

Une promenade aléatoire en probabilités

Alain MARUANI

École des Ponts ParisTech

Télécom ParisTech

alain.maruani@enpc.fr

alain.maruani@telecom-paristech.fr

Contenu

Quelques rencontres, au fil des années : regards rétrospectifs

Quelques points de vue pour les CPGE : regards prospectifs

Une question associée : Statut de la Mécanique Quantique ?

Au fil des années 1

Première affirmation

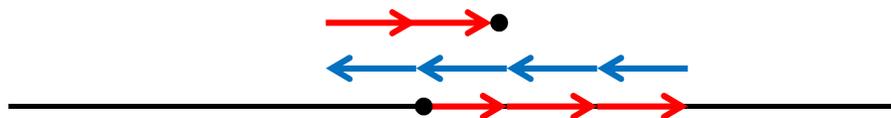
Probabilité = fréquence statistique, à travers des problèmes de dénombrement

Exemple : Probabilité qu'un paquet de quatre cartes pris dans un jeu dont on a retiré les trèfles contienne au moins un as rouge ?

Premiers développements

- Formule asymptotique de Wallis : $n! = n \ln n - n + O(\ln n)$
- Marche aléatoire, durée Δt , amplitude a : *la solution avant l'équation !*

$P(n, N)$ pour $N \rightarrow \infty$, $n \rightarrow \infty$ et $n/N \ll 1$.



$$t = N\Delta t$$

$$x = Na$$

$$D = \frac{a^2}{t}$$

$$P(x, t) = \frac{1}{\sqrt{\pi Dt}} \exp\left(-\frac{x^2}{4Dt}\right) \quad (\partial_t = D \partial_{x^2})$$

Au fil des années 2

En CPGE

Facteur de Boltzmann : $FB = \exp\left(-\frac{E}{kT}\right)$

Distribution de Maxwell $DM = \exp\left(-\frac{1}{2} \frac{mv^2}{kT}\right)$; équipartition, viriel.

Entropie et température : $dS = \frac{\delta Q}{T}$

La température est l'inverse du facteur intégrant de la quantité de chaleur, la forme de Pfaff associée est l'entropie.

Un peu plus tard

L'entropie :
$$S = -k' \ln \Omega = -k' \sum_n p_n \ln p_n$$
$$= -k' \int p(u) \ln [p(u)] du$$

Les statistiques quantiques $f_{FD} = \frac{1}{1 \pm \exp\left(\frac{E - \mu}{kT}\right)}$

Au fil des années 3

Encore plus tard (pêle-mêle)

La classification

La mesure

Les problèmes mal posés (radar, traitement d'image)

Des lois nouvelles (Weibull, ...)

Des techniques nouvelles (Transformée de Mellin, ...)

Des concepts nouveaux, ou rajeunis (la pseudo inverse) (réseaux de neurones, recuit simulé, ...)

Je n'ai pas cherché les probabilités, elles se sont insérées naturellement dans monpaysage.

Les probabilités sont probablement le seul enseignement commun à toutes les GE.

- *C'est donc une excellente raison de l'introduire en CPGE.*
- *C'est donc une excellente raison de ne pas l'introduire en CPGE.*

Un but de cet exposé

- Ne pas se préoccuper du contenu précis d'un éventuel programme (les dés sont-ils jetés ?)
- Proposer la présence d'une idée (sans se préoccuper de la technique qui pourrait l'accompagner) : méthodes bayésiennes.
- Rechercher des adhérences entre probabilités (au sens noble du terme : mesure, tribus boréliennes, ...) et autres éléments de cours de CPGE (versions dégradées) :

Physique

Chimie

Sciences industrielles (MP !)

Informatique

(Retour sur) La Mesure 1 : généralités

Signe de confiance dans l'intelligibilité de la Nature.

- Fidélité, justesse et précision.
- Résultats : nombres, courbe, affirmation d'existence (Neptune) ou jugement (théories).

Transducteurs, traitement de données

- Conversion d'une grandeur en une autre.

Il est souhaitable que le principe de fonctionnement de l'appareil soit différent du principe régissant le système étudié.

- Erreurs, inconsistances et problèmes mal posés sont le lot quotidien de la mesure. Théorie de Gauss.

Variabilités de l'objet, du système de mesure et de l'observateur, en supposant ces éléments séparables. On cherche un représentant moyen. Dans une classe, la moyenne représente la performance de l'élève prototype. Invocation de la loi des Grands Nombres.

- Inférence et théorie de la confirmation :

Peut-on quantifier la manière dont la découverte de Neptune a conforté la théorie newtonienne ?
Que vaut la performance de ce télépathe ?

(Retour sur) La Mesure 2 : questions

Inférence et théorie de la confirmation

- Comment modifier nos conclusions au fur et à mesure que les données s'accumulent ?
- Peut-on quantifier la confiance mise en un modèle ?
- Pourquoi des expériences indépendantes sont-elles censées mieux conforter un modèle que la répétition inlassable de la même expérience ?

Duhem : On ne saurait déduire les lois de Newton de l'observation du mouvement des planètes.

Newton : Deux révolutions du même objet ne sont jamais identiques !

Les degrés de croyance ne sont ni logiques ni objectifs, ils relèvent des lois de probabilité et du pari.

Hypothèse

Le crédit accordé à une hypothèse après l'acquisition de données expérimentales (probabilité *a posteriori*) est proportionnelle au produit de :

- la vraisemblance de cette hypothèse
- la probabilité d'obtenir ces résultats expérimentaux, en tenant l'hypothèse pour vraie (probabilité *a priori*)

Principe d'inférence (Bayes) 1

Une probabilité n'est pas le passage à la limite d'une fréquence, mais la traduction numérique d'un état de connaissance.

$$\begin{aligned}P(H \text{ et } e) &= P(H|e) \times P(e) \\ &= P(e|H) \times P(H)\end{aligned}$$

Ces règles résultent de

$$\text{Addition : } P(A \cup B | C) = P(A|C) + P(B|C) - P(A \cap B|C)$$

$$\text{Multiplication : } P(A \cap B) = P(A|B) \times P(B) = P(B|A) \times P(A)$$

La connaissance des causes possibles d'une observation permet de remonter aux causes.

Exemple : Vrai (V) ou Faux (F) $P(V|E) = 0,9$ $P(F|\bar{E}) = 0,95$

Théorème de Bayes : $P(E|V) = \frac{P(V|E)}{P(V)} \times P(E)$. Probabilité a priori.

Principe d'inférence (Bayes) 2

Modéliser l'attente en début de processus plutôt que de se fixer une méthode puis traiter les données.

- Possibilité de diminuer le poids des aprioris au fur et à mesure de l'acquisition des données.

Quoi utiliser ?

- Informations abondantes ou peu onéreuses : statistiques classiques.
- Autrement : méthodes bayésiennes.

