LE PHYSIQUE, LE MORPHOLOGIQUE, LE SYMBOLIQUE

REMARQUES SUR LA VISION

I. - CRITIQUE DU DUALISME SYMBOLIQUE/PHYSIQUE ET DU SOLIPSISME MÉTHODOLOGIQUE

1. Le dualisme symbolique/physique

1.1. Le paradigme classique — dit symbolique — des sciences cognitives actuelles est computationnel, symbolique et fonctionnaliste (pour une introduction, cf. les dossiers dans *Le Débat*, 1987 et *Préfaces*, 1988).

(i) Il postule d'abord l'existence de représentations mentales neurologiquement implémentées (et donc physiquement réalisées) dans des états mentaux. Il s'oppose sur ce point aux positions réductionnistes éliminationnistes et physicalistes qui considèrent que les représentations mentales ne sont que des artefacts de la description psychologique et ne possèdent pas d'existence objective en tant que telles (cf., par exemple, Churchland, 1984)[1].

(ii) Il postule ensuite que ces représentations mentales sont de nature symbolique, c'est-à-dire qu'elles appartiennent à un langage mental interne (le « mentalais » de Fodor) possédant la structure d'un langage formel (avec ses symboles, ses expressions, ses règles d'inférences, etc.). Il s'oppose sur ce point aux conceptions qui estiment qu'un certain nombre de résultats expérimentaux (par exemple, sur les rotations d'images mentales) plaident en faveur de représentations mentales *topologico-géométriques* non propositionnelles (cf. Kosslyn, 1980 et Shepard-Cooper, 1982).

(iii) Il postule enfin que, comme en informatique, on peut découpler les problèmes de matériel (hardware) de ceux de logiciel (software) et que les représentations mentales symboliques sont, en ce qui concerne leur structure formelle et leurs contenus informationnels, indépendantes de leur implémentation dans leur substrat physique (magnétique, neuronal,

1. Pour plus de précisions concernant les références placées entre parenthèses, dans cet article, se reporter à la Bibliographie, p. 180.

Revue de synthèse: IVe S. Nos 1-2, janvier-juin 1990.

etc.). Il s'oppose, sur ce point, aux conceptions *émergentielles* qui considèrent au contraire que l'on doit concevoir ces structures formelles comme des structures stables émergeant de processus dynamiques, coopératifs et statistiques sous-jacents (cf. Thom, 1972, 1980 ; Zeeman, 1977 ; PDP, 1986 ; Smolensky, 1988 ; Petitot, 1986b, 1989f,g,i). Une épistémologie de l'émergence interroge dans le paradigme symbolique une conception formaliste et « descendante » (*top-down* en jargon) du traitement de l'information et lui oppose une conception naturaliste et « ascendante » (*bottom-up* en jargon).

Pour le paradigme symbolique, les sciences cognitives doivent par conséquent se fonder dans une théorie computationnelle des manipulations formelles de représentations symboliques. Ces représentations traitent de l'information et, en particulier, de l'information issue du monde extérieur. Elles peuvent acquérir ainsi un contenu sémantique. Mais la causalité naturelle des opérations qui agissent sur elles et leur permettent d'agir (par exemple, sur des comportements à travers des contenus intentionnels d'attitudes propositionnelles) est une causalité strictement formelle et syntaxique. Autrement dit, en tant qu'états et processus mentaux, elles sont fermées à leur sémantisme.

1.2. Le mentalisme computationnel du paradigme classique est inséparable, en ce qui concerne l'information servant d'*input* aux calculs mentaux, d'un objectivisme physicaliste standard. Selon ce dernier, ce qu'il y a d'objectif dans l'environnement se réduit à ce qu'enseigne la physique fondamentale standard : atomes, rayonnement, ondes sonores, etc. On en arrive ainsi à un véritable *dualisme* (fortement réminiscent des dualismes philosophiques traditionnels) entre *le symbolique et le physique*. Dans son ouvrage de référence *Computation and Cognition*, Zenon Pylyshyn a excellemment exposé celui-ci. L'information externe étant conçue de façon physicaliste, elle est *a priori sans signification* pour le système cognitif. Elle se trouve soumise à une transduction par des modules périphériques (ces modules comprennent les récepteurs sensoriels comme la rétine ou la cochlée mais peuvent se prolonger à des transducteurs compilés), transduction qui la convertit en information interne (fréquences de *firing* de neurones) computationnellement significative. Il existe évidemment une corrélation causale nomologiquement descriptible entre l'information physique externe et l'information computationnelle interne produite par la transduction. Mais cela n'implique pas pour autant l'existence d'une science nomologique du rapport *significatif* que le sujet entretient avec son environnement. D'une part, en effet, la transduction décrite physiquement et causalement est cognitivement opaque. Sa fonction est non symbolique. Elle fait partie de l'architecture fonctionnelle qui contraint formellement la structure des algorithmes mentaux. D'autre part, la signification est le résultat des opérations effectuées par les représentations mentales symboliques et celles-ci ne sont pas causalement déterminées par le contenu physique objectif des états de choses externes. D'où, selon Pylyshyn, un dualisme physico-symbolique strict. Il existe une coupure irréductible entre le cognitif interne et le physique externe. Il existe un langage physique universel, cohérent et unificateur, composé de termes physiques. Mais il n'existe pas de descriptions physiques, dans ce langage, de ce qui est significatif dans l'environnement pour un sujet cognitif (cf. Petitot, 1989f). On pourrait aligner les citations concernant cette « strongest constraint » et cet « extremely serious problem » : « the relevant aspects of the environment are generally not describable in physical terms », « psychological regularities are attributable to *perceived*, not physically described properties », « the general failure of perceptual psychology to adequately describe stimuli in physical terms », etc. (Pylyshyn, 1986, p. 166-167). Il faut donc disposer de concepts perceptuels et cognitifs fonctionnels. Mais ceux-ci sont *sans* contenu physique. Le lexique physique et le lexique cognitif ne s'apparient pas naturellement. Ils ne sont compatibles qu'à travers les transductions.

On remarquera que de telles affirmations ne sont acceptables que sous certaines hypothèses :

(i) ce qui existe d'objectif dans l'environnement se réduit à ce que décrit la physique fondamentale standard ;

(ii) ce qui est significatif doit, pour être significatif, être au préalable représenté ;

(iii) la représentation s'identifie à un calcul : l'esprit est computationnel.

1.3. Comme l'ont noté de nombreux auteurs (Putnam, Searle, Dreyfus, etc.), deux grands problèmes demeurent énigmatiques dans le paradigme classique.

(i) Du côté du sujet, le problème *du sens et de l'intentionalité.* Comment des représentations mentales symboliques peuvent-elles acquérir un sens, une interprétation, une dénotation, une orientation intentionnelle vers le monde externe ? Comment un système cognitif peut-il agir en fonction du sens des symboles et des expressions symboliques alors qu'il ne possède de relations causales qu'avec la forme (logico-syntaxique) de ceux-ci ? Il ne suffit pas de dire que le sens est le résultat d'une « interaction » sujet-monde puisque cette interaction n'est pas nomologiquement descriptible et explicable.

(ii) Du côté du monde, le problème *de la manifestation qualitative et* *morphologique des phénomènes.* Comme Ray Jackendoff y a beaucoup insisté, on ne peut se borner à poser que le monde phénoménologique de l'expérience est un simple résultat des opérations de « l'esprit computationnel » (Jackendoff, 1987). Encore faut-il comprendre la part de ces opérations, en général opaques pour la conscience phénoménologique, qui se trouve devenir constitutive de la structuration qualitative du monde en choses, états de choses, événements, processus, etc., perceptivement appréhendables et linguistiquement descriptibles. En effet, le processus computationnel est inconscient. Seules quelques-unes des structures qu'il produit sont conscientes. On peut alors, comme Jackendoff, adopter un point de vue « projectiviste » faisant du monde phénoménologique un monde « projeté » résultant d'une « projection » de constructions cognitives, poser que la plus grande partie de la structure interne des constituants du langage mental (ce que Jackendoff appelle la « structure conceptuelle ») n'est pas projetable et faire de la « conscience » phénoménologique (différente, donc, de l'esprit computationnel) le corrélat (en un sens proche de celui de la corrélation noèse/noème chez Husserl) de ce monde projeté (le « Mind-Mind problem »). Mais on peut également, comme nous le ferons plus bas, utiliser les résultats scientifiques théoriques et expérimentaux qui démontrent l'existence de structures morphologiques et qualitatives *objectives émergeant*, par un processus dynamique (auto)organisateur, des substrats physiques. Ce point de vue proprement *« morpho-génétique »* s'oppose au point de vue *« morpho-projectif ».* Il prend appui sur l'existence démontrée d'un niveau de réalité morphodynamique que l'on pourrait appeler un niveau *« phéno-physique »* (expression phénoménologique du niveau de réalité proprement physique à travers un processus *naturel* objectif, non cognitif, de phénoménalisation des substrats matériels).

2. Les limites épistémologiques du cognitivisme symbolique : la non-prise en compte de la dimension morphodynamique

2.1. Fortement tributaire des recherches en Intelligence Artificielle (IA) dont il a hérité du point de vue computo-représentationnel, le cognitivisme symbolique, dans son rapport aux neurosciences, à la psychologie et à la philosophie de l'esprit, a permis des progrès décisifs dans la compréhension et la formalisation des mécanismes mentaux constitutifs du « sens commun » (applications de règles en fonction du contexte, inférences, décisions, représentations des connaissances, rôle causal du contenu intentionnel des attitudes propositionnelles dans le comportement et l'action, etc.). Se voulant science des états et des processus mentaux, son projet est de comprendre les sujets cognitifs en tant que

« systèmes symboliques physiques » et de *naturaliser* l'esprit, le langage et le sens.

Pour comprendre à quel point son statut épistémologique est toutefois délicat et problématique, il suffit de remarquer qu'il reprend l'ensemble des problèmes de la tradition sémantique (logique, philosophie analytique, etc.) en termes computationnels, qu'il les relie aux neurosciences et que, sur cette base, il transforme les descriptions noético-noématiques de l'expérience phénoménologique en sciences naturelles.

On peut s'étonner par conséquent du fait que, dans l'ensemble des débats (fort vifs) qui se sont développés à son sujet, les concepts ontologiques, théoriques et épistémologiques les plus fondamentaux — comme ceux de matière, de réalité physique, d'idéalité symbolique, de causalité, etc. — soient utilisés de façon non critique dans leur acception souvent la plus banale.

Par exemple, une des raisons principales du rejet des conceptions émergentielles par le cognitivisme symbolique vient d'une incompréhension de l'épistémologie de l'émergence. Lorsqu'un système est un système à *deux* niveaux d'organisation, par exemple un niveau qualitatif « macro » et un niveau physique « micro » sous-jacent, le niveau supérieur « macro » est causalement (au sens de la causalité matérielle) réductible au niveau inférieur. Mais cela ne l'empêche évidemment pas de posséder des éléments de structure très largement indépendants de la structure fine « micro » sous-jacente. Ces éléments possèdent une certaine autonomie objective. Cela est tout à fait banal en physique (phases, transitions de phases, défauts dans les cristaux liquides, etc.). Comme l'a souligné Searle, ce n'est que si l'on identifie un phénomène à sa genèse causale — autrement dit, si l'on passe subrepticement d'un réductionnisme causal, justifié, à un réductionnisme ontologique matérialiste, dogmatique et donc injustifié — que l'on est conduit à dénier l'autonomie et la réalité objective des niveaux supérieurs.

