 2.2.1** (Estrategia de comportamiento pública). Una estrategia de comportamiento  $\sigma_i$  es *pública* si para cada período  $t$ , la estrategia sólo depende de la historia pública  $h^t \in \mathcal{H}$  y no en la información privada de  $i$ . Es decir,  $\sigma_i$  es pública si

$$\forall h_i^t, \hat{h}_i^t \in \mathcal{H}_i \text{ tq } y^\tau = \hat{y}^\tau \quad \forall \tau \leq t-1, \text{ se tiene } \sigma_i(h_i^t) = \sigma_i(\hat{h}_i^t)$$

Diremos que una estrategia es *privada* si no es pública.

Para cada estrategia de comportamiento público es posible, al igual que en 2.1.3, definir un autómata que la represente, para ello basta cambiar el dominio de la función de transición  $\tau$ . Antes  $\tau$  dependía de las posibles acciones  $A$ , ahora lo hace de  $Y$ , i.e.  $\tau : \mathcal{W} \times Y \rightarrow \mathcal{W}$ , donde  $\mathcal{W}$  es el conjunto de posibles estados del autómata. Así, tal como en la sección anterior es posible definir un equilibrio de Nash acorde al contexto:

**Definición 2.2.2** (Equilibrio Público Perfecto). Un perfil de estrategias públicas  $\sigma$  es un *equilibrio público perfecto* (PPE) si para cada historia pública  $h^t$ ,  $\sigma$  es un equilibrio de Nash perfecto en subjuegos del juego repetido. Esto es, para todo  $t$  y todo  $h^t \in Y^t$ ,  $\sigma|_{h^t}$  es un equilibrio de Nash del juego repetido.

Un PPE es *estricto* si cada jugador prefiere estrictamente  $\sigma_i$  frente a cualquier otra estrategia pública.

Dado  $\delta$  denotamos al conjunto de pagos alcanzables con equilibrios públicos perfectos como  $\mathcal{E}(\delta)$  (o simplemente  $\mathcal{E}$ ).

Admirablemente este equilibrio, al igual que antes, permite calcularse sin considerar *todas* las historias públicas  $h^t$ :

**Teorema 2.2.3** (Principio de la desviación simple). *Un perfil de estrategias puras  $\sigma$  es un PPE ssi no hay desviaciones simples rentables, es decir,  $\sigma$  es un PPE ssi para toda historia  $h^t \in Y^t$ ,  $\sigma(h^t)$  es un equilibrio de Nash para el juego definido en forma normal por las funciones de pago*

$$g_i(a) = (1 - \delta)u_i(a) + \delta \sum_{y \in Y} U_i(\sigma|_{h^t, y}) \rho(y | a), \text{ para } i = 1, \dots, n.$$

**Teorema 2.2.4** (Equivalencia con un juego estático). *Sea  $\sigma$  un perfil de estrategias públicas descrita por el autómata  $(\mathcal{W}, w^0, f, \tau)$ , y sea  $V_i(w)$  el valor descontado (normalizado) de un juego que comienza en el estado  $w$ . Entonces,  $\sigma$  es un PPE ssi  $\forall w \in \mathcal{W}$  accesible desde  $w^0$ ,  $f(w)$  es un equilibrio de Nash para el juego descrito por los pagos*

$$g_i^w(a) = (1 - \delta)u_i(a) + \delta \sum_{y \in Y} V_i(\tau(w, y)) \rho(y | a), \text{ para } i = 1, \dots, n.$$

**Ejemplo 2.2.2.** Tal como en el ejemplo 2.1.4, nos interesa estudiar para el Dilema del Prisionero (ahora con información incompleta) en qué caso es posible que ambos jugadores jueguen una estrategia distinta a jugar el equilibrio de Nash  $(A, A)$  en todos los periodos.

Analizaremos para que condiciones sobre las señales y el descuento, la estrategia de Tolerancia Cero es un equilibrio. Esta estrategia es similar a la presentada dentro del caso con información perfecta, pero ahora dependerá de las señales y no de las acciones. Más precisamente, esta estrategia tiene una representación en un autómata

con dos estados,  $\mathcal{W} = \{w_{EE}, w_{AA}\}$ , estado inicial  $w_{EE}$ , función de decisión  $f(w_{II}) = II$  ( $I = E, A$ ) y función de transición

$$\tau(w, y) = \begin{cases} w_{EE} & , \text{ si } w = w_{EE} \wedge y = \bar{y} \\ w_{AA} & , \text{ si no} \end{cases}$$

donde  $y$  es la señal del periodo anterior. Definiendo el autómata así, la estrategia corresponde a comenzar esforzándose y continuar haciéndolo mientras salga una buena señal  $\bar{y}$ , una vez que sale una mala señal, se juega eternamente el equilibrio de Nash  $(A, A)$ . El autómata se presenta en la figura 2.6.

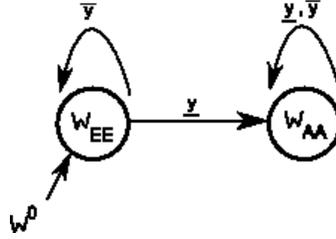


Figura 2.6: Autómata para la estrategia de Tolerancia Cero, en el caso de información incompleta.

Las condiciones buscadas las encontraremos utilizando el teorema 2.2.4, por lo cual necesitamos las utilidades descontadas  $V$  en cada estado. Omitiendo los subíndices, pues los pagos son simétricos, rápidamente obtenemos  $V(w_{AA}) = 0$  y para  $V(w_{EE})$  se tiene:

$$V(w_{EE}) = (1 - \delta)2 + \delta(pV(w_{EE}) + (1 - p)V(w_{AA})) \quad (2.5)$$

$$\Rightarrow V(w_{EE}) = \frac{2(1 - \delta)}{1 - \delta p} \quad (2.6)$$

Donde la primera línea viene de la definición de  $V$  y la segunda de ocupar el valor obtenido para  $V(w_{AA})$ . Ahora consideramos los pagos (omitiendo nuevamente los subíndices):

$$g^w(a) = (1 - \delta)u(a) + \delta \sum_{y \in \{\bar{y}, \bar{y}\bar{y}\}} V(\tau(w, y))\rho(y | a)$$

para  $w \in \mathcal{W}$ . Utilizando nuevamente la definición de  $V$  y los valores ya obtenidos podemos mostrar los pagos para  $w_{AA}$  y  $w_{EE}$ , en el cuadro 2.7.

En el estado  $w_{AA}$ ,  $f(w_{AA}) = AA$  es trivialmente un equilibrio de Nash. Por lo anterior, sólo basta chequear que en el estado  $w_{EE}$ ,  $f(w_{EE}) = EE$  es un equilibrio de Nash, para ello se debe cumplir

$$V(w_{EE}) \geq (1 - \delta)3 + \delta(qV(w_{EE}) + (1 - q)V(w_{AA})) \quad (2.7)$$

$$V(w_{EE}) \geq \frac{3(1 - \delta)}{1 - \delta q} \quad (2.8)$$

$w = w_{EE}$	
$a = EE$	$: g_E^w(a) = V(w_{EE})$
$a = EA, AE$	$: g_E^w(a) = (1 - \delta)3 + \delta(qV(w_{EE}) + (1 - q)V(w_{AA}))$
	$: g_A^w(a) = (1 - \delta)(-1) + \delta(qV(w_{EE}) + (1 - q)V(w_{AA}))$
$a = AA$	$: g_A^w(a) = (1 - \delta)0 + \delta(rV(w_{EE}) + (1 - r)V(w_{AA}))$
$w = w_{AA}$	
$a = AA$	$: g_A^w(a) = (1 - \delta)0 + \delta V(w_{AA})$
$a = EA, AE$	$: g_E^w(a) = (1 - \delta)(-1) + \delta V(w_{AA})$
	$: g_A^w(a) = (1 - \delta)0 + \delta V(w_{AA})$

Cuadro 2.7: Pagos en cada estado del juego estático equivalente para el Dilema del Prisionero con información incompleta.  $g_I$  indica el pago para quien jugó  $I$ .

$$\frac{2(1 - \delta)}{1 - \delta p} \geq \frac{3(1 - \delta)}{1 - \delta q} \quad (2.9)$$

$$\Leftrightarrow \delta(3p - 2q) \geq 1 \quad (2.10)$$

Luego, si  $\bar{y}$  es una señal lo suficientemente fidedigna, en el sentido de que

$$p > \frac{1}{3} + \frac{2}{3}q \quad \Leftrightarrow \quad (3p - 2q) > 1$$

obtenemos que la estrategia Tolerancia Cero es un equilibrio público perfecto para jugadores suficientemente pacientes, i.e. para  $\delta$  suficientemente grande.

Hagamos explícito que si la restricción de compatibilidad de incentivos (2.7) se cumple, el valor de equilibrio viene dado por (2.6). Ahora notemos que a medida que  $p \rightarrow 1$ , de manera que el perfil  $EE$  prácticamente garantiza  $\bar{y}$ , el valor del equilibrio se aproxima a 2, exactamente el mismo valor que en el caso de monitoreo perfecto. Más aún, si  $p = 1$  y  $q = 0$  de manera que  $\bar{y}$  es una señal inequívoca de que el oponente se esforzó, la estrategia Tolerancia Cero para el caso de información incompleta se ve como la misma estrategia que en el caso de información perfecta, más aún, en estas condiciones la restricción (2.10) es equivalente a la restricción  $\delta \geq 1/3$  encontrada en el ejemplo 2.1.4.

Notar que una importante característica de este juego, y en general de los juegos con información incompleta, es que puede que se castigue a un jugador (e.g. que ambos jugadores se salgan del perfil buscado) sin que necesariamente uno de los jugadores se haya aprovechado del otro, esto debido a que el ruido que posee la señal pública no permite ver con completa claridad las acciones de los jugadores.

### Construcción.

Utilizaremos el mismo método que fue utilizado para caracterizar los pagos en equilibrios para el caso de información completa (recordar 2.1.3).

**Definición 2.2.5** (Acción vinculable). Sea  $\mathcal{W}$  un subconjunto de  $\mathbb{R}^n$ , un perfil de acciones mixtas  $\alpha$  es *vinculable* en  $\mathcal{W}$  si existe una función  $\gamma : Y \rightarrow \mathcal{W}$  tal que,  $\forall i = 1, \dots, n, \forall a'_i \in A_i$  se tiene que

$$\begin{aligned} V_i(\alpha, y) &:= (1 - \delta)u_i(\alpha) + \delta \sum_{y \in Y} \gamma_i(y)\rho(y | \alpha) \\ &\geq (1 - \delta)u_i(a'_i, \alpha_{-i}) + \delta \sum_{y \in Y} \gamma_i(y)\rho(y | a'_i, \alpha_{-i}) \end{aligned}$$

Diremos en este caso que  $\gamma$  *vincula a*  $\alpha$  en  $\mathcal{W}$ .

Al igual que en 2.1.3, interpretaremos  $\gamma$  como la promesa de seguir el perfil  $\alpha$ , entonces pedirle a  $\alpha$  que sea vinculable, es básicamente el requerimiento de que sea compatible en incentivos.

**Definición 2.2.6** (Descomposición). Un vector de pagos  $v \in \mathbb{R}^n$  es *descomponible* en  $\mathcal{W}$  si existe un perfil de acciones mixtas  $\alpha \in \prod_{i=1}^n \Delta(A_i)$  vinculada por  $\gamma$  en  $\mathcal{W}$ , tal que

$$v_i = V_i(\alpha, \gamma)$$

**Definición 2.2.7** (Conjunto descomponible en otro,  $\mathcal{B}(\cdot)$ ). Dado un conjunto  $\mathcal{W} \in \mathbb{R}^n$  definimos el conjunto  $\mathcal{B}(\mathcal{W})$  como el mayor conjunto descomponible en  $\mathcal{W}$ , es decir

$$\mathcal{B}(\mathcal{W}) := \{v \in \mathbb{R}^n : v = V(\alpha, \gamma) \text{ para algún } \alpha \text{ vinculable en } \mathcal{W} \text{ por } \gamma\}$$

**Definición 2.2.8** (Conjuntos Auto-Generantes). Un conjunto de pagos  $\mathcal{W} \subset \mathbb{R}^n$  es *autogenerante* si  $\mathcal{W} \subset \mathcal{B}(\mathcal{W})$ .

Con las definiciones anteriores podemos ver que  $\mathcal{W}$  es autogenerante si todos los pagos que contiene son descomponibles en él, es decir, la definición es igual que en el contexto de información perfecta, pero notemos que  $\mathcal{B}(\mathcal{W})$  podría ser estrictamente más grande que  $\mathcal{W}$ .

**Teorema 2.2.9** (Conjuntos Auto-Generantes son Equilibrios). *Sea  $\mathcal{W}$  un conjunto acotado. Entonces si  $\mathcal{W}$  es autogenerante,  $\mathcal{B}(\mathcal{W}) \subset \mathcal{E}$ .*

Más aún, sólo para  $\mathcal{E}$  la desigualdad anterior se puede escribir en la otra dirección, se formaliza esto en el siguiente teorema.

**Teorema 2.2.10.**  $\mathcal{E} = \mathcal{B}(\mathcal{E})$ .

## Capítulo 3

# Juegos Repetidos a Tiempo Continuo con Información Incompleta.

En este capítulo introduciremos un modelo para juegos repetidos planteado en Sannikov [5]. Este modelo se asemeja a lo planteado anteriormente en que existe una etapa que se repite (aunque ahora se repite continuamente), también se asemeja en el sentido de que las acciones de los jugadores no son percibibles públicamente: sólo se perciben sus efectos sobre una señal.

Con el modelo que presentaremos podremos caracterizar el conjunto de pagos alcanzables en equilibrio, antes también lo habíamos logrado hacer pero ahora será posible explicitar cómo es la dinámica para un equilibrio óptimo, óptimo en el sentido de alcanzar el mayor conjunto de pagos posibles. En efecto, a pesar de que Abreu, Pearce y Stachetti [1] hacen manejable el problema de encontrar equilibrios, no queda claro como los pagos que los jugadores perciben se *mueven*, con el modelo que presentaremos podremos explicitar a través de una ecuación diferencial cómo es la frontera de los pagos alcanzables en equilibrio y cómo los jugadores se mueven en ella.

En los años recientes, el contexto continuo que presentaremos ha dado cabida a una serie de trabajos que lo estudian en casos particulares, nombraremos dos, en el primero se ha estudiado el problema del agente-principal en un contexto continuo (Sannikov, [6]) y el segundo corresponde a una variante del primero que incluye seguritización (DeMarzo y Sannikov, [2]). Incluso, recientemente se han estudiado contextos continuos, pero donde las acciones son tomadas en forma discreta (y frecuente) y la información, además de tener una componente como un movimiento Browniano, también tiene una componente como un proceso de Poisson (Skrzypacz y Sannikov, [7]).

### 3.1. El Modelo

Consideramos dos jugadores que participan en un juego repetido con información imperfecta. En todo momento  $t \in [0, \infty)$ , cada jugador  $i$  escoge una acción  $A_t^i$  desde un conjunto finito  $\mathcal{A}^i$ , notemos que esto es similar a lo presentado en el capítulo anterior, salvo que antes las acciones se elegían en  $t \in \{0, 1, \dots\}$ . Tal como en 2.2, los jugadores no observan directamente la acción del otro, en cambio, ven señales públicas  $d$ -dimensionales:

$$X_t = \int_0^t \mu(A_s) ds + \Sigma Z_t$$

Donde  $Z$  es un movimiento Browniano  $d$ -dimensional,  $\Sigma$  es una matriz simétrica definida positiva y  $\mu : \mathcal{A}^1 \times \mathcal{A}^2 \rightarrow \mathbb{R}^d$  es una función de *drift*. Es decir, los jugadores ven alguna función de sus acciones,  $\mu$ , más un cierto ruido o *shock* exógeno. El ruido de cada componente de la señal, puede estar o no correlacionado con el ruido de las demás componentes, por ejemplo, si  $\Sigma$  es diagonal, cada componente es independiente de las otras; en la misma línea se podrá concluir que una componente tiene más o menos varianza que las otras dependiendo de la estructura y valores de  $\Sigma$ . La información se representa por la filtración  $\{\mathcal{F}_t\}$ , que corresponde a la filtración generada por el movimiento Browniano<sup>1</sup>.

