Cournot-Nash Equilibria and Optimal transport

Guillaume Carlier[a] **and Adrien Blanchet**[b] **.**

Mean Field Games and Related Topics (2 !),

Padova, September 2013.

a. CEREMADE, Université Paris Dauphine
b. Toulouse School of Economics

Cournot-Nash Equilibria

Setting : type space X (metric compact) endowed with a probability measure $\mu \in \mathcal{P}(X)$, action space Y (metric compact). Cost : $C(x, y, \nu)$ where $\nu \in \mathcal{P}(Y)$ represents the distribution of actions (anonymous game).

Unknown : $\gamma \in \mathcal{P}(X \times Y)$: $\gamma(A \times B)$ is the probability that an agent has her type in A and takes an action in B. Then define

Definition 1 *A Cournot-Nash equilibrium (CNE) is a* $\gamma \in \mathcal{P}(X \times Y)$ *such that* $\Pi_{X\#}\gamma = \mu$ *and*

$$\gamma\Big(\{(x, y) \ : \ C(x, y, \nu) = \min_{z \in Y} C(x, z, \nu)\}\Big) = 1$$

where $\nu := \Pi_{Y\#}\gamma$.

Easy case

Theorem 1 (Mas-Colell, 1984) *If*

$$\nu \mapsto C(.,.,\nu) \textit{ is continuous from } (\mathcal{P}(Y), w-*) \textit{ to } C(X \times Y)$$

then there exists CNE.

Proof : Consider $C := \{\gamma = \mu \otimes \gamma^x\} = \{\gamma \ : \ \Pi_{X\#}\gamma = \mu\}$. For $\gamma = \mu \otimes \gamma^x \in C$ let $\nu := \Pi_{Y\#}\gamma$ and set

$$F(\gamma) = \{\mu \otimes \theta^x, \ \theta^x \in \mathcal{P}(\text{argmin}\, C(x,.,\nu))\}.$$

Since F has a closed graph and is convex-compact valued it has a fixed point $\gamma \in F(\gamma)$ i.e. γ is a CNE.

- the assumption is extremely strong : rules out congestion/purely local effects,
- what about uniqueness, characterization, explicit or numerically computable solutions ?

We shall restict ourselves to the separable case :

$$C(x, y, \nu) = c(x, y) + V[\nu](y) \tag{1}$$

and shall further impose that $\nu \in L^1(m_0)$ with m_0 a given reference measure on Y. Can be viewed as a simplified (static) version of the Mean-Field Games Theory of Lasry and Lions.

Example : Doctors location choice : $x \in X$ doctor's type (specialty, where she/he graduated from...), $y \in Y$ doctor's location choice. Total cost= commuting cost $c(x,y)$ +congestion cost+interaction cost. Congestion cost : ν absolutely continuous with respect to some reference measure m_0, $\nu(dy) = \nu(y)m_0(dy)$, congestion cost $f(y, \nu(y))$ with f increasing in its second argument. Interaction cost : probability to interact with other doctors around y : $\int_Y \psi(d(y,z))d\nu(z)$ with ψ increasing.

Benchmark : $\nu \in \mathcal{P}(Y) \cap L^1(m_0)$ (m_0 : fixed reference measure according to which congestion is measured)

$$V[\nu](y) = f(y, \nu(y)) + \int_Y \phi(y, z_1, \cdots, z_m) d\nu^{\otimes m}(z_1, \cdots, z_m).$$

Due to the first term, the previous fixed-point argument does not work.

Domain

$$\mathcal{D} := \{\nu \in L^1(m_0) \ : \ \int_Y |V[\nu]| d\nu < +\infty\}.$$

Outline

① Connection with optimal transport

② A variational approach

③ Hidden convexity : dimension one

④ Hidden convexity : quadratic cost

⑤ A PDE for the equilibrium

Connections with optimal transport

Again $m_0 \in \mathcal{P}(Y)$ fixed reference measure, $\mathcal{D}$ domain of the cost, CNE are then defined by

Definition 2 *$\gamma \in \mathcal{P}(X \times Y)$ is a Cournot-Nash equilibria if and only if its first marginal is μ, its second marginal, ν, belongs to $\mathcal{D}$ and there exists $\varphi \in C(X)$ such that*

$$c(x,y)+V[\nu](y) \geq \varphi(x) \ \forall x \in X \text{ and } m_0\text{-a.e. } y \text{ with equality } \gamma\text{-a.e.} \tag{2}$$

A Cournot-Nash equilibrium γ is called pure whenever it is of the form $\gamma = (\mathrm{id}, T)_{\#}\mu$ for some Borel map $T : X \to Y$.