De même, lorsque certains auteurs s'essaient à dépasser le dualisme du physique et du symbolique pour développer un monisme naturaliste, ils le font en général à partir d'un matérialisme vulgaire ou d'un physicalisme ne tenant aucun compte de récents résultats fondamentaux de certaines disciplines physiques. Par exemple, on cherchera à développer un behaviorisme physicaliste faisant des contenus mentaux de simples réponses de l'organisme à des états de choses. Ou bien on posera, au contraire, l'identité entre les états mentaux et des états cérébraux, quitte à affronter les multiples difficultés qui en découlent.

De même encore, pour en revenir au dualisme, le ***solipsisme méthodologique*** est la conséquence directe d'une certaine conception de l'objectivité physique. Selon Fodor par exemple, il est impossible d'introduire dans une psychologie scientifique le rapport significatif qu'un sujet cognitif entretient avec son environnement. En effet, ce rapport n'est pas, nous l'avons vu, nomologiquement légalisable dans l'état actuel des connaissances. On ne pourrait donc l'introduire qu'en termes, non scientifiques, de sens commun. D'où la légitimité de la morale provisoire solipsiste : seuls les contenus « étroits » (*de dicto* et non *de re*, ne dépendant que du sujet, de son langage mental interne et non pas de sa relation contextuelle à l'environnement) interviennent dans l'individuation et l'identification des états mentaux et possèdent des capacités causales (cf. Jackendoff, 1987).

2.2. Tout cela pour dire que l'ensemble du débat actuel sur la cognition dépend de façon déterminante de la préconception que se font les cognitivistes de l'objectivité physique. Un de leurs préjugés fondamentaux est *qu'il n'existe pas de physique qualitative des formes, de physique morphologique, de phéno-physique.* Or ce préjugé n'est justifié que pour la physique fondamentale (relativité générale et mécanique quantique incluses). *Il ne l'est absolument plus* si l'on prend en compte les résultats, profonds, nombreux et convergents, de l'ensemble des disciplines physiques qui se sont intéressées ces dernières années aux phénomènes d'(auto)organisation des substrats matériels.

Nous avons longuement commenté ailleurs ces travaux mathématiques et physico-mathématiques remarquables (cf., par exemple, Petitot, 1982, 1986b, 1989g et, surtout, leurs bibliographies) : théorie qualitative de la structure et de la stabilité structurelle des systèmes dynamiques non linéaires, de leurs attracteurs et de leurs bifurcations, attracteurs étranges et chaos déterministe, théorie des singularités et de leurs déploiements universels, théorie des phénomènes critiques (transitions de phases, etc.) et des phénomènes de rupture de symétrie dans les phases mésomorphes, structures dissipatives, etc. Ces résultats ont montré expérimentalement et démontré mathématiquement que, dans de nombreux systèmes naturels organisés à (au moins) deux niveaux (cf. plus haut), le niveau « macro » (global, grossier, en général finiment descriptible) émergeant, à travers des comportements collectifs ordonnés et coopératifs, du niveau « micro » sous-jacent (local, complexe, en général non finiment descriptible) est essentiellement organisé autour des *singularités* des processus physiques « micro ». Les singularités *structurent morphologiquement* les phénomènes. Elles sont *phénoménologiquement dominantes* et soumises à des contraintes *abstraites et universelles* (« platoniciennes ») mathématiquement formulables et largement indépendantes de la physique « fine » des substrats.

Le concept de physique qualitative des formes, de physique morphologique, de phéno-physique, *appartient désormais au concept de réalité* *objective.* Ce fait a, selon nous, des conséquences incalculables, à la fois théoriques et épistémologiques, pour le cognitivisme. En effet, comme nous le verrons, *le morphologique constitue un moyen terme entre le physique et le symbolique* : il est d'origine physique (émergent) mais sans être pour autant matériel, il est formel mais sans être pour autant symbolique ; il est *topologiquement et géométriquement* formel et non pas *logiquement* formel. Sa prise en considération rend légitime la double hypothèse suivante :

(i) il existe une information morphologique et qualitative présente dans le monde externe qui, tout en étant d'origine physique, est néanmoins de nature phénoménologique et, à ce titre, intrinsèquement significative ;

(ii) cette information morphologique est reconstituée après transduction et sert de base aux processus proprement symboliques de traitement de l'information.

Selon nous, la plupart des difficultés (voire des apories et des paralogismes) du cognitivisme classique proviennent du fait qu'il cherche à engendrer le morphologique à partir d'une conception logico-combinatoire (somme toute encore logiciste et analytique) du syntaxique et du sémantique alors que cela est pourtant clairement impossible, puisque les dimensions intrinsèquement spatio-temporelles et dynamiques du morphologique *ne sont pas* d'ordre formel au sens logico-symbolique (bien que physiquement réalisées, elles ne sont pas non plus d'origine physique). Comme y insiste Jackendoff, des représentations sémantiques propositionnelles ne peuvent pas être mises au fondement d'une expérience des formes.

2.3. Le problème philosophique qui intervient ici est considérable (cf. Petitot, 1982, 1986a, 1989f). Notre propos n'est pas de le reprendre. Mais nous ne saurions trop insister sur la limite fondamentale que constitue l'orientation dogmatiquement propositionaliste du cognitivisme symbolique. Une telle orientation n'est, en effet, légitime que dans le cadre d'un objectivisme logique, d'une sémantique formelle et/ou d'une logique phénoménologique des essences. Elle est *incompatible* avec une thèse *naturaliste* quelle qu'elle soit, car il n'existe pas de formes symboliques dans la nature externe ou interne. Il ne peut exister tout au plus que des formes géométriques et dynamiques. Toute naturalisation de l'esprit, du langage et du sens présuppose donc une révolution dans la conception du formel héritée du formalisme logique. Elle présuppose catégoriquement que les formes de l'esprit, du langage et du sens soient des formes géométriques et dynamiques. Ces formes doivent évidemment être symboliquement traductibles et manipulables à des niveaux cognitifs supérieurs de représentation. Mais leur *type d'objectivité* ne peut pas, pour des raisons de principe, être originairement celui de l'objectivité symbolique.

Disons brièvement que, si elle est *naturelle*, la « formellité » de l'esprit, du langage et du sens ne peut pas être symbolique. Pour la décrire et l'expliquer, il faut passer en quelque sorte d'une symbologie à une topologie.

Paraphrasant un aphorisme de Kant (« les intuitions sans concepts sont aveugles et les concepts sans intuitions sont vides »), on pourrait dire que le cognitivisme symbolique est « aveugle » et « vide » dans la mesure où il n'arrive pas à élaborer une authentique phénoménologie de la perception. En vérité, aucun passage du physique au symbolique n'est envisageable tant que l'on ne tient pas compte du fait :

(i) que le physique est spatio-temporellement conditionné (ce que Kant appelait l'Esthétique transcendantale) ;

(ii) que ce conditionnement spatio-temporel de la physique fondamentale est prolongeable aux dimensions topologiques, géométriques et dynamiques de la phéno-physique morphologique ;

(iii) que le symbolique constitue un niveau formel de surface par rapport aux infrastructures morphologiques.

3. La thèse de la morphodynamique cognitive et le principe de double émergence

Les thèses sous-jacentes à notre réflexion sont donc les suivantes.

(i) Entre le physique et le symbolique il existe la médiation du morphologique. Sans elle, il est impossible de dépasser le dualisme du physique et du symbolique et d'accéder à une théorie naturaliste intégrée (moniste mais non réductionniste) de leur unité ontologique.

(ii) Les structures morphologiques sont de façon générale les produits de processus dynamiques d'organisation des susbtrats (physiques ou mentaux). Elles émergent des substrats et sont phénoménologiquement dominées par les discontinuités qualitatives issues des singularités, des bifurcations, des instabilités structurelles, de ces processus dynamiques.

(iii) Les structures qualitatives émergentes existent aussi bien du côté du sujet cognitif que du côté du monde naturel.

(iv) L'information morphologique résiste à la transduction. Elle est encodée dans, et véhiculée par, les signaux lumineux et sonores, puis décodée-recodée par les transducteurs. Mais, au cours de cette opération, elle se reconstitue en restant en grande partie isomorphe à elle-même. Les discontinuités qualitatives sont « contagieuses » : elles se transfèrent de substrat à substrat.

Du côté du sujet cognitif, le programme de recherche d'une morphodynamique a pour vocation de développer une idée maîtresse introduite par R. Thom et Ch. Zeeman il y a déjà plus d'une vingtaine d'années, à savoir qu'une unité sémantique est identifiable à *la topologie d'un attracteur* d'une dynamique neuronale sous-jacente et que les structures combinatoires et logico-algébriques des automatismes de la compétence doivent par conséquent être interprétées comme des régularités émergentes stables. Cette idée a été extensivement développée en sémio-linguistique par l'école morphodynamique (cf. Thom, 1972, 1980, 1988 ; Wildgen, 1982 ; Brandt, 1986 ; Petitot, 1977, 1979, 1982, 1983, 1985, 1988, 1989a, c, d, f). Elle a été également — et indépendamment — développée dans les modèles connexionnistes du paradigme dit sub-symbolique (cf., par exemple, PDP, 1986 ; Smolensky, 1988 ; Amit, 1989). Le principal apport de ces modèles plus récents est d'avoir explicité les dynamiques « concrètes » qui intervenaient dans les modèles morphodynamiques généraux. Cela permet de spécifier ce que l'on entend par « substrat mental ». Mais, à part cela, les principaux concepts dynamiques du connexionnisme (attracteurs, bassins d'attraction, fonctions de Liapounov, stabilité structurelle, bifurcations d'attracteurs, quasi-attracteurs, ruptures de symétrie, dynamiques rapides et dynamiques lentes, phénomènes coopératifs et propriétés émergentes, etc.) sont les concepts de dynamique qualitative, de théorie de la bifurcation, de théorie des singularités, de thermodynamique statistique et de théorie des phénomènes critiques que les modèles morphodynamiques avaient déjà transférés (d'ailleurs dans l'incompréhension la plus générale) dans le domaine des disciplines psychologiques et sémio-linguistiques au début des années 1970.

Du côté du monde naturel, le programme de recherche d'une morphodynamique a pour vocation d'étudier les processus de phénoménalisation des substrats matériels (externes, non internes), de théoriser mathématiquement l'information morphologique qui en émerge, de comprendre comment cette information morphologique se trouve encodée dans, et véhiculée par, les signaux lumineux et sonores.

