

Suggestions concernant le problème du chat

BERNARD D'ESPAGNAT

Laboratoire de Physique Théorique, Université Paris-Sud, 91405 Orsay, France

RESUME. La théorie de la décohérence permet une meilleure compréhension de l'opération de mesure. Mais des questions restent pendantes lorsque, par la pensée, on intègre l'observateur lui-même, conçu comme un système physique conscient, dans l'instrument d'observation. Le présent article, de nature très spéculative, esquisse les contours d'une approche possible du problème. L'idée centrale en est qu'il pourrait exister, dans les systèmes physiques, une sorte de proto-conscience revêtant la forme d'"états de conscience intérieurs", mais que, au sein des plus simples d'entre eux, de tels états seraient privés et non prédictifs. Ce serait seulement dans le cas de systèmes complexes qu'ils engendreraient intersubjectivité et prévisibilité.

ABSTRACT. Decoherence theory allows for a better understanding of the quantum measurement riddle. But still, questions remain when the observer himself, viewed as a conscious physical system, is thought of as being an integral part of the whole observational device. The present paper outlines a tentative and highly unconventional approach to this problem. Its basic idea is that, within the physical systems, there may conceivably exist some kind of a protomentality having the form of "internal states of consciousness", but that, within the simplest systems, such states would be essentially private and nonpredictive. Only in the case of complex systems would they generate both intersubjectivity and predictivity.

1 Introduction

Que le chat ici mentionné soit celui de Schrödinger ne surprendra, bien sûr, aucun lecteur de ces Annales. Lorsque j'ai choisi le métier de physicien théoricien, le spectre du chat en question - inquiétant mais fascinant - hantait déjà la profession. Cinquante cinq ans après le voici toujours, profilé dans l'ombre. Au prix de quelles évolutions conceptuelles nous laissera-t-il en repos? Le présent article tente témérairement d'en esquisser une.

Ce courage m'a probablement été insufflé par l'exemple qu'Olivier Costa de Beauregard nous donne: car celui-ci n'a jamais hésité à proposer des idées neuves, quelque choquantes qu'elles pussent être aux yeux des "bien pensants" que nous tendons à être tous. En ce cas, ce sera la seconde fois que je

lui serai redevable car, avec Louis Leprince-Ringuet, il fut l'une des deux personnes qui m'encouragèrent le plus dans le choix ci-dessus évoqué d'un métier qui m'a apporté beaucoup. Ce sont là des choses que l'on n'oublie pas; pas plus que je ne saurais oublier l'attention, à sa manière pleine de sollicitude, avec laquelle Louis de Broglie pilota ma thèse tout en lâchant les rênes autant qu'il le fallait. Ecrire dans les Annales de la Fondation Louis de Broglie un article fêtant les soixante ans de la thèse d'Olivier Costa de Beauregard est donc pour moi un honneur que double un plaisir très particulier et sincère.

2 Conscience et apparence

S'il y a une idée qui nous est commune, à Olivier Costa de Beauregard et à moi-même (mais elle n'est pas notre apanage, et il s'en faut!), c'est bien qu'il est impossible de ne pas considérer l'esprit - ou "la conscience", ou "le mental", ou etc., bref, ce que les Anglo-Saxons appellent "the mind" - comme une donnée primitive. Primitive, du moins, dans le sens qu'elle s'avère non réductible à une "matière" connaissable qui serait "plus première encore". C'est là une prise de position à laquelle me paraissent contraints, en particulier, tous les positivistes conséquents, ou, en tout cas, tous ceux qui, à la suite de Schlick, posent en axiome que le sens d'une proposition *se réduit* à sa méthode de vérification (car toute vérification suppose, quelque part, un acte mental).