Diremos que  $A^i = \{A_t^i\}$  es una estrategia pública para el jugador  $i$ , si  $A^i$  es un proceso progresivamente medible<sup>2</sup> con respecto a  $\{\mathcal{F}_t\}$ , en concreto esto indica que la estrategia dirá qué jugar dependiendo sólo de lo que ya ocurrió (*no podemos medir  $A_t^i$ , si no se ha alcanzado  $t$* ). Dado un perfil de estrategias públicas  $A = (A^1, A^2)$ , el pago estocástico descontado (normalizado) para el jugador  $i$  esta dado por:

$$\begin{aligned} r \int_0^\infty e^{-rt} (c_i(A_t^i) + b_i(A_t^i)^\top \mu(A_t)) dt &+ r \int_0^\infty e^{-rt} b_i(A_t^i)^\top \Sigma dZ_t \\ &= r \int_0^\infty e^{-rt} (c_i(A_t^i) + b_i(A_t^i)^\top \mu(A_t)) dt + b_i(A_t^i)^\top \Sigma dZ_t \end{aligned}$$

<sup>3</sup>Para algunas funciones  $c_i : \mathcal{A}^i \rightarrow \mathbb{R}$ ,  $b_i : \mathcal{A}^i \rightarrow \mathbb{R}^d$ , con  $r$  siendo la tasa de descuento común para ambos jugadores. Es decir, a cada jugador su acción lo beneficia a través de  $c_i$  y además cada jugador es susceptible a la señal pública a través de  $b_i(A_t^i)$ . Denotamos por  $g_i$ , la esperanza del pago instantáneo para el jugador  $i$  en el momento  $t$ :

$$g_i(A_t) = c_i(A_t^i) + b_i(A_t^i)^\top \mu(A_t)$$

<sup>1</sup>En el trabajo original también se considera el caso en que la filtración es estrictamente mayor que la generada por  $Z$ .

<sup>2</sup>Un proceso  $X_t$  es progresivamente medible con respecto a una filtración  $\{\mathcal{F}_t\}$  si para cada  $t$   $X : [0, t] \times \Omega \rightarrow \mathbb{R}$  es  $\mathcal{B}(\mathbb{R}) \otimes \mathcal{F}_t$ -medible.

<sup>3</sup>En la segunda ecuación hemos mezclado integrales estocásticas y de Lebesgue, para acortar la notación. Además notamos por  $b^\top$  al traspuesto del vector  $b$ .

Haremos dos supuestos,: el primero tiene relación con que los dos jugadores son *estadísticamente distinguibles* a partir de la señal pública, mientras el segundo es más bien técnico. Para establecer estos supuestos, introduciremos la siguiente notación:

**Definición 3.1.1** (Matriz de señales comparadas). Dado un perfil de acciones  $a = (a_1, a_2) \in \mathcal{A}^1 \times \mathcal{A}^2$  definimos la *matriz de señales comparadas* para el jugador  $i$ ,  $M_i(a)$ , como la matriz de dimensiones  $d \times (|\mathcal{A}^i| - 1)$  con columnas  $\mu(a'_i, a_{-i}) - \mu(a)$ ,  $a'_i \neq a_i$ .

**Supuesto 1** (Identificabilidad de a pares). Todos los perfiles de acciones  $a \in \mathcal{A}^1 \times \mathcal{A}^2$  son *identificables de a pares*, es decir, el subespacio vectorial generado por las columnas de la matriz  $M_1(a)$  y el subespacio vectorial generado por las columnas de la matriz  $M_2(a)$  se intersectan sólo en el origen.

**Ejemplo 3.1.1.** Supongamos la señal es bidimensional, y que  $\mu(a_1, a_2)^T = (\frac{a_1+a_2}{2}, \frac{a_1+a_2}{2})$ . Vemos que si  $\mu$  se define así, la señal pública  $X$  no nos permitirá distinguir entre lo que hace un jugador y el otro, por tanto no cumplirá el supuesto de indentificabilidad de a pares. En efecto, tomemos  $\mathcal{A}_1 = \mathcal{A}_2 = \{1, 2\}$ , luego:

$$M_1((1, 1)) = \begin{pmatrix} -0,5 \\ -0,5 \end{pmatrix}, \quad M_2((1, 1)) = \begin{pmatrix} -0,5 \\ -0,5 \end{pmatrix}$$

Con los valores anteriores, es claro que los subespacios generados por  $M_1((1, 1))$  y  $M_2((1, 1))$  son los mismos.

**Supuesto 2.** Al menos una de las siguientes afirmaciones es verdadera:

1. Para todo  $i = 1, 2$  y  $a_i \in \mathcal{A}^i$ , la mejor respuesta estática para  $a_i$  es única, esto es, el conjunto  $\operatorname{argmax}_{a_j} g(a_i, a_j)$  es un singleton.
2. Para todo  $a \in \mathcal{A}$ , los generados por  $M_1(a)$  y  $M_2(a)$  son ortogonales.

**Definición 3.1.2** (Esperanza del Pago Descontado). El *pago descontado esperado después de una historia pública al momento  $t$*  (o *valor de continuar*) está definido como:

$$\begin{aligned} W_i^t(A) &= \mathbb{E}_t \left[ \int_t^\infty e^{-r(s-t)} (c_i(A_s^i) + b_i(A_s^i)^T \mu(A_s) ds + b_i(A_s^i)^T \Sigma dZ_s) | A_s, s \in [t, \infty) \right] \\ &= \mathbb{E}_t \left[ \int_t^\infty e^{-r(s-t)} g_i(A_s) ds | A_s, s \in [t, \infty) \right] \end{aligned}$$

Esta definición, tomando la esperanza condicionada en la información/historia pública al tiempo  $t$ , hace hincapié en que las acciones afectan los pagos *explícitamente* a través de  $g_i(A_s)$  y, además, *indirectamente* pues  $\{A_s : s \in [t, \infty)\}$  determina la distribución de probabilidad sobre los futuros caminos de  $X$ .

**Definición 3.1.3** (Equilibrio Público Perfecto, PPE). Un perfil de acciones  $A$  es un *equilibrio público perfecto* (PPE, por su siglas en inglés) ssi para  $i = 1, 2$ ,  $A^i$  maximiza el pago descontado esperado para  $i$ , dada la estrategia  $A^j$  del jugador  $j$  y cualquier historia pública.

Como antes, dado una tasa de descuento  $r$ , denotemos por  $\mathcal{E}(r)$  el conjunto de los pagos alcanzables a través de equilibrios públicos perfectos, a veces, le llamaremos  $\mathcal{E}(r, \Sigma)$  para hacer explícita la dependencia en el parámetro  $\Sigma$ .

### Conjuntos Importantes y Relación con el Capítulo I

En esta sección hacemos la conexión entre este modelo y lo presentado en el capítulo 1. La etapa del juego,  $G$ , tiene un conjunto de dos jugadores  $P = \{1, 2\}$ , un conjunto de acciones  $\mathcal{A}^i$  para cada jugador y finalmente funciones de pago  $g_i$ .

Denotamos al conjunto de perfiles de la etapa como  $\mathcal{A} = \mathcal{A}_1 \times \mathcal{A}_2$  y el conjunto de equilibrios de Nash en estrategias puras como  $\mathcal{A}^P$ . Denotamos por  $\mathcal{V}$  y  $\mathcal{N}$  las envolturas convexas de los pagos posibles<sup>4</sup> y de los equilibrios de Nash, es decir,

$$\mathcal{N} := \text{co}\{g(a) | a \in \mathcal{A}^P\}, \quad \mathcal{V} := \text{co}\{g(a) | a \in \mathcal{A}\}$$

El pago minmax en estrategias puras para el jugador  $i$  es

$$\underline{v}_i := \min_{a_j} \max_{a_i} g_i(a_i, a_j)$$

Recordamos que el jugador  $i$  puede garantizarse a si mismo su pago minmax en estrategias puras. Definimos  $\mathcal{V}^*$  como el subconjunto de  $\mathcal{V}$  en el cual cada jugador alcanza al menos su pago minmax<sup>5</sup>, es decir  $\mathcal{V}^*$  es el conjunto de pagos racionales y queda dado por

$$\mathcal{V}^* = \{v \in \mathcal{V} | v_i \geq \underline{v}_i, \forall i = 1, 2\}.$$

**Ejemplo 3.1.2.** Para ejemplificar las definiciones anteriores, podemos considerar un caso de bienes complementarios con colusión. Definiremos las señales a partir de los precios (esperados) de cada bien, para ello definamos:

$$\begin{aligned} \mu_1(a) &:= P_1(a) = 24 - 2a_1 + a_2 \\ \mu_2(a) &:= P_2(a) = 24 - 2a_2 + a_1 \end{aligned}$$

A partir de lo anterior definimos la señales como

$$X_t(A) = \int_0^t (P_1(A_s) \ P_2(A_s))^T ds + \int_0^t dZ_s$$

<sup>4</sup> $\mathcal{V} = \mathcal{F}^\dagger$ , según la notación del capítulo anterior.

<sup>5</sup>Así,  $\mathcal{V}^* = \mathcal{F}^{\dagger*}$ , según la notación del capítulo anterior.

donde  $A = (A_t^1, A_t^2)_t$  es un perfil de acciones y  $Z_t$  es un movimiento Browniano estándar 2 dimensional. Es decir, las señales corresponden al precio más un shock exógeno. Inspirados en esto, diremos que el ingreso para cada jugador será la cantidad producida por el precio *real*, o sea, el pago descontado para el jugador  $i$  viene dado por

$$r \int_0^t e^{-rs} A_s^i dX_s = r \int_0^t e^{-rs} A_s P_i(A_s) ds + r \int_0^t e^{-rs} A_s \Sigma dZ_s$$

donde  $r$  es la tasa de descuento común para ambos jugadores. Así, el pago instantáneo esperado se define como  $g_i(A_t) = A_t^i (24 - 2A_t^i + A_t^j)$ .

Consideraremos  $\mathcal{A}^1 = \mathcal{A}^2 = \{8, 9, 12\}$ , lo que entrega la matriz de pagos instantáneos esperados mostrada en el cuadro 3.1. Se observa que el único Nash es  $(8, 8)$ , i.e.  $\mathcal{A}^P = \{(8, 8)\}$ , alcanzando un pago de  $(128, 128)$ , es decir  $\mathcal{N} = \{(128, 128)\}$ . En el mismo cuadro es posible observar que el pago minmax para cada jugador es 128, el conjunto  $\mathcal{V}^*$  asociado se muestra de color verde oscuro en la figura 3.1, para finalizar, al costado de dicha figura, se muestra el conjunto  $\mathcal{V}$ .

		Jugador 2		
		8	9	12
Jugador 1	8	(128,128)	(136,126)	(160,96)
	9	(126,136)	(135,135)	(162,108)
	12	(96,160)	(162,108)	(144,144)

Cuadro 3.1: Pagos instantáneos esperados para el juego de bienes complementarios.

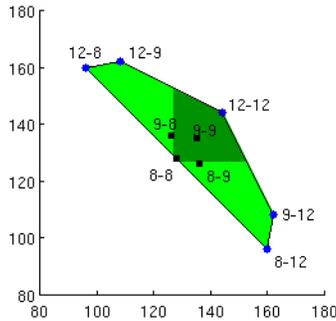


Figura 3.1:  $\mathcal{V}^*$  se obtiene al asegurar a cada jugador su pago mínimo, en este caso 128.

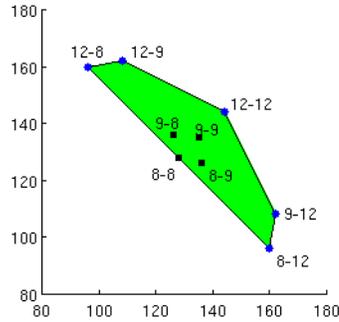


Figura 3.2: Pago instantáneo esperado de las acciones en el juego de bienes complementarios. Los puntos cuadrados pertenecen al interior de  $\mathcal{V}$ , los circulares pertenecen a la frontera de dicho conjunto.

## 3.2. Caracterización de Equilibrios Públicos Perfectos

Es posible encontrar otra representación para el valor de continuar, esta vez en términos explícitos de lo que ya ocurrió hasta el momento  $t$ , en contraposición a la definición a través de esperanzas condicionales.

**Teorema 3.2.1** (Representación del Valor de Continuar). *Un proceso estocástico  $W_t^i$  es el valor de continuar,  $W_t^i(A)$ , del jugador  $i$  bajo el perfil de estrategias  $A$  ssi existen procesos  $\beta^i = (\beta^{i1} \beta^{i2} \dots \beta^{id})$  en  $\mathcal{L}^{*6}$  tal que para todo  $t \geq 0$ ,  $W_t^i$  pertenece a  $\mathcal{V}^*$  y satisface*

$$W_t^i = W_0^i + r \int_0^t (W_s^i - g_i(A_s)) ds + r \int_0^t \beta_s^i (dX_s - \mu(A_s)ds) \quad (3.1)$$

Podemos reescribir (3.1) como:

$$dW_s^i = r (W_s^i - g_i(A_s)) ds + r \beta_s^i (dX_s - \mu(A_s)ds) \quad (3.2)$$

Interpretamos  $\beta_t^i$  como una medida de la exposición de  $W_t^i$  a la *volatilidad*  $\Sigma dX_s - \mu(A_s)ds$ . Por otro lado, podemos pensar en  $\beta_t^i$  como una medida del castigo (o premio) entregado por desviarse desde la señal  $dX_s$ , en este sentido, los procesos  $\beta$  permiten crear incentivos más allá de los dados por los pagos.

**Teorema 3.2.2** (Compatibilidad de Incentivos). *La estrategia  $A^i$  del jugador  $i$  es óptima, para todo tiempo, en respuesta a la estrategia  $A^j$  del jugador  $j$  ssi la condición de compatibilidad de incentivos:*

$$\forall a'_i \in \mathcal{A}^i, \quad g_i(A_t) + \beta_t^i \mu(A_t) \geq g_i(a'_i, A_t^j) + \beta_t^i \mu(a'_i, A_t^j) \quad (3.3)$$

*se cumple para todo  $t$ . Luego, un par de estrategias  $A$  es un PPE ssi (3.3) se cumple para ambos jugadores*

El teorema 3.2.2 es el análogo continuo del Principio de Desviación Simple que habíamos visto en los capítulos anteriores (2.1.2, 2.2.2). En efecto, si el jugador se desvía a una acción alternativa  $a'_i$  en el tiempo  $t$ , la esperanza del pago instantáneo varía en  $g_i(a'_i, A_t^j) - g_i(A_t)$  y la deriva (o *drift*) de  $X$  cambia por  $\mu(a'_i, A_t^j) - \mu(A_t)$ . Como  $\beta_i$  es la sensibilidad que el jugador tiene a  $dX$ , el valor de continuar tras la nueva acción cambia en  $\beta_t^i (\mu(a'_i, A_t^j) - \mu(A_t))$ . Así las cosas, si la compatibilidad de incentivos se tiene, la desviación instantánea es no rentable pues se tiene que

$$g_i(a'_i, A_t^j) - g_i(A_t) + \beta_i (\mu(a'_i, A_t^j) - \mu(A_t)) \leq 0$$

---

<sup>6</sup> $\mathcal{L}^*$  es el conjunto de todos los procesos  $B_t$  tal que  $\mathbb{E} \left[ \int_0^T B_t^T dt \right] < \infty$  para todo  $T < \infty$ , si  $B_t$  es multidimensional, consideramos la definición anterior para cada componente.

Notar que  $\beta^i$  es un vector fila.

Se podría pensar que desviaciones instantáneas se pierden dentro de la integral, pero dado que las desviaciones influyen en la distribución de probabilidad de los pagos descontados esperados se muestra que estas desviaciones sí se arrastran a lo largo del tiempo y con ello sí se afecta la integral (Sannikov, [5, proposición 2]).