Ayant traité ailleurs des relations entre la morphodynamique et le connexionnisme (Petitot, 1989f, i), nous nous focaliserons ici sur le problème *du type mathématique de l'information morphologique*. La possibilité d'élaborer une *phénoménologie de la perception* satisfaisante constituant un enjeu décisif dans les débats que nous avons évoqués, nous nous limiterons à l'exemple de la perception visuelle. De façon à pouvoir être suffisamment précis tout en demeurant à l'intérieur de limites raisonnables, nous nous bornerons à *un* problème *très particulier* (mais fondamental), celui de la reconstruction des objets à partir de leurs contours apparents. Qui plus est, nous dialoguerons avec des théories particulières, mais généralement acceptées (bien que parfois controversées sur certains points), nommément celles de David Marr et de Jan Koenderink. Cela nous permettra d'expliciter certaines des thèses proposées.

II. - INFORMATION MORPHOLOGIQUE ET THÉORIE DES SINGULARITÉS EN PERCEPTION VISUELLE

Des quatre domaines fondamentaux des sciences cognitives : perception, langage, inférence, action, nous choisissons donc, pour notre exemple, le premier. Des deux points de vue traditionnels : celui concernant le développement et celui concernant l'organisme adulte, nous choisissons le second. Des quatre niveaux d'analyse : biologique (mécanismes neurophysiologiques), psychologique (processus fonctionnels de détection, représentation, stockage, utilisation finalisée d'informations, etc.), computationnel (modélisation algorithmique), mathématique (propriétés formelles de la compétence), nous choisissons le troisième et le quatrième, mais dans une optique non symbolique. Nous allons, en fait, esquisser de façon brève et relativement peu technique quelques éléments de morphodynamique qui permettent d'analyser mathématiquement les contraintes topologiques, géométriques et optiques qui contraignent de façon essentielle la formation des images visuelles et leur traitement computationnel.

1. Processus modulaires et processus centraux. Traitement ascendant et traitement descendant

La vision computationnelle est la discipline théorique qui élabore des modèles mathématiques pour les processus de construction de représentations tridimensionnelles (3D) distales à partir d'images rétiniennes bidimensionnelles (2D) proximales. Elle doit donc élucider théoriquement et modéliser mathématiquement :

(i) les processus physiques de constitution de scènes externes morphologiquement organisées ;

(ii) les processus optiques d'encodage et de propagation de ces informations morphologiques ;

(iii) le processus physico-géométrique de formation des images par projection ;

(iv) le processus sensoriel périphérique d'analyse du signal (transduction) ;

(v) la façon dont l'information morphologique ainsi décodée et recodée contraint de façon essentielle la construction des représentations ;

(vi) les rapports (par exemple de compilation) entre les niveaux successifs de représentation (du topologico-géométrique vers le symbolique) ;

(vii) la façon dont les représentations de niveau supérieur (3D et au-delà) possèdent ou non un contenu *objectif.*

Il existe au moins deux grandes conceptions de la vision computationnelle. Pour les expliciter brièvement, reprenons l'opposition fodorienne entre processus périphériques modulaires et processus centraux non modulaires (cf. Fodor, 1984). La thèse est qu'il existe (au moins) deux types très différents de systèmes cognitifs. Les premiers sont les systèmes périphériques modulaires. Ils ont pour fonction de transformer les informations neuronales périphériques fournies par les transducteurs en représentations possédant un format propositionnel adéquat pour les calculs symboliques mentaux. Ce sont des transducteurs compilés, fonctionnant automatiquement et de façon strictement ascendante (« bottom-up » : du périphérique vers le central) comme des réflexes computationnels. Ils sont spécifiques et informationnellement cloisonnés (c'est-à-dire insensibles aux croyances, aux connaissances, aux attentes, etc., du sujet). Ils formulent des hypothèses et effectuent des inférences non démonstratives permettant aux stimuli sensoriels proximaux d'être transformés en représentations sur des objets distaux.

Mais il y a également les systèmes cognitifs centraux, qui sont non modulaires, non spécifiques, non cloisonnés, descendants, interprétatifs (et donc sensibles aux croyances, connaissances, attentes, etc.). Dans la mesure où il n'existe aucun contrôle nomologique de leur fonctionnement, ils ne sont pas, selon Fodor, traitables scientifiquement : c'est le problème du holisme sémantique. Ils sont « isotropes » (toute croyance, toute connaissance, toute attente est partiellement pertinente pour le traitement et l'interprétation de toute sortie des modules) et « quiniens » (l'ensemble des croyances, etc., influe sur chaque traitement, etc.). D'où d'ailleurs, chez Fodor, une critique de l'Intelligence Artificielle et des systèmes experts qui traitent les systèmes centraux *comme si* ils étaient modulaires, spécifiques, non isotropes et non quiniens.

Un des aspects du holisme sémantique est précisément le solipsisme méthodologique débattu plus haut.

Dans une approche « descendante » (« top-down ») inspirée de l'IA, on considère que le traitement de l'information rétinienne se réduit essentiellement à des processus *d'interprétation* des images, processus *inférentiels* régis par des connaissances. Mais une telle approche n'est pas directement applicable à la vision *naturelle.* Pour celle-ci, l'environnement est trop complexe, trop fluctuant et trop peu contraint pour être traitable à partir de mécanismes de détection de traits et d'applications de règles. Dans la vision naturelle, il existe une partie considérable du traitement de l'information qui est modulaire et « ascendante » (« bottom-up »). Plusieurs modules fonctionnels spécifiques, indépendants et fonctionnant en parallèle coopèrent dans le processing visuel précoce et leur produit intégré sert de base aux niveaux supérieurs (centraux) de représentation et d'interprétation.

La théorie de David Marr qui nous servira de base de discussion est modulaire et ascendante. Comme l'explique le collègue de Marr, Tomaso Poggio, elle considère que la tâche centrale de la vision computationnelle est de résoudre un *problème inverse.* Il existe un processus de projection des scènes 3D sur des images 2D. Le problème inverse est celui de la reconstruction des scènes 3D à partir des images 2D. *Mais l'on voit que ce problème inverse est double, à la fois cognitif et objectif.* Il est *objectif* dans la mesure où l'on peut le traiter de façon purement géométrique et optique, sans aucune référence à un esprit computationnel. Il est également *cognitif* dans la mesure où l'on peut le traiter en termes computationnels. La thèse est que *le problème inverse objectif contraint et finalise le problème inverse cognitif.* Autrement dit, il est impossible d'expliciter les algorithmes de la vision computationnelle si l'on ne connaît pas au préalable de façon précise le *type mathématique* des structures informationnelles à traiter.

Un tel point de vue est *néo-écologique.* Rappelons que l'on appelle « écologisme » la thèse *réaliste* de James Gibson selon laquelle, dit en termes plus actuels :

(i) il existe dans l'environnement des structures qualitatives et cognitivement significatives qui sont objectives sans être pour autant strictement physiques (ce que nous avons appelé le phéno-physique) ;

(ii) le système visuel détecte et extrait ces invariants phéno-physiques et construit sur cette base objective ses inférences et ses interprétations.

L'écologisme s'oppose au solipsisme méthodologique. Selon lui, les représentations symboliques représentant l'information ont pour fonction *d'expliciter* certains aspects de celle-ci.

2. Les trois niveaux de la théorie de Marr et leurs corrélats objectifs

La théorie de Marr concerne la vision computationnelle. On en trouvera une analyse conceptuelle et épistémologique dans Kitcher, 1988. Pour une introduction générale à la théorie de la vision, on pourra consulter, par exemple, les excellents Pinker, 1984, Brady, 1982, Ballard-Brown, 1982, Ullman, 1984, Stillings *et al.*, 1987.

Selon Marr, la « quintessence » de la vision comme traitement d'information est d'extraire, par corrélation, de l'information sur les objets du monde objectif externe à partir de la façon dont la lumière réfléchie par les surfaces physiques engendre des patterns 2D I(x,y) de luminance. A travers la transduction rétinienne effectuée par les photorécepteurs, ces patterns se trouvent discrétisés (comme les pixels d'un écran). La seule information *explicite* est, à l'entrée du système, I(x,y). A la sortie opèrent des représentations de haut niveau effectuant des tâches cognitives supérieures : différenciation d'objets, repérage de positions, appréhension de mouvements, perception des dimensions, formes et textures des surfaces, reconnaissance d'objets, regroupement par classes de ressemblance (catégorisation), etc. Comment s'opère donc, dans une théorie ascendante comme celle de Marr, le passage vers ce que G. Miller appelait « the crowning intellectual accomplishment of the brain », à savoir le monde réel ?

Marr introduit plusieurs niveaux de représentation explicitant certains aspects de l'information encodée dans les patterns I(x,y). Parmi ceux-ci trois sont fondamentaux.

(i) Le premier niveau, dit niveau 2D du « primal sketch » ou de *l'esquisse primaire,* est celui du traitement du signal I(x,y). Il s'agit *d'en expliciter la morphologie et l'organisation géométrique* de façon à pouvoir opérer des segmentations qui serviront de support aux phases intermédiaires et aux phases finales, cognitives et inférentielles, d'interprétation, de reconnaissance, de compréhension, etc. Ce niveau se décompose lui-même en (au moins) deux sous-niveaux.

(i)-a. Au niveau du « raw primal sketch », il s'agit essentiellement d'une analyse *locale* du pattern I(x,y) en termes de discontinuités qualitatives : segments de bords, terminaisons de bords, discontinuités d'orientation de bords (coins), petits domaines fermés (« blobs »), petits segments de barres, etc.

(i)-b. Au niveau du « full primal sketch », ces éléments locaux (souvent en mouvement) se trouvent agrégés et organisés, ce qui engendre des effets gestalistes bien connus : bords virtuels, etc.

(ii) Le second niveau, dit niveau 2-1/2D (pour bien montrer qu'il est intermédiaire entre le niveau 2D et le niveau 3D), est le niveau essentiel de la théorie de Marr. Nous y reviendrons plus loin. C'est un niveau unitaire globalement organisé où convergent et s'intègrent plusieurs computations modulaires effectuées sur l'esquisse primaire : les contours des surfaces visibles, les textures, la stéréopsie, le mouvement, l'ombrage, etc. Il représente le monde externe comme composé de surfaces visibles remplies de qualités sensibles et se déplaçant dans $\mathbb{R}^3$. Il n'est ni sensoriel (puisque les surfaces sont distales) ni objectif (puisque les apparences sont encore subjectives). C'est le niveau de *l'apparaître phénoménologique*. Comme nous allons le voir, il est d'essence proprement *morphologique*.

(iii) Le troisième niveau, dit niveau des modèles 3D, est celui, proprement objectif, des choses réelles, des volumes matériels et de leurs propriétés réales. C'est à partir de lui qu'opèrent les tâches cognitives supérieures et les constituants de la structure conceptuelle au sens de Jackendoff, par exemple la décomposition hiérarchique de formes en parties, la constitution de prototypes, etc. On peut faire l'hypothèse que la perception est un processing ascendant « 2D→2-1/2D→3D→ Structure conceptuelle » possédant des feed-back descendants (anticipations, inférences, interprétations, etc.) « Structure conceptuelle →3D→2-1/2D ». Le niveau 2-1/2D serait donc la fin du processing perceptif proprement ascendant. Comme le dit Marr, *c'est celui de la « perception pure »* (d'où son importance).