Les résultats sont asymptotiquement identiques.

Le test bayésien utilisé pour déterminer la plausibilité d'une distribution par rapport à des observations est asymptotiquement convergent avec celui des statistiques classiques (grand nombre d'observations).

Principe d'ignorance maximale. *Principe d'entropie maximale.*

Principe d'inférence (Bayes) 3 : exemples

La boîte A contient 30 biscuits au chocolat et 10 biscuits ordinaires.

La boîte B contient 20 biscuits au chocolat et 20 biscuits ordinaires.

Le biscuit choisi au hasard dans une boîte au hasard est au chocolat.

De quelle boîte a-t-il le plus de chances d'être issu, et avec quelle probabilité ?

Intuitivement : de la boîte A.

H_X = la proposition « le gâteau vient de la boîte X ». C = « le biscuit est au chocolat ».

$$P(H_A) = P(H_B) = 0,5.$$

$$P(C | H_A) = 30 / (30 + 10) = 0,75$$

$$P(C | H_B) = 20 / (20 + 20) = 0,5$$

Après le choix, cette probabilité est $P(H_A | C) = P(C | H_A) \times P(H_A) / P(C) = 0,75 \times 0,5 / (50/80) = 0,6$.

Le gain de vraisemblance (1,76 dB) ne dépend pas de la probabilité a priori des boîtes.

Autre exemple

Des résultats de double diffraction électronique contredisant la théorie relativiste de Dirac n'ont pas entraîné une remise en cause de cette équation : cette dernière se révèle robuste par rapport à une différence entre le résultat obtenu et le résultat attendu.

Information et surprise. Quantification de l'information

La quantité d'information est une fonction

Décroissante de la probabilité p (croissante de la surprise)

Additive

$$I_m = k \log_b(1/p_m)$$

$$I = \sum p_m I_m$$

C'est le manque d'information avant toute observation.

I est la seule quantité satisfaisant à des exigences de bon sens (Kinchine : jeu équiprobable, p rationnel, p réel)

L'incertitude devient une grandeur mesurable.

Que dire lorsque l'on ne sait rien ?

$$\textit{Entropie } S = -\sum p_m \log(p_m)$$

Principe d'ignorance maximale : $p_m = p = 1/W$

$S = -I$ est maximale (convexité)

$$I = \sum p_m I_m = k \sum (1/W) \log(1/W)$$

$$S = k \cdot \log(W)$$

Maxent : loi 1

1) Dé (ou dés)

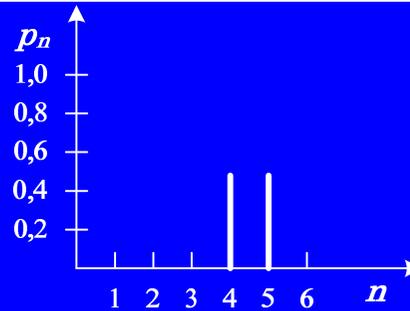
$$E = 4,5 \quad (E_{\text{loyal}} = 3,5)$$

Maximiser I , sous la contrainte

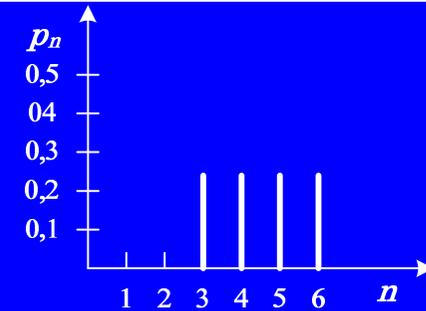
$$E_{\text{triche}} = 4,5$$

$$S_{\text{maxent}} = 1,61$$

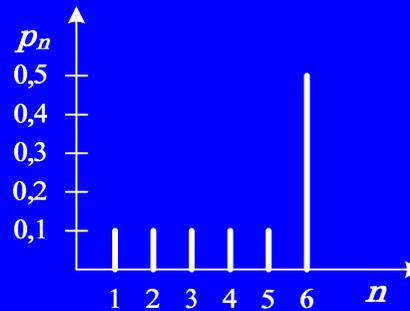
$$S_{\text{uniforme}} = 1,79$$



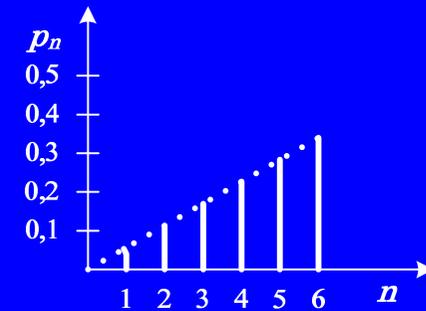
Juste, mais ne semble pas raisonnable.



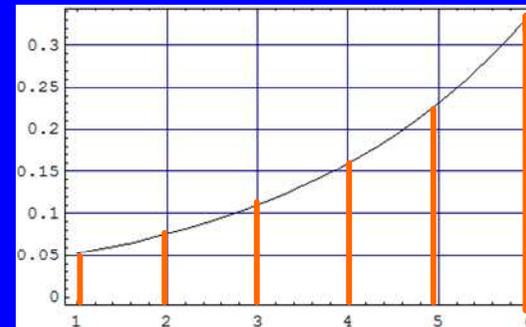
Moins injuste, mais peu satisfaisant.



Juste, mais le poids du 6 semble exagéré.
Toujours aussi arbitraire.



Différence constante.
Pas de solution pour $E = 4,7$.



La solution optimale

Maxent : loi 2 : le facteur de Boltzmann

Énergie individuelle \mathcal{E} , valeur moyenne E imposée.

Quelle est la loi de probabilité ?

Entropie maximale pour

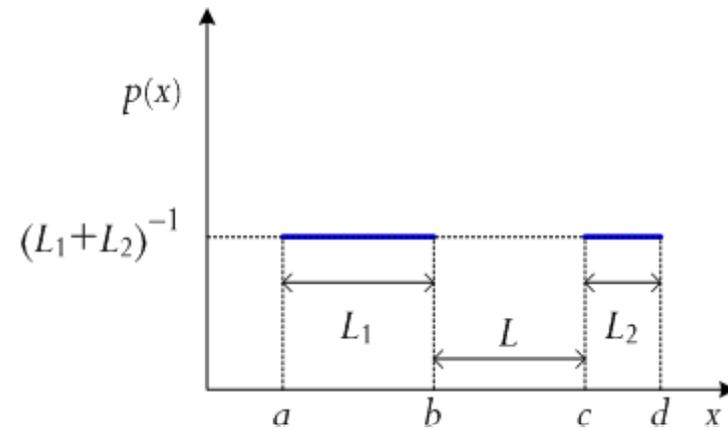
$$E = \sum_n p_n \mathcal{E}_n$$

$$S = -\sum_n p_n \ln p_n$$

$$\frac{\partial(S - \lambda E)}{\partial p_n} = -\ln p_n - 1 - \lambda \mathcal{E}_n = 0$$

$$p_n = A \exp(-\lambda \mathcal{E}_n) \quad \left\{ \begin{array}{l} \sum_n p_n = 1 \\ \sum_n p_n \mathcal{E}_n = E \end{array} \right. \Rightarrow A \left(= \frac{1}{Z} \right) \text{ et } \lambda \left(= \frac{1}{kT} \right)$$