Pour ma part, je ne me sens pas porté à spontanément adhérer à des axiomes *a priori*, et pas plus à celui de Schlick qu'à aucun autre. Mais je me fonde sur la physique quantique et comme celle-ci, du moins dans toutes ses versions "orthodoxes", n'est pas ontologiquement interprétable, elle me conduit, par un cheminement qui est, pour moi, *a posteriori*, à, justement, la prise de position que je viens de dire. Il est vrai (et je l'ai toujours reconnu) qu'en toute logique cet argument n'est pas strictement contraignant puisque des modèles cohérents existent qui fournissent les mêmes prédictions vérifiables que la mécanique quantique tout en étant ontologiquement interprétables. L'un de ceux-ci, le plus ancien sans doute mais sans doute aussi le moins artificiel de tous, est celui de l'onde-pilote de Louis de Broglie [1], généralisé par David Bohm [2] et reformulé par John Bell en des termes d'une remarquable clarté [3]. Mais tous ces modèles, y compris le dernier cité, présentent des traits qui font que, à mes yeux comme à ceux de beaucoup d'autres physiciens, ils ne sont aucunement scientifiquement convaincants. Mes raisons d'en juger ainsi ont été exposées en grand détail dans mon dernier livre [4] et ne seront pas reproduites ici.

D'un autre côté, cependant, l'esprit humain est ainsi fait que lorsqu'il cherche à raisonner de façon précise à partir de son expérience communicable, c'est en faisant appel à un langage et à des images réalistes qu'il y parvient le plus commodément. Et ceci a pour conséquence que même si - et c'est mon cas - on ne considère pas tel ou tel modèle réaliste comme étant l'expression d'une vérité ontologique, on peut trouver avantage à faire appel à lui en tant que "laboratoire théorique". Autrement dit, lorsque l'on a à affronter un problème conceptuel ardu, on peut se tourner vers le modèle en question en vue d'y puiser une suggestion de solution. "Les images - disait en substance Louis de Broglie - sont une grande aide pour l'imagination". C'est ainsi que j'utiliserai le modèle Broglie-Bohm dans ce qui suit.

C'est un fait que ce qui est de beaucoup le plus solide dans la mécanique quantique - ce qui, dirai-je, s'est montré jusqu'ici totalement solide - c'est l'ensemble de ses règles de prédiction d'observations. A l'heure actuelle celles-ci n'ont jamais été contredites par un résultat expérimental, alors que les tentatives pour les interpréter en termes descriptifs d'une réalité extérieure ont, au contraire, eu à faire face à beaucoup de difficultés en ce qui concerne la cohérence. Ceci plaide fortement en faveur d'une interprétation strictement instrumentaliste de cette mécanique, et effectivement j'estime pour ma part que, en dernière analyse, c'est là la position la plus prudente et la meilleure. Elle l'est à ce point que certains pourraient même y voir une manière vraiment trop facile d'esquiver les difficultés. A première vue ils pourraient, en effet, penser qu'en nous abstenant ainsi de voir dans la mécanique quantique une description de quoi que ce soit, en la considérant comme un simple jeu de prédictions observationnelles, nous échappons à tous les problèmes conceptuels. Curieusement, cependant, tel n'est pas le cas. Demeure, en effet, le problème bien connu du chat de Schrödinger dès que l'on admet que ce chat est, comme nous, un être conscient. Et la solution ne peut consister à poser - en suivant Descartes! - qu'il n'est qu'une machine car, comme Wigner l'a pertinemment remarqué, le problème ressurgit dès que, par la pensée, nous remplaçons le chat par un être humain (par un de nos amis, propose Wigner: un ami qui éprouve ceci ou cela selon que la particule qui, éventuellement, l'atteint est dans une branche ou l'autre de la fonction d'onde¹). Lorsque la fonction d'onde ψ du système particule-Ami est considérée comme étant une réalité, la difficulté inhérente à ce problème est évidente et bien connue. Elle réside dans le fait que l'Ami prend subitement

¹Inutile de préciser que le cyanure et le décès ne sont pas, dans l'allégorie, des ingrédients indispensables!