On remarquera que les niveaux 2D et 3D possèdent des corrélats objectifs. Les corrélats objectifs (non cognitifs) du niveau 2D relèvent, par exemple, de l'optique ondulatoire, de la photométrie, de l'analyse spectrale et de l'analyse de Fourier, de la théorie du signal, etc., c'est-à-dire des théories physico-mathématiques permettant de comprendre la formation d'images. Les corrélats objectifs (non cognitifs) du niveau 3D sont non moins évidents. Ils relèvent par exemple de la géométrie de l'espace, de la structure du groupe de Lie SO(3) des rotations de $\mathbb{R}^3$, de la mécanique du mouvement des solides, de la représentation des volumes, etc. Et il est clair que les théories objectives de ces corrélats objectifs contraignent et finalisent les algorithmes opérant sur ces niveaux puisqu'elles *déterminent* le *type* de l'information qui doit être explicitée et la nature des tâches computationnelles à effectuer.

Or, curieusement, on n'admet pas en général que le niveau 2-1/2D puisse posséder également des corrélats objectifs. Toujours fidèle au dualisme physique/symbolique, on postule une simple complémentarité entre le traitement numérique de l'image (analyse du signal et théorie de l'information) et son interprétation symbolique (structures sémantiques, inférences, etc.). Entre le numérique et le sémantique, on n'introduit pas en général ce qui est pourtant le caractère le plus manifeste de la perception visuelle, à savoir d'être une perception de formes. Cela est d'autant plus étrange que les théories géométriques qui permettent d'analyser les formes comptent parmi les plus profondes, les plus vastes et les plus prestigieuses de toute la géométrie. *Cette méconnaissance théorique constitue selon nous la limite principale des théories actuelles de la vision computationnelle.* Notre thèse est que :

(i) le niveau 2-1/2D de Marr *possède bien pour corrélat objectif un niveau de réalité* ;

(ii) ce niveau est précisément le niveau *morphologique* de la « phénophysique » ;

(iii) la théorie *objective* (physico-mathématique) de ce niveau —

théorie qui existe — contraint donc et finalise de façon essentielle tous les algorithmes envisageables au niveau 2-1/2D.

3. Le niveau 2D et le concept de discontinuité qualitative

La façon dont Marr conçoit le niveau 2D de l'esquisse primaire est exemplaire de sa conception. A ce niveau se nouent trois dimensions :
(i) les données de la neurophysiologie ;
(ii) le traitement du signal (transduction) ;
(iii) la finalisation des algorithmes rétiniens par le problème inverse objectif (au sens exposé plus haut).

3.1. Les données de la neurophysiologie.

Rappelons très brièvement et très sommairement quelques éléments de la structure générale du système visuel (cf. Buser et Imbert, 1987).

La rétine réalise une énorme *compression* de l'information visuelle et cela essentiellement grâce à l'organisation *antagoniste* centre-périphérie des champs récepteurs des *cellules ganglionnaires* dont les axones constituent le nerf optique. Ces neurones visuels sont sous-jacents aux photorécepteurs superficiels. Ils répondent essentiellement aux discontinuités. La compression de l'information rétinienne fait passer d'environ 160 millions de photorécepteurs à environ 1 million de fibres dans le nerf optique. L'image est ainsi traitée de façon modulaire et organisée en traits distinctifs (arêtes rectilignes contrastées, courbure de bord, mouvement d'un contour selon une direction donnée). Il existe des champs de fibres — des modules — spécialisées dans certaines opérations et opérant sur l'ensemble de la rétine. D'où une *cartographie* du message rétinien relayée avec une bonne *rétinotopie* (une bonne préservation des relations topographiques) jusqu'au cortex visuel primaire. Le relais fondamental est *le corps genouillé latéral* dont les cellules sont analogues aux ganglionnaires rétiniennes et encore plus sensibles au contraste local. Les représentations cartographiques s'y superposent en couches (en registres). D'où une organisation modulaire en *colonnes*, dites *colonnes de projection*, associées à une même zone du champ visuel. Les différences d'organisation et de physiologie des cellules rétiniennes se traduisent dans ces structures supérieures post-rétiniennes par l'innervation de couches différentes. Les opérations des différentes classes fonctionnelles de cellules rétiniennes (en particulier ganglionnaires) sont donc maintenues *séparées* (modularité).

Après le corps genouillé latéral, les radiations optiques traversent la substance blanche et arrivent à l'aire visuelle primaire occipitale (aire striée) : aire principale 17 et aires secondaires 18 et 19. Les colonnes genouillées sont projetées avec préservation de la rétinotopie. Le cortex strié est organisé lui aussi modulairement en colonnes (superposition de couches, cf. les travaux de Hubel et Wiesel) ce qui permet de représenter avec une bonne rétinotopie sur la *surface* du cortex non seulement la *position* dans le champ visuel mais également *d'autres* variables comme la dominance oculaire et l'orientation. L'existence de colonnes de dominance oculaire et de colonnes d'orientation dont les ensembles sont *indépendants* et *transversaux* l'un à l'autre implique que l'aire primaire soit décomposée en *hypercolonnes* (d'environ 1 mm^2 de section) dont chacune traite les contours contrastés (les discontinuités qualitatives) dans toutes les directions de vision binoculaire d'un domaine spatial. On peut donc faire l'hypothèse que le cortex strié visuel sert à extraire de façon *topographique* des attributs visuels caractéristiques et stables comme la couleur, l'orientation, la direction, la vitesse. Ces attributs seraient alors *redistribués* de façon globale (*non* topographique) dans les aires secondaires afin d'y être analysés.

La transduction s'opère au niveau des photorécepteurs, évidemment au moyen d'intermédiaires photochimiques. Des pigments rétiniens (chromoprotéines comme la rhodopsine) absorbent l'énergie lumineuse dans les récepteurs photiques. Leur isomérisation déclenche une chaîne d'événements dans le cytoplasme de ces récepteurs, chaîne aboutissant au blocage du courant dans la membrane plasmique et, donc, à une variation du potentiel membranaire.

La rétine contient, entre les photorécepteurs et les cellules ganglionnaires, d'autres couches de cellules (bipolaires, horizontales, amacrines). En ce qui concerne l'analyse morphologique des stimuli (la couleur pose d'autres problèmes), c'est *l'organisation spatiale des champs récepteurs* (c'est-à-dire la surface de l'espace visuel et de la rétine à laquelle une cellule réagit) *qui est essentielle.* La plupart des neurones rétiniens possèdent une organisation *concentrique et antagoniste* de leur champ récepteur. Ils sont, par exemple, Centre-ON et Périphérie-OFF si un stimulus lumineux ponctuel appliqué au centre du champ récepteur conduit à une activation du centre et à une inhibition de la périphérie.

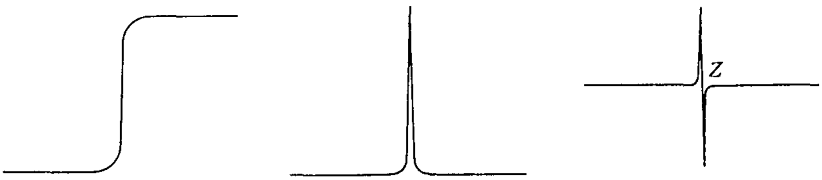
Les cellules ganglionnaires sont essentielles car elles constituent le terme de la transduction. C'est à travers elles (à travers leurs axones constituant, nous l'avons vu, le nerf optique) qu'est transmis le message rétinien aux niveaux post-rétiniens. Elles sont ON, OFF ou ON-OFF et, en ce qui concerne leur réponse temporelle, soit *toniques* (répondant pendant toute la durée du stimulus), soit *phasiques* (répondant seulement à une discontinuité temporelle du stimulus). Elles se regroupent en trois classes fonctionnelles principales X, Y et W. Les cellules X sont énergétiques et toniques. Leur gradient d'antagonisme centre/périphérie est fort, leur résolution spatiale élevée et leur résolution temporelle faible. Ce sont des analyseurs de contrastes spatiaux, et donc de formes. A l'inverse, les cellules Y sont des détecteurs de mouvements et des analyseurs de structures temporelles.

3.2. L'Analyse du signal : critère de zero-crossing et ondelettes.

Évidemment, il existe des interactions subtiles et compliquées entre les différents neurones rétiniens : mécanismes de renforcement et d'inhibition latérale, combinaisons de contrastes spatiaux et chromatiques, etc. Mais l'on voit néanmoins apparaître clairement un certain nombre de faits massifs. Le plus massif est sans doute que, de par la structure de leur champ récepteur et leur caractère tonique, les cellules X ont pour fonction de détecter des contrastes, c'est-à-dire des discontinuités qualitatives de la luminance.

Marr a formalisé ce contenu fonctionnel en introduisant le dispositif de détection de discontinuités qu'il a appelé le critère de *zero-crossing.* L'idée en est simple.

Considérons une fonction différentiable d'une variable réelle f(x). La traversée d'une discontinuité se caractérise par un pic de la dérivée première (distribution δ de Dirac) et par un double pic — un pic positif et un pic négatif séparés par une traversée de 0, c'est-à-dire un « zero-crossing » — de la dérivée seconde (cf. figure 1).



(a) (b) (c)

Figure 1. *Le critère de zero-crossing (d'après Marr, 1982, p. 54). (a) discontinuité de la fonction f, (b) pic de la dérivée f', (c) double pic de la dérivée seconde f''.*

Il s'agit de généraliser à deux dimensions. Pour ce faire on va :

(i) lisser localement le pattern d'intensité I(x,y) à une certaine échelle, par exemple en opérant une convolution $G*I$ avec une gaussienne centrée en un certain point : $G(r) = \exp(-r^2/2\pi\sigma^2)$ (r = distance au point considéré) ;

(ii) considérer les dérivées secondes, c'est-à-dire le laplacien $\Delta(G*I)$.

Marr remarque alors les deux choses suivantes :

(i) Comme $\Delta(G*I) = \Delta G*I$, on peut effectuer la double opération de lissage et de dérivation *en effectuant la convolution du signal avec le laplacien d'une gaussienne.*

(ii) Le profil des champs récepteurs des cellules ganglionnaires X est *précisément* celui du laplacien d'une gaussienne (cf. figure 2).

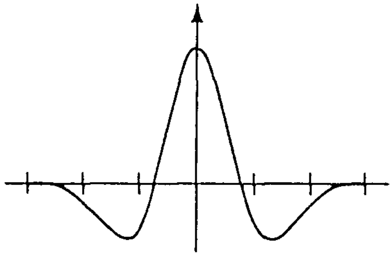


Figure 2. *Le « profil récepteur » du laplacien d'une gaussienne (d'après Marr, 1982, p. 55).*

Si une cellule X ON et une cellule X OFF voisines sont activées ensemble, cela détecte un « zero-crossing » et donc une discontinuité. Ce dispositif a été amélioré après Marr et a suscité de nombreux travaux et discussions (cf., par exemple, Haralick, 1984 ; Grimson-Hildreth, 1985 ; Richter-Ullman, 1986). Son accord avec l'expérience est remarquable.