**Definición 3.2.3** (Acción vinculable). Un perfil de acciones  $a \in \mathcal{A}$  será *vinculable* si existe una matriz  $B$  con dimensiones  $2 \times d$  con la forma:

$$B = \begin{bmatrix} \beta^1 \\ \beta^2 \end{bmatrix} = \begin{bmatrix} \beta^{11} & \dots & \beta^{1d} \\ \beta^{21} & \dots & \beta^{2d} \end{bmatrix}$$

tal que se tiene:

$$\forall i = 1, 2 \forall a'_i \in \mathcal{A}^i, \quad g_i(a) + \beta^i \mu(a) \geq g_i(a'_i, a_j) + \beta^i \mu(a'_i, a_j)$$

En este caso diremos que  $B$  vincula a  $a$ .

Podemos resumir los dos resultados obtenidos en el siguiente teorema:

**Teorema 3.2.4** (Caracterización de PPE). *En cualquier PPE  $A = (A_t^1, A_t^2)_{t \in \mathbb{R}^+}$ , el par de valores de continuar  $W = W(A)$  es un proceso en  $\mathcal{V}^*$  que satisface*

$$W_t = W_0 + r \int_0^t (W_s - g(A_s)) ds + r \int_0^t B_s (dX_s - \mu(A_s) ds) \quad (3.4)$$

Donde  $B$  es un  $2 \times d$  proceso en  $\mathcal{L}^*$  tal que  $B_t$  vincula a  $A_t$  para todo  $t$ .

Conversamente, si  $W$  es un proceso bidimensional en  $\mathcal{V}^*$ , que satisface (3.4) para  $A$  y  $B$  como arriba, entonces  $W$  es el par de valores de continuar en el equilibrio perfecto público  $A$ .

Notemos que si existe un equilibrio de Nash para el juego con pagos  $g$ , entonces, la estrategia que juega en cada tiempo ese equilibrio de Nash es un PPE con matriz  $B$  nula.

Siguiendo con las analogías con lo realizado para tiempo discreto, en el contexto de tiempo continuo también es posible caracterizar el conjunto de pagos alcanzable en PPE a través de la noción de *autogeneración*.

**Definición 3.2.5** (Conjunto Autogenerante). Un conjunto  $\mathcal{W} \subset \mathbb{R}^2$  es *autogenerante* ssi para todo punto  $W_0 \in \mathcal{W}$  existe un proceso  $W$  que comienza en  $W_0$ , se queda siempre en  $\mathcal{W}$  y satisface (3.4) para algunos procesos  $A$ ,  $B$  que cumplen el teorema 3.2.4.

**Corolario 3.2.6.**  $\mathcal{E}(r)$  es el conjunto autogenerante acotado más grande.

### 3.3. Equilibrios Públicos Perfectos con valores extremos

**Definición 3.3.1** (Vector Vinculante). Un vector de  $\phi \in \mathbb{R}^d$  vinculará en la tangente  $T = (t_1, t_2)^T$  a  $a \in \mathcal{A}$  si la matriz

$$B = T\phi^T = \begin{bmatrix} t_1\phi_1 \cdots t_1\phi_d \\ t_2\phi_1 \cdots t_2\phi_d \end{bmatrix}$$

vincula a  $a$ . De todos los vectores  $\phi$  que vinculan a  $a$  en la tangente  $T$ , sea  $\phi(a, T)$  el de menor norma, si  $a$  no es vinculable en la tangente  $T$ , usamos  $\phi(a, T) = \infty$ .

Con la definición anterior es posible dar una importante propiedad que posee el borde de  $\mathcal{E}(r)$ , los métodos de resolución que ocuparemos en los siguientes capítulos dependen de este resultado. Denotaremos el borde de  $\mathcal{E}(r)$  como  $\partial\mathcal{E}(r)$ .

**Teorema 3.3.2** (Caracterización del borde de  $\mathcal{E}(r)$ ).  $\mathcal{E}(r)$  es el subconjunto cerrado de  $\mathcal{V}^*$  con curvatura

$$\kappa(w) = \max_{a \in \mathcal{A} \setminus \mathcal{A}^N} \frac{2N(w)^T(g(a) - w)}{r|\Sigma\phi(a, T(w))|^2} \quad (3.5)$$

en todos los puntos  $w \notin \mathcal{N}$  de su borde, donde  $T(w)$  y  $N(w)$  son los vectores tangente y normal (exterior) en  $w$ . Llamamos a (3.5) la ecuación de optimalidad.

Debido a muchos tecnicismos involucrados en la demostración de este teorema, en las subsecciones 3.3.1 y 3.3.2 daremos sólo una intuición acerca de por qué la curvatura tiene esa forma funcional. En 3.3.3, mostramos el resultado final para PPE con valores extremos que vincula el teorema anterior con los otros teoremas que se presentaron.

#### 3.3.1. El valor de continuar tiene una trayectoria curva

Consideremos un proceso

$$dW_t = r(W_t - g(A_t))dt + rT\phi_t^T\Sigma dZ_t \quad (3.6)$$

con  $W_t$  en una curva convexa  $\mathcal{C}$  en el momento  $t$ , donde  $T$  es el vector tangente a  $\mathcal{C}$  en el punto  $W_t$ . Sea  $N$  la normal exterior unitaria de  $\mathcal{C}$  en  $W_t$ . Consideremos  $(x, f(x))$  una parametrización de  $\mathcal{C}$  en coordenadas tangencial y normal -notar que  $(T, N)$  se encuentran en el origen del plano-, con ellas definamos

$$D_{t+\varepsilon} = N^T W_{t+\varepsilon} - f(T^T W_{t+\varepsilon})$$

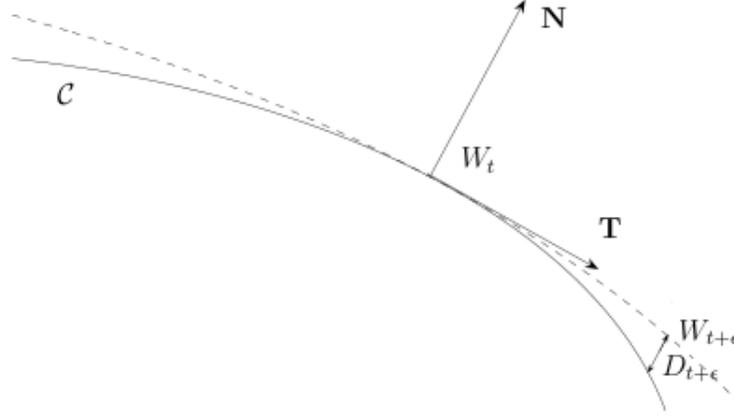


Figura 3.3: La línea punteada muestra al proceso  $W_t$  mientras la línea sólida muestra a  $\mathcal{C}$ . Notar que el vector  $N$  y la línea que identifica a  $D_{t+\epsilon}$  son paralelas. Aquí los ejes  $(T, N)$  están trasladados desde el origen.

como una medida de la distancia desde  $W_{t+\epsilon}$  a la curva  $\mathcal{C}$ , se puede ver gráficamente esta definición en 3.3<sup>7</sup>.

El objetivo es demostrar que para que  $W_t$  se mantenga en  $\mathcal{C}$ , o equivalentemente para que  $D_t \equiv 0$ , es necesario que la curvatura,  $\kappa$ , de  $\mathcal{C}$  tenga la forma planteada en (3.5). Notemos que por definición  $D_0 = 0$  por lo cual para mostrar que  $D_t$  sea siempre 0, calcularemos que ecuación diferencial cumple  $D_t$  y la igualaremos a 0.

Comencemos notando que como  $\mathcal{C}$  es convexa,  $\mathcal{C}$  alcanza su máximo -para los ejes  $(T, N)$ - exactamente en el punto que define a  $T$ , es decir en el punto  $W_t$ , que en nuestras coordenadas es el punto  $(T^T W_t, f(T^T W_t))$ ; con lo anterior es fácil ver que  $f'(T^T W_t) = 0$  y que  $f''(T^T W_t) < 0$ , con ello la curvatura en  $W_t$  queda dada por

$$\kappa = \frac{|f''(T^T W_t)|}{(1 + f'(T^T W_t)^2)^{3/2}} = -f''(T^T W_t)$$

Ahora proyectemos (3.6) en el eje tangente, lo que da

$$d(T^T W_t) = rT^T (W_t - g(A_t)) dt + r\phi_t^T \Sigma dZ_t \quad (3.7)$$

Usando el lema de Itô, se concluye

$$d(T^T W_t) = f'(T^T W_t)d(T^T W_t) + f''(T^T W_t)\frac{r^2|\Sigma\phi_t|^2}{2}d\langle Z \rangle_t \quad (3.8)$$

$$= -\kappa\frac{r^2|\Sigma\phi_t|^2}{2} \quad (3.9)$$

<sup>7</sup>Éstas y las siguientes imágenes se deben a Sannikov [5, sección 6].

Análogamente, proyectando (3.6) en el eje normal se obtiene

$$d(N^T W_t) = rN^T(W_t - g(A_t))dt \quad (3.10)$$

Uniando (3.9) y (3.10) se obtiene la ecuación buscada

$$dD_t = d(N^T W_t - f(T^T W_t)) = \left( rN^T(W_t - g(A_t)) + \kappa \frac{r^2 |\Sigma \phi_t|^2}{2} \right) dt$$

Si imponemos que  $dD_t \equiv 0$ , se concluye que

$$\kappa = \frac{2N^T(W_t - g(A_t))}{r|\Sigma \phi_t|^2}.$$

### 3.3.2. Ecuación de optimalidad y puntos extremos de $\mathcal{E}$

En esta subsección argumentamos que la curvatura planteada anteriormente efectivamente corresponde a la curvatura del borde  $\mathcal{E}$  y además, esta curvatura se alcanza maximizando sobre los perfiles de acciones que no son equilibrios de Nash del juego estático, tal como se plantea en el teorema 3.3.2.

Primero demostremos que se tiene

$$k(W_t) \leq \frac{2N^T(g(A_t) - W_t)}{r|\Sigma \phi_t|^2} \quad (3.11)$$

donde  $k(W_t)$  es la curvatura de  $\partial\mathcal{E}$  en el punto  $W_t$  y  $W_t$  es el valor de continuar para la estrategia  $(A_t)_{t \in \mathbb{R}_+}$ , es decir,  $W_t$  cumple (3.6) con la salvedad ahora de que  $\phi_t$  vincula a  $A_t$  en la tangente  $T$ . Notemos que si  $A_t \in \mathcal{A}^N$ , entonces  $\phi_t$  puede ser 0 pues la compatibilidad de incentivos (3.3) se obtiene directamente usando  $\phi = 0$  y el hecho de que el perfil  $A_t$  domina a los otros posibles perfiles; con el mismo argumento es posible demostrar que si  $A_t \notin \mathcal{A}^N$ , entonces  $\phi_t \neq 0$ . Con lo anterior en mente, asumamos que  $A_t \notin \mathcal{A}^N$ , así la cota planteada en 3.11 esta bien definida.

Supongamos que la curvatura de  $\partial\mathcal{E}$  en el punto  $W_t$  es mayor que el lado derecho de 3.11, luego como las trayectorias del valor de continuar tendrían una curvatura menor, inmediatamente escaparían de  $\mathcal{E}$ . Este fenómeno se muestra en la figura 3.4.

Ahora veamos una cota inferior para la curvatura  $\partial\mathcal{E}$  pero en cualquier punto y cualquier estrategia:

$$k(w) \geq \frac{2N^T(g(a) - w)}{r|\Sigma \phi(a, T)|^2} \quad (3.12)$$

donde  $k(w)$  es la curvatura de  $\partial\mathcal{E}$  en el punto  $w \in \partial\mathcal{E} \setminus \mathcal{N}$ ,  $a \notin \mathcal{A}^N$  y  $N$  junto a  $T$ , son los vectores unitarios normal y tangente en el punto  $w$ , respectivamente. Si la desigualdad en (3.12) falla, el valor de continuar asociado a este perfil de acciones  $a$

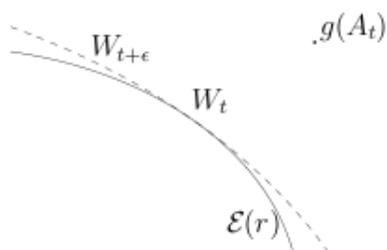


Figura 3.4: La línea punteada muestra al proceso  $W$  mientras la línea sólida muestra el borde de  $\mathcal{E}$ . Como el proceso  $W$  tiene menos curvatura que  $\partial\mathcal{E}$  en el punto  $W_t$ ,  $W$  se escapa de  $\mathcal{E}$ .

vinculado por  $\phi(a, T)$  (recordar (3.6)) tendría una mayor curvatura que  $\partial\mathcal{E}$  en ese punto. Intuitivamente, esto quiere decir que es posible llegar a  $w$  desde el interior de  $\mathcal{E}$ , tal como se muestra en la figura 3.5. Es decir, es posible garantizar  $w$  con valores de continuación en el interior de  $\mathcal{E}$ . Entonces si se mueven los valores de continuar que llevan a  $w$  en la dirección normal, se podrá llegar a un par  $w + \epsilon N$  afuera de  $\mathcal{E}$ , lo que es una contradicción pues  $\mathcal{E}$  es el conjunto más grande de pagos alcanzables en equilibrios. Notemos que el argumento dado aquí es de optimalidad, a diferencia del argumento dado para demostrar la desigualdad (3.11), que corresponde a uno de factibilidad.

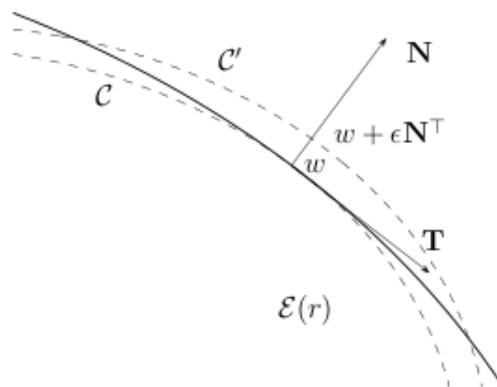


Figura 3.5: La línea punteada inferior muestra una curva desde la cual se podría llegar a  $w$  a través de un proceso de difusión como en (3.6), la otra línea punteada, es el desplazamiento de la anterior en la dirección normal. La línea sólida muestra el borde de  $\mathcal{E}$ .

Ya que se tienen las 2 desigualdes anteriores, (3.11) y (3.12), es posible concluir

que si  $\phi_t = \phi(A_t, T)$ , entonces (3.11) debe alcanzarse con igualdad, si no,  $w = W_t$  y  $a = A_t$  no cumplirían (3.12). Usando lo anterior, se tiene que al tomar máximo sobre todos los perfiles que no corresponden a equilibrios de Nash estáticos en (3.12) se alcanza la igualdad.

Con todo lo anterior, es claro que la ecuación de optimalidad, (3.5), se debe cumplir.

### 3.3.3. Conclusión sobre PPE con valores extremos

Podemos resumir lo discutido hasta ahora en el siguiente teorema, donde se consolidan los distintos teoremas que se vieron en este capítulo.

