Online Appendix to: The role of information in repeated games with frequent actions.

Yuliy Sannikov and Andrzej Skrzypacz

Appendix O-A: Constructing polygon approximation of *M*

In this appendix we construct a polygon that approximates M.

Lemma O-A: For any ε there exists a finite collection of directions $\{\tilde{N}_1, \tilde{N}_2 ..., \tilde{N}_L\}$ such that the distance between M and any point of the polygon

$$P = \bigcap_{l=1}^{L} H(\tilde{N}_l)$$

(which contains M) is at most $\varepsilon/2$.

Proof: Start with any direction N_0. Now, parameterize all directions going clockwise from N_0 back to N_0 by $p\in[0, 2\pi]$. Let K be a positive integer, and define N_k to be a direction corresponding to $p_k = 2\pi k/K$ for $k = 1, \dots K$. For each k, choose a point v_k on the boundary of M with the normal vector N_k. If M has no kink at v_k, let N_k be the unique normal vector at v_k. Draw a tangent line at v_k, and let H_k be the half-space containing M bounded by this tangent line. (Note that H_k may not be the same as $H(N_k)$: it may be strictly smaller – as at point v_k = (2,3) in our example from Section 4 with Brownian signals. However H_k is approximated arbitrarily closely by $H(N')$ with N' close to N_k (close in the Housdorff metric computed after intersecting the half-spaces with a circle of a large radius around the set of feasible payoffs). Indeed, since v_k is on the boundary of M, there is a sequence $H(N_m')$ such that the distance between v_k and the boundary of $H(N_m')$ converges to 0. Since M has no kink at v_k and since $M \subseteq H(N_m')$, it follows that $N_m' \to N_k$). If M has a kink at v_k, draw two extreme tangent lines through v_k, and let H_k' and H_k'' be the two half-spaces containing M bounded by the two tangent lines.

Consider the polygon P', which is an intersection of half-spaces H_k (if there is no kink at v_k) or H_k' and H_k'' (if there is a kink at v_k) for each k. Let us show that the distance between any point of P' and M decreases uniformly to 0 as K increases. Indeed, note that the points of P' furthest away from M are vertices. Indeed, while moving along any side of P' away from v_k, the distance to M weakly increases. Furthermore, any vertex w of P' either coincides with a kink point of M (i.e. $w = v_k$) or has an the angle greater than π - $2\pi/K$. In the latter case, consider the points v_k and v_{k+1} of M on the sides of P' adjacent to w. The distance between points v_k and v_{k+1} is at most $\overline{V}$. Because $\angle v_{k+1} w v_k$ is an obtuse angle, $v_k v_{k+1}$ is the longest segment of the triangle $v_{k+1} w v_k$, and so the segments $w v_k$ and $w v_{k+1}$ are of length less than $\overline{V}$. Because one of the angles $\angle w v_k v_{k+1}$ or $\angle w v_{k+1} v_k$ is less than π/K, the distance from w to the segment $v_k v_{k+1}$ is at most $\sin(\pi/K)\,\overline{V}$. This is an upper bound on the distance from any point on the boundary of P' to M. As K increases, this bound converges to 0.

Now, note that each side of P' (with a normal vector N) can be approximated arbitrarily closely by the hyperplane corresponding to $H(N')$ with N' close to N. Thus, the polygon P' can be approximated arbitrarily closely as an intersection of hyperplanes

$$P = \bigcap_{l=1}^{L} H(\tilde{N}_l)$$

This completes the proof. QED.

Appendix O-B: Destroying value with Brownian signals.

We would like to show that

$$\int (\omega(x,0)\cdot N)(f_a(x)-f_{a'}(x))dx \leq O(\Delta^{1.49999})$$

whenever

$$\omega(x,0)\cdot N \in [-\bar{V},0] \quad \text{and} \quad |E[\omega\cdot N \mid a]| \leq O(\Delta).$$

We adapt the arguments of Sannikov and Skrzypacz (2007) to prove this claim. Lemma O-B1, which is analogous to Lemma 3 from Sannikov and Skrzypacz (2007) shows that the solution to this problem involves a tail test, which triggers a punishment if and only if the likelihood ratio $f_{a'}(x)/f_a(x)$ becomes sufficiently high. Thereafter, Lemma O-B2 (which generalizes Lemma 2 from Sannikov and Skrzypacz (2007)) implies that a tail test that destroys value on the order of Δ per period creates incentives on the order of at most $\Delta^{1.49999}$.

Lemma O-B1. Suppose $D > 0$. Consider the problem

$$\max \int v(x)(f_a(x)-f_{a'}(x))dx$$

$$\text{s.t.} \quad \forall x \in \Re, v(x) \in [-\bar{V},0] \quad \text{and} \quad \int_{-\infty}^{\infty} v(x) f_a(x)dx \leq D\Delta.$$

The solution of this problem takes the form of a "tail test," i.e.

$$v(x) = \begin{cases} 0 & \text{if} f_a(x)/f_{a'}(x) > c \Leftrightarrow x\cdot(\mu(a)-\mu(a')) > c' \\ -\bar{V} & \text{if} f_a(x)/f_{a'}(x) \leq c \Leftrightarrow x\cdot(\mu(a)-\mu(a')) \leq c' \end{cases}$$

for some c and c'.

Proof: Write the Lagrangian for the maximization problem

$$L = \int_{-\infty}^{\infty} v(x)(f_a(x)-f_{a'}(x))dx + \rho_0\left(\int_{-\infty}^{\infty} v(x) f_a(x)dx - D\Delta\right)$$
$$+\int_{-\infty}^{\infty} \rho_1(x)(v(x)+\bar{V})dx - \int_{-\infty}^{\infty} \rho_2(x)v(x)dx,$$

where $\rho_1(x) > 0$ only if $v(x) = -\bar{V}$ and $\rho_2(x) > 0$ only if $v(x) = 0$. Taking first-order conditions with respect to $v(x)$ gives

$$f_a(x) - f_{a'}(x) + \rho_0 f_a(x) + \rho_1(x) - \rho_2(x) = 0$$

It follows that

$$\begin{cases} v(x) = 0 & \text{and } \rho_2(x) > 0 \quad \text{if } f_a(x) - f_{a'}(x) + \rho_0 f_a(x) > 0 \\ v(x) = -\bar{V} & \text{and } \rho_1(x) > 0 \quad \text{if } f_a(x) - f_{a'}(x) + \rho_0 f_a(x) < 0 \end{cases}$$

We have

$$f_a(x) - f_{a'}(x) + \rho_0 f_a(x) < 0 \Leftrightarrow \frac{f_{a'}(x)}{f_a(x)} > 1 + \rho_0$$

Now,

$$\frac{f_{a'}(x)}{f_a(x)} = \exp\left(\frac{-(x-\Delta\mu(a'))^2+(x-\Delta\mu(a))^2}{2\Delta}\right) = \exp\left(x(\mu(a')-\mu(a))+\tfrac{\Delta}{2}(\mu^2(a)-\mu^2(a'))\right)$$

so whether the ratio is larger or smaller than $1+\rho_0$ depends only on whether $x(\mu(a) - \mu(a'))$ is above or below a threshold Therefore, the solution to the optimization problem above takes the conjectured form. QED

To evaluate the efficiency of tail tests, we can assume that x is one-dimensional, because the likelihood ratio $f_{a'}(x)/f_a(x)$ stays constant along directions orthogonal to the line connecting $\mu(a)$ and $\mu(a')$.