On peut ainsi faire l'hypothèse qu'il existe des champs de cellules ganglionnaires dont la vocation fonctionnelle est la détection et l'explicitation de discontinuités qualitatives, *localement et à plusieurs échelles.* Ces champs ont une architecture uniforme et modulaire et ils calculent de façon massivement parallèle.

Il semble que l'algorithme de Marr soit un exemple, neurophysiologiquement implémenté, de ce que l'on appelle maintenant une analyse du signal par développement *en série d'ondelettes* (cf. Meyer, 1989). L'analyse en termes d'ondelettes est un processus multirésolution d'analyse de Fourier locale et multiéchelle qui consiste à développer une fonction f(x) (éventuellement très compliquée, fractale par exemple) appartenant à un certain espace fonctionnel (l'espace de Hilbert $L^2(\mathbb{R})$, par exemple) sur une base (orthonormée) d'ondelettes $\Psi_{j,k}$ construites à partir d'une seule fonction Ψ par dilatations et translations. On aura par exemple $\Psi_{j,k} = 2^{j/2}\Psi(2^j x-k)$ où $j,k \in Z$. Les coefficients $f_{j,k}$ du développement de f sur la base $(\Psi_{j,k})$ sont alors obtenus par convolution. Dans le dispositif de Marr, c'est ΔG — c'est-à-dire le profil d'un champ récepteur typique — qui joue le rôle d'ondelette.

3.3. La finalisation des algorithmes rétiniens par le problème inverse objectif.

Comme y insiste Y. Meyer, « une image contient une quantité énorme d'information et une grande partie de cette information est superflue ». Son analyse en termes d'ondelettes permet d'en extraire — d'en expliciter au sens de Marr — « diverses versions schématiques, simplifiées, dont le codage numérique et la transmission soient réalisables avec un coût raisonnable » (Meyer, 1989, p. 40). Il est remarquable que cette schématisation de l'information *coïncide avec une analyse morphologique objective de l'image.* La théorie mathématique des algorithmes de traitement de l'information et la vocation fonctionnelle de la base neurophysiologique implémentante rejoignent les contraintes et les finalisations imposées par les corrélats objectifs. Un « zero-crossing » stable à plusieurs échelles sera l'indice d'une discontinuité objective d'origine géométrique et physique. De telles discontinuités objectives seront préférentiellement traitées comme bords perceptuels. Dès les niveaux les plus précoces de la perception c'est donc son orientation vers les structures objectives (son intentionalité) qui domine. Et cette orientation n'est pas quelconque. Elle repose, insistons-y, sur une structuration *morphologique* du signal.

La base morphologique de la perception est donc imposée par la physiologie et les mathématiques. *Sa nécessité est d'origine à la fois informationnelle et objective. Une théorie mathématique morphologique doit donc être intégrée aux principes de la modélisation en perception visuelle.* C'est en particulier à partir d'elle — et cela pose un magnifique problème mathématique — qu'il faut retrouver les représentations symboliques opérant aux niveaux cognitifs supérieurs.

A propos de cette base morphologique, Marr remarque : « zero-crossing provides a natural way of moving from an analogue or continuous representation like the two-dimensional image intensity values I(x,y) to a discrete, symbolic representation » (p. 67). On ne saurait mieux exprimer le fait que le morphologique se situe entre le continu physique et le discret symbolique et que la vision *naturelle* le présuppose. Pour des systèmes naturels (où le discret symbolique ne peut pas exister d'emblée), *les discontinuités qualitatives morphologiques fournissent, une fois explicitées, la condition de possibilité de la constitution d'un niveau symbolique. En tant que singularités objectives encodées dans le signal, elles supportent l'information.*

> « The raw primal sketch is a very rich description of an image since it contains virtually all the information in the zero-crossings from several chanels. Its importance is that it is the first representation derived from an image whose primitives have a high probability of reflecting physical reality directly » (p. 71).

Comme l'explique T. Poggio :

> « Instead of raw numerical values of intensity, one seeks a more symbolic, compact and robust representation of the visual world : a description of the

world in which the primitive symbols — the signs in which the visual world is coded — are intensity variations » (Poggio, 1984, p. 72).

La structuration conceptuelle de l'image n'est donc pas, selon Marr, essentiellement descendante. Elle n'a pas à être entièrement inférée à partir de connaissances supplémentaires préalables. Elle est en grande partie reconstructible de façon ascendante à partir de la base morphologique extraite de ce que Marr appelle « the physics of the situation ». La connaissance supplémentaire nécessaire *n'est pas conceptuelle.* C'est une « general knowledge embeded in the early visual processes as general constraints, together with the geometrical consequences of the fact that the surfaces coexist in three-dimensional space » (p. 273).

4. L'esquisse 2-1/2D et le problème des contours apparents

Par globalisation, l'esquisse primaire « complète » explicite l'organisation morphologique de l'image. La question devient alors : *comment remonter de l'organisation morphologique 2D à des modèles 3D?* Il est nécessaire de passer par un niveau intermédiaire et l'un des principaux mérites de Marr est d'avoir compris ce point fondamental.

4.1. Le problème du contour comme problème central de la vision computationnelle.

Marr appelle, nous l'avons vu, esquisse 2-1/2D le niveau de cette « intermediate vision » qui constitue le « pivotal point » de toute sa théorie. C'est le niveau de la « pure perception ». C'est « an internal representation of objective physical reality that *preceded* the decomposition of the scene into " objects " ». Pré-conceptuel, modulaire et ascendant, il représente et explicite « what the photons are carrying information about ». A ce titre, « it provides the cornerstone for an overall formulation of the entire vision problem » (p. 269-272).

Comme nous l'avons vu, l'esquisse 2-1/2D intègre tout un ensemble de données issues des modules inférieurs et, en particulier, les données concernant les valeurs, les variations continues et les discontinuités de la profondeur (stéréopsie) et de l'orientation locale des surfaces. Énormément de travaux expérimentaux et mathématiques ont été consacrés à la façon dont les informations locales issues de la stéréopsie, de la texture, de l'ombrage, du mouvement et des contours coopèrent dans le processus de saisie perceptive d'une forme. Ce sont les problèmes « shape from stereo », « shape from texture », « shape from shading », etc. (cf., par exemple, Brady, 1982 ; Mingolla-Todd, 1986 ; Ikeuchi, 1984). Mais, selon nous, l'ensemble en est subordonné à la résolution d'un problème central. En effet, d'après nos principes épistémologiques, les algorithmes de l'esquisse 2-1/2 D doivent être finalisés par le problème inverse objectif. Or, quelle est la nature de celui-ci à ce niveau ?

Le problème est le suivant. Comment remonter de distributions de discontinuités 2 D à des objets 3 D ? Cela n'est possible que si :

(i) on sait interpréter certaines discontinuités comme des *contours apparents* ;

(ii) on sait remonter des contours apparents d'un objet à cet objet lui-même.

Le premier problème est proprement perceptif. Il suppose que, au moyen des données de profondeur fournies par la stéréopsie ou des données de courbure et d'orientation de surfaces fournies par l'ombrage, etc., on puisse *désambiguïser* les multiples projections 3 D $\rightarrow$ 2 D pouvant aboutir à la même morphologie 2 D (entre deux domaines homogènes contigus séparés par un bord, lequel est devant et lequel est derrière ?, etc.).

Le second problème est en revanche *strictement géométrique et objectif.* Nous l'appellerons *le problème du contour* : comment est-il possible de reconstruire une forme géométrique 3 D à partir de ses contours apparents 2 D ? Ce problème est le problème central du niveau 2-1/2 D. C'est le noyau du problème inverse objectif car c'est sur lui que se concentre le *saut dimensionnel* 2 D $\rightarrow$ 3 D. *Sa résolution mathématique* devrait donc contraindre et finaliser de façon essentielle *l'ensemble des algorithmes 2-1/2 D de la vision computationnelle.* Or cela est très loin d'être le cas actuellement, la plupart des théoriciens de la vision ignorant les éléments de géométrie différentielle et de théorie des singularités exigés. Il faut dire que ceux-ci sont profonds et sophistiqués.

Encore une fois, Marr fait ici partiellement exception. A propos du saut dimensionnel, il remarque : « when one reflects upon it, this is actually quite an amazing fact » (p. 215). Et il pose bien le problème du contour comme problème central. Mais sa méconnaissance de certains récents résultats mathématiques puissants le conduit à faire des hypothèses *ad hoc.*

Soit T un objet (une forme) dans $\mathbb{R}^3$ et C son contour apparent (CA) relativement à une certaine projection Π.

(i) Marr introduit — et cela est correct — une hypothèse de *généricité* : T est en position générale par rapport à Π.

(ii) Il définit ensuite — et cela est également correct — le générateur du contour G, c'est-à-dire (cf. plus bas) le lieu critique de Π (qui est une courbe se projetant sur C).

(iii) Mais, comme il veut pouvoir reconstruire T à partir d'un seul CA et, pour ce faire, appliquer un théorème simple, il introduit l'hypothèse, *ad hoc* et irréaliste, que le générateur G est *planaire* et que la forme T est un « cône généralisé », c'est-à-dire une surface engendrée en déplaçant une section variable le long d'une âme (cf. figure 3). Dans ce cas, en effet, C détermine bien T.

Cette hypothèse *ad hoc* est loin d'être innocente puisqu'elle conduit à décomposer les formes naturelles en cônes généralisés et, par conséquent, à imposer des contraintes non naturelles et non justifiées au niveau 3 D.

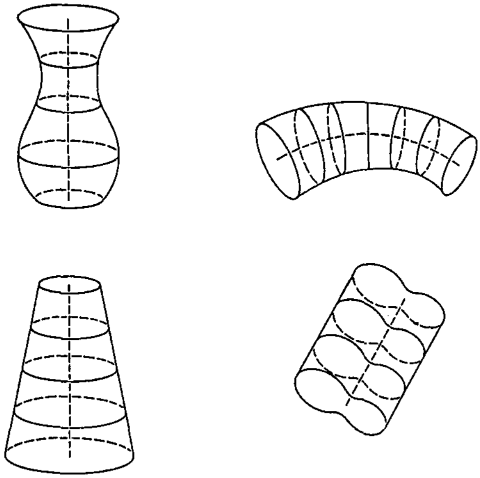


Figure 3. *Le concept de cône généralisé chez Marr (d'après Marr, 1982, p. 224).*

4.2. Le contenu géométrique du problème du contour.

Qu'est-ce que géométriquement le contour apparent (CA) d'un objet (d'une forme, d'une surface) T dans $\mathbb{R}^3$? Supposons pour fixer les idées que la surface T soit un tore (cf. figure 4). Se donner un CA de T consiste :

(i) à choisir dans $\mathbb{R}^3$ un plan de projection Δ ;

(ii) à choisir une direction de projection δ transverse à Δ ;

(iii) à considérer la projection Π de T sur Δ parallèlement à δ.

Le générateur du contour Γ est alors défini comme le *lieu critique* — ou le *lieu singulier* — de l'application $\Pi\,|_T : T \rightarrow \Delta$ restriction de Π à la surface T, c'est-à-dire comme le lieu des points $x \in T$ où la direction de projection δ est *tangente* à T. Le CA (géométrique) C est alors la projection $\Pi(\Gamma)$ de Γ (en situation perceptive réelle, C ne sera en général que partiellement visible) (cf. figure 4).