conscience de l'état dans lequel l'interaction considérée le met, qu'il réduit par là-même la fonction d'onde, et qu'il faudrait donc, semble-t-il, admettre que la simple prise de conscience d'un individu peut modifier - très appréciablement et sur des distances aussi grandes que l'on voudra - la réalité. Dans l'interprétation instrumentaliste de la fonction d'onde ici proposée cette difficulté n'existe pas, puisque la fonction d'onde n'y est qu'un simple outil de prédiction. Mais il s'en présente une autre, similaire. Elle consiste en ce que rien n'interdit d'imaginer que l'expérience est reproduite à un nombre d'exemplaires arbitrairement grand. On a alors affaire à un ensemble statistique de systèmes complexes particule-Ami le quel, considéré "de l'extérieur" (du point de vue du physicien, O, qui monte l'expérience), est un cas pur, constitué d'éléments tous identiques et décrit par ψ . Mais d'un autre côté les Amis présents dans chacun de ces éléments sont dans des états de conscience qui diffèrent les uns des autres, ce qui fait que l'ensemble doit être un mélange propre, m . La contradiction logique est manifeste et il est clair que, à elle seule, la théorie de la décohérence - qui explique si bien les *apparences* observées - est impuissante à la lever.

3 Réalité et connaissance

Face à ce genre de problèmes - ni mathématiques ni physiques mais essentiellement conceptuels - l'esprit se sent quelque peu désarmé et, dans sa quête d'un cheminement intellectuel susceptible de les résoudre, il est normal qu'il cherche des inspirations du côté de problèmes analogues qu'il sait traiter. C'est à ce titre que le modèle Broglie-Bohm peut éventuellement nous aider. Précisons bien, toutefois, dès le départ la position ici adoptée. Pour les raisons auxquelles allusion déjà a été faite, je tiens la réalité en soi - *alias*, la réalité indépendante, *alias* le "réel" ou le "surréal" - pour irrévocablement voilée, autrement dit non scientifiquement descriptible "telle qu'elle est vraiment". La science atteint seulement notre expérience humaine communicable, *alias* la "réalité empirique". Mais cela n'empêche pas qu'à mes yeux la notion de réalité en soi reste tout à fait essentielle. Je considère donc comme (inévitavelmente inadéquats mais) légitimes, intéressants et éventuellement suggestifs les efforts faits pour la représenter, et en particulier le modèle dont il s'agit. Je verrai par conséquent celui-ci comme une figuration - enfantine, balbutiante mais, encore une fois, éventuellement suggestive - de la réalité en soi, étant bien reconnu que, de cette réalité, nous ne pouvons atteindre que de fort trompeuses apparences (on connaît l'image, souvent utilisée par Bohm, de plusieurs gouttes d'encre encodées dans un magma visqueux puis décodées, et donnant par là l'illusion d'une goutte unique en mouvement; elle illustre bien, dans le cadre du modèle lui-même, l'étendue

du gouffre qui sépare la réalité empirique de la réalité en soi). Comme il est bien entendu qu'il ne s'agit que de cela - que je ne tiens pas le modèle en question pour l'expression de la vérité - je le prendrai sous sa forme la plus simple, non relativiste.

On sait que dans une expérience de pensée de fentes de Young la particule ne passe jamais, selon ce modèle, que par une seule fente. On sait aussi que, malgré cela, le modèle reproduit les prédictions quantiques c'est à dire, en l'espèce, l'apparition de franges sur l'écran. On sait enfin qu'entre ces deux données l'incompatibilité n'est qu'apparente. Elle est levée par le fait que la particule est pilotée par une onde - l'"onde pilote" - qui, elle, passe par les deux fentes et n'est donc pas la même si les deux fentes sont ouvertes ou si une seule l'est.