Incertitude ou dispersion ?



$$\Delta x = F(a, b, c, d) = \Phi(L_1, L_2, L)$$

$$S = \int p(x) \ln[p(x)] dx = \ln(L_1 + L_2)$$

Quelques avatars de l'entropie

Thermodynamique : Clausius 1865

$$dS = \frac{\delta Q}{T} \quad S(2) - S(1) = \int_{\text{Réversible}} \frac{\delta Q}{T}$$

Mécanique statistique : Boltzmann 1877

$$S = k \ln W$$

Mécanique quantique : von Neumann 1927

$$S = -k \text{Tr}(\rho \ln \rho)$$

Et après : Kullback, Kolmogorov, ...

Vers l'information : exorcismes

Démon de Maxwell : 1871

Réfutation Szilard : 1929

Hartley : 1928, information et surprise ($i = -\ln p$)

Shannon : H est la moyenne de l'information sur tous les messages

Cox, Jaynes, Kullback : L'entropie mesure ce que l'on ignore d'un système connu uniquement par sa loi de distribution.

Grandeurs déduites

Entropie mutuelle :

$$H(x, y) = - \sum_{i,j} p_{ij} \ln p_{ij}$$

Entropie conditionnelle :

$$H(x, y) = H(x) + H(y|x) = H(y) + H(x|y)$$

Information mutuelle (redondance)

$$\begin{aligned} I(x, y) &= H(x) + H(y) - H(x, y) \\ &= H(x) - H(x|y) \\ &\leq H(x) \end{aligned}$$

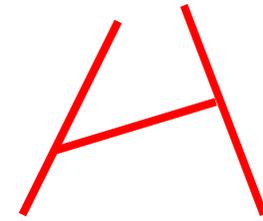
Un exemple

X = le temps qu'il fait

Y = le temps prédit

Pro		X (Il fait ...)				Paresseux		X (Il fait ...)		
		Beau	Pas beau				Beau	Pas beau		
Y (Il fera ...)	Beau	5/8	1/16	11/16	Y (Il fera ...)	Beau	13/16	3/16	1	
	Pas beau	3/16	1/8	5/16		Pas beau	0	0	0	
		13/16	3/16				13/16	3/16		
		Pr (erreur)	4/16				Pr (erreur)	3/16		
		H(X)	0,696				H(X)	0,696		
		H(Y)	0,896				H(Y)	0		
Amateur		X (Il fait ...)								
		Beau	Pas beau				H(X Amat.)	0,695		
Y (Il fera ...)	Beau	403/512	93/512	31/32			H(X Pro.)	0,605		
	Pas beau	13/512	3/512	1/32						
		13/16	3/16							
		Pr (erreur)	0,207							
		H(X)	0,696							
		H(Y)	0,200							

Qu'est ceci ?



- Statistiquement : A ?
- Structurellement : H ?

Principes de décisions

- Espace des formes Ω (tous les objets possibles $\mathbf{a}, a, \mathcal{A}, \mathfrak{a}, \mathfrak{a}, \dots$).
- Espace de représentations R (mesures, forme découpée en segments, graphes).
- Espace des classes $J (= \{0, 1, \dots, 9, ?\})$
- Fonction de classement idéale $C : \Omega \rightarrow J$
- *Fonction de classement* $d : R \rightarrow J$

Les méthodes statistiques

- **On connaît :**
 - La loi du vecteur d'état \mathbf{x} , $p(\mathbf{x})$
 - Le nombre de classes, c , notées ω_i
 - La probabilité *a priori* d'apparition de chaque classe $P(\omega_i)$
 - La probabilité conditionnelle $p(\mathbf{x} | \omega_i)$
- **On cherche à :**
 - évaluer les probabilités a posteriori $p(\omega_i | \mathbf{x})$

Les grandes familles

- Statistiques
 - Classification bayésienne
 - Cas gaussien
 - Estimation des lois
 - Estimation paramétrique
 - Estimation non paramétrique
- Méthodes fondées sur une partition
 - Cas bayésien
 - Bayésien gaussien
 - Discrimination paramétrique non bayésienne
 - Cas linéaire : perceptron
- Segmentation de l'espace : k-ppv
- Approche neuromimétique

Estimation des lois (cadre bayésien)

- Approche paramétrique : estimer les paramètres d'une loi affirmée.
- Estimation sans modélisation

L'inférence bayésienne

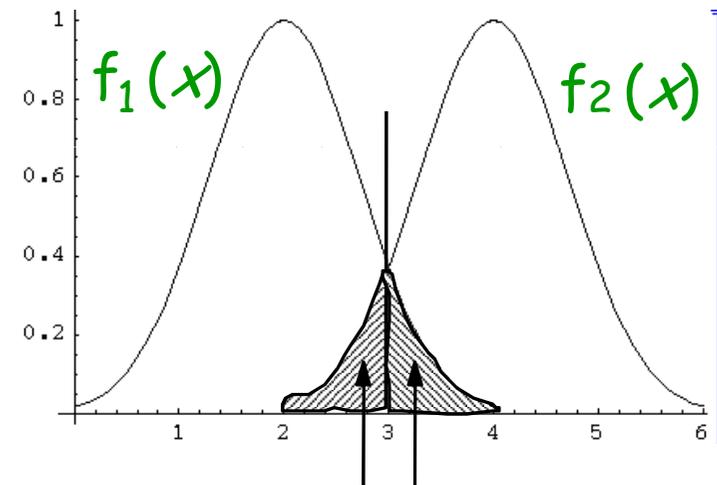
- L'étiquette
- La probabilité d'erreur

$$P(\omega_i | \mathbf{x}) = \frac{P(\omega_i) \times p(\mathbf{x} | \omega_i)}{p(\mathbf{x})}$$

$$R(d) = \int [1 - P(d(\mathbf{x}))] p(\mathbf{x}) d\mathbf{x}$$

Le but

- Min. erreur moyenne de classification.
 - Place la séparatrice au point d'intersection des gaussiennes
 - dans la zone hachurée de gauche, on attribuerait « 1 » à un objet « 2 ».
- Non paramétrique
 - les formes peuvent ne pas être gaussiennes.
- Minimiser l'aire hachurée



Image

Information véhiculée \rightarrow donnée des niveaux de gris en chaque site : zones, contours, structures définis par les contrastes, textures ...

Le niveau de gris en un site est en relation avec les pixels voisins.