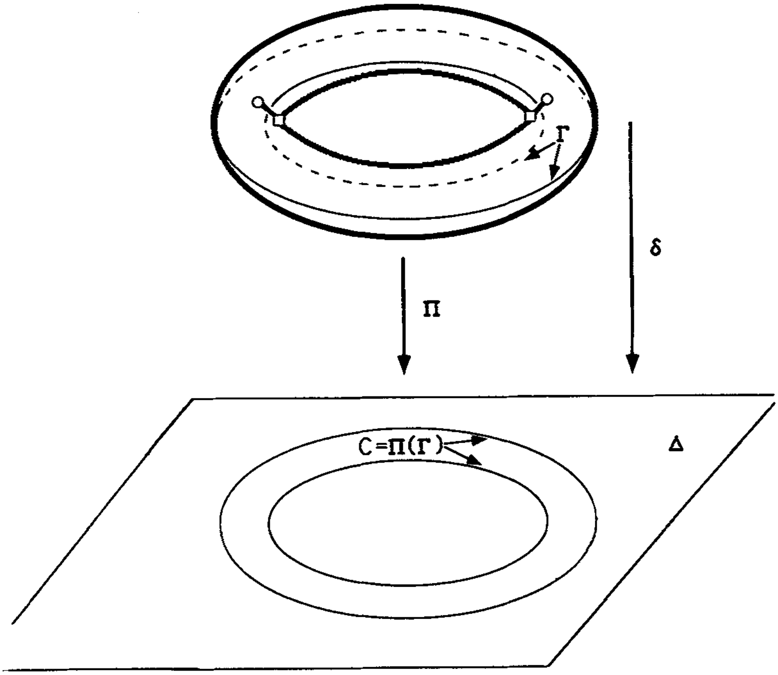


Figure 4. *Le contour apparent d'un tore T est l'ensemble des singularités de la projection Π : T → Δ parallèlement à la direction δ . Γ est le générateur du contour et son image C = Π (Γ) est le contour.*

On notera que le CA n'est pas seulement un ensemble C de courbes dans Δ. C'est un ensemble de courbes qui est un lieu critique, c'est-à-dire *un ensemble de singularités d'application.*

Le *type* de l'information morphologique que constitue un CA est donc loin d'être évident. Il reste incompréhensible en dehors de la théorie mathématique spécifique qui permet de le définir. *Les systèmes visuels sont des reconstructeurs de formes fondés sur des analyseurs de singularités eux-mêmes fondés sur des détecteurs de discontinuités.* Il s'agit là d'un fait fondamental.

Le programme de recherche d'une morphodynamique visuelle est donc bien défini.

(i) *Décrire et classer* les types de singularités locales pouvant (et devant) apparaître génériquement (stablement) dans les CA de surfaces.

(ii) Décrire et classer les singularités plus complexes pouvant (et devant) apparaître stablement dans des *déformations* génériques de CA.

(iii) Montrer qu'il existe *des formes normales algébriques* de ces singularités génériques. Cela est nécessaire pour pouvoir réduire la *géométrie* de celles-ci (qui, *a priori,* comprend une information infinie) *à une infor-*

mation numérique finie (pouvant donc être codée et transmise à un « coût raisonnable », cf. plus haut le problème analogue pour l'esquisse 2 D).

(iv) Reconstruire qualitativement *la géométrie différentielle* des surfaces à partir de la *famille* de leurs CA.

(v) Comprendre comment l'information morphologique 2-1/2 D *peut être encodée dans des champs 2 D de données numériques ponctuelles*. Cela est nécessaire à son calcul par des champs de processeurs ponctuels neuronalement implémentés.

(vi) Comprendre, enfin, comment les corrélats objectifs de cette information peuvent être encodés dans le signal lumineux. Une telle « optique morphologique » est nécessaire à la thèse réaliste, selon laquelle (a) il existe bien de tels corrélats objectifs et (b) l'esquisse 2-1/2 D explicite « ce sur quoi les photons véhiculent de l'information ».

Or, il se trouve qu'une partie considérable de ce programme de recherche est d'ores et déjà réalisée (pour des indications, cf. par exemple, Petitot, 1982, 1986 b et, surtout, leurs bibliographies). Il paraît donc légitime, souhaitable et urgent d'intégrer tous ces résultats fondamentaux à la théorie de la vision computationnelle.

4.3. Le théorème de Whitney-Thom.

La surface T est une variété différentiable de dimension 2 plongée dans $\mathbb{R}^3$. On peut évidemment la décrire par ses équations. Mais une telle description est *extrinsèque*. Si l'on souhaite une description intrinsèque de sa géométrie, au niveau de structure différentiable, alors, comme le faisait déjà Gauss au début du siècle dernier, on doit introduire des *coordonnées locales*. En effet, si T est une surface *régulière* (sans singularités), elle est, en chaque point x, *localement* identifiable, au niveau de structure différentiable, à un plan. Une telle identification est réalisée par la donnée de coordonnées locales (x_1, x_2) en chaque point. Ces systèmes (dits cartes locales) se recollent entre eux à travers des changements différentiables de coordonnées locales[2]. Si (x_1, x_2) (resp. (y_1, y_2)) est un système de coordonnées locales au voisinage de x (resp. $\Pi(x)$)[3] l'application $f = \Pi|_T : T \rightarrow \Delta$ est localement décrite par un système d'équations $(y_1 = f_1(x_1, x_2); y_2 = f_2(x_1, x_2))$ où f_1 et f_2 sont des fonctions différentiables de deux variables réelles à valeurs réelles. $\Pi|_T$ est donc un cas particulier d'application différentiable $f: M \rightarrow N$ entre deux surfaces différentiables M et N ($M = T$ et $N = \Delta$).

Pour décrire *qualitativement* la géométrie locale de f, l'idée fondamentale est de généraliser le concept classique de série de Taylor, c'est-à-dire d'approximations successives de f par des applications *polynômiales*

2. Il est évidemment impossible de rappeler ici serait-ce des rudiments de géométrie différentielle.

3. Comme Δ est un plan, (y_1, y_2) est aussi un système de coordonnées global.

(donc algébriques) de degré de plus en plus grand. Cette méthode sera adéquate dans les cas — dits de *détermination finie* — où l'on saura démontrer que le développement de Taylor $T^k(f)$ de f à un ordre fini k suffit à caractériser qualitativement la géométrie locale de f (autrement dit que l'adjonction de termes d'ordre supérieur à k ne modifie $T^l(f)$ pour $l > k$ que quantitativement et non pas qualitativement, ou encore que l'écart entre $T^k(f)$ et $T^l(f)$ est résorbable par un changement approprié de coordonnées locales).

Le développement de Taylor à l'ordre 1 correspond à ce que l'on appelle *l'application linéaire tangente* D_xf de f en x. Elle représente la façon dont f agit infinitésimalement sur les vecteurs tangents à M en x. Ces vecteurs constituent un espace vectoriel T_xM (de dimension égale à celle de M) dit *espace tangent* à M en x. D_xf est une application linéaire $D_xf : T_xM \longrightarrow T_{f(x)}N$ qui est la meilleure approximation linéaire de f en x (comme la tangente à une courbe en un point est sa meilleure approximation linéaire en ce point). Relativement aux bases de T_xM et $T_{f(x)}N$ associées au choix de coordonnées locales (x_1, x_2) et (y_1, y_2), la matrice de D_xf est donnée par la matrice des dérivées partielles de f_1 et f_2, dite *matrice jacobienne* :

$$[D_xf] = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} \end{bmatrix}$$

La considération du type de D_xf permet déjà d'obtenir de précieux renseignements sur la géométrie locale de f en x. Soit $\mathcal{D}$ l'espace vectoriel des matrices 2×2 $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ (dim $\mathcal{D} = 4$). Dans $\mathcal{D}$, il existe une stratification naturelle *par le rang*, c'est-à-dire une décomposition Σ de $\mathcal{D}$ en sous-variétés de dimensions décroissantes se recollant entre elles avec de bonnes propriétés d'incidence.

(i) La première strate est la strate Σ^0 des matrices régulières (c'est-à-dire des matrices de rang 2, des matrices de déterminant $|D| = ad\text{-}bc$ non nul, bref des matrices inversibles). Elle est topologiquement ouverte (donc de dimension 4) et dense dans $\mathcal{D}$. En effet, une matrice $D \in \mathcal{D}$ est en général de déterminant $|D| \neq 0$ et cette propriété est stable par petite perturbation des coefficients a, b, c, d.

(ii) La seconde strate Σ^1 est celle des matrices D de rang 1, c'est-à-dire des matrices (non identiquement nulles) de déterminant $|D| = 0$. C'est l'*hypersurface* de $\mathcal{D}$ (moins l'origine) d'équation $ad\text{-}bc = 0$. Elle est donc de dimension 3, ou encore — si on appelle *codimension* d'une sous-variété W plongée dans une variété ambiante V la différence de dimensions $\dim V - \dim W$ — *de codimension 1.*

(iii) La troisième strate Σ^2 se réduit à l'origine. Elle ne contient que la matrice nulle de rang 0 D = 0.

On remarquera que la frontière de Σ^0 est $\Sigma^1 \cup \Sigma^2$ et que la frontière de Σ^1 est Σ^2.

L'application linéaire tangente $D_x f$ détermine la structure locale de f en x au sens suivant. Notons $\Sigma^0(f)$, $\Sigma^1(f)$ et $\Sigma^2(f)$ les sous-ensembles de M constitués des $x \in M$ tels que $D_x(f) \in \Sigma^0, \Sigma^1, \Sigma^2$.

(i) Si $x \in \Sigma^0(f)$ (si $D_x f \in \Sigma^0$, c'est-à-dire si $D_x f$ est inversible), l'application f est *localement inversible.* C'est un difféomorphisme local de (M, x) sur (N, f (x)) et sa géométrie est donc *qualitativement triviale.*

(ii) Si $x \in \Sigma^1(f)$ (c'est-à-dire si $D_x f \in \Sigma^1$), alors il existe une direction δ de $T_x M$ qui se trouve *annulée* par $D_x f$. Autrement dit, le noyau $\mathrm{Ker}(D_x f)$ de $D_x f$ n'est pas trivial (il n'est pas réduit à 0).

De façon générale, on dit que x est un *point critique* de f si $D_x(f)$ n'est pas de rang maximal. On dit alors que f (x) est une *valeur critique* de f. L'ensemble $\Sigma(f)$ des points critiques de f est donc donné par $\Sigma(f) = \Sigma^1(f) \cup \Sigma^2(f)$. Nous allons voir que, sous certaines conditions, la géométrie locale reste déterminée à un ordre fini (et même à un ordre très bas). Mais notons d'abord que dans le cas particulier $f = \Pi|_T : T \to \Delta$, dire que $x \in \Sigma^1(f)$ revient à dire que la direction de projection δ appartient à $T_x M$ et qu'elle est donc tangente à T. Autrement dit, *$\Sigma^1(f)$ n'est dans ce cas rien d'autre que le générateur du contour* Γ.