For $\nu \in \mathcal{P}(Y)$, let $\Pi(\mu,\nu)$ denote the set of probability measures on $X \times Y$ having μ and ν as marginals and let $\mathcal{W}_c(\mu,\nu)$ be the least cost of transporting μ to ν for the cost c *i.e.* the value of the Monge-Kantorovich optimal transport problem :

$$\mathcal{W}_c(\mu,\nu) := \inf_{\gamma \in \Pi(\mu,\nu)} \iint_{X \times Y} c(x,y)\, \mathrm{d}\gamma(x,y)$$

let us also denote by $\Pi_o(\mu,\nu)$ the set of optimal transport plans i.e.

$$\Pi_o(\mu,\nu) := \{\gamma \in \Pi(\mu,\nu) \ : \ \iint_{X \times Y} c(x,y)\, \mathrm{d}\gamma(x,y) = \mathcal{W}_c(\mu,\nu)\}.$$

A first link between Cournot-Nash equilibria and optimal transport is based on the following straightforward observation.

Lemma 1 *If γ is a Cournot-Nash equilibrium and ν denotes its second marginal then $\gamma \in \Pi_o(\mu, \nu)$.*

Proof. Indeed, let $\varphi \in C(X)$ be such that (2) holds and let $\eta \in \Pi(\mu, \nu)$ then we have

$$\iint_{X \times Y} c(x, y) \, \mathrm{d}\eta(x, y) \geq \iint_{X \times Y} (\varphi(x) - V[\nu](y)) \, \mathrm{d}\eta(x, y)$$

$$= \int_X \varphi(x) \, \mathrm{d}\mu(x) - \int_Y V[\nu](y) \, \mathrm{d}\nu(y) = \iint_{X \times Y} c(x, y) \, \mathrm{d}\gamma(x, y)$$

so that $\gamma \in \Pi_o(\mu, \nu)$. $\square$

The previous proof also shows that φ solves the dual of $\mathcal{W}_c(\mu, \nu)$ i.e. maximizes the functional

$$\int_X \varphi(x)\,\mathrm{d}\mu(x) + \int_Y \varphi^c(y)\,\mathrm{d}\nu(y)$$

where φ^c denotes the c-transform of φ i.e.

$$\varphi^c(y) := \min_{x \in X}\{c(x, y) - \varphi(x)\} \tag{3}$$

In an euclidean setting, there are well-known conditions on c and μ which guarantee that such an optimal γ necessarily is pure whatever ν is :

Corollary 1 *Assume that $X=\overline{\Omega}$ where Ω is some open connected bounded subset of $\mathbb{R}^d$ with negligible boundary, that μ is absolutely continuous with respect to the Lebesgue measure, that c is differentiable with respect to its first argument, that $\nabla_x c$ is continuous on $\mathbb{R}^d \times Y$ and that it satisfies the twist condition :*

$$\textit{for every } x \in X, \textit{ the map } y \in Y \mapsto \nabla_x c(x,y) \textit{ is injective,}$$

then for every $\nu \in \mathcal{P}(Y)$, $\Pi_0(\mu,\nu)$ consists of a single element and the latter is of the form $\gamma = (\mathrm{id}, T)_{\#}\mu$ hence every Cournot-Nash equilibrium is pure (and fully determined by its second marginal).

Similar to what happens in mean-field games : monotonicity implies uniqueness (covers the case of pure congestion) :

Theorem 2 *If* $\nu \mapsto V[\nu]$ *is strictly monotone in the sense that for every* ν_1 *and* ν_2 *in* $\mathcal{P}(Y)$*, one has*

$$\int_Y (V[\nu_1] - V[\nu_2]) d(\nu_1 - \nu_2) \geq 0$$

and the inequality is strict whenever $\nu_1 \neq \nu_2$ *then all equilibria have the same second marginal* ν*.*