**Teorema 3.3.3** (PPE con valores extremos). *Recordemos que  $\partial\mathcal{E}(r)$  indica el borde de  $\mathcal{E}(r)$  y definamos por*

$$a : \partial\mathcal{E} \setminus \mathcal{N} \rightarrow \mathcal{A} \setminus \mathcal{A}^N$$

*el perfil de acciones que maximiza la ecuación (3.5). Cualquier valor  $W_0 \in \partial\mathcal{E}(r) \setminus \mathcal{N}$  es alcanzado por un PPE con las siguientes características:*

1. *El par de valores de continuar  $(W_t^1, W_t^2)_t$  bajo este PPE satisface la ecuación diferencial estocástica:*

$$W_t = W_0 + \int_0^t r(W_s - g(A_s))ds + \int_0^t rT(W_s) \phi^T(A_s, T(W_s)) \underbrace{(dX_s - \mu(A_s)ds)}_{\Sigma dZ_s} \quad (3.13)$$

*hasta el tiempo  $\tau$  cuando  $W_t$  llega al conjunto  $\mathcal{N}$ .*

2. *Para  $t < \tau$ , los jugadores toman acciones  $A_t = a(W_t)$ .*
3. *Si  $\partial\mathcal{E}(r) \cap \mathcal{N} = \emptyset$ , entonces  $\tau = \infty$ .*
4. *Si  $\tau < \infty$ , después de  $\tau$  los jugadores juegan un equilibrio de Nash estático con valor  $W_\tau$ . En este caso, los jugadores son absorbidos en un equilibrio de Nash estático con probabilidad 1.*

Con el teorema anterior es posible describir la dinámica para alcanzar el punto  $W_0 \in \partial\mathcal{E} \setminus \mathcal{N}$ . Apenas comienza el juego el valor de continuar empieza a moverse en el borde de  $\mathcal{E}$  según la ecuación de difusión (3.13). El punto  $W_t$  juega el rol de variable de estado, en el sentido que determina las acciones que los jugadores toman en cualquier instante y además determina la ley según la cual él mismo evoluciona, basado en las observaciones de la señal  $X$ . Si hay pagos alcanzables en equilibrio de Nash (estáticos) dentro de  $\partial\mathcal{E}$ , entonces con probabilidad 1 los jugadores alcanzarán uno, cuando eso pasa, los jugadores eligen seguir jugando el perfil de equilibrio

estático asociado para siempre. Si todos los equilibrios estáticos están al interior de  $\mathcal{E}$ , entonces el movimiento de los valores de continuar nunca se termina.

Los perfiles de acciones en equilibrio  $A_t$  vienen de la ecuación de optimalidad (3.5). El objetivo de esto es que para que se alcance el mayor conjunto de pagos se debe tener en cuenta el intercambio entre la extremalidad de los pagos y los incentivos necesarios para poder vincularlos. La extremalidad del pago, es medida por la ganancia en la dirección normal (recordar el numerador de (3.5)). Los incentivos son medidos por la varianza tangente instantánea de los valores de continuar (revisar el denominador de (3.5)). Recordemos que el perfil de acciones que entrega (3.5) vienen de alcanzar el máximo en esa ecuación y que sólo esos perfiles permiten mantenerse en  $\partial\mathcal{E}$ , pues perfiles que no alcancen el máximo, llevan los valores de continuar afuera de  $\mathcal{E}$ .

# Capítulo 4

## Cálculo Numérico de Equilibrios para el modelo a Tiempo Continuo

### 4.1. Marco Teórico

La ecuación de optimalidad (3.5) no es posible de abordar directamente, para poder usarla en el cálculo de  $\mathcal{E}(r)$  es necesario recordar que, dado una curva  $\mathcal{C}^1$ , la curvatura  $\kappa$  cumple:

$$\kappa(\theta) = d\theta/dl$$

donde  $l$  es la longitud de arco de la curva y  $\theta$  es el ángulo tangencial (es decir,  $T(\theta) = (-\sin(\theta), \cos(\theta))^T$  y  $N(\theta) = (\cos(\theta), \sin(\theta))^T$ ). Con lo anterior, siendo ligeros en la notación, concluimos

$$\frac{dw(\theta)}{dl} = T(\theta) \tag{4.1}$$

$$\frac{dw(\theta)}{d\theta} \frac{d\theta}{dl} = T(\theta) \tag{4.2}$$

$$\frac{dw(\theta)}{d\theta} = \frac{T(\theta)}{\kappa(\theta)} \tag{4.3}$$

Ahora, la ecuación (4.3) sí es resoluble numéricamente comenzando en cualquier condición inicial  $(\theta, w) \in [0, 2\pi) \times \mathcal{V}^*$ , tratando  $w$  como función de  $\theta$  y usando la ecuación de optimalidad:

$$\kappa(w) = \max_{a \in \mathcal{A} \setminus \mathcal{A}^N} \frac{2N(w)^T(g(a) - w)}{r|\Sigma\phi(a, T(w))|^2}$$

Aunque para calcular esta última cantidad es necesario obtener  $\phi(a, T(w))$ . Dada la definición de  $\phi(a, T(w))$ , es posible encontrar su valor al solucionar el problema

cuadrático:

$$\begin{aligned} |\phi(a, (t_1, t_2))|^2 &= \min_{\phi} |\phi|^2 \\ \text{s.a.} \quad g_i(a) + t_i \phi^T \mu(a) &\geq g_i(a'_i, a_j) + t_i \phi^T \mu(a'_i, a_j), \quad \forall i = 1, 2, \quad \forall a'_i \in \mathcal{A}^i \end{aligned}$$

Podemos reescribir las restricciones del problema cuadrático, usando la siguiente notación:

- Sea  $G_i(a)$  el vector de  $|\mathcal{A}^i| - 1$  componentes, tal que en cada una de ellas tiene el valor  $g_i(a) - g_i(a'_i, a_j)$ .
- Análogamente, sea  $U_i(a)$  la matriz de  $(|\mathcal{A}^i| - 1) \times d$  que en cada fila tiene el valor  $(\mu(a) - \mu(a'_i, a_j))^T$ .

Así, el problema es buscar el mínimo de  $|\phi|^2$  sujeto a

$$\begin{bmatrix} G_1(a) \\ G_2(a) \end{bmatrix} \geq - \begin{bmatrix} t_1 U_1(a) \\ t_2 U_2(a) \end{bmatrix} \phi$$

En el siguiente capítulo, el 5, usaremos los puntos 4.2 y 4.3.2 de este capítulo para resolver un ejemplo de colusión con bienes complementarios, en estos puntos explicamos como resolver la ecuación en un contexto simétrico y en un contexto asimétrico no  $\mathcal{C}^1$ . Por completitud también se explicará en 4.3.1 como resolver la ecuación en un contexto asimétrico  $\mathcal{C}^1$ . En 4.2 y 4.3.1 seguimos a Sannikov [5, sección 8].

## 4.2. Caso Simétrico

En esta sección se explica como resolver la ecuación (4.3) en el caso que los conjuntos de acciones para ambos jugadores son iguales. Además, las señales y los pagos instantáneos son simétricos, es decir, se tiene que

$$\mu_i(a_i, a_j) = \mu_j(a_j, a_i) \quad \wedge \quad g_i(a_i, a_j) = g_j(a_j, a_i)$$

Con todo esto, es posible notar que  $\mathcal{E}(r)$  debe ser a su vez simétrico, pues ambos jugadores se encuentran siempre bajo las *mismas* condiciones e incentivos. Al aceptar esta premisa, se llega a la conclusión que  $\mathcal{E}(r)$  cruza la diagonal del plano, es decir cruza la recta que pasa por el origen en  $45^\circ$ .

A partir de lo anterior, podemos resolver la ecuación diferencial que nos interesa usando como condición inicial  $(\pi/4, w_0)$  con  $w_0 \in \mathcal{V}^* \cap \{x_1 = x_2\}$ , donde hemos denotado por  $\{x_1 = x_2\}$  la diagonal en  $\mathbb{R}^2$ .

Con un método de bisección a lo largo de la diagonal, es posible ubicar los puntos  $w_0$  tal que la solución  $\tilde{w}(\theta)$  de

$$\begin{aligned}\frac{dw(\theta)}{d\theta} &= \frac{T(\theta)}{\kappa(\theta)}, & w \notin \mathcal{N} \\ w(\pi/4) &= w_0\end{aligned}$$

es una curva cerrada continua en  $\mathcal{V}^*$  y  $\mathcal{C}^1$  salvo en  $\mathcal{N}$ .

En los gráficos de la figura 4.1 se muestra parte de las iteraciones del método de bisección utilizado para encontrar el conjunto  $\mathcal{E}(2)$  en un caso simétrico (ver más adelante 5.1). En este caso se sabe que el conjunto de pagos alcanzables en equilibrios de Nash corresponde a un único punto, a saber  $\mathcal{N} = (128, 128)$ ; por ello se busca un punto en la diagonal tal que la solución encontrada llegue a dicho punto, luego por la simetría sabemos que resolviendo la ecuación *en el otro sentido* también llegaremos al mismo punto, resultando una curva cerrada y  $\mathcal{C}^1$  salvo en  $\mathcal{N}$ .

Vale la pena destacar que este algoritmo, puede entregar más de una solución dependiendo de las reglas que se incorporen al método de bisección. Si ocurre esto, el conjunto  $\mathcal{E}(r)$  estará dado por la solución que encierre el área más grande dentro de  $\mathcal{V}^*$ .

### 4.3. Caso Asimétrico

En esta sección dejamos atrás la hipótesis de que conocemos una línea y un ángulo tal que  $\mathcal{E}$  cruza dicha recta con ese ángulo, en la sección anterior asumíamos que  $\mathcal{E}$  cruzaba la diagonal con un ángulo de  $45^\circ$ .

Dividiremos el análisis en dos partes, primero analizaremos el caso en que el conjunto de pagos alcanzables por PPE no contiene a los pagos alcanzables por equilibrio de Nash estáticos, i.e. si denotamos por  $\partial\mathcal{V}^*$  la frontera de  $\mathcal{V}^*$  entonces primero veremos el caso en que  $\partial\mathcal{V}^* \cap \mathcal{N} = \emptyset$ , en contraposición, la segunda parte corresponde al caso  $\partial\mathcal{V}^* \cap \mathcal{N} \neq \emptyset$ .

#### 4.3.1. Conjuntos sin Nash en la frontera de $\mathcal{V}^*$

La forma propuesta para resolver la ecuación en el caso en que todo  $\mathcal{N}$  queda en el interior de  $\mathcal{V}^*$ , corresponde al siguiente algoritmo (Sannikov, [5, subsección 8.2]):

1. Se toma  $w_0^1$  como un punto arbitrario en la frontera de  $\mathcal{V}^*$ . Sea  $\underline{\theta}^1$  tal que la normal en ese punto, señala hacia *afuera* del conjunto  $\mathcal{V}^*$ .
2. Se resuelve la ecuación para condiciones iniciales  $(w_0^1, \theta_0)$  con  $\theta_0 = \underline{\theta}^1$ .

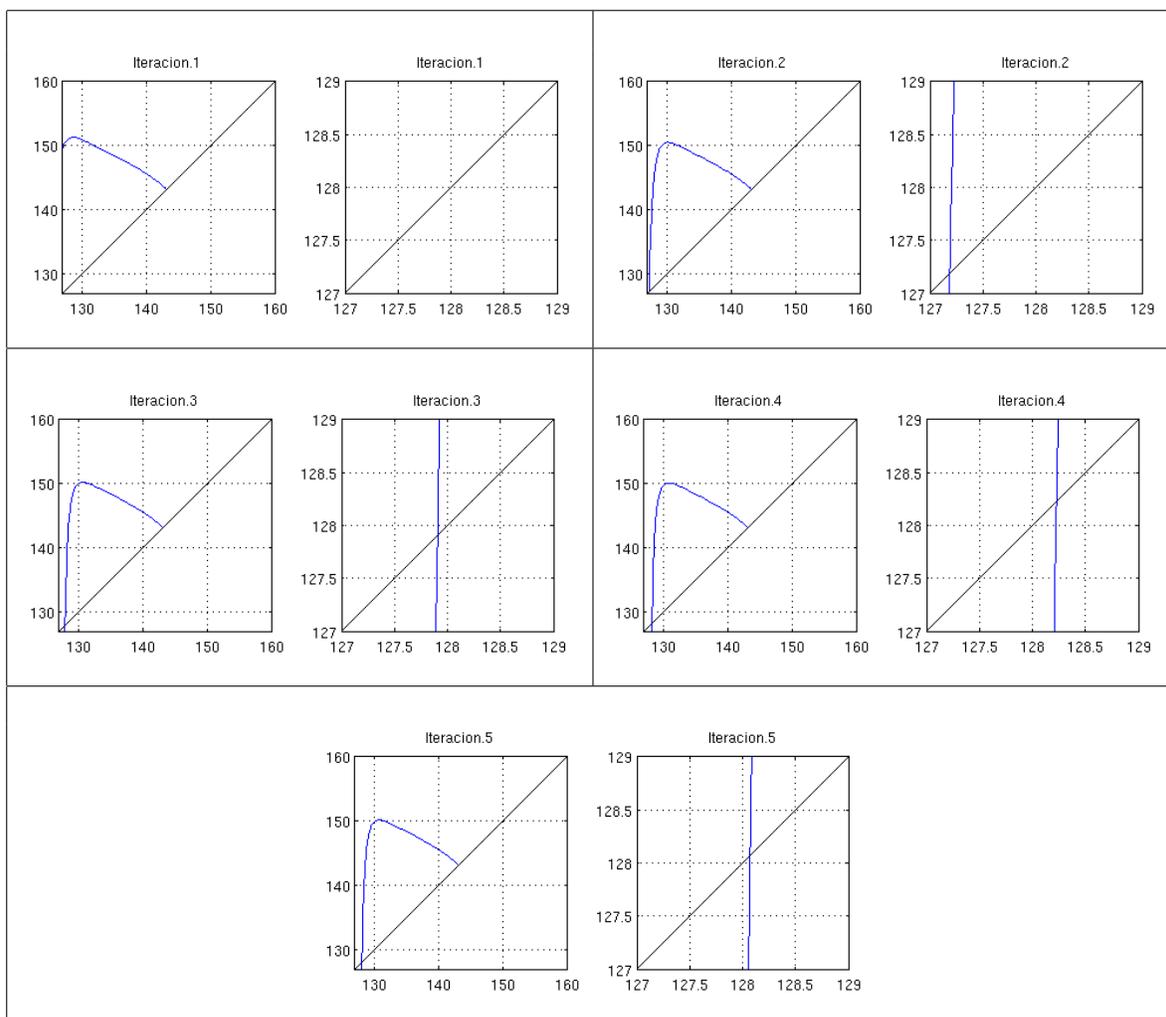


Figura 4.1: Iteraciones del método de bisección para un caso simétrico con un único equilibrio de Nash cuyo pago queda en la esquina inferior izquierda de  $\mathcal{V}^*$ , en este caso, esa esquina corresponde al  $(128, 128)$ . Una vez encontrado  $w_0$  tal que la solución llegue a  $(128, 128)$ , se resolverá la ecuación desde ese punto pero en el otro sentido, (i.e. en  $-d\theta$ ), hasta que la solución llegue a  $(128, 128)$ .

3. Se varía continuamente  $\theta_0 \in [\underline{\theta}^1, 2\pi)$  hasta que la solución llegue nuevamente a  $w_0^1$ . Notemos que la definición de  $\underline{\theta}^1$  asegura que la solución siempre *entrará* hacia  $\mathcal{V}^*$ .
4. Tomamos  $w_0^2$  el punto en la solución encontrada tal que tiene normal  $(1, 0)^T$  (i.e. el ángulo tangencial  $\theta$  en ese punto es 0).
5. Volvemos a resolver la ecuación desde  $(w_0^2, \theta_0)$ , aunque ahora  $\theta_0$  varía continuamente entre  $[0, 2\pi)$ , hasta que la solución vuelva a tocar a  $w_0^2$ .
6. Se repite 4 y 5, hasta que el método converga.

Notemos que las soluciones, cerradas, de la ecuación de optimalidad encontradas en las distintas etapas del algoritmo deben encerrar a  $\mathcal{E}(r)$ . Si no, en algún punto, la frontera de  $\mathcal{E}(r)$  debe intersectar con la solución encontrada; como variamos el ángulo  $\theta_0$  continuamente, alguna solución *parcial* hubiera sido tangente con la frontera de  $\mathcal{E}(r)$  y como ambas curvas satisfacen la misma ecuación, por unicidad de soluciones<sup>1</sup>, ambas curvas serían la misma, lo que es imposible, por las propiedades que debe cumplir  $\partial\mathcal{E}(r)$ . Este argumento viene a justificar por qué el algoritmo anterior debería servir (Sannikov, op. cit.).