Il faut se convaincre que la complexité d'une application différentiable $f: M \to N$ peut être prodigieuse. Par exemple, on peut montrer (théorème dû à Borel) que si F est un fermé de M (et un tel F peut être d'une complexité infinie, fractale par exemple), il existe une application différentiable $f: M \to \mathbb{R}$ telle que $F = f^{-1}(0)$. Il est donc *impossible* de classer les types qualitatifs des f. Pour accéder malgré cela à une possibilité de classification, on applique la stratégie *de la stabilité structurelle.* Soit $\mathfrak{F}(M, N)$ l'espace fonctionnel des f. Sur $\mathfrak{F}$ il existe une topologie $\mathfrak{T}$ (dite topologie de Whitney) naturellement adaptée au niveau de structure différentiable (intuitivement, c'est la topologie de la convergence uniforme des fonctions et de toutes leurs dérivées partielles sur les compacts de M, avec en plus une contrainte d'identité « à l'infini », c'est-à-dire sur le filtre des complémentaires des compacts). D'autre part, sur M et sur N il existe les changements de coordonnées *globaux* que sont les difféomorphismes. Il est évident que deux applications f, $g \in \mathfrak{F}$ sont *qualitativement (différentiablement) équivalentes,* si elles sont conjuguées par de tels difféomorphismes, autrement dit s'il existe $\varphi \in \mathrm{Diff}(M)$ et $\psi \in \mathrm{Diff}(N)$ tels que $g = \psi \circ f \circ \varphi^{-1}$. Autrement dit, le groupe $G = \mathrm{Diff}M \times \mathrm{Diff}N$ opère sur F *comme un groupe de relativité* et les orbites de son action sont *les types qualitatifs* des éléments $f \in \mathfrak{F}$.

On dit alors que f est *structurellement stable* s'il existe un voisinage de f (pour la topologie $\mathfrak{T}$) dont tous les éléments g sont qualitativement équivalents à f (autrement dit, si le type qualitatif de f résiste aux petites perturbations). La stratégie de la stabilité structurelle, introduite par Whitney en 1955 et considérablement développée par Thom, Arnold et d'autres, consiste :

(i) à analyser d'abord la géométrie locale des applications *structurellement stables* ;

(ii) à analyser ensuite celle des applications instables, mais en introduisant *progressivement* des degrés de plus en plus grands d'instabilité (cela suppose évidemment que l'on ait explicité les causes possibles d'instabilité structurelle).

La structure locale des applications structurellement stables entre surfaces est entièrement connue. Elle est résumée dans le théorème suivant.

Théorème de Whitney-Thom :
1. Les applications structurellement stables $f : M \rightarrow N$ sont *génériques* dans $\mathfrak{T}$: toute application $g \in \mathfrak{T}$ est approximable aussi près que l'on veut par une application f structurellement stable.
2. Si f est structurellement stable, sa géométrie locale est équivalente à celle de l'un des trois modèles locaux algébriques suivants :
(a) $y_1 = x_1$, $y_2 = x_2$: *point régulier* (f est un difféomorphisme local),
(b) $y_1 = x_1^2$, $y_2 = x_2$: *point pli,*
(c) $y_1 = x_1^3 + x_1x_2$, $y_2 = x_2$: *point fronce.*

Ce théorème montre que, sous l'hypothèse de stabilité structurelle, la géométrie locale de f est *déterminée à l'ordre 2* (c'est-à-dire par T^2 (f)). Ce résultat est fondamental pour ce qui nous occupe ici puisqu'il ne s'agit de rien de moins que d'un théorème *de réduction d'une information morphologique à une information algébrique finie.* Il existe des *modèles locaux algébriques universels* pour la géométrie locale. En hommage à Marr, nous les appellerons des modèles 2-1/2 D.

La structure géométrique d'un pli est évidente (cf. figure 5).

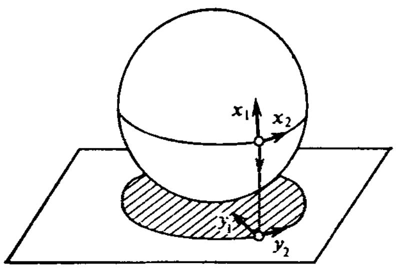


Figure 5. *La structure d'un point pli (d'après Arnold, 1986, p. 16).*

Celle d'une fronce est un peu plus complexe. Mais elle est facile à dériver de son modèle algébrique. Pour la visualiser, considérons le graphe de $y_1 = x_1^3 + x_1 x_2$ dans l'espace $\mathbb{R}^3$ de coordonnées (x_1, $x_2 = y_2$, y_1). f est la projection de ce graphe sur le plan (y_1, y_2) (cf. figure 6).

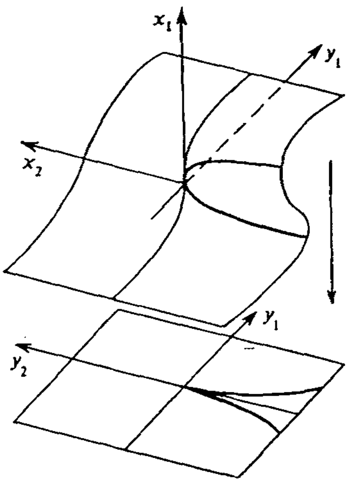


Figure 6. *La structure d'un point fronce (d'après Arnold, 1986, p. 16).*

La matrice jacobienne de f en x est $[D_x f] = \begin{bmatrix} 3x_1^2 + x_2 & x_1 \\ 0 & 1 \end{bmatrix}$. Donc $x \in \Sigma^1(f)$ si $D_x f = 3x_1^2 + x_2 = 0$ (équation d'une parabole Γ dans le plan (x_1, x_2) : le générateur du contour). $C = f(\Gamma)$ est donc la parabole semi-cubique du plan (y_1, y_2) d'équations paramétriques $y_1 = -2x_1^3$, $y_2 = -3x_1^2$ (car $x_2 = -3x_1^2$ sur Γ). Elle possède à l'origine un point de rebroussement appelé un point *cusp.* On voit qu'aux points plis de Γ, le noyau $\mathrm{KerD}_x f$ (la direction de projection) est *transverse* à Γ. En revanche au point fronce $x = 0$, $\mathrm{KerD}_x f$ est au contraire *tangente* à Γ.

Autrement dit, en un point pli x, f est singulière, $x \in \Sigma^1(f)$, mais la restriction $f \mid \Sigma^1(f)$ de f à $\Sigma^1(f) = \Gamma$ est, elle, *régulière.* C'est pourquoi il est naturel de noter $\Sigma^{1,0}(f)$ l'ensemble des points plis de f. En revanche en un point fronce x de f, $x \in \Sigma^1(f)$ mais la restriction $f \mid \Sigma^1(f)$ est *singulière.* On note donc $\Sigma^{1,1}(f)$ l'ensemble des points fronces de f.

On remarquera d'autre part que, dans ces modèles locaux, $\Sigma^2(f) = \varnothing$. Nous allons voir qu'il s'agit là d'une *nécessité* sous l'hypothèse de stabilité structurelle.

On voit ainsi apparaître l'idée fondamentale de *types* de points singuliers *finiment* descriptibles et d'une *hiérarchie* de ces types. Les lieux singuliers des applications stables sont des *stratifications,* des « empilements » de lieux singuliers de restrictions à des lieux singuliers : Sing(f), Sing(f | Sing(f)), Sing(f | Sing(f | Sing(f))), etc., la stabilité structurelle *bornant* ce type d'itération par les *dimensions* de M et de N.

4.4. La théorie des jets et le processing ponctuel des géométries locales.

Bien que les termes *homogènes* du développement de Taylor d'une application ne possèdent pas de *signification géométrique intrinsèque* (indépendante du choix, conventionnel, des coordonnées locales), on peut montrer que les développements $T^k(f)$ jusqu'à un rang donné possèdent, eux, une signification géométrique intrinsèque. Cela a permis à Charles Ehresmann d'élaborer *la théorie des jets* qui fournit une réponse au problème fondamental *de l'encodage d'une géométrie locale par un champ de données numériques ponctuelles* (problème 4.2.(v)).

L'idée généralise cette opération bien connue qui consiste, étant donné une courbe $y = f(x)$, à considérer le *champ* de ses tangentes T_x en chaque point et à reconstruire la courbe comme *enveloppe* de ses tangentes.

Soit $f : M \rightarrow N$. En chaque point $x \in M$, le développement de Taylor au premier ordre est constitué de trois groupes de données *ponctuelles* :

(i) $x \in M$: deux coordonnées : x_1, x_2;

(ii) $y = f(x) \in N$: deux coordonnées : $y_1 = f_1(x_1, x_2), y_2 = f_2(x_1, x_2)$;

(iii) $D_x f \in \mathcal{D}$: quatre coordonnées : $a = \frac{\partial f_1}{\partial x_1}(x_1, x_2)$, $b = \frac{\partial f_1}{\partial x_2}(x_1, x_2)$, $c = \frac{\partial f_2}{\partial x_1}(x_1, x_2)$, $d = \frac{\partial f_2}{\partial x_2}(x_1, x_2)$.

Ces *huit* données numériques constituent le *1-jet* de f en x, 1-jet noté $j^1f(x)$. $j^1f(x)$ habite naturellement dans un espace à huit dimensions qui, localement, est le produit direct $M \times N \times \mathcal{D}$. Lorsque l'on globalise, ces produits directs se recollent en un fibré vectoriel de base $M \times N$, appelé espace (ou fibré) des 1-jets des applications différentiables $f : M \rightarrow N$ et noté $J^1(M, N)$. Si $f \in \mathcal{F}$, on lui associe son 1-jet j^1f qui est l'application de M dans $J^1(M, N)$ définie par le *champ* des 1-jets $j^1f(x)$:

$$j^1f : M \rightarrow J^1(M, N)$$
$$x \rightarrow j^1f(x).$$

Mais nous avons vu que, dans les fibres $\mathcal{D}$ de $J^1(M, N)$ — et donc dans $J^1(M, N)$ — il existe une stratification naturelle $\Sigma = (\Sigma^0, \Sigma^1, \Sigma^2)$. Il est clair que l'on a $\Sigma^0(f) = (j^1f)^{-1}(\Sigma^0)$, $\Sigma^1(f) = (j^1f)^{-1}(\Sigma^1)$, $\Sigma^2(f) = (j^1f)^{-1}(\Sigma^2)$. *La stratification $\Sigma(f)$ de la source M opérée par f au moyen du rang de l'application linéaire tangente D_xf n'est donc rien d'autre que l'image réciproque de la stratification universelle Σ par le 1-jet $j^1(f)$ de f.*