Les interactions locales entre niveaux de gris voisins définissent les différentes régions de l'image.

Elles permettent d'utiliser le formalisme markovien en restauration, segmentation, analyse ...

Principe : définir des énergies locales entre groupes de sites reflétant les interactions entre niveaux de gris.

L'énergie globale est reliée à la probabilité d'apparition de l'image

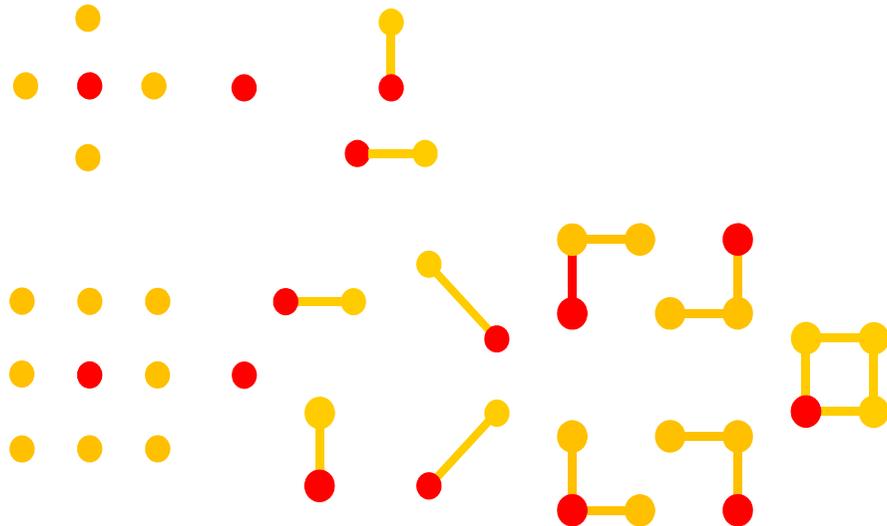
Image

L'image : ensemble fini S de sites s_i correspondant aux pixels.

$S \in \mathbb{Z}^2$ est un réseau discret fini, muni de la notion de voisinage V .

Site \rightarrow descripteur de son état (niveau de gris, ...) à valeurs dans E .

Clique : ensemble de sites voisins.



$$U = \sum_{c \in C} U_c \quad U_s = \sum_{c \in C / s \in c} U_c$$

Image : modélisation probabiliste

Site $s \rightarrow$ variable aléatoire X_s , valeurs x_s dans E .

Le champ aléatoire $X = (X_s, X_t, \dots)$ est à valeurs dans $\Omega = E^{|S|}$.

L'image est une réalisation x de $X : P(X = x)$, est une sorte de vraisemblance de l'image.

- L'hypothèse markovienne permet d'évaluer le lien statistique entre un niveau de gris et le reste de l'image.

Configuration de l'image sans le site $s : x^s = (x_t)_{t \neq s}$

X est un champ de Markov ssi la probabilité conditionnelle locale en un site n 'est fonction que de la configuration du voisinage du site considéré

$$P(X_s = x_s | x^s) = P(X_s = x_s | x_t, t \in V_s)$$

Mesure de Gibbs

Pour l'énergie $U : \Omega \rightarrow \mathbb{R}$: la probabilité P sur Ω

$$P(X = x) = \frac{1}{Z} \exp[-U(x)] = \frac{1}{Z} \exp \left[- \underbrace{\sum_{c \in C} U_c(x)}_{\text{Somme d'énergies locales}} \right]$$

Image binaire 512×512 : $\text{card}(\Omega) = 2^{262114}$

Champ de Gibbs de potentiel associé au système de voisinage V : c'est le champ aléatoire X dont la probabilité est une mesure de Gibbs associée au système de voisinage V .

Théorème de Hammersley-Clifford

- S dénombrable,
- Système de voisinage V borné,
- Espace des états discret

X est un champ de Markov relativement à V et $P(X = x) \forall x \in \Omega$ ssi X est un champ de Gibbs de potentiel associé à V .

Exemple de la 4-connexité

$$U(x) = \sum_{c=(s) \in C_1} U_c(x_s) + \sum_{c=(s,t) \in C_2} U_c(x_s, x_t)$$

Après quelques calculs (non triviaux), le résultat, remarquablement simple, ne fait intervenir que les potentiels des cliques contenant le site s :

$$\begin{aligned} \text{Prob. cond. loc. : } P(X_s = x_s | X^s = x^s) &= \frac{\exp\left[-\sum_{c \in C/s \neq c} U_c(x) - U_s(x_s | V_s)\right]}{\sum_{x_s \in E} \exp\left[-\sum_{c \in C/s \neq c} U_c(x) - U_s(x_s | V_s)\right]} \\ &= \frac{\exp[-U_s(x_s | V_s)]}{\sum_{x_s \in E} \exp[-U_s(x_s | V_s)]} \end{aligned}$$

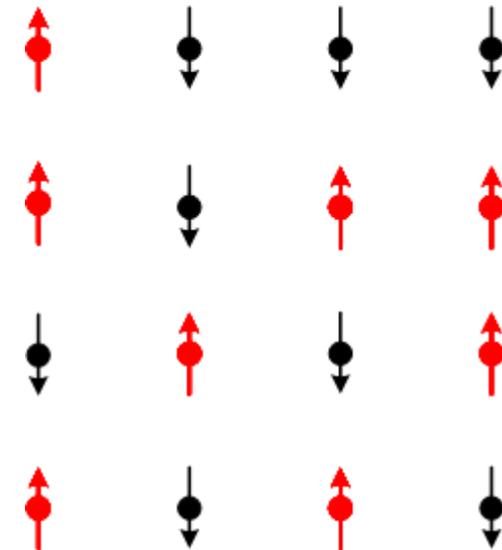
Pas possible d'accéder à la probabilité d'une configuration x donnée (à cause de la constante de normalisation), mais possible de calculer en chaque site la probabilité conditionnelle locale.