The following lemma is a generalization of Lemma 2 from Sannikov and Skrzypacz (2007); the proof is virtually identical to the proof there.

Lemma O-B2. Fix $C_1 > 0$, $\varepsilon > 0$, $k>0$ and $\mu - \mu' > 0$. Consider a tail test with a critical region $(-\infty, c]$ of the hypothesis that $x \sim N(\Delta\mu, \Delta)$ against an alternative that $x \sim N(\Delta\mu', \Delta\sigma^2)$. Denote by $g(x)$ and $g'(x)$ the densities of these two distributions respectively. There exists a constant $C_2 > 0$ such that if the likelihood difference of this test is bigger than or equal to $C_1\Delta^k$, i.e.

$$\int_{-\infty}^{c} (g'(x)-g(x))dx \geq C_1\Delta^k$$

then for small Δ, the probability of a false positive associated with this test is

$$\int_{-\infty}^{c} g(x)dx \geq C_2\Delta^{k-1/2+\varepsilon}.$$

Proof. Without loss of generality, $c < \mu$ (or else the probability of a false positive is at least ½). There are two ways of representing the likelihood difference of a tail test on a graph, as shown on Figure O1.

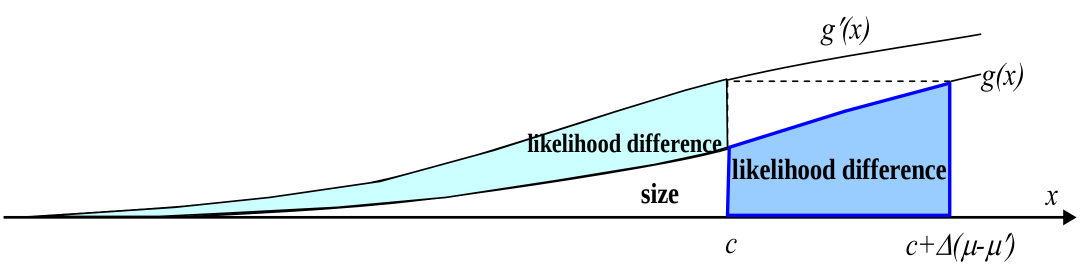

Figure O1

Using the area to the right of c, we see that there exists $x^* \in (c, c+\Delta(\mu-\mu'))$ such that

$$\text{likelihood difference } = \Delta(\mu - \mu')g(x^*) = \frac{\Delta(\mu-\mu')}{\sqrt{2\pi\Delta}}\exp\left(-\frac{(x^*-\mu)^2}{2\Delta}\right)$$

Then if the likelihood difference is greater than or equal to $C_1\Delta^k$:[1]

$$\frac{\Delta(\mu-\mu')}{\sqrt{2\pi\Delta}}\exp\left(-\frac{(x^*-\Delta\mu)^2}{2\Delta}\right) \geq C_1\Delta^k \quad \Rightarrow \quad \Delta\mu - x^* = \sqrt{-2\Delta\log\left(\sqrt{2\pi}\Delta^{k-1/2}\frac{C_1}{(\mu-\mu')}\right)}$$

Let $\alpha > 1$ be a number to be specified later. Let y^* satisfy $(\Delta\mu - y^*) = \alpha(\Delta\mu - x^*)$. Because $x^* - \Delta(\mu - \mu') < c$, the probability of making type I error (i.e. the size of the test) is greater than the shaded area in the Figure O2, i.e.

$$(x^* - \Delta(\mu'-\mu) - y^*)g(y^*) = \left((\alpha-1)(\Delta\mu - x^*) - \Delta(\mu-\mu')\right)\frac{1}{\sqrt{2\pi\Delta}}\exp\left(-\frac{\alpha^2(x^*-\Delta\mu)^2}{2\Delta}\right) =$$

$$\left((\alpha-1)\sqrt{-2\Delta\log\left(\sqrt{2\pi}\Delta^{k-1/2}\frac{C_1}{(\mu-\mu')}\right)} - \Delta(\mu-\mu')\right)\frac{1}{\sqrt{2\pi\Delta}}\left(\frac{\sqrt{2\pi}C_1\Delta^{k-1/2}}{(\mu-\mu')}\right)^{\alpha^2} > O\left(\Delta^{\alpha^2(k-1/2)}\right)$$

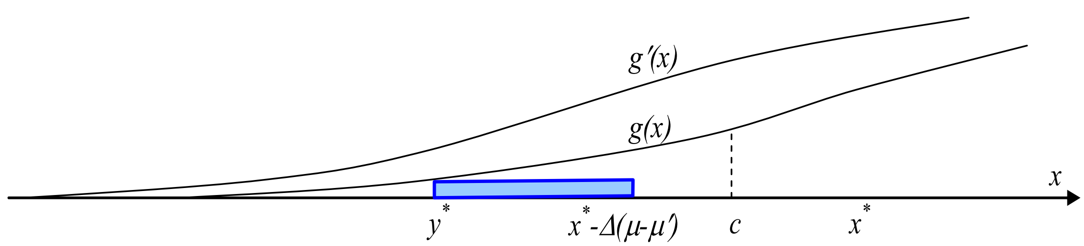

Figure O2

Taking α sufficiently close to 1 proves the Lemma. QED

[1] From now on we take $x^* < \Delta\mu'$. Otherwise, the size is strictly positive as $\Delta \to 0$ and the Lemma is trivially satisfied.

Appendix O-C: Proof of Theorem 2.

Here we prove a technical Proposition O-C1, used in the proof of Theorem 2 (the rest of the proof follows FLM).

Proposition O-C1. *Suppose W is a smooth set in the interior of M_{-}. For any v on the boundary of W there exists a neighborhood $N_\delta(v)$ of v with radius δ_v, a discount rate r_v and period length Δ_v such that any extreme point of W in this neighborhood* ($w \in \text{ext } W \cap N_\delta(v)$) *is generated by W for all discount rates and period lengths not exceeding r_v and Δ_v.*

***Proof of Proposition* O-C1:** Let T and N be unit tangent and normal vectors to W at v. Consider the problem of generating $w \in \text{ext } W$ in a neighborhood of v. By definition (see Section 5.1), w is generated by W if there is an admissible pair (a,ω), i.e. current-period action profile a and a map $\omega(x,(j_y))$ from signals to continuation-value transitions, that satisfies the feasibility constraint $w + \omega(x,(j_y)) \in W$, the promise-keeping constraint

$$w = g(a) + \frac{e^{-r\Delta}}{1-e^{-r\Delta}} E[\omega(x,(j_y)) \mid a] \tag{O.1}$$

and the incentive compatibility constraints:

$$(g_i(a) - g_i(a')) + \frac{e^{-r\Delta}}{1-e^{-r\Delta}} (E[\omega_i(x,(j_y)) \mid a] - E[\omega_i(x,(j_y)) \mid a']) \geq 0 \tag{O.2}$$

for any a' (such that $a_j' = a_j$ and $a_i' \in A_i$).

We first show a natural way to construct an *approximately* admissible pair, so that promise-keeping holds, incentive compatibility (O.2) hold strictly, and the feasibility condition is nearly satisfied (all for small r and Δ). Then we show how to adjust ω in such a way that all three constraints hold.