Ainsi, la particule du modèle est à tout instant en un lieu bien défini alors même qu'elle intervient dans un processus typiquement quantique. A cet égard elle nous fait penser à l'Ami, qui est à tout instant dans un état de conscience bien défini alors qu'il intervient, lui aussi, dans un processus typiquement quantique. Et étant donné que notre problème provient du fait que l'Ami en question a une conscience, il est naturel que, dans le modèle, nous attribuions une conscience à la particule. Pour rendre l'idée, non pas, certes, plausible mais un tout petit peu moins invraisemblable, je remplacerai à partir d'ici la particule par une molécule d'ADN. Dans le cadre du modèle et d'une expérience de fentes de Young utilisant de telles molécules comme projectiles n'importe laquelle de celles-ci constatera alors, à un certain instant, qu'elle est à l'intérieur d'une certaine fente, bien spécifiée. C'est là un état de conscience de la molécule, et comme, dans le modèle, la position de la molécule est un élément de la réalité en soi, cet état de conscience doit lui-même être considéré, dans le modèle ainsi "perfectionné", comme un élément de la réalité en soi.

Puisque que cet état de conscience est tel, autrement dit, est une donnée objective au sens fort, il paraîtrait à première vue tout naturel – et même presque inévitable – de dire qu'il doit être considéré comme appartenant au “domaine public de la connaissance”. En d'autres termes, il semblerait au premier abord que l'on ait à considérer cette prise de conscience, par la molécule, de sa position comme pleinement équivalente à tous égards à une mesure – éminemment objective - que ferait de cette position un observateur extérieur par le moyen d'un instrument enregistreur. Et cependant, la thèse d'une telle équivalence n'est pas soutenable puisque les franges sont observées alors que toute détermination, par l'observation, de la fente par laquelle passe chaque molécule empêche ces franges de se produire.

4 Réalité et objectivité

Faut-il conclure de là que l'idée d'attribuer un "état de conscience" à la molécule se heurte à une impossibilité logique? Non, car le raisonnement qui précède est moins irréprochable qu'il ne semble. En effet, l'état de conscience en question est, par hypothèse, intimement lié à la variable "position de la molécule". Or n'oublions pas que, dans le modèle, celle-ci est une variable cachée, un élément de la réalité en soi, une variable du domaine de l'ordre implicite et non pas du domaine public, alors qu'au contraire toute mesure relève de ce dernier par définition. Pour échapper à l'aporie il suffit donc de poser que l'état de conscience de la molécule est lui-même un élément de la réalité en soi, autrement dit, est ce que j'appellerai un "état de conscience intérieur", échappant au domaine public.

A première vue, l'idée paraît paradoxale: on a l'habitude de considérer que le "domaine public de la connaissance" est plus objectif que les états de conscience intérieurs. On tend à penser que les positions des aiguilles sur les cadrans, observées lors d'une mesure, sont des réalités en soi alors que les états de conscience, au contraire, étant subjectifs, semblent ne pas être de telles réalités. Mais, déjà, une analyse des fondements conceptuels de la mécanique quantique montre (cf.[4]) que la première de ces deux idées n'est pas exacte: que les grandeurs dont il s'agit sont du domaine de la réalité empirique et non de la réalité en soi. Et, dans le cadre du modèle, cette conclusion demeure puisque les grandeurs en question relèvent de l'ordre explicite et non pas de l'ordre implicite. Ci-dessus nous considérons la mesure effectuée par un appareil enregistreur comme un paradigme d'objectivité. C'était oublier que de telles mesures sont à objectivité seulement faible. A la réflexion l'idée ici considérée échappe par conséquent au reproche d'incohérence. Nous poserons donc que l'état de conscience intérieur de la molécule est bien une réalité en soi. Que pour la molécule elle-même, l'énoncé : "je suis maintenant à tel endroit" est objectif au sens fort, mais qu'il n'est pas nécessairement extrapolable au domaine public de la connaissance et que, par conséquent, l'opérateur O qui a monté toute l'expérience n'a pas l'obligation de prendre en compte son existence dans ses prévisions observationnelles. En d'autres termes, les états de conscience intérieurs ici introduits n'ont pas, en général, de rôle prédictif.