Application : Modèle d'Ising 2D, par une approche markovienne

- Réseau carré de spins interagissant

$$U = \sum_{(rs)} U_{rs}$$

$$U_{rs} = \begin{cases} -J\sigma_r\sigma_s & (\sigma_i = \pm 1) & \text{Ising-Heisenberg} \\ -K\delta_K(x_r, x_s) & (x_i = 0,1) & \text{Potts } (K = 2J) \end{cases}$$



- Prototype des champs de Markov aléatoires :

Probabilité spin en $s = x_s$ **conditionnellement au réseau**, S_s
 =
 Probabilité spin en $s = x_s$ **conditionnellement au voisinage**, N_s

Hammersley-Clifford + Bayes

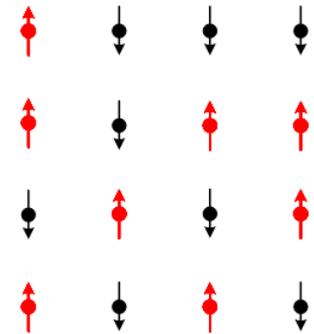
- Réseau carré de spins interagissant

Prob. cond. loc. : $P(x_s | S_s) = P(x_s | N_s)$

$$= \frac{1}{\underbrace{\sum_{\{x_s\}} \exp\left[-\frac{U(x_s | N_s)}{kT}\right]}_{=Z_{N_s}}} \exp\left[-\frac{U(x_s | N_s)}{kT}\right]$$

- Règle de Bayes

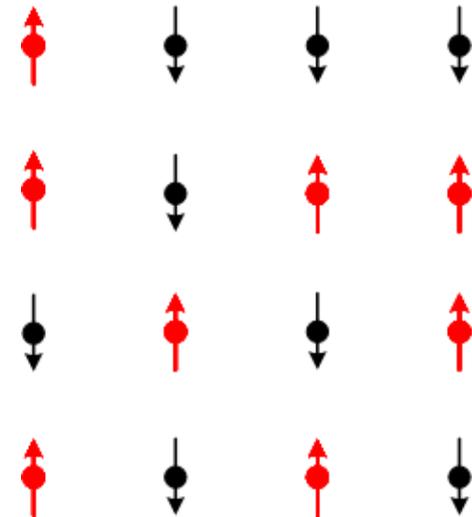
$$\begin{aligned} p &= P(x_s = 1) \\ &= \sum_{\{N_s\}} P(x_s = 1 | N_s) P(N_s) \\ &= \sum_{\{N_s\}} P(N_s) \frac{1}{Z_{N_s}} \exp\left[-\frac{U(x_s | N_s)}{kT}\right] \end{aligned}$$



Équation résolvente

- Notation : $\frac{K}{kT} \mapsto \kappa$
- **Les spins de N_s ne sont pas corrélés**
- Probabilité invariante par translation

$$\begin{aligned}
 p &= \frac{\binom{4}{0} \exp(4\kappa)}{1 + \exp(4\kappa)} p^4 + \frac{\binom{4}{1} \exp(3\kappa)}{\exp(\kappa) + \exp(3\kappa)} p^3 (1-p) \\
 &+ \frac{\binom{4}{2} \exp(2\kappa)}{\exp(2\kappa) + \exp(2\kappa)} p^2 (1-p)^2 \\
 &+ \frac{\binom{4}{3} \exp(\kappa)}{\exp(\kappa) + \exp(3\kappa)} p (1-p)^3 + \frac{\binom{4}{4}}{1 + \exp(4\kappa)} (1-p)^4 \\
 &= \varphi(p)
 \end{aligned}$$



Solution (invariances)

$$p = \frac{\exp(4\kappa)}{1 + \exp(4\kappa)} p^4 + \frac{4\exp(2\kappa)}{1 + \exp(2\kappa)} p^3(1-p) + 3p^2(1-p)^2 + \frac{4}{1 + \exp(2\kappa)} p(1-p)^3 + \frac{1}{1 + \exp(4\kappa)} (1-p)^4$$

$$= \varphi(p)$$

$$p = \frac{1}{2} \quad \forall \kappa$$

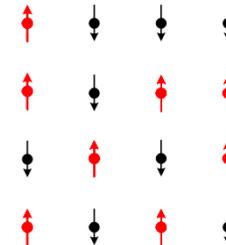
Centre de symétrie $\left(\frac{1}{2}, \frac{1}{2}\right)$

$$\{x_s \leftrightarrow 1 - x_s\} \quad \{p \leftrightarrow 1 - p\} \quad \{U_{rs} \leftrightarrow U_{sr}\}$$

$$[a] \quad p = \varphi(p) \Leftrightarrow [b] \quad 1 - p = \varphi(1 - p)$$

$$[a] + [b] \Rightarrow \varphi(p) + \varphi(1 - p) = 1: d^\circ \varphi \text{ impair } (= 3) !$$

$$[a] - [b] \Rightarrow \mu = 2p - 1 = \varphi(p) - \varphi(1 - p)$$



$$\Rightarrow \left\{ \begin{array}{l} \left[p^2 + (1-p)^2 \right] \tanh(2\kappa) + 4p(1-p) \tanh \kappa = 1 \\ p = \frac{1}{2} \end{array} \right.$$

Analyses (équation de degré 2 !)

$$\left[p^2 + (1-p)^2 \right] \tanh(2\kappa) + 4p(1-p) \tanh \kappa = 1$$

- Constante critique : $\frac{1}{2}$ racine double.

$$\frac{1}{2} \tanh(2\kappa) + \tanh \kappa = 1$$

$$\kappa_c = 0,653$$

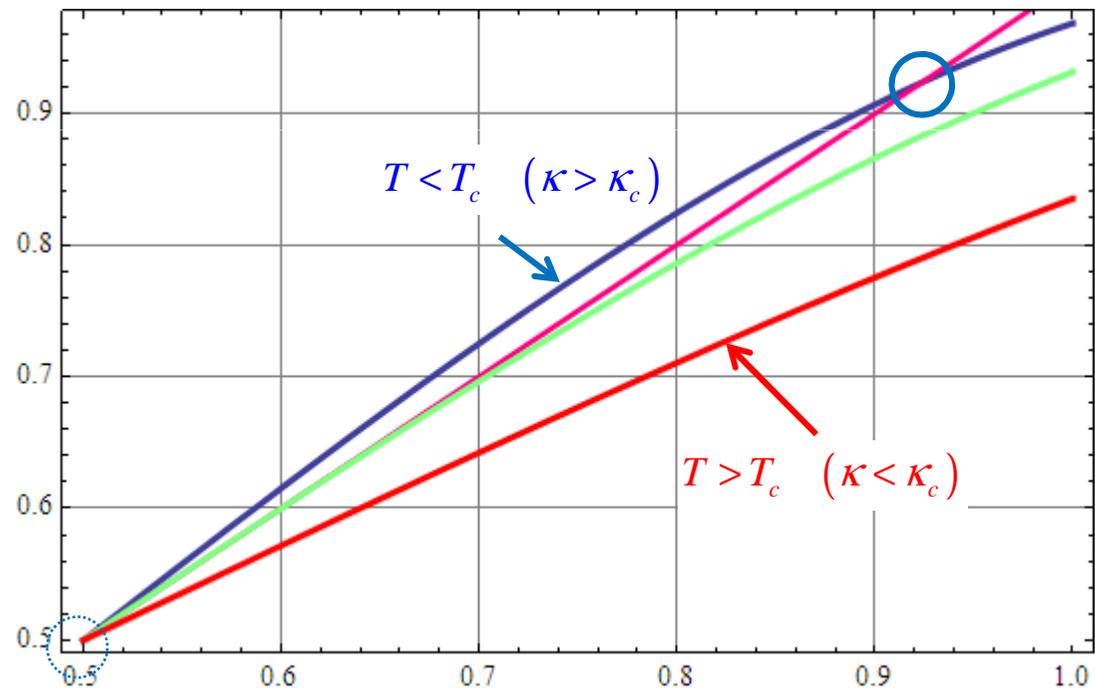
$$(\kappa_c)_{\text{Onsager}} = \ln(1 + \sqrt{2}) \approx 0,882$$

$$(\kappa_c)_{\text{CM}} = 0,5$$

$$\frac{T_c}{J} = 3,063$$

$$\left(\frac{T_c}{J} \right)_{\text{Onsager}} = 2,269$$

$$\left(\frac{T_c}{J} \right)_{\text{CM}} = 4,0$$



Magnétisation spontanée

Constante critique : $\frac{1}{2}$ racine double.