Following the definition of $M(\varepsilon)$, define $D(N,\varepsilon)$ as the solution to the linear-programming problem:

$$D(N,\varepsilon) = \max_{a,\beta,d(y)} \left(g(a) + \sum_{y \in Y} d(y)\lambda(y \mid a) \right) \cdot N \text{ s.t. } d(y) \cdot N \leq 0 \tag{O.3}$$

and $$g_i(a) - g_i(a') + \beta(\mu(a) - \mu(a'))T_i + \sum_{y \in Y} d_i(y)(\lambda(y \mid a) - \lambda(y \mid a')) + \varepsilon \geq 0, \tag{ICε}$$

Since v is an extreme point of W, which is in the interior of M_- (recall that $M\_$ is the limit of $M(\varepsilon)$ as we take ε to zero from below), there exists a constant $\varepsilon<0$ such that for all $w \in W$:

$$w \cdot N \leq v \cdot N \leq D(N, \varepsilon) + \varepsilon \,. \qquad \text{(O.4)}$$

Let $\{a, \beta, d\}$ be the instruments that solve the maximization problem (O.3). Fix a and consider the following construction of ω from the β and d:

$$\omega(x,(j_y)) = \frac{1-e^{-r\Delta}}{\Delta e^{-r\Delta}}\left(\beta x T + \sum_y j_y d(y)\right) + \varpi \qquad \text{(O.5)}$$

for some $\varpi \in R^2$. See Figure O3 for illustration of such ω. With such continuation payoffs the incentive compatibility (O.2) constraints hold strictly for sufficiently small Δ (independently of w):

$$(g_i(a) - g_i(a')) + \frac{e^{-r\Delta}}{1-e^{-r\Delta}}(E[\omega_i \mid a] - E[\omega_i \mid a'])$$
$$= g_i(a) - g_i(a') + \beta(\mu(a) - \mu(a'))T_i + \sum_y d_i(y)\big(\lambda(y \mid a) - \lambda(y \mid a')\big) + O(\Delta) \geq -\varepsilon$$

The promise-keeping constraint is satisfied by an appropriate choice of ϖ.
However, the feasibility constraint does not hold for at least two reasons. First, since x and j_y are unbounded, the payoffs are unbounded. Second, because the set W is curved, we cannot move on a tangent at v, as then continuation payoffs get outside the set.

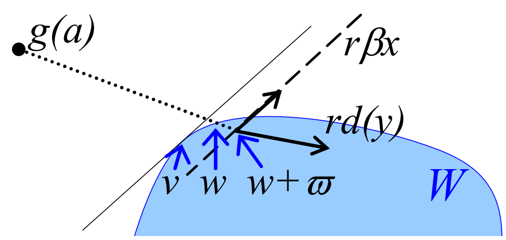

Figure O3: First construction of ω and violation of feasibility constraint.

Nevertheless, since T is nearly the tangent vector to the boundary in a neighborhood of v, we are able to adjust ω to satisfy feasibility, without violating promise-keeping and affecting IC by less than ε. In fact, we adjust ω by (1) cutting off the tails of x, (2) ignoring multiple arrivals of the jumps and (3) forcing the payoffs inside the curvature of W.

We need to introduce two functions to be able to adjust ω for the curvature of W. In the tangential and normal coordinates near point v, let $f(\theta)$ be the parameterization of the boundary of W, where θ represents the tangential coordinate and the point $(0, f(0))$ corresponds t o v. Fix a small positive constant α, to be specified later. Because W is a smooth set, for any $\alpha > 0$ we can find $\delta > 0$ such that for all $\theta \in (-\delta, \delta)$, $|f'(\theta)| \leq \alpha$ and there exists a constant κ such that $|f''(\theta)| \leq \kappa$. Then for all $\theta,\ \vartheta\ \in (-\delta/2, \delta/2)$,

$$f(\theta + \vartheta) \geq f(\theta) + f'(\theta)\vartheta - \tfrac{\kappa}{2}\vartheta^2.$$

This inequality is illustrated in Figure O4.

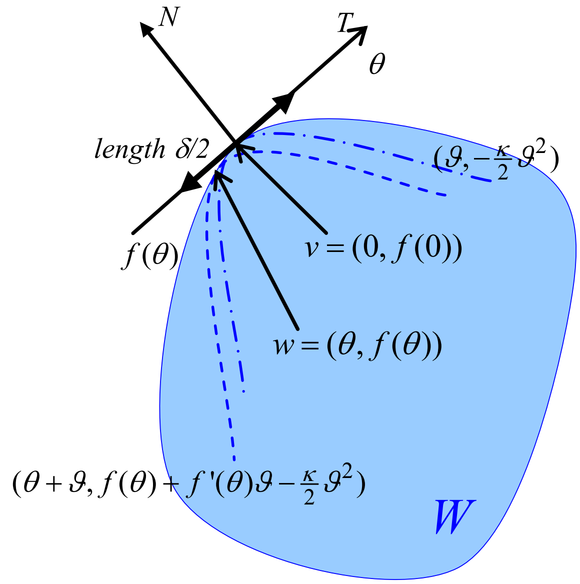

Figure O4: Set W and the points $(\theta+\vartheta, f(\theta)+f'(\theta)\vartheta-\frac{\kappa}{2}\vartheta^2)$

Moreover, it is convenient to use the following inequalities for all θ, $\vartheta_1+\vartheta_2+\vartheta_3 \in (-\delta/2, \delta/2)$:

$$f'(\theta)(\vartheta_1+\vartheta_2+\vartheta_3)-\tfrac{\kappa}{2}(\vartheta_1+\vartheta_2+\vartheta_3)^2 \geq f'(\theta)(\vartheta_1+\vartheta_2+\vartheta_3)-\tfrac{3\kappa}{2}(\vartheta_1^2+\vartheta_2^2+\vartheta_3^2) \Rightarrow$$
$$f(\theta+\vartheta_1+\vartheta_2+\vartheta_3) \geq f(\theta)+h_\theta(\vartheta_1)+h_\theta(\vartheta_2)+h_\theta(\vartheta_3)$$

where $h_\theta(\vartheta)=f'(\theta)\vartheta-\frac{3\kappa}{2}\vartheta^2$.

Note that $(\theta+\vartheta_1+\vartheta_2+\vartheta_3, f(\theta)+h_\theta(\vartheta_1)+h_\theta(\vartheta_2)+h_\theta(\vartheta_3)-\varpi_n) \in W$ for all $\varpi_n \in [0, A_\delta]$ for some strictly positive constant A_δ (and all θ, $\vartheta_1+\vartheta_2+\vartheta_3 \in (-\delta/2, \delta/2)$). We will use functions h_θ to introduce normal components of ω for every tangential component to guarantee that feasibility constraint is satisfied.