Proof. Let $(\nu_1, \gamma_1, \varphi_1)$, $(\nu_2, \gamma_2, \varphi_2)$ be such that

$$V[\nu_i](y) \geq \varphi_i(x) - c(x,y), \ i = 1, 2,$$

for every x and m_0-a.e. y with an equality γ_i-a.e., using the fact that $\gamma_i \in \Pi(\mu, \nu_i)$, we get

$$\int_Y V[\nu_i] d\nu_i = \int_X \varphi_i d\mu - \int_{X\times Y} c d\gamma_i, \ i = 1, 2$$

$$\int_Y V[\nu_i] d\nu_j \geq \int_X \varphi_i d\mu - \int_{X\times Y} c d\gamma_j, \text{ for } i \neq j$$

substracting, we get $\int_Y V[\nu_1] d(\nu_1 - \nu_2) \leq \int_{X\times Y} cd(\gamma_2 - \gamma_1)$ and $\int_Y V[\nu_2] d(\nu_2 - \nu_1) \leq \int_{X\times Y} cd(\gamma_1 - \gamma_2)$ and monotonicity thus gives $\nu_1 = \nu_2$.

$\square$

A variational approach

Take $V[\nu](y) = f(y, \nu(y)) + \int_Y \phi(y,z) d\nu(z)$ with $f(y,.)$ continuous nondecreasing (+ power or logarithm growth) and ϕ continuous and symmetric i.e. $\phi(y,z) = \phi(z,y)$. Then define $F(y,\nu) := \int_0^\nu f(y,s) ds$ and

$$E[\nu] = \int_Y F(y, \nu(y)) dm_0(y) + \frac{1}{2} \iint_{Y \times Y} \phi(y,z)\, \mathrm{d}\nu(y)\, \mathrm{d}\nu(z)$$

then $V[\nu] = \frac{\delta E}{\delta \nu}$ in the sense that for every $(\rho, \nu) \in \mathcal{D}^2$, one has

$$\lim_{\varepsilon \to 0^+} \frac{E[(1-\varepsilon)\nu + \varepsilon\rho] - E[\nu]}{\varepsilon} = \int_Y V[\nu]\, \mathrm{d}(\rho - \nu).$$

Equilibria may be obtained by solving

$$\inf_{\nu \in \mathcal{D}} J_\mu[\nu] \quad \text{where} \quad J_\mu[\nu] := \mathcal{W}_c(\mu, \nu) + E[\nu]. \tag{4}$$

Theorem 3 (Minimizers are equilibria) *Assume that* $X = \overline{\Omega}$ *where* Ω *is some open bounded connected subset of* $\mathbb{R}^d$ *with negligible boundary, that* μ *is equivalent to the Lebesgue measure on* X *(that is both measures have the same negligible sets) and that for every* $y \in Y$, $c(., y)$ *is differentiable with* $\nabla_x c$ *bounded on* $X \times Y$. *If* ν *solves* (4) *and* $\gamma \in \Pi_o(\mu, \nu)$ *then* γ *is a Cournot-Nash equilibrium. In particular there exist CNE.*

optimality condition for (4) : there is a constant M such that

$$\begin{cases} \varphi^c + V[\nu] \geq M \\ \varphi^c + V[\nu] = M \qquad \nu\text{-a.e. }, \end{cases} \tag{5}$$

If E is convex : equivalence between minimization and being an equilibrium. If E strictly convex : uniqueness (of ν). The congestion term is convex and forces dispersion whereas the interaction term is nonconvex and rather fosters concentration. It may be the case that the congestion term dominates so as to make E convex but this is more the exception than the rule. There is hidden convexity (McCann's displacement convexity) in the problem as we shall see now. The following ideas are initially due to Robert J. McCann and the notion of convexity that we will us is a slight variant of McCann's displacement convexity due to Ambrosio, Gigli and Savaré to deal with the nonconvexity of the squared-2-Wasserstein distance.

Hidden convexity : dimension one

Intuition is easy to understand in dimension one : the functional J_μ is not convex with respect to ν but it is with respect to T, the optimal transport map from μ to ν. Let us take $X = Y = [0,1]$, m_0 is the Lebesgue measure on $[0,1]$, μ is absolutely continuous with respect to the Lebesgue measure, and assume that $V[\nu]$ takes the form :

$$V[\nu](y) = f(\nu(y)) + V(y) + \int_{[0,1]} \phi(y,z)\, \mathrm{d}\nu(z)$$

the corresponding energy reads

$$E(\nu) := \int_0^1 F(\nu(y))\, \mathrm{d}y + \int_0^1 V(y)\, \mathrm{d}\nu(y) + \frac{1}{2}\int_{[0,1]^2} \phi\, \mathrm{d}\nu^{\otimes 2}$$

(with $F' = f$).