### 4.3.2. Conjuntos con Nash en la frontera de $\mathcal{V}^*$

Para presentar este método, nos inspiramos en la situación que se ve en el gráfico 3.1 del capítulo anterior (página 33), en él podemos observar que  $\mathcal{N}$  ocupa una esquina de  $\mathcal{V}^*$ . En este caso, el algoritmo de la sección anterior no es posible de aplicar pues como  $\mathcal{N} \subset \mathcal{E}(r) \subset \mathcal{V}^*$ , se tiene que si  $\mathcal{N}$  toca la frontera de  $\mathcal{V}^*$  también toca la frontera de  $\mathcal{E}(r)$ . Por esto último, no encontraremos ningún punto  $w_0$  en  $\mathcal{V}^* \setminus \mathcal{N}$  tal que la solución de la ecuación se cierre e incluya a todo  $\mathcal{N}$ ; como en  $\mathcal{N}$  la ecuación no aplica, no se puede hacer nada más en la dirección del algoritmo anterior.

Propondremos un algoritmo para resolver el problema anterior que involucrará una búsqueda en 2 direcciones, pero en un contexto acotado, por ello haremos el siguiente supuesto:

**Supuesto 3** (Unicidad de Equilibrios de Nash). Existe un único equilibrio de Nash para el juego de los pagos instantáneos esperados. Más aún, el conjunto  $\mathcal{N}$  asociado está en la frontera de  $\mathcal{V}^*$ .

Con lo anterior, el algoritmo queda como sigue

<sup>1</sup>En [5] se demuestra que la ecuación de optimalidad asegura unicidad de soluciones y dependencia continua en las condiciones iniciales.

1. Se identifica una recta de interés que cruce  $\mathcal{V}^*$ , e.g. la recta que une  $\mathcal{N}$  con el pago máximo colusivo, es decir, la recta que une  $\mathcal{N}$  con  $\max_a u_1(a) + u_2(a)$ . Llamamos a esa recta  $D$ .<sup>2</sup>
2. Se resuelve la ecuación con condiciones iniciales  $(w_0^1, \theta_0)$ , donde  $w_0^1 \in D$  y  $\theta_0$  varía continuamente entre  $\Theta_0^- := [0, \pi)$  ó  $\Theta_0^+ := [\pi, 2\pi)$ . De los posibles conjuntos para  $\theta_0$ , escogemos el cual asegure que la solución se cierre hacia dentro de  $\mathcal{V}^*$  y hacia  $\mathcal{N}$ .
3. Se deja de buscar hasta que una solución llegue a  $\mathcal{N}$  en cierto ángulo  $\theta_N^1$  para algún valor  $\theta_0^1$ , es decir,  $\theta_0$  deja de variar una vez encontrada una solución que llegue a  $\mathcal{N}$ , y en ese caso, esta solución estará definida para  $\theta \in [\theta_0^1, \theta_N^1)$ .
4. Se resuelve la ecuación con las mismas condiciones iniciales  $(w_0^1, \theta_0^1)$ , pero en el sentido contrario hasta un ángulo  $\hat{\theta}$ , en el cual alguna de las siguientes cosas ocurra:
  - $[\theta_0^1, \theta_N^1) \cup [\theta_0^1, \hat{\theta}] = [0, 2\pi)$ , aquí hemos sido ligeros en la notación. Lo que se plantea es que ya no queden ángulos en los cuales resolver.
  - La solución sale de  $\mathcal{V}^*$ .
  - La solución llega a  $\mathcal{N}$ .
5. Si la solución llegó a  $\mathcal{N}$ , hemos encontrado  $\mathcal{E}(r)$ , pues tenemos una curva cerrada continua y  $\mathcal{C}^1$  salvo en  $\mathcal{N}$ .  
Si no, se establecen reglas para un método de bisección en  $w$ , éstas reglas pueden venir de la experiencia obtenida al resolver la ecuación en distintos puntos  $w$  (pero igual  $\theta$ ). Con dichas reglas, se escoge  $w_0^2 \in D$  y se repiten los puntos 2, 3 y 4 hasta que sea necesario.

Obs 1. Para hacer más eficiente el método, se puede implementar un método de bisección en  $\theta$ , en vez de variar *continuamente*.

Obs 2. Para aumentar la eficiencia y estabilidad, tras elegir el siguiente punto inicial  $w_0$ , se puede acotar el ángulo dentro del cual se varía continuamente  $\theta_0$ , e.g. una regla podría ser que si la distancia entre  $w_0^{n+1}$  y  $\mathcal{N}$  es menor que la distancia entre  $w_0^n$  y  $\mathcal{N}$ , entonces el conjunto para  $\theta$  queda como  $[0, \theta_0^n)$ .

En la figura 4.2 es posible observar parte de la iteración para la resolución de  $\mathcal{E}(3, \begin{pmatrix} 1.3 & 0 \\ 0 & 0.7 \end{pmatrix})$  para un juego de bienes complementarios (ver 5.2). Para la resolución se tiene que el pago máximo colusivo corresponde a 144 para cada jugador, además se tiene que  $\mathcal{N} = \{(v_1, v_2) = (128, 128)\}$ . Por lo anterior, se utiliza  $\Theta_0^- : [0, \pi)$  y la dinámica de resolución, que se muestra en 4.2, es como sigue:

<sup>2</sup>En vez de  $\mathcal{N}$ , también se podría probar usar la esquina de  $\mathcal{V}^*$ . Por su definición,  $\mathcal{V}^*$  siempre tiene una esquina que corresponde a  $(v_1, v_2)$ , los pagos minmax de ambos jugadores.

1. Se encuentra  $\theta_0^1$  tal que se llega a  $\mathcal{N}$ , como  $w_0^1$  esta muy cerca del pago máximo, no es posible resolver la ecuación en el otro sentido. Por lo cual se *baja*  $w$  para la siguiente iteración. Además, el conjunto para  $\theta_0$  se acotará por arriba con  $\theta_0^1$ . Aquí  $w_0^1 = (143, 5, 143, 5)$  y  $\theta_0^1 = 0,9117$ .
3. Ahora sí es posible resolver en el otro sentido, pero aún no alcanzamos a acercarnos a  $\mathcal{N}$ , por lo que se sigue bajando  $w$  y se vuelve a acotar  $\Theta_0^-$  por arriba con  $\theta_0^3$ . En este paso se tiene  $w_0^3 = (143, 125, 143, 125)$  y  $\theta_0^3 = 0,7541$  (notar  $\theta_0^3 < \theta_0^1$ ).
5. Finalmente es posible resolver en todo  $\theta$  pero aún así no se alcanza  $\mathcal{N}$  por lo que se *sube*  $w$ , ahora el conjunto  $\Theta_0^-$  se acota por abajo por  $\theta_0^5$ . Los valores obtenidos en este paso son  $w_0^5 = (143, 0312, 143, 0312)$  y  $\theta_0^5 = 0,7192$  (aquí  $\theta_0^5 < \theta_0^3$ ).
7. Nuevamente se cae fuera de  $\mathcal{V}^*$  pero esta vez con un ángulo distinto, de todas maneras,  $w$  debe bajar y  $\Theta_0^-$  se debe acotar por arriba por  $\theta_0^7$ . Ahora  $w_0^7 = (143, 1016, 143, 1016)$  y  $\theta_0^7 = 0,7426$ .
11. Se alcanza el conjunto  $\mathcal{N}$  por ambos lados y por dentro de  $\mathcal{V}^*$ . Finalmente se obtiene  $w_0^{11} = (143,0913, 143,0913)$  y el ángulo  $\theta_0^{11} = 0,7401$ . Se concluye que  $w_0^{11} \in \partial\mathcal{E}$ , más aún, el ángulo tangente de  $\partial\mathcal{E}$  en el punto  $w_0^{11}$  corresponde a  $\theta_0^{11}$ .

### Nota sobre otro método de resolución

Antes de probar el método explicado más arriba, se probó haciendo otra bisección en dos dimensiones aunque esta vez  $w$  se movía, en vez de en una recta que cruce gran parte de  $\mathcal{V}^*$ , en una pequeña línea cercana a  $\mathcal{N}$ . Entonces se hacía lo mismo: primero se buscaba un ángulo que permitiera llegar a  $\mathcal{N}$  desde el punto donde se encontraba  $w$  y luego se resolvía en el otro sentido con la esperanza de tocar nuevamente a  $\mathcal{N}$ . Si no ocurría eso, se movía  $w$  y se repetía.

A pesar de que la idea es la misma entre este método y el anterior, no se pudo nunca resolver un ejemplo usando este último método. Esto puede ser posible debido al gran error numérico que se acumula al resolver en el otro sentido partiendo desde muy cerca de  $\mathcal{N}$ , además al estar muy cerca de  $\mathcal{N}$  se necesita mucha precisión para poder mover  $w$ , lo que también crea error numérico.

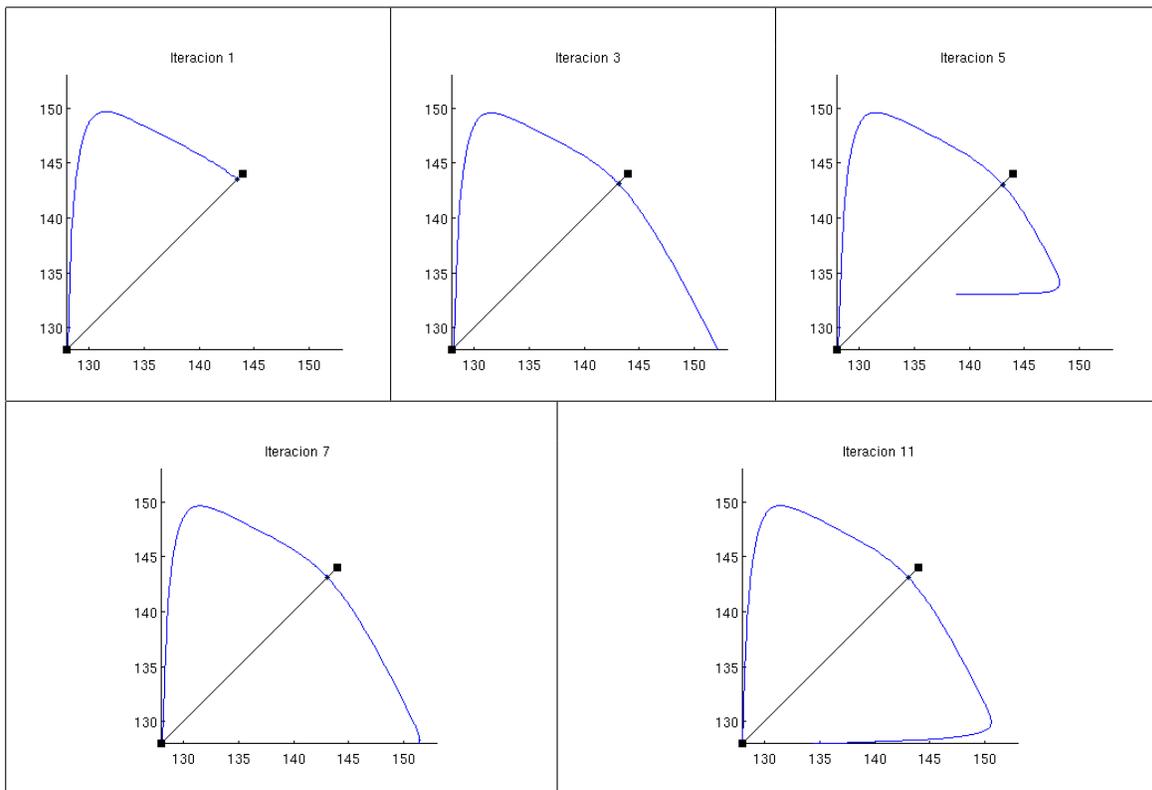


Figura 4.2: Iteraciones del método de bisección en dos dimensiones para un caso asimétrico con un único equilibrio de Nash cuyo pago queda en la esquina inferior izquierda de  $\mathcal{V}^*$ , en este caso, esa esquina corresponde al  $(128,128)$ . En negro se muestra  $D$ , que en este caso coincide con la diagonal.

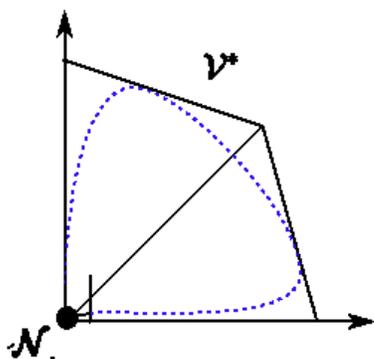


Figura 4.3: Método alternativo planteado. Se muestra la línea sobre la cual se mueve  $w$ , ésta línea corresponde a la línea sólida vertical a la derecha de  $\mathcal{N}$ . La línea punteada corresponde a lo que debe recorrer la solución al resolver la ecuación en *el otro sentido*.

# Capítulo 5

## Resolución y Estudio Comparativo de Casos

A lo largo de este capítulo veremos distintas variantes de un caso de Duopolio con Bienes Complementarios. Nuestro caso base será el siguiente: existen dos bienes y dos productores, uno para cada bien; la cantidad de un bien afecta el precio del otro aumentándolo, es decir, son complementarios; los productores sólo pueden observar los precios de ambos bienes y sus utilidades corresponden a la cantidad vendida por su precio. Más precisamente, consideraremos que las acciones corresponden a la cantidad producida y que los precios (esperados) vienen dados por

$$\begin{aligned}P_1(a) &= B - Ca_1 + \gamma a_2 \\P_2(a) &= B - Ca_2 + \gamma a_1\end{aligned}$$

donde  $a = (a_1, a_2)$  indica la cantidad producida por el productor/jugador 1 y por el jugador 2, respectivamente. Las constantes  $B$ ,  $C$  y  $\gamma$  son dadas y positivas, aquí  $\gamma$  indica que tan complementarios son los bienes, por ejemplo, si  $\gamma = 0$  el bien de un productor no tiene importancia para el precio del otro.

A pesar que los precios esperados están definidos como arriba, existen *shocks* exógenos que afectan el precio *real* u observado. Luego, dado un perfil de acciones  $A = (A_t^1, A_t^2)_t$ , la señal que finalmente se observa es bidimensional y corresponde a

$$X_t(A) = \underbrace{\int_0^t (P_1(A_s) \ P_2(A_s))^T ds}_{\text{precios esperados}} + \underbrace{\int_0^t \Sigma dZ_s}_{\text{ruido o shocks}}$$

con  $\Sigma$  una matriz simétrica definida positiva de  $2 \times 2$  y  $Z$  un movimiento Browniano estándar en 2 dimensiones. Dependiendo de la forma que  $\Sigma$  tenga, los ruidos pueden o no estar correlacionados entre sí, asimismo la variabilidad que ellos tengan también dependerán de  $\Sigma$ . Por ejemplo, para modelar que el *shock* del precio del bien 1 es

independiente del *shock* para el bien 2, usaremos una matriz  $\Sigma$  diagonal; más aún, si el primer *shock* presenta mayor variabilidad que el segundo, entonces la primera componente de la diagonal será mayor que la segunda. Para poner la definición de la señal pública según el contexto de lo presentado en los dos capítulos anteriores, hace falta definir, dado un perfil  $a$ ,

$$\mu(a) := (P_1(a), P_2(a))^T \quad (5.1)$$

con lo que la señal pública queda definida tal como queremos:

$$X_t(A) = \int_0^t \mu(A_s) ds + \int_0^t \Sigma dZ_s$$

Por último, el pago descontando del jugador  $i$ , al momento  $t$ , queda dado por

$$r \int_0^t e^{-rs} A_s^i dX_s = r \int_0^t e^{-rs} A_s P_i(A_s) ds + r \int_0^t e^{-rs} A_s \Sigma dZ_s$$

Según la notación de los capítulos anteriores, el pago descontando del jugador  $i$ , al momento  $t$ , era

$$r \int_0^\infty e^{-rt} (c_i(A_t^i) + b_i(A_t^i)^T \mu(A_t)) dt + b_i(A_t^i)^T \Sigma dZ_t$$

en nuestro caso  $c_i(a_i) \equiv 0$ ,  $b_i(a_i) = a_i$  y  $\mu$  como en 5.1. Así, el pago instantáneo es  $g_i(A_t) = A_t^i (B - CA_t^i + \gamma A_t^i)$ .