We now claim that for any such $\theta \in (-\delta/2, \delta/2)$, the payoff pair w with coordinates (θ, $f(\theta)$) is generated by W (and hence the radius can be picked to be $\delta_v=\delta/2$) if α is chosen appropriately small. Let $1_{|x|\leq c}$ denote the event that |x|<c and 1_y denote the event that exactly one jump arrives (of type y). Consider:

$$\omega(x,(j_y))=\tfrac{1-e^{-r\Delta}}{\Delta e^{-r\Delta}}\left((1_{|x|\leq\Delta^{1/3}}\beta x)T+\sum_y 1_y d(y)+\varpi_T T\right)$$
$$+h_\theta(\tfrac{1-e^{-r\Delta}}{\Delta e^{-r\Delta}}(1_{|x|\leq\Delta^{1/3}}\beta x))N+\sum_y 1_y h_\theta(\tfrac{1-e^{-r\Delta}}{\Delta e^{-r\Delta}}d(y)\cdot T)N+h_\theta(\tfrac{1-e^{-r\Delta}}{\Delta e^{-r\Delta}}\varpi_T)N-\varpi_N N \quad \text{(O.6)}$$

Note that we have truncated the linear βx at $|x|\leq\Delta^{1/3}$, used h_θ to assure that the tangential transfers are accompanied by the necessary normal transfers of payoffs to stay within W. Finally, constants ϖ_T and ϖ_N are given by the promise-keeping constraint:

$$\varpi_T = \Delta(w - g(a))\cdot T - E\left[(1_{|x|\le\Delta^{1/3}}\beta x) + \sum_y 1_y d(y)\cdot T \mid a\right]$$

$$\varpi_N = \tfrac{1-e^{-r\Delta}}{e^{-r\Delta}}(g(a) - w + \tfrac{1}{\Delta}E[\sum_y 1_y d(y) \mid a])\cdot N +$$

$$E\left[h_\theta(\tfrac{1-e^{-r\Delta}}{\Delta e^{-r\Delta}}(1_{|x|\le\Delta^{1/3}}\beta x)) - \sum_y 1_y h_\theta(\tfrac{1-e^{-r\Delta}}{\Delta e^{-r\Delta}} d(y)\cdot T) \mid a\right] + h_\theta(\tfrac{1-e^{-r\Delta}}{\Delta e^{-r\Delta}}\varpi_T)$$

We now prove that for small Δ and r, the incentive compatibility (O.2) and feasibility constraints hold as well. Let us evaluate terms involved in these constraints. For all action profiles a' (including a) we have

$$E[1_{|x|\le\Delta^{1/3}}\beta x \mid a'] = \beta E[x \mid a'] + O(\Delta^2) = \beta\Delta\mu(a') + O(\Delta^2), \quad \text{and}$$

$$E\left[\sum_y 1_y d(y) \mid a'\right] = \sum_y d(y)\lambda(y \mid a')\Delta + O(\Delta^2)$$

Let us introduce a constant $A > 0$ independent of Δ (for small Δ), r or α, such that

$$|E[1_{|x|\le\Delta^{1/3}}\beta x \mid a']| \le A\Delta, \; E[1_{|x|\le\Delta^{1/3}}(\beta x)^2 \mid a'] \approx Var[\beta x] + O(\Delta^2) \le A\Delta,$$

$$|\lambda(y \mid a)d(y)\cdot T|, \; |\lambda(y \mid a)(d(y)\cdot T)^2| \; \le A, \text{ and}$$

$$\varpi_T = \Delta(w - g(a))\cdot T - E\left[(1_{|x|\le\Delta^{1/3}}\beta x) + \sum_y 1_y d(y)\cdot T \mid a\right] \le A\Delta$$

Then

$$E\left[h_\theta\left(\tfrac{1-e^{-r\Delta}}{\Delta e^{-r\Delta}}1_{|x|\le\Delta^{1/3}}\beta x\right) \mid a'\right] \approx rf'(\theta)E[1_{|x|\le\Delta^{1/3}}\beta x \mid a'] + \tfrac{3\kappa}{2}r^2 E[1_{|x|\le\Delta^{1/3}}(\beta x)^2 \mid a'],$$

$$P[y \mid a]E[h_\theta(\tfrac{1-e^{-r\Delta}}{\Delta e^{-r\Delta}}d(y)\cdot T) \mid a] \approx \Delta\lambda(y \mid a)(rf'(\theta)d(y)\cdot T - \tfrac{3\kappa}{2}r^2(d(y)\cdot T)^2),$$

and $\quad h_\theta(\tfrac{1-e^{-r\Delta}}{\Delta e^{-r\Delta}}\varpi_T) \approx rf'(\theta)\varpi_T - \tfrac{3\kappa}{2}r^2\varpi_T{}^2$

all have absolute values bounded by $\alpha rA\Delta + \tfrac{3\kappa}{2}r^2 A\Delta$ for all $\theta \in (-\delta, \delta)$.

Hence for any ε and A we can pick α, r_ν and Δ_ν small enough so that

$$\begin{aligned}&(g_i(a) - g_i(a')) + \frac{e^{-r\Delta}}{1-e^{-r\Delta}}(E[\omega_i \mid a] - E[\omega_i \mid a']) \ge g_i(a) - g_i(a') + \beta\big(\mu(a) - \mu(a')\big)T_i \\ &+\sum_y d_i(y)\big(\lambda(y \mid a) - \lambda(y \mid a')\big) - O(\Delta) - (|Y|+2)(\alpha A + \tfrac{3\kappa}{2}rA) \ge 0\end{aligned} \qquad \text{(O.7)}$$

And hence the incentive compatibility constraints (O.2) are satisfied (recall, that β and d satisfy the ICε constraints for some ε<0, so that for small enough r, Δ and α, the additional terms in (O.7) have less impact on the constraints (O.2) than ε).

For feasibility, for small enough Δ and r, all the variables in h_θ terms in (O.6) are less then $\delta/2$ so that it is sufficient to check that ϖ_N is not too large. For that note:

$$\varpi_N = \tfrac{1-e^{-r\Delta}}{e^{-r\Delta}}(g(a) - w + \tfrac{1}{\Delta} E[\sum_y 1_y d(y) \mid a]) \cdot N +$$

$$E\left[h_\theta(\tfrac{1-e^{-r\Delta}}{\Delta e^{-r\Delta}}(1_{|x|\leq \Delta^{1/3}} \beta x)) - \sum_y 1_y h_\theta(\tfrac{1-e^{-r\Delta}}{\Delta e^{-r\Delta}} d(y) \cdot T) \mid a \right] + h_\theta(\tfrac{1-e^{-r\Delta}}{\Delta e^{-r\Delta}} \varpi_T)$$

$$\geq r\Delta \mid \varepsilon \mid -(\mid Y \mid +2)(\alpha r A \Delta + \tfrac{3\kappa}{2} r^2 A \Delta) \geq 0$$

when α and r are sufficiently small (and by taking Δ or r to be small ϖ_N can be bounded from above by A_δ). Thus, for any w in the neighborhood of v of radius δ/2 the constructed (a, ω) are an admissible pair generating w. This completes the proof of the proposition. QED

Appendix O-D: Generically, $M\_=M$.

Before proving that generically $M\_=M$, we present a singular example in which M and $M\_$ are different. Consider the following two-player partnership game. Each player can choose between three effort levels $a_i = 0$, 1 or 2. Actions are private and players equally share the continuous stream of revenue

$$4dX_t = 4(\mu(a_1, a_2)dt + dZ_t),$$

where $\mu(a_1, a_2) = a_1 + a_2$ and Z is a Brownian motion. The cost of effort is $c_i(a_i) = - a_i^2$, so expected stage-game payoffs are

$$g_i(a_1, a_2) = 2(a_1 + a_2) - a_i^2.$$

The matrix of expected stage-game payoffs is

	0	1	2
0	0,0	2,1	4,0
1	1,2	3,3	5,2
2	0,4	2,5	4,4

The static Nash equilibrium of this game is (1,1). If X_t is the only public signal in this game, then vectors $(\beta^1, \beta^2) = (\beta T_1, \beta T_2)$ enforce action profiles such that each player maximizes $g_i(a_1, a_2) + \beta^i(a_1 + a_2) = (2+ \beta^i)(a_1 + a_2) - a_i^2$. Therefore, β^i enforces the following action:

$$a_i = \begin{cases} 0 & \beta^i < -1 \\ 1 & \beta^i \in [-1,1] \\ 2 & \beta^i \geq 1 \end{cases}$$

The expected stage-game payoffs enforced by various vectors (β^1, β^2) are illustrated in the left panel of Figure O5. The right panel of Figure O5 shows the sets M and $M\_$ constructed with the help of the left panel.