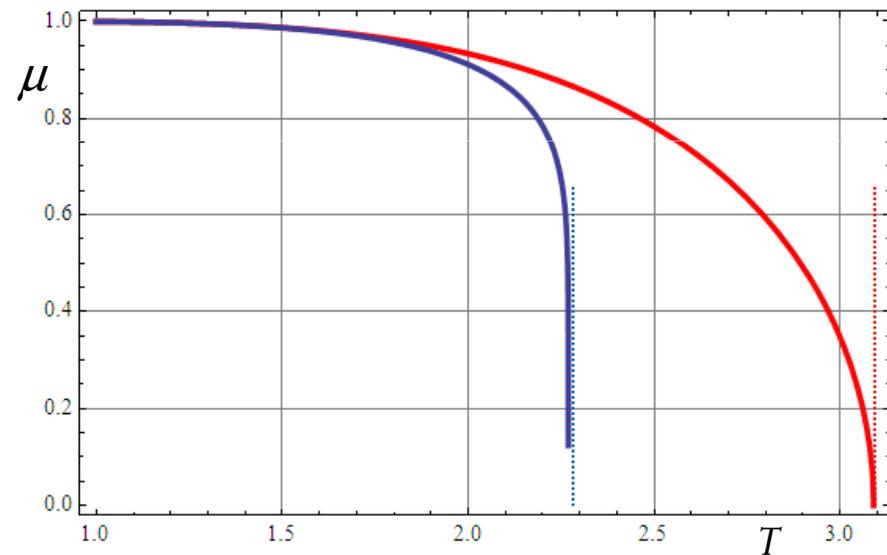
$$\left[p^2 + (1-p)^2 \right] \tanh(2\kappa) + 4p(1-p) \tanh \kappa = 1$$

$$\frac{1}{2}(1+\mu^2) \tanh(2\kappa) + (1-\mu^2) \tanh \kappa = 1$$

$$x = \exp(-\kappa)$$

$$\mu = \sqrt{1 + \frac{4x^4(1+x^2)}{(1-x^2)^3}} = 1 - 2x^4 - 8x^6 - 20x^8 + O(x^{10})$$

$$(\mu)_{\text{Yang}} = \left[\frac{(1+x^2)\sqrt{1-6x^2+x^4}}{(1-x^2)^2} \right]^{\frac{1}{4}} = 1 - 2x^4 - 8x^6 - 34x^8 + O(x^{10})$$



Du côté de la Mécanique quantique

- Le sens de la mesure (une projection)
- Une théorie linéaire. **Ce n'est pas une approximation.**
- Une fonction d'onde réellement complexe.
- Des corrélations bizarres.
- De belles inégalités (Weyl, Bell, ...)
- Des applications immédiates (calcul quantique, cryptographie quantique)
- Les techniques mathématiques sont familières. Difficultés :
 - quelles notions garder (espace et temps)
 - et quelles notions redéfinir (à peu près tout).

Densités de probabilité

- Univocité de la définition (forme bilinéaire définie positive, ...).
- Addition cohérente de tous les processus possibles.
- Pas de précédent en Physique non quantique. Image.
- Les fonctions d'onde sont, aussi bien, des vecteurs d'un espace normé.

Notation de Dirac

ket $|f\rangle$ représente la fonction d'onde.

bra $\langle g|$ représente son adjoint hermitique.

Bracket $\langle f|g\rangle$ représente le produit scalaire et $\langle f|g\rangle = (\langle g|f\rangle)^*$

$$f(x) = \sum_n c_n \psi_n(x)$$

$$|f\rangle = \sum_n c_n |\psi_n\rangle \quad c_n = \langle \psi_n | f \rangle$$

$$|f\rangle = \sum_n |\psi_n\rangle \langle \psi_n | f \rangle = \underbrace{\left(\sum_n |\psi_n\rangle \langle \psi_n | \right)}_{=\text{Id}} |f\rangle$$

Produit tensoriel $|f\rangle \otimes |g\rangle \mapsto |f\rangle |g\rangle$

Pas de clonage !

$$|f\rangle \rightarrow |f\rangle|f\rangle$$

$$|g\rangle \rightarrow |g\rangle|g\rangle$$

$$(|f\rangle + |g\rangle) \rightarrow (|f\rangle + |g\rangle)(|f\rangle + |g\rangle) \quad (\text{définition})$$

$$\neq |f\rangle|f\rangle + |g\rangle|g\rangle \quad (\text{linéarité})$$

Intrigante intrication ...

Si $\{S_1 + S_2\}$ décrit par produit tensoriel entre un état de S_1 et un état de S_2 : état dit *séparable* ou *factorisable* (S_1 non altéré par mesures sur S_2).

$$|\Psi_{1+2}\rangle = |\psi_1\rangle \otimes |\psi_2\rangle = |\psi_1\rangle |\psi_2\rangle$$

$$|\Psi_{\text{sep}}\rangle = \frac{1}{\sqrt{2}} (|+\rangle_1 |-\rangle_2 - |-\rangle_1 |-\rangle_2) = \frac{1}{\sqrt{2}} (|+\rangle_1 - |-\rangle_1) \otimes |-\rangle_2$$

État intriqué : non séparable $|\Psi_{\text{int}}\rangle = \frac{1}{\sqrt{2}} (|+\rangle_1 |-\rangle_2 - |-\rangle_1 |+\rangle_2)$

Impossible de séparer conceptuellement les systèmes, de parler de « l'état de S_1 » : seul le système global $\{S_1 + S_2\}$ a un état défini.