El juego en forma normal descrito por los pagos  $g_1$  y  $g_2$  tiene un único equilibrio de Nash, que resulta ser simétrico, y es cuando ambos jugadores producen

$$a_N := \frac{B}{2C - \gamma}$$

Si en el mismo juego, los jugadores se coluden para alcanzar el máximo de  $g_1(a) + g_2(a)$ , ambos deberán jugar

$$a_M := \frac{B}{2(C - \gamma)}$$

Ahora bien, si ambos jugadores están coludidos jugando  $a_M$ , la mejor desviación que un jugador podría hacer es

$$a_D := \frac{B(2C - \gamma)}{4C(C - \gamma)}$$

Es decir,  $a_D$  es la mejor respuesta contra  $a_M$ , y lo es no importando el jugador.

Es fácilmente demostrable que si  $\gamma < C$ , entonces  $a_N \leq a_D \leq a_M$ , lo que diferencia este juego de duopolio de los juegos tradicionales de monopolio, donde los jugadores se coluden para producir menos que lo que producirían en competencia perfecta (o en el equilibrio de Nash). Lo anterior se debe a que los jugadores internalizan el efecto positivo que su producción tiene sobre el precio del otro jugador.

A menos que se diga lo contrario, usaremos como parámetros  $B = 24$ ,  $C = 2$  y  $\gamma = 1$ , luego las cantidades críticas definidas anteriormente resultan ser  $a_N = 8$ ,  $a_D = 9$ ,  $a_M = 12$ . Por lo anterior, usaremos como conjunto de acciones para ambos jugadores sólo estas cantidades, es decir,  $\mathcal{A}_1 = \mathcal{A}_2 = \{a_N, a_D, a_M\}$ . Podemos ver los pagos alcanzables y los pagos racionales para este juego en las figuras 5.1 y 5.2, respectivamente.

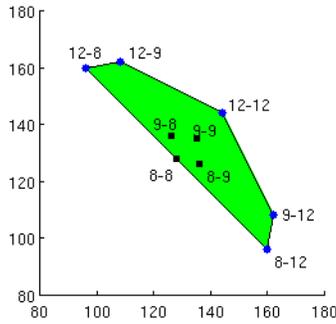


Figura 5.1: Pago instantáneo de las distintas acciones en el juego de referencia. Los puntos cuadrados pertenecen al interior del conjunto de los pagos posibles, los circulares pertenecen a la frontera de los posibles pagos.

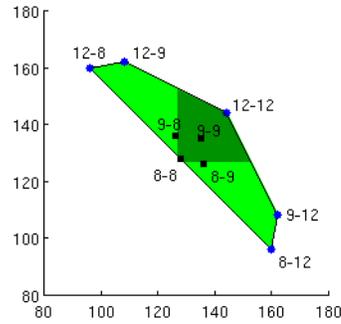


Figura 5.2: Se muestra  $\mathcal{V}^*$  en color oscuro. Este conjunto se obtiene al asegurar a cada jugador su pago minmax, en este caso 128, que corresponde al pago cuando ambos jugadores juegan el equilibrio de Nash  $(a_N, a_N)$ .

## 5.1. Caso Simétrico

Tomamos en este caso

$$\Sigma = \sigma \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad (5.2)$$

con lo que el juego a tiempo continuo queda completamente simétrico, luego ocuparemos el algoritmo mostrado en 4.2.

En la figura 5.3 es posible ver al conjunto de pagos alcanzables en equilibrio públicos perfectos cuando la tasa de descuento es  $r = 2$  y la varianza viene dada por  $\sigma = 1$  (ver (5.2)) i.e. vemos  $\mathcal{E}(2)$ .

Podemos observar que se alcanza prácticamente todo el conjunto  $\mathcal{V}^*$  (recordar figura 5.2), explicamos esto como sigue: dado que los bienes son complementarios, los incentivos de ambos jugadores se encuentran alineados, luego es muy fácil que los

jugadores se puedan coordinar para obtener el pago máximo. A pesar de lo anterior, no es posible alcanzar el mejor pago posible (a saber, 144 para cada jugador, cuando ambos juegan 12) pues al haber información incompleta, eventualmente se harán castigos aunque ninguno de los jugadores efectivamente haya hecho *trampa*, tal como ocurría en el modelo a tiempo discreto (ver ejemplo 2.2.2).

La dinámica esperada para este caso es como sigue:

1. Ambos jugadores comienzan jugando su mejor opción conjunta, i.e., ambos juegan  $a_M = 12$ . Se mantienen ahí, hasta que eventualmente la señal  $X_t$  sale de la *zona de confianza*; esta zona corresponde a la línea discontinua bajo la etiqueta 12 – 12 en 5.3.
2. Una vez que se sale de la zona de confianza (y lo harán con probabilidad positiva), los jugadores caen a alguno de los costados de ella, esto demarcado por la línea continua bajo las etiquetas 12 – 9 ó 9 – 12. En aquella zona, el jugador que usa  $a_D = 9$  se está viendo beneficiado, obteniendo su máximo pago (instantáneo) al desviarse óptimamente desde la acción monopólica, mientras el otro jugador está siendo perjudicado por este desvío. Para ejemplificar lo que sigue digamos que el jugador 2 esta siendo beneficiado y por tanto estamos a la izquierda de la figura 5.3. Notemos que en esta zona, el jugador 2 alcanza su máximo pago posible en un PPE.
3. Dependiendo de la señal corresponderá una nueva etapa donde el jugador 2 puede:
  - a) Recompensar al jugador 1 y volver al máximo pago conjunto, esto es, el jugador 2 juega  $a_M = 12$ .
  - b) Castigar al jugador 1 jugando 8. Notemos que esto es costoso e implica, para el beneficiado, ganar menos que en la situación anterior.
4. Si el jugador 1 fue recompensado, volvemos al punto 1. Si fue castigado, como sostener que él siga jugando 12 requiere de muchos incentivos en términos de futuros *drifts* existe en esta zona gran volatilidad (matemáticamente, esto quiere decir que  $\|\phi(12, 8)\|$  es muy grande). Para disminuir dicha volatilidad es óptimo que el jugador 2 *ceda*, por ello él juega 8, por su parte para el jugador 1 también es óptimo ceder *algo*, por lo que juega su segunda mejor opción frente a 8, es decir el primer jugador juega 9; con lo anterior se disminuye la volatilidad.
5. Mientras el jugador 2 toma la acción 8 y el jugador 1 la 9, pueden retroceder al punto 3 o quedarse en el 4. Si no ocurre lo anterior, ambos jugadores llegan a  $\mathcal{N} = \{(128, 128)\}$ , como la única forma de que ambos alcancen 128 es que se queden ahí para siempre, eso es lo que hacen. En efecto, si  $W_t$  cae en  $\mathcal{N}$ , es decir, si el pago descontado esperado normalizado es 128 para cada jugador,

implica que si se movieran desde esa posición, por ejemplo, para ofrecer más, es necesario que exista la posibilidad de ofrecer menos, para así ser consistente con el pago esperado ofrecido de 128, pero como  $\mathcal{N}$  esta en la frontera *inferior* de  $\mathcal{V}^*$  no existen pagos menores -para ninguno de los jugadores- por lo que la única forma de ofrecer 128 y ser consistente en ello, es decidir que se jugará ese pago para siempre, pase lo que pase.<sup>1</sup>

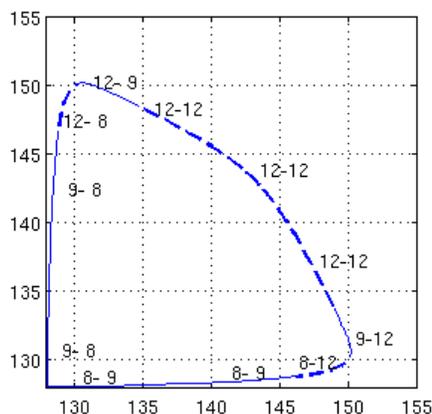


Figura 5.3: Conjunto  $\mathcal{E}(2)$  para el caso simétrico. Existen dos tipos de líneas para marcar  $\partial\mathcal{E}(2)$ : discontinua (más gruesa) y continua, el cambio en el tipo de línea indica un cambio en el perfil jugado. Se muestran los perfiles jugados en el formato  $a_1 - a_2$ .

En la figura 5.4 es posible ver el drift tangencial<sup>2</sup> en cada punto, es posible observar que el perfil 12-12 es un punto repelente (aunque no con mucha fuerza), y por ello el pago esperado descontado ofrecido al jugar ese perfil, es menor que el pago 144 que ambos jugadores reciben instantáneamente. Es interesante ver que las zonas donde se beneficia de sobre manera a un jugador, son altamente repelentes e impulsan a los jugadores a dejar esa zona, esto se condice con el hecho que el pago ofrecido es bastante en menor que el pago instantáneo que los jugadores están recibiendo (ver 5.1).

Análogamente, en la figura 5.5 es posible ver de qué tamaño<sup>3</sup> es la volatilidad, se muestran dos colores para separar las influencias de la señal 1 y de la señal 2. Para

<sup>1</sup>Notar que esta explicación es general aunque Sannikov [5, pág. 27] argumenta que siempre que los jugadores lleguen a un punto en  $\mathcal{N}$  se quedarán ahí para siempre.

<sup>2</sup>En nuestro contexto el drift tangencial en el punto  $W_t$  corresponde a  $(T^T(W_t - g(A_t)))T$ , es decir, a la proyección sobre  $T$  de  $W_t - g(A_t)$ .

<sup>3</sup>En la figura, todos los vectores han sido aumentados por una constante para poder visualizarlos, esto sin pérdida de generalidad pues nos interesa la comparación entre los distintos vectores y no el valor absoluto de su norma.

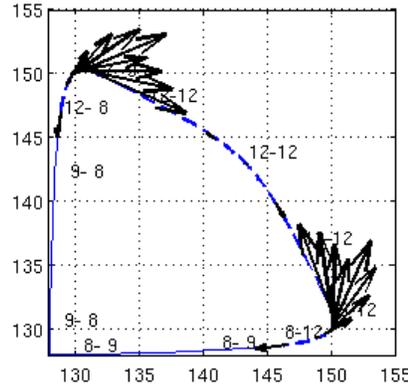


Figura 5.4: Conjunto  $\mathcal{E}(2)$  para el caso simétrico, las flechas indican los *drifts* tangenciales en algunos puntos.

precisar lo anterior, recordemos que la volatilidad se ve influenciada por la matriz:

$$T\phi(a, T)^T = \begin{pmatrix} t_1\phi_1(a, T) & t_1\phi_2(a, T) \\ t_2\phi_1(a, T) & t_2\phi_2(a, T) \end{pmatrix}$$

Luego, graficaremos los vectores  $v_1 = T\phi_1(a, T)$  y  $v_2 = T\phi_2(a, T)$ , ya que la volatilidad viene dada por:

$$T\phi(a, T)^T \Sigma dZ_t = \sigma \begin{pmatrix} t_1\phi_1(a, T)dZ_t^1 + t_1\phi_2(a, T)dZ_t^2 \\ t_2\phi_1(a, T)dZ_t^1 + t_2\phi_2(a, T)dZ_t^2 \end{pmatrix} = \sigma (v_1 dZ_t^1 + v_2 dZ_t^2)$$

Así las cosas, en la figura podemos observar las zonas donde existe mayor volatilidad, zonas que coinciden con las zonas donde los jugadores necesitan más incentivos para jugar las acciones acordadas.

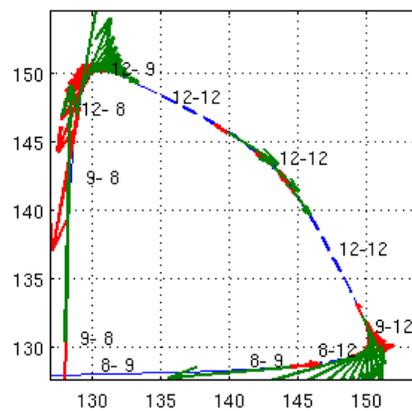


Figura 5.5: Conjunto  $\mathcal{E}(2)$  para el caso simétrico, las flechas indican las volatilidades en algunos puntos, las flechas rojas indican la volatilidad asociada a la primera componente de la señal y las verdes, la asociada a la segunda. La distribución es simétrica, en el sentido que en el lado bajo la diagonal del plano, las flechas verdes ocupan el lugar que las flechas rojas ocupan sobre la diagonal y viceversa (aunque no es posible graficar todas las flechas para mostrar explícitamente este fenómeno).

### 5.1.1. Efecto del descuento

Una pregunta interesante es saber que influencia tiene la tasa de descuento en la forma y tamaño del conjunto  $\mathcal{E}(r)$ , si en tiempo discreto cuando la información es completa se tiene que  $\mathcal{E}(r) \rightarrow \mathcal{V}^*$  cuando  $r \rightarrow 0$  (recordar 2.1.4), ¿existe algún resultado similar en nuestro contexto? La respuesta será que sí, aunque no mostraremos exactamente el mismo resultado. Aquí veremos experimentalmente que  $\mathcal{E}(r_1) \subset \mathcal{E}(r_2)$  si  $r_1 > r_2$ , es decir los equilibrios van creciendo a medida que los jugadores se hacen más pacientes.

Resolvemos numéricamente la ecuación, usando las constantes anteriores, pero cambiando la tasa de descuento  $r$ . Estos resultados los resumimos en la figura 5.6. En dicha figura es posible apreciar que efectivamente los conjuntos  $\mathcal{E}(r)$  se hacen más pequeños a medida que  $r$  se hace más grande, esto es de esperar, pues mayor  $r$ , mayor impaciencia, ergo los jugadores están menos susceptibles a incentivos que hablen de un mejor pago futuro enfocándose en un área más cercana al pago del Nash (128,128), área donde menos incentivos *futuros* se necesitan para coordinar a los jugadores.

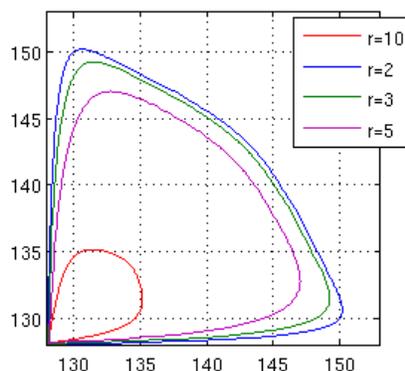


Figura 5.6: Conjunto  $\mathcal{E}(r)$  para el caso simétrico,  $r$  toma los valores 2, 3, 5 y 10.

### 5.1.2. Efecto de la incertidumbre

Tal como es de importancia preguntarse por los efectos de la tasa de descuento, también es de importancia saber que efecto tiene el *tamaño* de la incertidumbre en este contexto. Recordemos que en esta sección estamos usando

$$\Sigma = \sigma \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Por lo tanto usamos la notación  $\mathcal{E}(r, \sigma)$  para hacer explícita la dependencia en  $\sigma$ . Ahora la pregunta que nos hacemos es ¿ $\mathcal{E}(r, \sigma_1) \subset \mathcal{E}(r, \sigma_2)$ , si  $\sigma_1 > \sigma_2$ ?, es decir ¿si la

incertidumbre disminuye: el conjunto de pagos posibles en equilibrio públicos crece? Para esto, tal como antes, resolvemos la ecuación varias veces, pero esta ocasión usamos distintos valores de  $\sigma$  y mantenemos  $r$  fijo. Los resultados los podemos ver en la figura 5.7, donde se observa que la respuesta a la pregunta planteada es afirmativa, tal como la intuición permite suponer: entre mejor se pueden observar las acciones de los jugadores, es más fácil que se pongan de acuerdo y estén dispuestos a soportar pérdidas por no *aprovecharse*, en pos de mejores ganancias futuras.