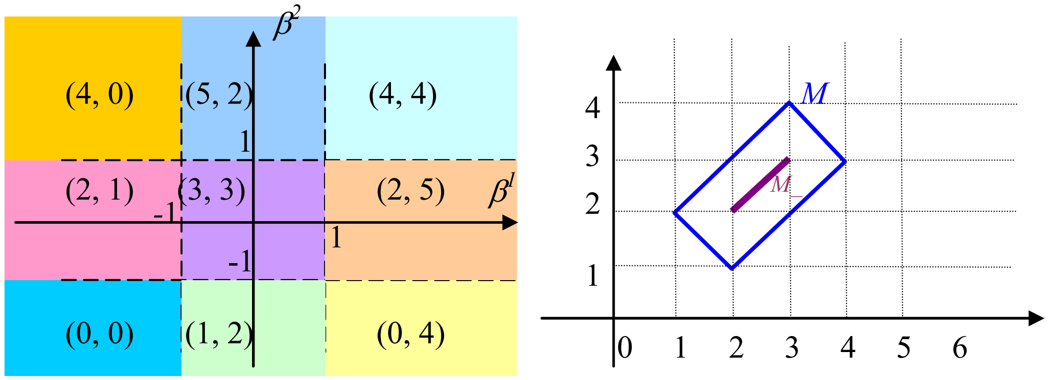

Figure O5: Enforceable action profiles and sets M and $M\_$.

From the left panel we can read which payoff pairs are enforceable (weakly or strictly) on each tangent hyperplane. For example, all payoffs *except* for (0, 0) and (4, 4) are *weakly* enforceable on the negative 45-degree tangent, such that $T_1 = -T_2$ and $\beta^1 = -\beta^2$. At the same time, only (4, 0), (3, 3) and (0, 4) are enforceable *strictly*. Also, only (0, 0), (3, 3) and (4, 4) are *strictly* enforceable on the positive 45-degree tangent, while additionally (5,2), (2,5), (2,1) and (1,2) are enforceable weakly. As a result, the maximal ε-strictly enforceable halfspaces in positive and negative 45-degree directions collapse as ε becomes positive. That is why M and $M\_$ in this example are different.

The singularity, which leads M and $M\_$ to be different in this example, is that as T passes the negative 45-degree direction, the maximal enforceable payoff profile switches between (2, 5) and (5, 2), both of which are enforced weakly. One of the central ideas of our proof is to show that such situations are non-generic.

Central ideas of the proof that generically $M = M\_$:

(i) The set M is defined as an intersection of half-spaces, one for each direction.

(ii) Each half-space has an action profile that generates it.

(iii) Whenever the action profile associated with a half-space can be enforced strictly (with constraints tightened a bit), the half-space changes continuously in T and ε.

(iv) Half-spaces collapse as we increase ε only at points where the action profile associated with the half-space is enforced weakly.

(v) Part (a) of the Proposition O-D below implies that that there are finitely many directions T in which the best action profile is enforced weakly, and Lemma O-D1 says that generically the second-best action profile is enforced strictly in neighborhoods of those directions.

(vi) Part (b) of the Proposition O-D below shows that when in some direction T the maximal half-space defined by the best action profile is discontinuous in ε, the second-best action profile at T becomes the best action for some directions in a neighborhood of T.

(vii) Steps (v) and (vi) imply that whenever, as we increase ε above 0, a maximal half-space in some direction T collapses, the collapse happens only up until the half-space defined by the second-best action profile at T, i.e. only up to the half-space which already bounds the set M in some directions near T (see Figure O6). Thus, as we increase ε above 0, for generic games the collapse of the half-space defined by the best action in any direction T does not cause the collapse of $M(\varepsilon)$.

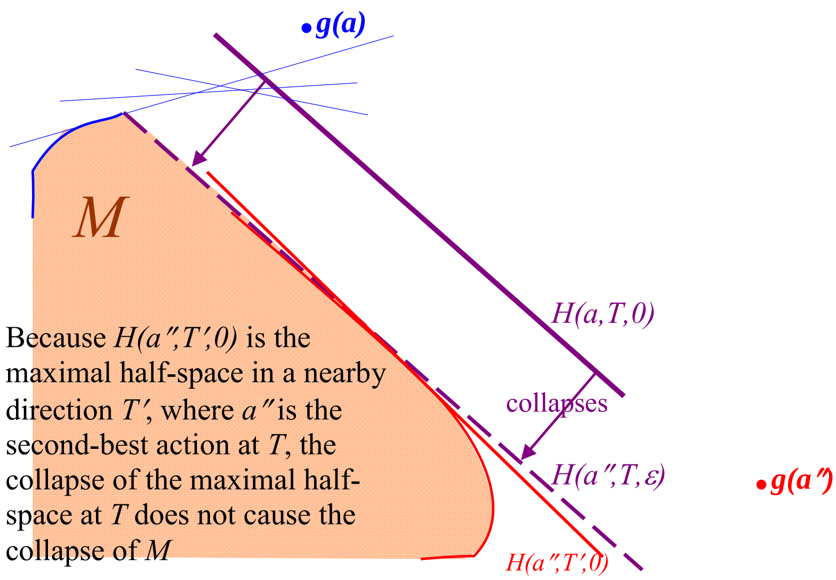

Figure O6: Collapse of the hyperplane in direction T does cause the collapse of M.

When we say that generically something is true, we mean that for any *game structure* (set of actions of each player, the set of possible Poisson jumps and the number of dimensions of the Brownian signal) the statement is true everywhere except for a set of game parameters of measure 0. For a given game structure, game parameters define for each action profile payoffs to each player, the mean of the Brownian signal and the intensity of each possible Poisson jump.

We can represent all unit tangent vectors T as points of a unit circle. Then, an *interval* of vectors T corresponds to an arc of a circle.

The following proposition looks at the incentive constraint for a given action profile, and shows that generically there are finitely many directions $T(\alpha) = (cos(\alpha), sin(\alpha))$ for which that action profile is enforceable but not strictly enforceable (we will call this *weakly enforceable)*. The proposition also shows that in any neighborhood of such an α, there are directions in which the action profile is not enforceable.

<u>Proposition O-D.</u> *Suppose for $k = 1, \ldots K', K'+1, \ldots K$, $g_k \in R$ are real numbers, $\mu_k \in R^N$, $\lambda_k \in R^M$ are vectors. For each $\alpha \in [0, 2\pi)$, consider the set $S(\alpha) \subseteq R^N \times [0,\infty)^M$ of vectors $(x, y) \in R^N \times [0,\infty)^M$ (with $x \in R^N$ and $y \in [0,\infty)^M$) that satisfy the constraints*

$g_k + (x \cdot \mu_k) \cos(\alpha) + (y \cdot \lambda_k) \sin(\alpha) \geq 0$ for all $k = 1, \ldots K'$ and
$g_k - (x \cdot \mu_k) \sin(\alpha) + (y \cdot \lambda_k) \cos(\alpha) \geq 0$ for all $k = K'+1 \ldots K$.
Then generically (i.e. for generic values of g_k, μ_k and λ_k):

(a) the set of values of $\alpha \in [0, 2\pi)$, for which the set $S(\alpha)$ is nonempty but has measure 0, has a finite number of elements and

(b) if $S(\alpha)$ is nonempty but has a measure 0, then in any arbitrarily small neighborhood $(\alpha-\varepsilon, \alpha+\varepsilon)$ of α, there is a point α' such that $S(\alpha')$ is empty and a point α'' such that $S(\alpha'')$ has positive measure.