Corrélation parfaite des mesures réalisées sur S_1 avec les mesures réalisées sur S_2 . Si S_1 donne « + » (50% des cas), le système total est projeté dans l'état $|+\rangle_1 \otimes |-\rangle_2$. La mesure de S_2 donnera « - » avec certitude.

$$\left[\begin{aligned} {}_1\langle + | \Psi_{\text{int}} \rangle &= \frac{1}{\sqrt{2}} \left(\underbrace{{}_1\langle + | + \rangle_1}_{=1} \otimes |-\rangle_2 - \underbrace{{}_1\langle + | - \rangle_1}_{=0} \otimes |+\rangle_2 \right) \\ &= \frac{1}{\sqrt{2}} |-\rangle_2 \end{aligned} \right]$$

Inégalité triviale, calcul exact

- Paire de particules:

$$P(0, \overline{90}) = P(0, \overline{45}, \overline{90}) + P(0, \overline{45}, \overline{90})$$

$$P(0, \overline{45}) + P(45, \overline{90}) \geq P(0, \overline{90}) \quad P(0, \overline{45}, \overline{90}) \leq P(45, \overline{90}) \quad P(0, \overline{45}, \overline{90}) \leq P(0, \overline{45})$$

$$P(\uparrow\uparrow) = P(\downarrow\downarrow) = \frac{1}{2} \sin^2\left(\frac{a-b}{2}\right)$$

$$P(\uparrow\downarrow) = P(\downarrow\uparrow) = \frac{1}{2} \cos^2\left(\frac{a-b}{2}\right)$$

$$\mathbf{a} \cdot \mathbf{b} = \begin{cases} \uparrow\uparrow = \downarrow\downarrow = 1 \\ \uparrow\downarrow = \downarrow\uparrow = -1 \end{cases}$$

$$E(a, b) = \langle \mathbf{a} \cdot \mathbf{b} \rangle = \sin^2\left(\frac{a-b}{2}\right) - \cos^2\left(\frac{a-b}{2}\right) = -\cos(a-b)$$

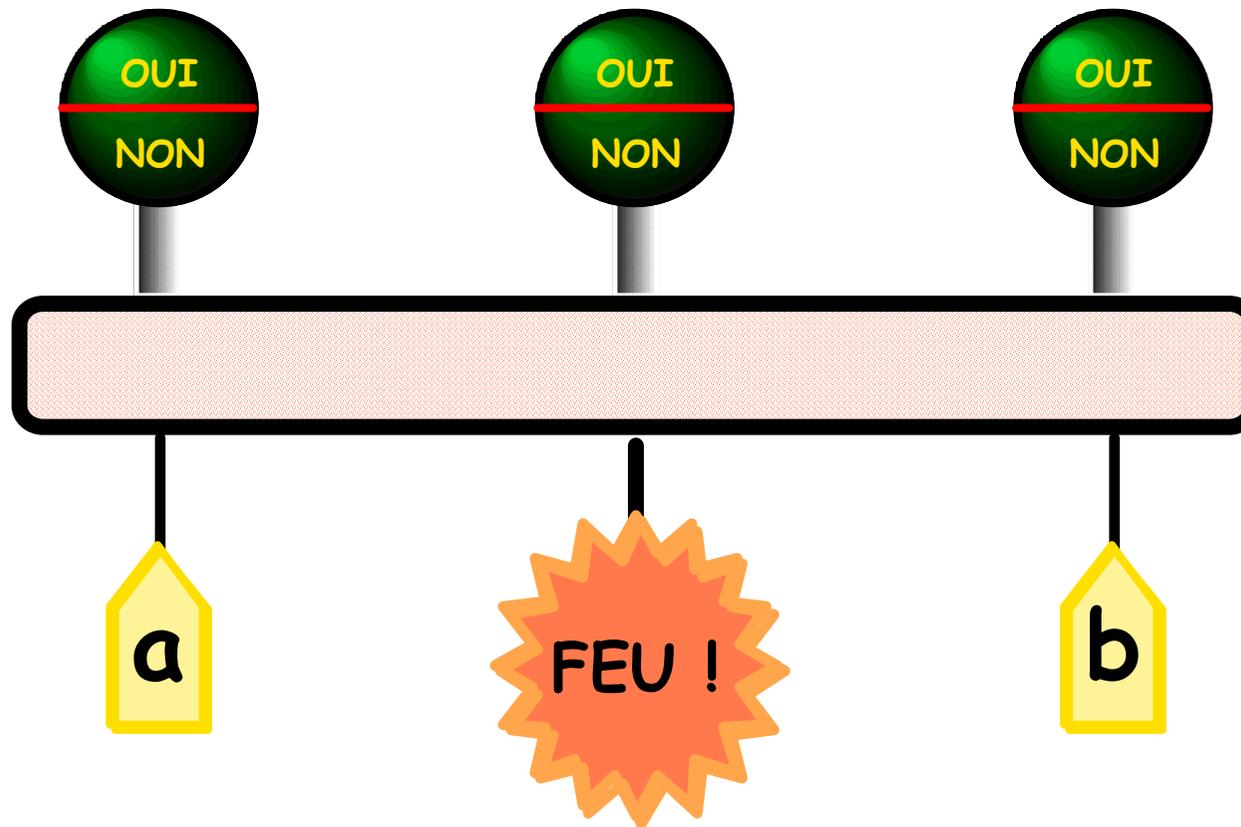
- Conflit ouvert

$$\underbrace{\frac{1}{2} \sin^2\left(\frac{\pi}{8}\right) + \frac{1}{2} \sin^2\left(\frac{\pi}{8}\right)}_{0,1464} \geq \underbrace{\frac{1}{2} \sin^2\left(\frac{\pi}{4}\right)}_{0,2500}$$

Hypothèses

- Aucune hypothèse sur l'appareillage.
- Aucune mention du déterminisme.
- Aucune référence particulière (spin ...).
- Aucune référence à la mécanique quantique.

Rien que des corrélations.



- $t = t_1$: Feu ! puis vérification (oui central).
- $t = t_1 + T - \delta$: injection de a et de b .
- $t = t_1 + T$: réponse des autres indicateurs.
- $c \delta \ll L$.

Théorème

- Résultats de A et B corrélés.
- Corrélations inexplicables localement ; il faut une explication.
- Si on peut identifier plusieurs facteurs causaux, les fluctuations résiduelles sont indépendantes.