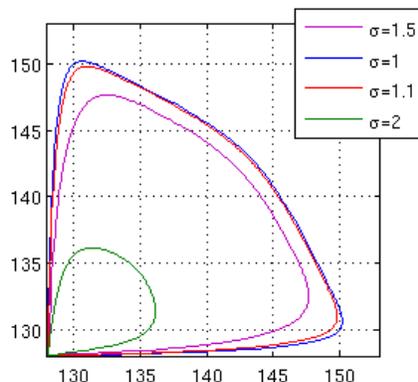


Figura 5.7: Conjunto  $\mathcal{E}(r)$  para el caso simétrico,  $\sigma$  toma los valores 1, 1.1, 1.5 y 2.

Hagamos explícitos que por su definición,  $\phi(a, T)$  no cambia si  $\sigma$  toma distintos valores, luego el efecto anterior se ve explicado sólo por el efecto en el aumento en la volatilidad que trae  $\sigma$ . Este aumento se traduce en menores pagos futuros, pues la absorción en  $\mathcal{N}$  se hace más probable.

## 5.2. Asimetría en la aleatoridad de las señales

En esta sección resolveremos el problema de referencia pero en el caso que las señales tienen distinta varianza, a saber consideramos

$$\Sigma = \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{pmatrix} \quad (5.3)$$

Ya que una de las esquinas de  $\mathcal{V}^*$  corresponde al pago obtenido en un equilibrio de Nash, usamos el procedimiento planteado en 4.3.2. Usaremos para los distintos cálculos  $r = 5$  y para los demás parámetros, los valores planteados en el preámbulo del capítulo.

En la figura (5.8) es posible ver la resolución de la ecuación de optimalidad usando  $\sigma_1 = 1,5$  y  $\sigma_2 = 0,5$ . Observamos que básicamente las acciones escogidas son

las mismas que en el caso simétrico, más aún, en el mismo orden, por lo que no hay cambios en la dinámica del juego. Un cambio que sí se puede notar fácilmente, es que los pagos que se le ofrecen al jugador 1, quien es quién tiene la señal con mayor variabilidad, en general son mayores que los pagos ofrecidos al jugador 2; se podría decir que al jugador 1 se le ofrecen, y obtiene, más ganancias al desviarse que el jugador 2 cuando él es quién se desvía. Podemos interpretar lo anterior como *quien puede esconder sus acciones puede beneficiarse con ello*, a diferencia del caso cuando ambos ven influida su varianza donde los dos jugadores se ven perjudicados por un aumento en la variabilidad.

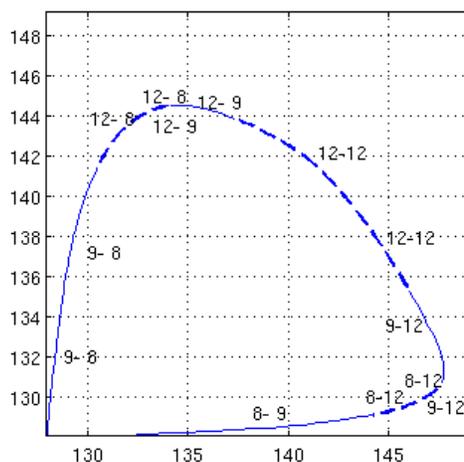


Figura 5.8: Conjunto  $\mathcal{E}(5)$  para el caso no simétrico en las señales, con  $\sigma_1 = 1,5$  y  $\sigma_2 = 0,5$ . Como antes, se muestran las acciones jugadas por cada jugador en el formato  $a_1 - a_2$ , y el cambio en el tipo de línea indica un cambio en el perfil jugado.

### 5.2.1. Comparación con el caso simétrico

Ya vimos que  $\mathcal{E}(r, \Sigma)$  deja de ser simétrico privilegiando al jugador que puede esconder su señal, pero no queda claro si el jugador beneficiado por esta asimetría, esta mejor o peor que el caso simétrico. Para responder esta pregunta se compara la solución obtenida en la sección anterior con dos casos simétricos (ver figuras 5.9 y 5.10), tras la comparación se llega a la conclusión que el jugador beneficiado esta efectivamente mejor que en el caso simétrico.

Si definimos la norma de  $\Sigma$  como

$$\|\Sigma\| := \sqrt{\sigma_1^2 + \sigma_2^2}$$

podemos comparar el caso asimétrico de antes, con un caso simétrico donde la matriz de varianza tiene la misma norma, esto lo vemos en 5.10, incluso en este escenario

se llega a la misma conclusión: ambos jugadores se ven afectados más allá de lo que podrían alcanzar en el caso simétrico.

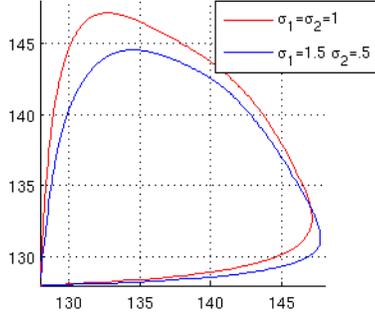


Figura 5.9: Comparación de  $\mathcal{E}(5, \sigma_1 = 1.5, \sigma_2 = 0.5)$  con  $\mathcal{E}(5, \sigma_1 = 1, \sigma_2 = 1)$ .

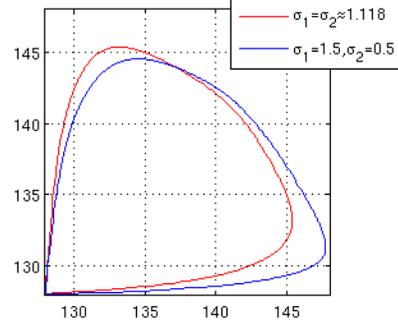


Figura 5.10: Comparación de  $\mathcal{E}(5, \sigma_1 = 1.5, \sigma_2 = 0.5)$  con  $\mathcal{E}(5, \sigma_1 \approx 1.118, \sigma_2 \approx 1.118)$ .

### 5.3. Asimetría en las señales

Otra forma de asimetría en la información es que las señales de ambos jugadores sean funcionalmente distintas. Por ejemplo, que la señal asociada al jugador 1 sea una función de su precio y del precio del otro, mientras la señal del jugador 2 se mantiene como su propio precio. Se podría pensar que en este caso el jugador 1, tal como en la sección anterior, podría esconder sus acciones y obtener dividendos de esto. Para analizar este fenómeno, consideramos como señales públicas, dado un perfil de estrategias  $A$ , a las siguientes variables estocásticas:

$$\begin{aligned} X_t^1(A) &= \int_0^t \frac{P_1(A_s) + P_2(A_s)}{2} ds + \int_0^t dZ_s^1 \\ X_t^2(A) &= \int_0^t P_2(A_s) ds + \int_0^t dZ_s^2 \end{aligned} \quad (5.4)$$

Es decir, ahora  $\mu = \left( \frac{P_1(A_s) + P_2(A_s)}{2}, P_2(A_s) \right)^T$ .

Resolvemos entonces  $\mathcal{E}(3)$  ( $\Sigma$  simétrico como en (5.2) con  $\sigma = 1$ ) considerando la estructura de señales anterior. Podemos ver los valores obtenidos en la figura 5.11. Al igual que en la sección anterior es posible observar que las acciones básicamente no cambian, pero sí lo hace la forma que tiene el conjunto  $\mathcal{E}(r)$ . Éste deja de ser simétrico para pasar a privilegiar, i.e. ofrecer mayores pagos, al jugador que puede esconder su señal mezclando su precio con el del otro jugador.

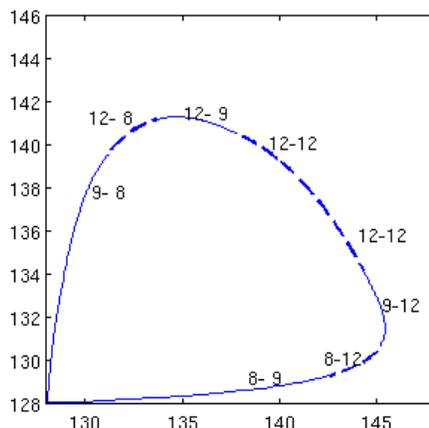


Figura 5.11: Conjunto  $\mathcal{E}(3)$  para el caso no simétrico con señales como (5.4).

### 5.3.1. Comparación con caso simétrico y caso asimétrico en la varianza

Vale la pena preguntarse cómo se compara este caso con el caso simétrico y también cómo se compara con el caso asimétrico de la sección anterior. Recordemos que en la sección anterior se mostró que a pesar que para un jugador se disminuían los pagos con respecto al caso simétrico, para el otro jugador se aumentaban los pagos, entonces es menester de esta sección revisar si el mismo fenómeno ocurre cuando la estructura de señales es como en (5.3).

Para comenzar se puede observar en la figura 5.12 que ambos jugadores se ven perjudicados por este cambio en la estructura de señales, restringiéndose considerablemente el conjunto  $\mathcal{E}(3)$ . Podemos seguir con la figura 5.13 donde se ve cuál es la diferencia entre el escenario asimétrico en las señales y el escenario donde son asimétricas *sólo* las aleatoridades de ellas, nuevamente observamos una gran disminución en el tamaño de  $\mathcal{E}(3)$ .

Ambas disminuciones las atribuimos al aumento en la volatilidad de las señales, notemos que al cambiar la estructura de señales (en particular, acabamos de cambiar la definición de  $\mu$ ), los vectores  $\phi(a, T)$  cambian, a pesar que las acciones no han cambiado, esto es directo al recordar que en las restricciones que se usan para encontrar dichos vectores se encuentra  $\mu$ . Más intuitivamente podemos decir que en este caso, dada la estructura de señales, se deben ofrecer más incentivos en términos de volatilidades para que los jugadores se comprometan con la estrategia, esto conlleva una gran disminución en el tamaño de  $\mathcal{E}$  pues mayor volatilidad indica absorción más rápida en el equilibrio de Nash. Podemos ver un gráfico similar al de la sección 5.1, en la figura 5.14, al compararlo con 5.5 notamos el gran aumento en la norma de los vectores que influyen sobre la volatilidad.

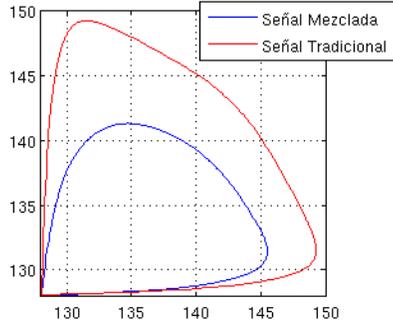


Figura 5.12: Comparación de  $\mathcal{E}(3)$  caso simétrico y caso asimétrico en las señales.

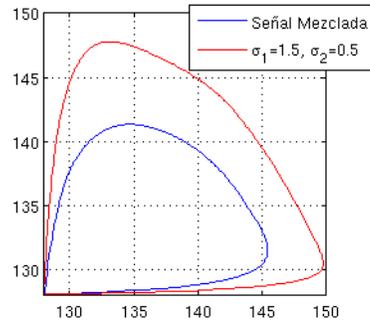


Figura 5.13: Comparación de  $\mathcal{E}(3)$  con señales como en (5.4) y  $\mathcal{E}(3, \sigma_1 = 1,5, \sigma_2 = 0,5)$ .

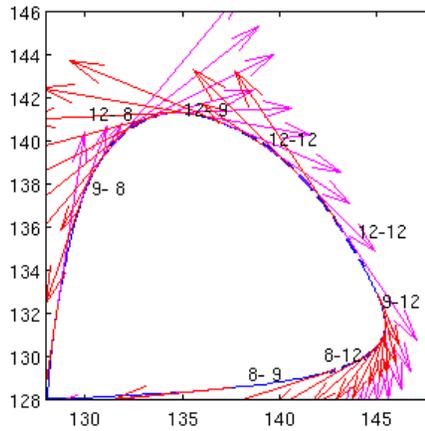


Figura 5.14: Conjunto  $\mathcal{E}(3)$  para el caso asimétrico, las flechas indican las volatilidades en algunos puntos, las flechas rojas indican la volatilidad asociada a la primera componente de la señal y las verdes, la asociada a la segunda. Las flechas se encuentran escaladas por la misma constante que en 5.5, por lo que ambos gráficos son comparables.

## 5.4. Asimetría en la complementaridad

Tal como se cambió la estructura de señales en la sección anterior para hacerla asimétrica, es posible hacer que los pagos sean asimétricos para los jugadores. En particular, nos interesará hacer que uno de los jugadores sea más beneficiado que el otro por la complementaridad de los bienes, en efecto, si los precios son de la forma

$$\begin{aligned} P_1(a) &= B - Ca_1 + \gamma_1 a_2 \\ P_2(a) &= B - Ca_2 + \gamma_2 a_1 \end{aligned} \quad (5.5)$$

estudiaremos que ocurre si  $\gamma_1 \neq \gamma_2$ . Notar que las cantidades críticas ahora dependen del jugador ya que los precios, y con ellos los pagos, dejaron de ser simétricos, algunos cálculos simples nos llevan a los siguientes valores:

<i>Situación</i>	<i>Jugador 1</i>	<i>Jugador 2</i>
Equilibrio de Nash	$a_N^1 = \frac{B(\gamma_1 + 2C)}{(2C)^2 - \gamma_2\gamma_1}$	$a_N^2 = \frac{B(\gamma_2 + 2C)}{(2C)^2 - \gamma_2\gamma_1}$
Monopolio	$a_M^1 = \frac{B}{2C - \gamma_1 - \gamma_2}$	$a_M^2 = \frac{B}{2C - \gamma_1 - \gamma_2}$
Desviación óptima desde monopolio	$a_D^1 = \frac{B}{2C} \frac{2C - \gamma_2}{2C - \gamma_1 - \gamma_2}$	$a_D^2 = \frac{B}{2C} \frac{2C - \gamma_1}{2C - \gamma_1 - \gamma_2}$

Con lo anterior en mente, definiremos los conjuntos de acciones para los jugadores como  $\mathcal{A}_i = \{a_N^i, a_D^i, a_M^i\}$ . Las señales y los pagos no cambiarán de estructura pero sí de valores, ya que dependen del precio. Inspirados en los cálculos anteriores usaremos  $B = 24, C = 2, \gamma_2 = 1$  y *moveremos*  $\gamma_1$  para observar el fenómeno que nos interesa, para hacer explícito lo anterior a lo largo de esta sección en vez de escribir  $\mathcal{E}(r)$  escribimos  $\mathcal{E}(r; \gamma_1)$ .

En la figuras 5.15 y 5.17 es posible ver los pagos que se pueden alcanzar utilizando  $\gamma_1 = 0,8$  y  $\gamma_1 = 1,2$ , respectivamente. Al costado de ambas imágenes (figuras 5.16 y 5.18), es posible ver los conjuntos  $\mathcal{V}^*$  para cada caso; vale la pena destacar que en ambos escenarios, el conjunto  $\mathcal{V}^*$  se construye asegurándole a cada uno de los jugadores el pago alcanzado en el único equilibrio de Nash del juego estático, tal como en el caso simétrico (ver gráfico 5.2). En las figuras nombradas más arriba es posible hacer las primeras conclusiones respecto al efecto de la asimetría de la complementaridad, podemos resumirlas en los siguientes puntos:

1. El jugador que posee el mayor  $\gamma$  puede alcanzar mayores pagos, esto es de esperar pues para quién tiene mayor  $\gamma$ , las acciones del otro jugador lo benefician más, entregándole un mayor precio.
2. Entre mayor sean ambos  $\gamma$ 's, mayor es el pago que reciben ambos jugadores en colusión. Esto ocurre pues, como ya se dijo, para el jugador  $i$  mayores valores de  $\gamma_i$  significa que las acciones del otro jugador pueden aumentar en mayor

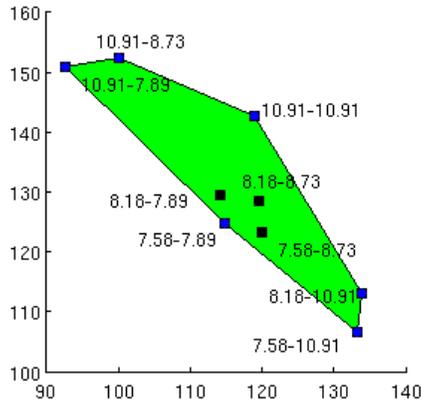


Figura 5.15: Pagos instantáneos que se pueden alcanzar cuando  $\gamma_1 = 0,8$  y  $\gamma_2 = 1$ .