Proof. In order to not interrupt the flow, we provide the proof (via lemmas p1-p7), which is quite involved, at the end of this online Appendix.

Lemma O-D1 implies that for any direction α in which the best action profile is weakly enforced (by the proposition there are finitely many such directions), the second-best action profile is generically strictly enforced. The best action profile is defined as the arg max in the maximization of $D(N(\alpha))$. The second best profile is defined as the argmax of the same problem with the first best profile removed from the choice set (since some profiles may require $d_N>0$, the ranking is not necessarily the same as the ranking of stage-game payoffs corresponding to those profiles).

Lemma O-D1. Consider a direction $T=T(\alpha)$ in which the best action a is weakly enforceable. Then generically the second-best action a'' is strictly enforceable in direction T.

Proof. Note that there is at least one enforceable action profile in the direction T besides a, because the Nash equilibrium is strictly enforceable generically in all directions. Thus, a'' is enforceable in the direction T.

Action profiles a and a'' have at least one component that's different between players 1 and 2. Without loss of generality, let's say that player 2's action is different. Then changing $g_1(a'')$ does not affect the incentive constraints for enforcing profile a (since $a' = (a_1'', a_2'')$ cannot result from player 1's deviation from $a = (a_1, a_2)$). Let us show that generically there is at most one value of $g_1(a'')$ for which the action profile a'' is weakly enforceable. We can make this conclusion if we show that, if a'' is weakly enforceable for a given value of $g_1(a'')$, then a'' is strictly enforceable for all larger values of $g_1(a'')$.

Definition. Denote by $B_i(a,T)$ the set of (β, d_T, $d_N \geq 0$) that satisfy

$$g_i(a) - g_i(a') + \beta\,(\mu(a) - \mu(a'))\,T_i + \Sigma_y\,(d_T(y)\,T_i + d_N(y)\,N_i)\,(\lambda(y|a) - \lambda(y|a')) \geq 0$$

for all deviations a' of player i. $B_i(a,T)$ is a convex set, as it is defined as an intersection of a finite number of linear constraints (= half-spaces).

If a'' is weakly enforceable on T, it means that $B_1(a'',T)$ and $B_2(a'',T)$ intersect. If we increase $g_1(a'')$ a little bit, $B_1(a'',T)$ grows in all directions and at the original intersection point all constraints of player 1 become slack. It turns out that $B_2(a,T)$ is generically nonempty (see Lemma O-D2 below). Since $B_2(a,T)$ does not change as we increase g_1-*(a)*, we can move away from the original point of intersection in the direction where all constraints of player 2 become slack. Thus, if we increase $g_1(a'')$ a little bit, it is possible to satisfy all constraints strictly. QED

Lemma O-D2 was used in the proof of Lemma O-D1:

Lemma O-D2. Generically, for all a, $B_i(a,T)$ is empty or has nonempty interior for all regular hyperplanes.[2]

Proof. Let us call a set *degenerate* if it is nonempty but has empty interior.
Fix a regular hyperplane T. Let us show that the set of games for which $B_i(a,T)$ is degenerate has measure 0. Let us vary $g_i(a)$ while keeping all the other parameters of the game fixed. Then there is at most one value of $g_i(a)$ for which $B_i(a,T)$ is degenerate. Indeed, if (β, d_T, d_N) satisfies the constraints of player i weakly for a given value of $g_i(a)$, then at (β, d_T, d_N) satisfies all constraints strictly for all larger values of $g_i(a)$.
Now, if $B_i(a,T)$ has nonempty interior for one regular hyperplane T, then $B_i(a,T')$ has nonempty interior for all regular T' in the same quadrant, since $B_i(a,T')$ is obtained from $B_i(a,T)$ by stretching along some dimensions and squeezing along other dimensions.
QED

Lemma O-D3 implies that the maximal hyperplane (i.e. the one corresponding to $D(T,\varepsilon)$) changes continuously in both T and ε in the range where the best action profile is enforced strictly. Together Lemmas O-D1 and O-D3 imply that in the neighborhood of the direction in which the best action profile is enforced weakly, generically the second-best hyperplane changes continuously in both T and ε.

Lemma O-D3 If we change T and ε continuously in the range where action profile a can be enforced on tangent T with constraints tightened by ε, the minimal amount of value required to be destroyed to enforce a (that is, $\Sigma_y\, d_N(y)\, \lambda(y|a)$) is continuous (and weakly increasing in ε).

Proof. Denote by $B_i(a,T,\varepsilon)$ the set of $(\beta, d_T, d_N \geq 0)$ that satisfy the incentive constraints tightened by ε:

$$g_i(a) - g_i(a') + \beta\,(\mu(a) - \mu(a'))\, T_i + \Sigma_y\, (d_T(y)\, T_i + d_N(y)\, N_i)\, (\lambda(y|a) - \lambda(y|a')) \geq \varepsilon$$

for all deviations a' of player i.

The set $B_i(a,T,\varepsilon)$ is shrinking continuously as ε increases, as all half-spaces that define it become continuously smaller. When T changes, the set $B_i(a,T,\varepsilon)$ is also changing continuously: it gets stretched in some dimensions and shrinks in other dimensions. As a

[2] All hyperplanes but coordinate ones are regular. Coordinate hyperplanes are those parallel to one of the axes.

result, the intersection $B_1(a,T,\varepsilon) \cap B_2(a,T,\varepsilon)$ is changing continuously in T and ε (and it is shrinking continuously as ε increases). Since the function $\Sigma_y\, d_N(y)\, \lambda(y|a)$ is continuous, the minimum of this function over the intersection $B_1(a,T,\varepsilon) \cap B_2(a,T,\varepsilon)$ is changing continuously in T and ε, and it is weakly increasing continuously as ε increases. QED

Now, as ε increases the best half-spaces that define the set M(ε) near directions in which the best action profile is weakly enforceable collapse. If hyperplanes near direction T, on which the best action profile a is weakly enforceable, collapse, then the second-best action profile a' (which defines the half-spaces of M to one side of T – by part (b) of the Proposition) is generically strictly enforceable near T by Lemma O-D1. By Lemma O-D3, the half-space generated by a' changes continuously near T. Because of that, $M(\varepsilon)$ does not collapse but shrinks continuously at $\varepsilon = 0$. See Figure O6.

Proof of the Proposition O-D. We carry out the proof of the proposition for $\alpha \neq 0, \pi/2, \pi, 3\pi/2$ (i.e. we carry it only for regular hyperplanes). For coordinate hyperplanes, it is enough to observe that generically $S(\alpha)$ is empty or has positive measure for any fixed α. The reason is that, as we change g_k for all $k = 1, \ldots K$ by the same constant $const \in R$, there is at most one value of $const$ for which $S(\alpha)$ is a nonempty set with measure 0. Apart from these special values of α, the proposition follows from a sequence of lemmas. The first lemma derives a necessary condition for the set $S(\alpha)$ to be nonempty of zero measure.