Assume

- the transport cost c is of the form $c(x,y) = C(x-y)$ where C is strictly convex and differentiable,
- f is convex increasing (+growth condition),
- V is convex on $[0,1]$ and ϕ is convex, symmetric, differentiable and has a locally Lipschitz gradient.

Let $(\rho, \nu) \in \mathcal{P}([0,1])^2$ then there is a unique optimal transport map T_0 (respectively T_1) from μ to ν (respectively from μ to ν) for the cost c and it is nondecreasing. For $t \in [0,1]$, let us define :

$$\nu_t := T_{t\#}\mu \text{ where } T_t := ((1-t)T_0 + tT_1)$$

then the curve $t \mapsto \nu_t$ connects $\nu_0 = \nu$ to $\nu_1 = \rho$. A functional $J : \mathcal{P}(Y) \to \mathbb{R} \cup \{+\infty\}$ is called displacement convex whenever $t \in [0,1] \mapsto J(\nu_t)$ is convex (for every choice of endpoints ν and ρ), it is called strictly displacement convex when, in addition $J(\nu_t) < (1-t)J(\nu) + tJ(\rho)$ when $t \in (0,1)$ and $\rho \neq \mu$.

We claim that J_μ is strictly displacement convex ; indeed, take (ν, ρ) two probability measures in the domain of E (which is convex by convexity of F), define ν_t as above and, let us consider the four terms in J_μ separately.

By definition of $\mathcal{W}_c$, ν_t and the strict convexity of C we have

$$\begin{aligned}\mathcal{W}_c(\mu,\nu_t) &\leq \int_0^1 C(x-((1-t)T_0(x)+tT_1(x)))\,\mathrm{d}\mu \\ &\leq (1-t)\int_0^1 C(x-T_0(x))\,\mathrm{d}\mu + t\int_0^1 C(x-T_1(x))d\mu \\ &= (1-t)\mathcal{W}_c(\mu,\nu)+t\mathcal{W}_c(\mu,\rho)\end{aligned}$$

with a strict inequality if $t\in(0,1)$ and $\nu\neq\rho$.

By construction

$$\int_0^1 V\ \mathrm{d}\nu_t = \int_0^1 V(T_t(x))d\mu(x) = \int_0^1 V((1-t)T_0(x)+tT_1(x))d\mu(x)$$

which is convex with respect to t, by convexity of V. Similarly

$$\int_{[0,1]^2} \phi\ \mathrm{d}\nu_t^{\otimes 2} = \int_{[0,1]^2} \phi(T_t(x), T_t(y))\ \mathrm{d}\mu(x)\ \mathrm{d}\mu(y)$$

is convex with respect to t, by convexity of ϕ, The convexity of the remaining congestion term is more involved. Since $\nu_t = T_{t\#}\mu$ and T_t is nondecreasing, at least formally we have $\nu_t(T_t(x))T_t'(x) = \mu(x)$, by the change of variables formula we also have

$$\int_0^1 F(\nu_t(y))\,\mathrm{d}y = \int_0^1 F(\nu_t(T_t(x)))T_t'(x)\,\mathrm{d}x = \int_0^1 F\Big(\frac{\mu(x)}{T_t'(x)}\Big)T_t'(x)$$

and we conclude by observing that $\alpha \mapsto F(\mu(x)\alpha^{-1})\alpha$ is convex and that $T_t'(x)$ is linear in t.

All this yields :

Theorem 4 *Under the assumptions above, optima and equilibria coincide and there exists a unique equilibrium (which is actually pure).*

Hidden convexity : quadratic cost

The arguments of the previous paragraph can be generalized in higher dimensions when the transport cost is quadratic.
Throughout this section, we will assume the following :

- $X = Y = \overline{\Omega}$ where Ω is some open bounded convex subset of $\mathbb{R}^d$,
- μ is absolutely continuous with respect to the Lebesgue measure (that will be the reference measure m_0 from now on) and has a positive density on Ω,
- c is quadratic *i.e.*

$$c(x, y) := \frac{1}{2}|x - y|^2, \ (x, y) \in \mathbb{R}^d \times \mathbb{R}^d,$$

– V again takes the form

$$V[\nu](y) = f(\nu(y)) + V(y) + \int_{Y^m} \phi(y,.)\,\mathrm{d}\nu^{\otimes m}$$

where V is convex, f nondecreasing (+growth conditions) and $\phi \in \mathcal{C}(\mathbb{R}^{d(m+1)})$ is symmetric.