5 Conscience et prédictif

Ce raisonnement a été essentiellement conduit à l'intérieur du modèle Broglie-Bohm qui, je le répète, n'est pas à mes yeux scientifiquement convaincant mais que je considère comme un fort bon "inspirateur". Je suggère donc de, certes, mettre ce modèle à l'écart, mais tout en conservant

l'essentiel de la conception générale qu'il vient de nous suggérer. Plus précisément je dirai donc qu'il n'est pas illégitime de poser que certains systèmes physiques, même éventuellement microscopiques, peuvent avoir des états de conscience - ou de "protoconscience" - intérieurs, qui sont des éléments de la réalité en soi, mais cela à la condition de ne pas faire intervenir cette hypothèse "métaphysique" dans les calculs de prédiction d'observations.

La question qui se pose ensuite est de savoir s'il a des circonstances dans lesquelles des énoncés correspondant à de tels états de conscience intérieurs peuvent, malgré tout, être étendus au domaine public de la connaissance événementielle (laquelle, encore une fois, est à objectivité seulement faible). Pour tenter de répondre, considérons, à nouveau, le modèle qui nous a servi. On peut, soit imaginer que l'expérience est faite avec des projectiles "plus macroscopiques" que des molécules, soit considérer des systèmes conscients plus complexes encore.

Dans le cadre du premier choix la décohérence intervient et a pour conséquence que, relativement à toutes les prévisions relatives aux expériences que l'opérateur O est, en pratique, capable d'effectuer, celui-ci peut très bien remplacer, dans tous ses calculs, le cas pur ψ par le mélange propre m . Ce dernier a pour composantes des fonctions d'onde, disons ψ_1 et ψ_2 (si, pour faire simple, on ne considère que deux fentes), auxquelles - en pratique toujours! - il est parfaitement légitime d'attribuer le rôle prédictif normalement dévolu aux fonctions d'onde dans le formalisme quantique. Il en résulte que, dans la pratique, les états de conscience intérieurs - qui sont, je le répète, des éléments de la réalité en soi - peuvent alors être identifiés à ces éléments de la réalité empirique que sont les états de conscience prédictifs: ceux auxquels nous faisons allusion quand nous disons que tel résultat de mesure a été - ou pourrait être - observé.

Le second des deux choix évoqués plus haut consiste à revenir au cas d'une expérience du type "chat de Schrödinger" ou "ami de Wigner". Dans un tel cadre, la décohérence joue bien entendu, aussi, les termes d'interférence sont quasiment supprimés et, comme ci-dessus, les cas purs ne peuvent pratiquement plus être distingués des mélanges propres. Ici aussi, dès lors, l'extension de l'"intérieur" au "prédictif" peut, en pratique, être faite. L'état de conscience intérieur de l'Ami - que j'appellerai A - peut, pratiquement, être identifié à un "état de conscience prédictif", c'est-à-dire finalement à une donnée (du type: " A a observé a ") appartenant au domaine public. Et, cela étant, de nouveau on peut mettre à l'écart le modèle Broglie-Bohm tout en conservant l'essentiel de la conception générale qu'il suggère. Plus précisément, on peut donc dire que dans l'expérience du chat - entendons, bien sûr: dans celle de l'Ami substitué au chat - il est légitime de considérer simultanément l'ensemble statistique des systèmes particule-Ami

- voire celui des systèmes particule-Ami-environnement - comme étant un cas pur, *et* chaque Ami comme ayant un état de conscience intérieur bien défini. Cela est non contradictoire à condition de considérer les états de conscience en question comme des éléments de la réalité en soi, donc comme étant personnels ou, autrement dit, généralement cachés au public; et, en conséquence, de ne pas voir dans leur existence une donnée devant, pour d'impérieuses raisons de principe, être prise en considération dans le cadre d'une discussion de la cohérence intersubjective du formalisme de la mesure.