<u>Lemma p1</u>. If $S = \{x : g_i + x^T \beta_i \geq 0,\ i = 1, \ldots m\}$ is a nonempty set of zero measure, then for any $x \in S$, among vectors (g_i, β_i) there is a subset of $dim(x)+1$ (or fewer) linearly dependent vectors, such that each of these vectors corresponds to a constraint that binds at x.

Proof. First, let us show that if S is nonempty but has measure 0, then one of the inequalities must hold with equality on the entire set S. In other words, for some $j = 1, \ldots m$, for all $x \in S$, $g_j + x^T \beta_j = 0$. If not, then for all $i = 1, \ldots m$, there exists $x_i \in S$ such that $g_i + x_i^T \beta_i > 0$. Then, since S is convex, $x = (x_1 + x_2 + \ldots x_m)/m \in S$, and $g_i + x^T \beta_i > 0$ for all $i = 1, \ldots m$. A contradiction.

Now, if $g_j + x^T \beta_j = 0$ for some $j = 1, \ldots m$, then the problem

$$\max_x\ g_j + x^T \beta_j$$
$$\text{s.t. } g_i + x^T \beta_i \geq 0,\ i \neq j$$

has value 0 and the solution set S. Consider any $x \in S$. Then by Kuhn-Tucker theorem, there are Lagrange multipliers $\eta_i \geq 0$, $i \neq j$ such that

$$\beta_j = \Sigma_{i \neq j}\, \eta_i\, \beta_i,$$

where $\eta_i > 0$ only if $g_i + x^T \beta_i = 0$. This equation represents β_j as a linear combination of other β_i (such that $g_i + x^T \beta_i = 0$ and $i \neq j$). We can always represent β_j as a linear combination of at most dim*(x)* of these β_i's as follows

$$\beta_j = \Sigma_{i \in I'} \eta_i' \beta_i,$$

where $|I'| \leq \dim(x)$.[3] Multiplying both sides by $-x^T$ (and using that the value of the problem is 0), we get:

$$g_j = -x^T \beta_j = -\, x^T \Sigma_{i \in I'} \eta_i' \beta_i = -\Sigma_{i \in I'} \eta_i' g_i.$$

Let $I = I' \cup \{j\}$. Then the vectors (g_i, β_i), $i \in I$ are linearly dependent. QED

Lemma p1 implies that a *necessary* condition for $S(\alpha)$ to be nonempty of zero measure is that among 1+N+M-dimensional vectors

$$\begin{pmatrix} g_i \\ \mu_i \cos\alpha \\ \lambda_i \sin\alpha \end{pmatrix},\ i = 1, \ldots k;\ \begin{pmatrix} g_i \\ -\mu_i \sin\alpha \\ \lambda_i \cos\alpha \end{pmatrix},\ i = k{+}1, \ldots K,\ \begin{pmatrix} 0 \\ 0 \\ e_m \end{pmatrix},\ m = 1, \ldots M \qquad \text{(O.8)}$$

there is a subset of at most $1 + N + M$ linearly dependent vectors, where e_m is an M-dimensional vector with entry 1 in location m, and entries 0 in all other locations. The next lemma allows us to focus on subsets of exactly 1 + N + M vectors when we discuss instances when $S(\alpha)$ is nonempty of measure 0.

<u>Lemma p2</u>: Generically, any subset of fewer than $1 + N + M$ vectors among vectors (O.8) are independent for all $\alpha \neq 0, \pi/2, \pi, 3\pi/2$.

Proof. Consider any subset of $L < 1 + N + M$ vectors among (O.8), and let us show that generically they are independent for all $\alpha \neq 0, \pi/2, \pi, 3\pi/2$. Consider the $1 + N + M \times L$ matrix composed of these vectors. Generally this matrix could include M' columns of the form ($0\ 0\ e_m$). If we exclude these columns and corresponding rows (with 1 from ($0\ 0\ e_m$)) to get a smaller $1 + N + (M - M') \times (L - M')$ matrix A, then whenever the columns of the original $1 + N + M \times L$ matrix are dependent, the columns of matrix A will also be dependent. Let us show that generically, the columns of matrix A are independent for all $\alpha \neq 0, \pi/2, \pi, 3\pi/2$.

Note that whenever the columns of A are dependent, then any (L – M') rows of A are also dependent. Consider the square matrix B composed of the last (L – M') rows of A. Rows of B are dependent if and only if the determinant of B is zero. Since

[3] If a set of vectors (β's) span a linear space of dimension $\leq \dim(x)$, then we can pick a subset of these vectors (not more than $\dim(x)$) that form the basis of this subspace.

$$\sin(\alpha) = \frac{2\tan(\frac{\alpha}{2})}{1+\tan^2(\frac{\alpha}{2})} \quad \text{and} \quad \cos(\alpha) = \frac{1-\tan^2(\frac{\alpha}{2})}{1+\tan^2(\frac{\alpha}{2})},$$

the determinant of B is a rational function (i.e. a ratio of two polynomials) of tan(α/2). Generically, this rational function is not identically 0, so it becomes 0 only at finitely many values of α. For any such value of α, the last (L-M') rows of A span a subspace of $R^{L-M'}$ of dimension less than L-M'. Generically, the first row of A, which is of the form $(g_{k_1},...g_{k_{L-M'}})$, will not be in this subspace. QED

The next lemma proves part (a) of the proposition.

<u>Lemma p3.</u> Generically, there are finitely many values of α at which any subset of $1 + N + M$ vectors among (O.8) are dependent.

Proof. A subset of $1 + N + M$ vectors among (O.8) are dependent if and only if the determinant of the square matrix composed of these vectors is zero. Since the determinant is a rational function of $\tan(\alpha/2)$, generically there are finitely many values of α that set this function to zero. QED

To prove part (b) of the proposition, we first show that generically, whenever $S(\alpha)$ is nonempty of measure 0, there is exactly ***one*** subset of $1 + N + M$ dependent vectors among (O.8).

<u>Lemma p4.</u> Generically, there is no value of α for which two different subsets of $1 + N + M$ vectors among (O.8) are dependent.

Proof. Consider two subsets S_1 and S_2 of $1 + N + M$ vectors, and suppose that either $S_1 \backslash S_2$ or $S_2 \backslash S_1$ has a vector among the first K vectors in the collection (O.8). Let's say that it is vector $v \in S_2 \backslash S_1$. Generically there are finitely many α for which the set S_1 is dependent, because those α correspond to zeros of the determinant of S_1, which is a rational function of $\tan(\alpha/2)$. Let's show that for any one of these α, generically the collection of vectors S_2 are independent. Note that $S_2 \backslash v$ are generically independent for all $\alpha \neq 0, \pi/2, \pi, 3\pi/2$, by Lemma p2. Now, v is dependent on $S_2 \backslash v$ if and only if v belongs to the $N + M$ –dimensional subspace of the $1 + N + M$ –dimensional space spanned by $S_2 \backslash v$. Since each component of v is determined by a different parameter (and note that these parameters are separate from those parameters that define α and $S_2 \backslash v$), it follows that generically v will not be in the span of $S_2 \backslash v$.