Again denoting by F the primitive of f that vanishes at 0, the corresponding energy reads

$$E[\nu] = \int_Y F(\nu(y))\,\mathrm{d}(y) + \int_Y V\,\mathrm{d}\nu + \frac{1}{m+1}\int_{Y^{m+1}} \phi\,\mathrm{d}\nu^{\otimes(m+1)}.$$

Brenier's Theorem implies the uniqueness and the purity of optimal plans γ between μ and an arbitrary ν (and the optimal map is of the form $T = \nabla u$ with u convex). Variational problem

$$\inf_{\nu \in \mathcal{P}(\overline{\Omega})} J_\mu[\nu] \quad \text{where} \quad J_\mu[\nu] := \frac{1}{2}\mathcal{W}_2^2(\mu,\nu) + E[\nu] \tag{6}$$

with $\mathcal{W}_2^2(\mu,\nu)$ is the squared-2-Wasserstein distance between μ and ν. Structural assumptions to guarantee the (strict) convexity of J_μ along (generalized) geodesics are McCann's condition :

$$\nu \mapsto \nu^d F(\nu^{-d}) \text{ is convex nonincreasing on } (0,+\infty), \tag{7}$$

and ϕ convex and smooth (C^1 with a locally Lipschitz gradient).

Under these conditions, again equilibria coincide with minimizers and there is uniqueness.

A PDE for the equilibrium

For computational simplicity, take $V = 0$ and $f(\nu) = \log(\nu)$ (satisfies McCann's condition and ensures that the mass remains positive everywhere). Optimality condition :

$$\log(\nu(y)) + \varphi^c(y) + \int_{Y^m} \phi(y, .) d\nu^{\otimes m} = 0 \qquad (8)$$

Optimal transport map (Brenier) $T = \nabla u$ between μ and ν : Monge-Ampère equation

$$\mu(x) = \det(D^2 u(x))\, \nu(\nabla u(x)), \quad \forall x \in \Omega \qquad (9)$$

which has to be supplemented with the natural sort of boundary condition

$$\nabla u(\Omega) = \Omega. \qquad (10)$$

On the other hand $\varphi(x) = \frac{1}{2}|x|^2 - u(x)$, $\varphi^c(y) = \frac{1}{2}|y|^2 - u^*(y)$ so

$$\varphi^c(\nabla u) = \frac{1}{2}|\nabla u|^2 - u^*(\nabla u) = \frac{1}{2}|\nabla u|^2 - x \cdot \nabla u + u$$

substituting $y = \nabla u(x)$ in (8), using

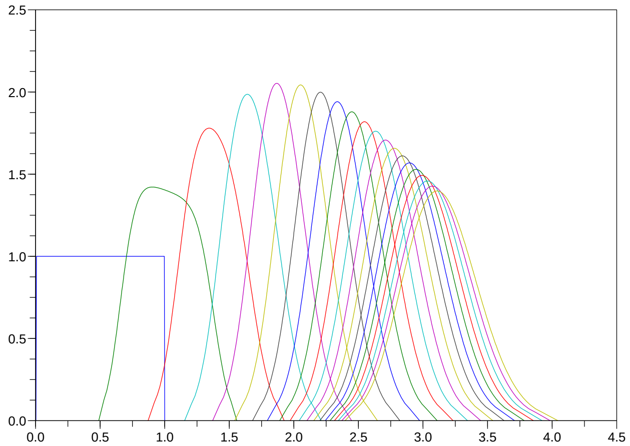
$$\int_{Y^m} \phi(\nabla u(x), .)\, \mathrm{d}\nu^{\otimes m} = \int_{\Omega^m} \phi(\nabla u(x), \nabla u(x_1), \ldots, \nabla u(x_m))\, \mathrm{d}\mu^{\otimes m}\,.$$

and eliminating ν thanks to (9), we get

$$\mu(x) = \det(D^2 u(x)) \exp\left(-\frac{1}{2}|\nabla u(x)|^2 + x \cdot \nabla u(x) - u(x)\right) \times$$
$$\exp\left(-\int_{\Omega^m} \phi(\nabla u(x), \nabla u(x_1), \ldots, \nabla u(x_m))\, \mathrm{d}\mu^{\otimes m}(x_1, \ldots x_m)\right)\,. \tag{11}$$

The equilibrium problem is therefore equivalent to a non-local and nonlinear partial differential equation.