Bien entendu il reste vrai, d'une part, que - à la différence de nos molécules supposées - les Amis eux-mêmes ne commettent aucune faute s'ils font un usage prédictif de de leurs états de conscience intérieurs, et d'autre part, que, la décohérence aidant, l'opérateur O, lorsqu'il songe aux les états de conscience intérieurs des Amis, peut, en pratique et sans aucun risque d'erreur, les concevoir par la pensée comme s'ils étaient prédictifs. Ainsi, au total, peut-être n'est-il pas absurde de se demander s'il n'y a pas là, en définitive, l'idée directrice d'une amorce de solution au fameux paradoxe du chat de Schrödinger.

6 Remarque

Considérons une expérience de corrélation à distance du "type Aspect", dans laquelle les deux observateurs distants - A et B, disons - mesurent la polarisation de leurs photons respectifs selon la même direction. Leurs états de conscience intérieurs sont des éléments de la réalité en soi et sont strictement corrélés mais on sait que cette corrélation n'est aucunement attribuable à des "causes communes" qui seraient des événements ayant pris place à la source. Dans le modèle Broglie-Bohm il résulte d'un calcul effectué par John Bell [3] que la corrélation en question résulte du fait que ce qui se passe dans l'instrument utilisé par A influence directement - via l'onde pilote non locale - le comportement du photon observé par B et vice-versa. Mais ce qui se passe c'est que A et B tendent tout naturellement à interpréter leurs états de conscience intérieurs comme étant des états de conscience prédictifs et, corrélativement, à penser leur corrélation selon un modèle mécanistique, par causes communes à la source; alors qu'aussi bien cette nature prédictive des états de conscience que ce modèle mécanistique ne sont que de simples apparences. Ce dernier modèle n'est ainsi qu'un élément de la "réalité empirique" construite par nous à partir de notre expérience intersubjective, et l'on voit bien sur cet exemple que, dans le cadre des vues ici décrites, il est légitime de parler d'une co-émergence de la conscience et de la réalité empirique à partir du réel en soi. Avec, "en prime" une précision importante, qui est

que lorsque l'on s'exprime ainsi, ce qu'il faut entendre par conscience, ce n'est pas ce qu'il y a de plus profond dans la pensée, ce n'est même pas le jeu des états de conscience intérieurs, mais c'est seulement celui des états de conscience prédictifs.

7 Conclusion

Pour conclure, je noterai un certain parallélisme entre les idées et les résultats qui précèdent d'une part et certaines des vues de Kant d'autre part. Selon Kant en effet - et c'est là, d'après Putnam [5], un "point vraiment subtil" de sa pensée - le fait que notre perception ne nous fait rien connaître pouvant être appelé un "objet en soi" ne concerne pas seulement les objets du "sens externe". Il porte également sur les objets du "sens interne", autrement dit sur les sensations. Alors que pour Berkeley ces dernières appartenaient, tout comme l'esprit lui-même, au domaine du réel en soi, pour Kant elles ne sont elles-mêmes que des phénomènes. Or dans l'approche ci-dessus décrite il nous a fallu dédoubler la notion d'état de conscience. Nous avons dû considérer, d'une part les "états de conscience intérieurs" et d'autre part les "états de conscience prédictifs". Il n'est pas inintéressant de remarquer que, par là même (et sans l'avoir cherché!) nous avons donné - partiellement - raison aussi bien à Berkeley qu'à Kant. A Berkeley en ce qui concerne les états de conscience intérieurs, puisque ceux-ci sont, selon nous, des éléments de la réalité en soi, et à Kant en ce qui concerne les état de conscience prédictifs, dont nous avons reconnu le statut de simples phénomènes.

References

- [1] L. de Broglie, *J.Phys.* **5**, 225 (1927).
- [2] D.Bohm, *Phys.Rev.* **85** 166 (1952).
- [3] J.S.Bell, *Speakable and Unspeakable in Quantum Mechanics*, Cambridge University Press, London, 1987, chapitre 15.
- [4] B.d'Espagnat, *Trait de physique et de philosophie*, Fayard, 2002.
- [5] H. Putnam, *Reason, Truth and History*, Cambridge University Press, London 1981.

Reçu le lundi 6 janvier 2003.