Now, suppose that both $S_1 \backslash S_2$ and $S_2 \backslash S_1$ have only vectors of the form $(0\ 0\ e_m)^T$. Denote by I_1 the set of indices where one of the vectors in $S_1 \backslash S_2$ has a 1. Similarly define I_2. Denote by I_3 the set of indices for which one of the vectors in $S_1 \cap S_2$ of the form (0 0 e_m) has a 1. Take the matrix consisting of vectors $S_1 \cap S_2$ as columns, and let's exclude from this matrix columns of the form (0 0 e_m) and rows from the set I_3. We end up with a $1+N+M - |I_3| - |I_1| \times 1+N+M - |I_3|$ matrix. Note that $1+N+M - |I_3| - |I_1|$ rows of this matrix, excluding the rows from I_1, are dependent, and another set of rows of the same size, excluding the rows from I_2 (note that $|I_1| = |I_2|$), are dependent. This situation is nongeneric, but the same argument as in the first paragraph of this proof applies. QED

Lemma p5. Generically, whenever $S(\alpha)$ is a nonempty set of zero measure, it consists of a single point, at which $1 + N + M$ inequality constraints bind and the rest of the constraints are slack. The set of $1 + N + M$ vectors that correspond to these constraints are linearly dependent.

Proof. If $S(\alpha)$ is nonempty of zero measure, then by Lemmas p1 and p2, generically there is a subset of $1 + N + M$ dependent vectors among (O.8), and by Lemma p4, generically all other vectors are independent of these $1 + N + M$. Lemma p1 also implies that there is $(x,y) \in S(\alpha)$ at which the $1 + N + M$ constraints are binding. All other constraints must be slack at (x,y), since a binding constraint, i.e. a hyperplane passing through the same point (x,y), would correspond to a vector that is dependent on these $1 + N + M$. It remains to be shown that $S(\alpha)$ consists of a single point. Suppose there is another point (x', y'). Then Lemmas p1 and p2 imply that the same $1 + N + M$ constraints must bind at (x',y') as at (x,y) (because some constraints, which correspond to linearly dependent vectors, must bind at (x',y'), and generically these are not different from the ones that bind at (x,y)).

Towards a contradiction, note that the first $N + M$ vectors of these, (g_i, β_i), $i \in I$, are independent, and so the matrix with columns β_i is invertible, and so the system of equations $g_i + x\, \beta_i = 0$ has a unique solution. Therefore, there can be at most one point at which $N + M$ of these constraints are binding. QED

Part (b) of the proposition follows from Lemma p6.

Lemma p6. Generically, whenever $S(\alpha^*)$ is a nonempty set of zero measure, there is $\varepsilon > 0$ such that in a neighborhood $(\alpha^*, \alpha^*+\varepsilon)$, $S(\alpha)$ has positive measure and in a neighborhood $(\alpha^*-\varepsilon, \alpha^*)$, $S(\alpha)$ is empty, or *vice versa*.

Proof. Consider $1 + N + M$ constraints that define $S(\alpha^*)$. Then, since all other constraints are slack, by continuity $S(\alpha)$ equals the intersection of these $1+N+M$ half-spaces in a neighborhood of α^*. Let's put the vectors that correspond to these half-spaces into a matrix

$$M(\alpha) = \begin{pmatrix} g_{k_0} & \dots & g_{k_{N+M}} \\ \mu_{k_0}\cos\alpha & \dots & -\mu_{k_{N+M}}\sin\alpha \\ \lambda_{k_0}\sin\alpha & \dots & \lambda_{k_{N+M}}\cos\alpha \end{pmatrix}.$$

We consider the case when $M(\alpha)$ does not have columns of the form $(0\ 0\ e_m)$, but that is without loss of generality because such columns do not contain free parameters affecting the determinant of $M(\alpha)$.

$S(\alpha)$ goes from an empty set to a set of positive measure as α passes through α^* if and only if the determinant of $M(\alpha)$ *changes sign* as α passes through α^*.[4] That happens if and only if the determinant of the following matrix

$$\begin{pmatrix} g_{k_0} & \dots & g_{k_{N+M}} \\ \mu_{k_0} & \dots & -\mu_{k_{N+M}} \sin\alpha/\cos\alpha \\ \lambda_{k_0} \sin\alpha/\cos\alpha & \dots & \lambda_{k_{N+M}} \end{pmatrix} = \begin{pmatrix} g_{k_0} & \dots & g_{k_{N+M}} \\ \mu_{k_0} & \dots & -\mu_{k_{N+M}} x \\ \lambda_{k_0} x & \dots & \lambda_{k_{N+M}} \end{pmatrix} = M'(x)$$

changes sign as x passes through $\sin\alpha^*/\cos\alpha^*$ (here we are assuming that α is not 0, $\pi/2$, π or $3\pi/2$). But that is true generically: the fact that the determinant of a matrix $M'(x)$ has only single roots follows from Lemma p7 below. QED

Lemma p7. Consider a square matrix

$$A(x) = \begin{pmatrix} a_{11} & \dots & a_{1n}x \\ \vdots & & \vdots \\ a_{n1}x & \dots & a_{nn} \end{pmatrix},$$

where x's (x is some real variable) could multiply any entries in the matrix. Then, for generic coefficients a_{11} through a_{nn}, the determinant of the matrix has no double roots (except for, possibly, $x = 0$).

Proof. We prove this statement by induction on n. Clearly, both a_{11} and $a_{11}x$ have no double roots.

The statement follows for n if we show that for generic coefficients a_{12} through a_{nn} (all coefficients but a_{11}) the set of values of a_{11} for which the determinant has double roots (except for, possibly, $x = 0$) consists of isolated points.

Note that the determinant of $A(x)$ is of the form $f(x) + a_{11}\, g(x)$, where $f(x)$ and $g(x)$ are polynomials. Let's see what happens to the roots of the determinant as we change a_{11}. If x^* is an isolated root, it moves with speed $g(x^*)/f'(x^*)$, which is bounded. So, in a neighborhood of a_{11} single roots cannot merge into double roots. If x^* is a double root, it disappears or bifurcates (or becomes a single root in case of a triple root, etc.) if $g(x^*) \neq 0$ or $g(x^*) = 0$ and $g'(x^*) \neq 0$. Thus, if there is a double root x^* at a_{11}, there are no double roots in a neighborhood of a_{11} unless x^* is a double root of $g(x)$.

[4] The intuition is that, first, Lemma p5 implies that generically $S(\alpha)$ is a nonempty set of zero measure if and only if the columns of matrix $M(\alpha)$ are dependent, i.e. $\det M(\alpha) = 0$. If $S(\alpha)$ goes from empty to positive measure near α^*, then for arbitrarily small perturbations of parameters g_{k0} through g_{kN+M}, $S(\alpha)$ still goes from positive measure to empty. However, if the determinant of $M(\alpha)$ did not change sign at α^*, then by adjusting the parameters of the first row of $S(\alpha)$ arbitrarily slightly, the determinant of $M(\alpha)$ can be made to stay strictly positive (or negative) in a neighborhood of α^*, contradicting that $S(\alpha)$ goes from empty to positive measure.

However, $g(x)$ is the determinant of

$$\begin{pmatrix} a_{22} & \dots & a_{2n}x \\ \vdots & & \vdots \\ a_{n2}x & \dots & a_{nn} \end{pmatrix}$$

or the determinant of this matrix times x. In both cases, $g(x)$ has no double roots (except for possibly $x = 0$) by the inductive hypothesis. QED

This completes the proof of Proposition O-D. QED