$$P(A, B | a, b) \neq P_1(A | a) \times P_2(B | b)$$

$$P(A, B | a, b, \lambda) = P_1(A | a, \lambda) \times P_2(B | b, \lambda)$$

$$P(A, B | a, b) = \int P(A, B | a, b, \lambda) f(\lambda) d\lambda$$

$$I(a, b, a', b') = E(a, b) + E(a, b') + E(a', b) - E(a', b')$$

$$|I(a, b, a', b')| \leq 2$$

Cas particulier

$$I(a,b,a',b') = E(a,b) + E(a,b') + E(a',b) - E(a',b')$$
$$|I(a,b,a',b')| \leq 2$$

$$a = 0, a' = \pi/2, b = \pi/4 \text{ et } b' = -\pi/4$$

$$I(a,b,a',b') = -3\cos\left(\frac{\pi}{4}\right) + \cos\left(\frac{3\pi}{4}\right) = -2\sqrt{2}$$

$$|I(a,b,a',b')| \approx 2,828$$

$$2,828 > 2$$

$$E(a,b) = P_{a,b}(\text{oui}, \text{oui}) + P_{a,b}(\text{non}, \text{non}) - P_{a,b}(\text{oui}, \text{non}) - P_{a,b}(\text{non}, \text{oui})$$

$$P(A,B|a,b,\lambda) = P_1(A|a,\lambda) \times P_2(B|b,\lambda) \quad P(A,B|a,b) = \int P(A,B|a,b,\lambda) f(\lambda) d\lambda$$

$$E(a,b) = \int \bar{A}(a,\lambda) \bar{B}(b,\lambda) f(\lambda) d\lambda \quad \begin{cases} \bar{A}(a,\lambda) = P_1(\text{oui}|a,\lambda) - P_1(\text{non}|a,\lambda) \\ \bar{B}(a,\lambda) = P_2(\text{oui}|b,\lambda) - P_2(\text{non}|b,\lambda) \end{cases}$$

$$E(a,b) \pm E(a,b') = \int \bar{A}(a,\lambda) [\bar{B}(b,\lambda) \pm \bar{B}(b',\lambda)] f(\lambda) d\lambda$$

$$0 \leq P_1 \leq 1 \Rightarrow |\bar{A}(a,\lambda)| \leq 1 \quad 0 \leq P_2 \leq 1 \Rightarrow |\bar{B}(b,\lambda)| \leq 1$$

$$|E(a,b) \pm E(a,b')| \leq \int \underbrace{[\bar{B}(b,\lambda) \pm \bar{B}(b',\lambda)]}_{\leq 1} f(\lambda) d\lambda \quad |E(a',b) \pm E(a',b')| \leq \int \underbrace{[\bar{B}(b,\lambda) \mp \bar{B}(b',\lambda)]}_{\leq 1} f(\lambda) d\lambda$$

$$\int f(\lambda) d\lambda = 1 \Rightarrow |E(a,b) \pm E(a,b')| + |E(a',b) \pm E(a',b')| \leq 2$$

Relations entropiques d'inégalité ?

- Bach : pas de staccato sur les basses : $\Delta f \Delta t \geq c$
- Gabor (1940) : durée de signal et largeur spectrale

$$G(\omega) = \frac{1}{\sqrt{2\pi}} \int g(t) \exp(-i\omega t) dt \quad g(t) = \frac{1}{\sqrt{2\pi}} \int G(\omega) \exp(i\omega t) dt$$
$$\langle t \rangle = \frac{\int t \bar{g}(t) g(t) dt}{\int \bar{g}(t) g(t) dt} \quad \langle \omega \rangle = \frac{\int \omega \bar{G}(\omega) G(\omega) d\omega}{\int \bar{G}(\omega) G(\omega) d\omega}$$

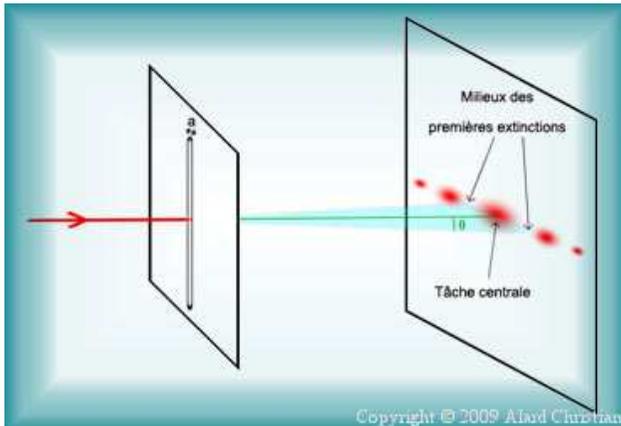
$$\Delta \omega \Delta t \geq 1$$

$$\psi(x,t) \mapsto p(x,t) = |\psi(x,t)|^2$$
$$\varphi(p,t) \mapsto p(x,t) = |\varphi(x,t)|^2$$

Relations entropiques d'inégalité ?

$$-\frac{a}{2} \leq x \leq \frac{a}{2} \quad \psi(x) = \frac{1}{\sqrt{a}} \quad \varphi(p) = \sqrt{\frac{a}{2\pi}} \frac{\sin\left(\frac{ap}{2}\right)}{\frac{ap}{2}} \quad \Delta p_y \geq \frac{2\pi\hbar}{\Delta x} = \frac{2\pi\hbar}{a} \quad \theta \approx \frac{p_y}{p_x} = \frac{2\pi\hbar}{a(\hbar k_y)} = \frac{\lambda}{a}$$

$$\Delta x = \frac{a}{2\sqrt{3}} \quad \Delta p = \infty$$

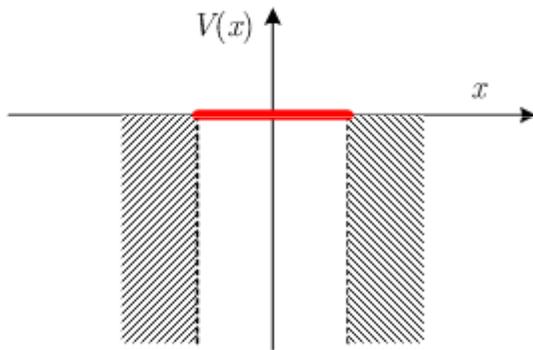


$$S_p = - \int_{-\infty}^{\infty} \frac{a}{2\pi} \left[\frac{\sin(ap/2)}{ap/2} \right]^2 \ln \left\{ \frac{a}{2\pi} \left[\frac{\sin(ap/2)}{ap/2} \right]^2 \right\} dp = \frac{J}{\pi} + \ln \left(\frac{\pi}{2a} \right)$$

$$S_x = - \int_{-\frac{a}{2}}^{\frac{a}{2}} \frac{1}{a} \ln \left(\frac{1}{a} \right) dx = \ln a$$

$$S_p + S_x = C$$

Relations entropiques d'inégalité ?



$$-a \leq x \leq a$$

$$u_n^+ = \frac{1}{\sqrt{a}} \cos \left[\pi \left(n - \frac{1}{2} \right) \frac{x}{a} \right] \quad E_n = \frac{\pi^2 \hbar^2}{2ma^2} \left(n - \frac{1}{2} \right)^2$$
$$(\Delta x)^2 = \frac{a^2}{3} \left[1 - \frac{3}{2\pi^2 \left(n - \frac{1}{2} \right)^2} \right] \quad (\Delta p)^2 = \frac{\pi^2 \hbar^2}{a^2} \left(n - \frac{1}{2} \right)^2$$

$$\Delta x \Delta p = F(n)$$

$$S_p + S_x = Cte$$

Ubiquité de l'entropie ?

Thermodynamique

Information

Quantique