The problem can be solved numerically in dimension 1.



Convergence and stabilisation toward the equilibrium in the case of a logarithmic congestion, cubic interaction, and a potential $V(x) := (x-5)^3$ with uniform measure on $[0, 1]$ as initial guess

Cost of anarchy

The equilibrium is the unique minimiser of the functional J_μ. It would therefore be tempting to interpret this result as a kind of welfare theorem. A simple comparison between J_μ and the total social cost tells us however that the equilibrium is not efficient. Indeed, the total social cost $\mathrm{SC}(\nu)$ is the sum of the transport cost $W_2^2(\mu,\nu)/2$ and the additional cost $\int_Y V[\nu](y)\,\nu(y)\,\mathrm{d}y$ *i.e.*

$$\mathrm{SC}[\nu] = \frac{1}{2}W_2^2(\mu,\nu) + \int_Y f(\nu)\nu + \int_Y V\,\mathrm{d}\nu + \int_{Y^{m+1}} \phi\,\mathrm{d}\nu^{\otimes(m+1)}\,.$$

The second term represents the total congestion cost and the fourth one the total interaction cost.

The functional J_μ whose minimiser is the equilibrium has a similar form, except that in its second term $f(\nu)\nu$ is replaced by $F(\nu)$ (with $F' = f$) and the interaction term is divided by $m+1$. The equilibrium corresponds indeed to the case where agents selfishly minimise their own cost

$$c(x,.) + V[\nu] = c(x,.) + f(\nu(.)) + V(.) + \int_{Y^m} \phi(.,z)\nu^{\otimes m} \, .$$

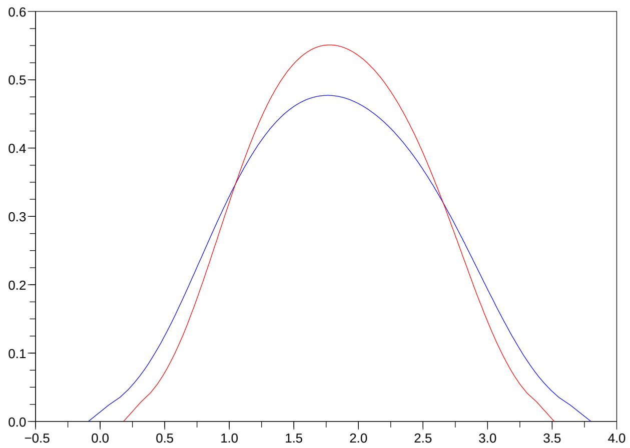
Natural way to restore efficiency of the equilibrium : proper system of tax/subsidies which, added to $V[\nu]$, will implement the efficient configuration. A tax system that restores the efficiency is easy to compute (up to an additive constant) :

$$\mathrm{Tax}[\nu](y) = f(\nu(y))\,\nu(y) - F(\nu(y)) + m \int_{Y^{m+1}} \phi(y,z)\,\mathrm{d}\nu^{\otimes m}(z).$$

Similar inefficiency of equilibria, arises in the slightly different framework of congestion games, where it is usually referred under the name *cost of anarchy*, which has been extensively studied in recent years (Roughgarden). In our Cournot-Nash context, we may similarly define the cost of anarchy as the ratio of the worst social cost of an equilibrium to the minimal social cost value :

$$\text{Cost of anarchy} := \frac{\max\{\mathrm{SC}[\nu^e] \ : \ \nu^e \text{ equilibrium}\}}{\min_\nu \mathrm{SC}[\nu]} \ .$$

The computation of the equilibrium and the optimum can be done numerically in dimension 1



In the previous numerical example, both the equilibrium and the optimum are unique and the cost of anarchy can be numerically computed as being approximately 1.8.