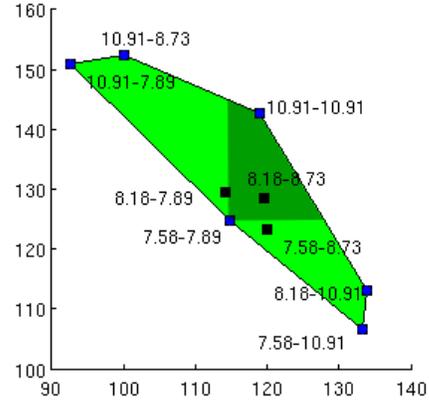


Figura 5.16:  $\mathcal{V}^*$  se muestra en color más oscuro. Este conjunto se obtiene al asegurar a cada jugador su pago mínimo, en este caso el pago del equilibrio de Nash  $(a_N^1, a_N^2)$ .

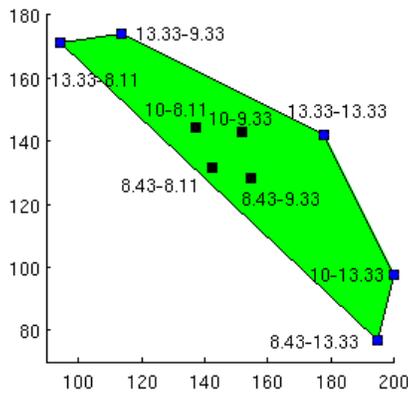


Figura 5.17: Pagos instantáneos que se pueden alcanzar cuando  $\gamma_1 = 1,2$  y  $\gamma_2 = 1$ .

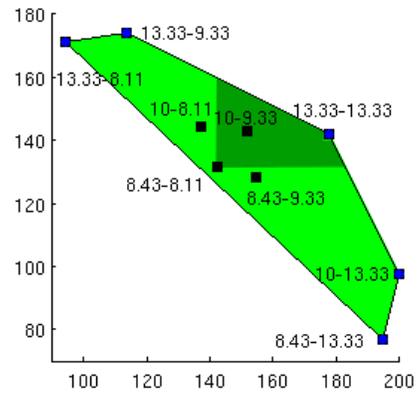


Figura 5.18:  $\mathcal{V}^*$  se encuentra en color más oscuro. El conjunto en cuestión se obtiene al asegurar a cada jugador su pago mínimo, en este caso el pago del equilibrio de Nash  $(a_N^1, a_N^2)$ .

medida el precio, incluso puede llegar a quitar el efecto negativo sobre su precio que tiene la producción propia de  $i$ .

En 5.19 y 5.20 se muestran  $\mathcal{E}(5; 0,8)$  y  $\mathcal{E}(5; 1,2)$ , respectivamente. Podemos notar que la dinámica no cambia con respecto a los perfiles jugados, en efecto, los jugadores comienzan jugando  $(a_M^1, a_M^2)$ , luego al salir de la *zona de confianza*, juegan  $(a_M^1, a_D^2)$  o bien  $(a_D^1, a_M^2)$ . Por simplicidad supongamos, sin pérdida de generalidad, que los jugadores siguen moviéndose hacia la izquierda, luego jugarán  $(a_M^1, a_N^2)$  para finalmente terminar jugando  $(a_N^1, a_N^2)$ . Es claro que esta es la dinámica que hemos visto en todos nuestros ejemplos, sólo que antes los superíndices no eran necesarios, pues las acciones eran las mismas para ambos jugadores.

Volviendo a las figuras, al igual que al ver los pagos estáticos, se concluye que quién se beneficia más de las acciones del otro jugador, alcanza los mayores pagos. Algo que no era esperado es que cuando la complementaridad total (dada por  $\gamma_1 + \gamma_2$ ) se empequeñece, no sólo los pagos estáticos se achican, si no,  $\mathcal{E}(r; \gamma_1)$  también se hace más pequeño en relación a  $\mathcal{V}^*$ ; es posible observar esto al comparar la figura 5.19 con 5.16 y 5.20 con 5.18, en el primer caso  $\mathcal{E}$  alcanza sólo una pequeña fracción de  $\mathcal{V}^*$ , en cambio, en el segundo caso,  $\mathcal{E}$  alcanza casi todo el conjunto  $\mathcal{V}^*$ .

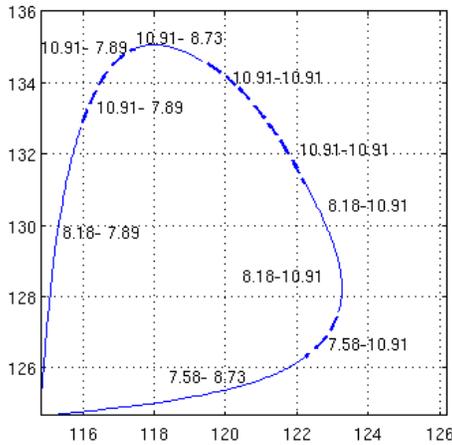


Figura 5.19:  $\mathcal{E}(5)$  cuando la complementaridad beneficia más al jugador 2 y  $\max_i \gamma_i \leq 1$ , a saber  $\gamma_1 = 0,8$ .

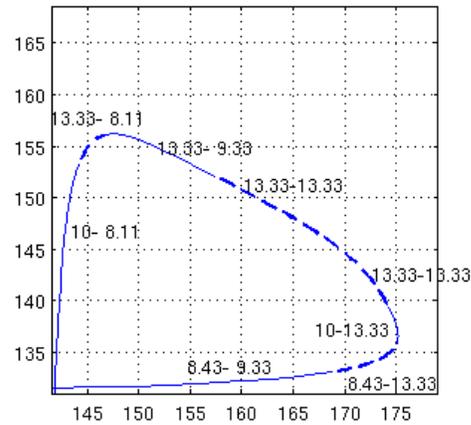


Figura 5.20:  $\mathcal{E}(5)$  cuando la complementaridad beneficia más al jugador 1 y  $\max_i \gamma_i > 1$ , a saber  $\gamma_1 = 1,2$ .

Argumentamos que lo anterior se debe a que al disminuir  $(\gamma_1 + \gamma_2)$ , es decir cuando la influencia de un jugador sobre el otro disminuye, al jugador que no esta jugando su óptimo, le interesan menos las acciones del otro jugador, pues su precio no está tan influenciado por ellas, esto se traduce en que los incentivos, en términos de cambio de la volatilidad, deben ser mayores en este caso. Lo anterior conlleva, como en los ejemplos anteriores, a una absorción más rápida en  $\mathcal{N}$  lo que disminuye

los pagos esperados de ambos jugadores. Es posible ver las volatilidades para ambos valores de  $\gamma_1$  en las figuras 5.21 y 5.22.

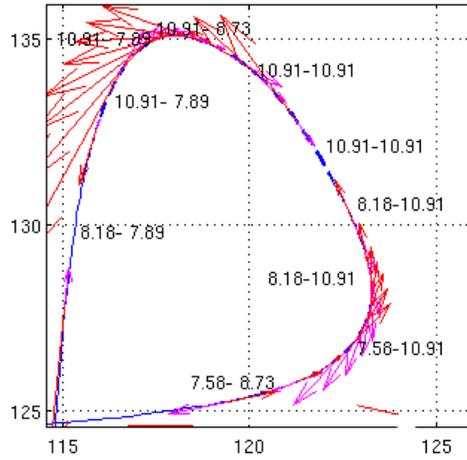


Figura 5.21:  $\mathcal{E}(5; 0,8)$ . Las flechas indican las volatilidades en algunos puntos, las flechas rojas indican la volatilidad asociada a la primera componente de la señal y las verdes, la asociada a la segunda. Las flechas se encuentran escaladas por la misma constante que en 5.5, por lo que ambos gráficos son comparables.

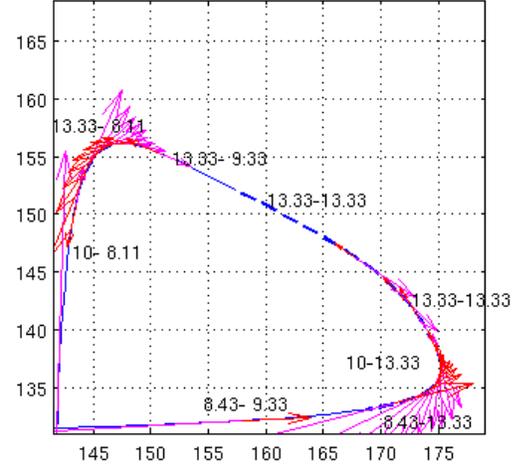


Figura 5.22:  $\mathcal{E}(5; 1,2)$ . Las flechas indican las volatilidades en algunos puntos, las flechas rojas indican la volatilidad asociada a la primera componente de la señal y las verdes, la asociada a la segunda. Las flechas se encuentran escaladas por la misma constante que en 5.5, por lo que ambos gráficos son comparables.

# Capítulo 6

## Conclusiones

Al final de este Trabajo de Título concluimos que el modelo presentado por Sannikov [5] es un modelo flexible que permite estudiar diversas situaciones, aunque sólo en juegos que involucran dos jugadores. De la misma manera, el modelo es interesante por la gran cantidad de información que uno puede obtener de la situación estudiada, información en el sentido de *drifts*, de volatilidades, de las acciones tomadas y de los pagos ofrecidos a cada jugador.

De las comparaciones realizadas en el capítulo 5, podemos conjeturar que los resultados tradicionales de juegos repetidos también se pueden obtener en el contexto continuo. En efecto, pudimos ver que entre más paciente es el jugador, más grande es el conjunto de equilibrios que se pueden alcanzar, es decir, tal como es natural esperar, mientras menos nos interese el futuro menos nos interesa emprender empresas en conjunto.

Otro efecto interesante estudiado, es el efecto de la información, pudimos ver que entre menos información sobre las acciones entrega la señal, menos equilibrios son posibles de alcanzar. Esto también es natural de esperar, pues entre menos información tengamos sobre nuestro oponente/compañero, hay menos sobre lo que tener cualquier acuerdo.

### 6.1. Preguntas Abiertas

Una primera primera pregunta que debería resolverse es ¿existe alguna forma eficiente de resolver el modelo a tiempo continuo?. La pregunta es importante ya que si  $n$  es la cantidad de acciones que ambos jugadores pueden tomar y sólo existe un Nash estático, entonces se deben resolver  $(n - 1)^2$  problemas de optimización ¡por cada punto en que se resuelve la ecuación!

En relación a la pregunta anterior, notemos que lo que hacemos es un problema inverso, en el sentido de que buscamos una observación (e.g. un punto en la diagonal)

de manera que la ecuación resuelta con esa condición inicial alcance otra observación que tenemos apriori (e.g.  $\mathcal{N}$ ). Lo anterior es de gran interés porque permitiría una manera, quizás, más rápida y con menos error de encontrar el conjunto  $\mathcal{E}(r)$ .

Una pregunta de índole matemática sería ¿es posible recuperar analíticamente los resultados acerca del efecto de la información e impaciencia obtenidos? En la misma dirección, ¿será posible recuperar los resultados que se obtuvieron cuando la situación era asimétrica?

Por último, económicamente parece interesante resolver algunos problemas particulares a través de modelos parecidos a los presentados en las distintas secciones del capítulo 5. Por ejemplo, ¿cómo se relacionan una empresa proveedora de un gigante de computadoras y dicho gigante? Para tener una idea, recordamos lo presentado en 5.3 donde la señal acerca de un jugador era el promedio entre el precio de un jugador y el otro, en efecto, si la empresa A le vende discos duros a la empresa B, sólo es posible ver el precio al cual la empresa A vende sus discos y al cual la empresa B vende el computador *junto* al disco duro que la empresa A le vendió. En un contexto menos monetario, también podríamos estudiar el caso en que un niño es muy impredecible y el padre no sabe que hacer para que él este feliz (y con él, el padre también), esto nos lleva a la sección 5.2 donde una señal es más ruidosa que la otra ya que el padre nunca tendrá claridad acerca del niño, pues este último, si es muy pequeño, no puede comunicarse claramente.

También en la línea económica, ¿será posible extender los resultados presentados aquí a más jugadores? Los resultados ocupan en gran parte la geometría que es posible obtener en  $\mathbb{R}^2$ , por lo que, la extensión a  $\mathbb{R}^3$  sería la más natural, aunque en rigor no es claro cómo se podría hacer. Desde  $\mathbb{R}^3$  a  $\mathbb{R}^n$ , probablemente otro tipo de herramientas sería necesario, pero vale la pena hacer el intento pues el espectro de nuevas situaciones que se podrían estudiar crecería enormemente.

# Bibliografía

- [1] D. Abreu, D. Pearce, and E. Stacchetti. Toward a theory of discounted repeated games with imperfect monitoring. *Econometrica*, 58(5):1041–1063, 1990. [17](#), [29](#)
- [2] P. DeMarzo and Y. Sannikov. Optimal security design and dynamic capital structure in a continuous-time agency model. *Journal of Finance*, 2006. [29](#)
- [3] G. J. Mailath and Larry Samuelson. *Repeated Games and Reputations. Long-Run Relationships*. Oxford University Press, 2006. [3](#), [9](#), [20](#), [24](#)
- [4] J. Nash. Non-cooperative games. *Annals of Mathematics*, 54:289–295, 1951. [5](#)
- [5] Y. Sannikov. Games with imperfectly observable actions in continuous time. *Econometrica*, 75(5):1285–1329, 2007. [29](#), [35](#), [37](#), [43](#), [44](#), [46](#), [54](#), [67](#)
- [6] Y. Sannikov. A continuous-time version of the principal-agent problem. *Review of Economic Studies*, 75(3):957–984, 07 2008. [29](#)
- [7] Y. Sannikov and A Skrzypacz. The role of information in repeated games with frequent actions. *Econometrica*, 78, 2010. [29](#)
- [8] M. Sipser. *Introduction to the theory of computation*. Thomson Course Technology, 2006. [13](#)

# Índice alfabético

- accesible, estado, 14
- acción, 3
  - mixta, 4
  - vinculable, 18, 28, 35
- autómata, 13
- Conjunto
  - autogenerante, 18, 28, 35
  - de historias, 7, 24
  - de historias públicas, 24
  - de los pagos alcanzables, 4, 32
  - de los pagos para las etapas, 4
  - de los pagos racionales, 6, 32
- descomposición, 18, 28
- Desviación
  - simple, 11
  - simple rentable, 11
- Dilema del Prisionero
  - información completa, 6
  - información incompleta, 23
- Equilibrio
  - de Nash, 5, 9
  - público perfecto, 25, 32
  - perfecto en subjuegos, 11
- Estrategia, 3
  - de comportamiento, 7, 24
  - de comportamiento pública, 24
  - de continuación, 8
  - mixta, 7, 24
  - pública, 30
  - pura, 7, 24
- etapa, 3
- historia, 7, 24
  - pública, 23
- Identificabilidad de a pares, 31
- juego continuado, 7
- Mejor Respuesta, 4
  - en acciones mixtas, 5
- Pago
  - de la etapa, 4
  - descontado, esperanza del, 31
  - dominante, 4
  - eficiente, 4
  - ex ante, 23
  - ex post, 23
  - individualmente racional, 5
  - instantáneo, esperanza del, 30
  - minmax, 5, 32
- perfil, 3, 4
  - de estrategias, 8
  - minmax, 5
- proceso progresivamente medible, 30
- promesa, 15
- resultado, 8
- valor de continuar, 31
- vector vinculante, 36