Le théorème de Whitney montre que les 2-jets $j^2(f)$ *suffisent* pour reconstruire qualitativement la géométrie locale de *toutes* les applications *structurellement stables*. En généralisant aux 2-jets (cela est trop technique pour être exposé ici) les constructions précédentes, on en arrive à la conclusion que, génériquement, la géométrie locale de f est descriptible *à partir de l'image inverse, par les 2-jets $j^2(f)$, de stratifications universelles* (indépendantes de f) des espaces de jets. Autrement dit, *la géométrie locale est calculable au moyen des champs de données numériques que sont les jets d'ordre ≤ 2. La théorie des jets est donc bien fondamentale pour la vision computationnelle puisqu'elle explique comment des champs de processeurs ponctuels possédant une bonne rétinotopie peuvent calculer de la géométrie, c'est-à-dire traiter de l'information morphologique.*

Il est d'ailleurs étrange que les spécialistes de la vision aient été aussi peu attentifs jusqu'ici à l'une des idées les plus profondes et les plus fécondes de toutes les sciences, à savoir celle de la dialectique du local et du global. L'idée est que les contraintes (lois de la nature, etc.) se décrivent au niveau local par des équations sur des jets et que, par intégration, elles admettent pour solutions des entités globales. Par exemple, une équation différentielle ordinaire consiste à se donner en chaque point d'un espace M (espace de configurations ou espace de phases d'un système mécanique, etc.) un vecteur tangent $X(x) \in T_xM$ et à chercher les trajectoires intégrales d'un tel champ. De même, un feuilletage (un système de Pfaff) consiste à se donner en chaque point $x \in M$ un sous-espace vectoriel $P(x)$ de T_xM et à chercher les variétés intégrales. De même encore, une équation aux dérivées partielles est une équation dans un espace de jets convenable. Par exemple, une équation de diffusion (à une dimension) comme $\frac{\partial f}{\partial t} = \frac{\partial^2 f}{\partial x^2}$ s'exprime par l'équation $a = m$ dans l'espace de jets $J^2(\mathbb{R}, \mathbb{R})$ de coordonnées $(x\,;\, y = f(x)\,;\, a = \frac{\partial f}{\partial t}, b = \frac{\partial f}{\partial x}\,;\, k = \frac{\partial^2 f}{\partial t^2}, l = \frac{\partial^2 f}{\partial x \partial t}, m = \frac{\partial^2 f}{\partial x^2})$.

L'importance des théories évoquées plus haut est d'avoir montré que l'analyse morphologique peut se ramener à de tels champs de données numériques, champs dont les formes sont en quelque sorte des solutions intégrales. Elle donne une nouvelle dimension à l'intuition initiale des Gestalt-théoriciens.

Certes, de très nombreux modèles de vision computationnelle consistent à reconstruire des formes à partir de champs de données locales. Par exemple, on cherchera à associer à chaque point de l'image une orientation locale de surface (ce qui est équivalent à une direction normale : on cherche à reconstruire l'application de Gauss de la surface, cf. p. 172) obtenue à partir des informations sur la stéréopsie, l'ombrage, le gradient, la texture, etc. (cf., par exemple, les articles, déjà cités, Brady, 1982 ; Ikeuchi, 1984 ; Mingolla-Todd, 1986). Mais de tels modèles restent très en deçà des ressources actuelles de la géométrie différentielle.

Revenons à la théorie des jets. D'après un théorème fondamental de Thom, dit *théorème de transversalité,* la stabilité structurelle s'exprime par des propriétés de transversalité des applications jets $j^k(f)$ sur les stratifications universelles des espaces de jets $J^k(M, N)$. Cela implique que, lorsque l'on prend les images réciproques de ces stratifications, leur structure géométrique soit préservée autant qu'il est possible. Cela implique à son tour *une borne drastique à la complexité* des singularités génériques. Considérons, par exemple, la strate Σ^2 de $J^1(M, N)$. Elle est de codimension 4. Comme dim $M = 2$, l'image $j^1f(M)$ de M dans $J^1(M, N)$ par j^1f est (au plus) de dimension 2. La stabilité implique la transversalité, et la transversalité implique à son tour, pour une simple raison de dimension ($2 < 4$), que $j^1f(M)$ *évite* Σ^2. C'est pourquoi, si f est structurellement stable, on a nécessairement $\Sigma^2(f) = \emptyset$. On montre aussi que, sous la même hypothèse, $\Sigma^1(f) = \Gamma = (j^1f)^{-1}(\Sigma^1)$ est une courbe régulière de M.

4.5. La solution du problème inverse objectif.

Sur le plan global, on peut montrer que si $f : M \rightarrow N$ est structurellement stable et si M est compacte, alors le générateur Γ est une courbe régulière et le CA $C = f(\Gamma)$ ne peut présenter comme singularités que des cusps isolés et des croisements normaux. Il existe alors des relations précises entre le nombre de cusps et la structure globale de M et de N (par exemple, leur caractéristique d'Euler-Poincaré).

D'autre part, on peut aussi classifier et mettre sous forme normale les singularités qui apparaissent stablement lors de *déformations* génériques de CA. La plus importante est *la queue d'aronde* où un point pli devient instable et se stabilise en engendrant deux cusps (cf. figure 7 pour l'exemple du tore). Le nombre de types qualitatifs de CA que peut présenter une forme T, ainsi que leurs relations d'incidence, fournit un renseignement fondamental sur *la complexité morphologique* de T.

L'ensemble de ces résultats (que nous n'avons fait qu'esquisser de façon très élémentaire) permet de résoudre le problème inverse *objectif.* Celui-ci impose à la résolution computationnelle du problème inverse *cognitif* les contraintes suivantes.

(i) Il doit exister des dispositifs de détection et de représentation (d'explicitation) des lignes de discontinuités (projections de points plis), des points d'arrêt de telles lignes (points cusps dont en général une des branches de points plis sera occultée si la surface est opaque), et des croisements de telles lignes (en général, une partie de la ligne pli arrière sera occultée et le croisement sera donc en forme de T). *L'algorithme de Marr correspond au premier de ces dispositifs.* Il faut donc en généraliser le principe aux deux autres cas.

(ii) Il faut pouvoir associer à ces primitives morphologiques 2 D (plis,

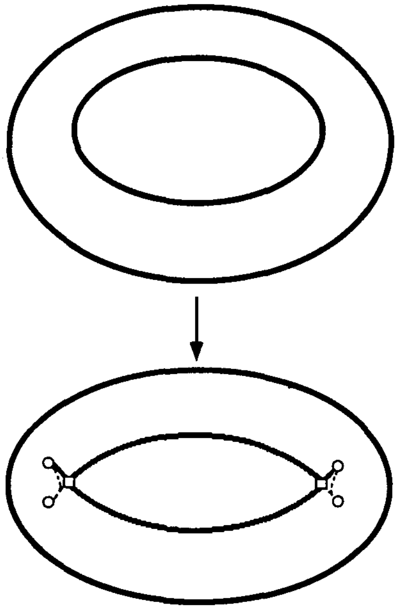


Figure 7. *Singularité de transition « queue d'aronde » pouvant apparaître stablement dans une déformation générique de contour apparent. Un point pli dégénère et engendre deux cusps (petits cercles) et un croisement (petit carré).*

cusps, croisements) les modèles locaux 2-1/2 D correspondants. Nous avons vu comment cela était possible.

5. Les travaux de Jan Koenderink

Jan Koenderink est l'un des rares spécialistes de la vision qui ait compris tout le bénéfice que la vision computationnelle peut tirer de l'usage des théories mathématiques évoquées plus haut pour résoudre le problème du « jump between logical levels (i.e. from the physical to the semantic domain) » (Koenderink, 1987, p. 367). Dans une série d'articles remarquables, parus pour la plupart dans *Biological Cybernetics,* il les a appliquées à tout un ensemble de problèmes[4].

5.1. Le point de vue épistémologique.

Adoptant une perspective « écologiste », Koenderink considérait dès 1976 :

---

4. Je remercie S. Thorpe de m'avoir récemment signalé ces travaux, apparemment inconnus jusqu'ici dans les milieux mathématiques pourtant directement concernés.

« that prior to the study of visual shape perception or visual egocentric localization an inventory of the invariants of the optical input under voluntary deplacements of the observer, has to be made. Such invariants pertain to objective geometrical properties of the environment. »

Mais il ajoutait aussitôt : « However, a comprehensive quantitative theory of the geometrical invariants of optical stimulation does not exist » (Koenderink, 1976, p. 51). Il introduisait alors l'idée directrice que l'information pertinente est concentrée dans les singularités des projections visuelles et que c'est donc la théorie des singularités qui permet de fonder mathématiquement un « écologisme » scientifiquement légitime.

5.2. La résolution du problème du contour.

Une des premières réussites de J. Koenderink est d'avoir explicitement utilisé dans sa théorie la résolution du problème du contour. Soit encore une fois notre forme T plongée dans $\mathbb{R}^3$. Ce que nous avons dit reste essentiellement valable si, au lieu de considérer une projection parallèle Π, nous considérons la projection Π_p de T sur Δ à partir d'un point de vue p extérieur à T. Soit $F_p : T \rightarrow \mathbb{R}$ la fonction *distance* d(p, x) de p à $x \in T$. Les points critiques de F_p sont ceux pour lesquels d(p, x) est stationnaire, c'est-à-dire ceux pour lesquels la direction px est orthogonale à T. Génériquement, ce sont des minima, des maxima et des cols. Soit w(x) le vecteur unitaire de la direction px d'origine x. Soit $v(x) \in T_xT$ sa composante tangentielle : v(x) définit un champ de vecteurs tangents sur T dont les trajectoires sont les lignes de pente de la distance. Les points critiques de F_p sont les points critiques ($v(x) = 0$) de ce champ.

En couplant ce champ au CA, on obtient ce que Koenderink appelle *un aspect*. « The aspect is a Gestalt-like feature of the visual input. » Il détermine le pattern d'excitation corticale. Il est constitué des CA orientés, des occultations de bords, des cusps, des croisements de lignes pli, des points où les lignes de pente de F_p touchent un bord occluant, des points critiques de F_p avec leur type (minimum, maximum, col), des séparatrices des directions de lignes de pente, des lignes de pente issues des cusps et des lignes de pente touchant un bord occluant (cf. figure 8).

La considération des aspects et de leurs déformations lors des changements de points de vue ou des déplacements d'objets permet non seulement de reconstruire la topologie de la surface T et sa structure différentiable, mais également de reconstruire partiellement ses propriétés riemaniennes (donc métriques). Cela signifie la chose suivante. On sait (depuis Gauss) que si l'on considère une surface plongée dans $\mathbb{R}^3$ comme une variété riemanienne, sa structure métrique est localement elliptique, hyperbolique ou parabolique. En un point hyperbolique, il

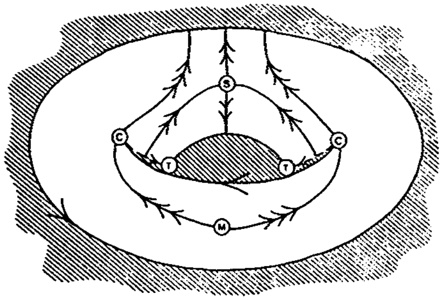


Figure 8. *Le concept d'aspect chez Koenderink.* → : *contour;* →→ : *séparatrice;* →→→ : *chemin passant par un cusp; C : point cusp; M : minimum de la fonction distance; S : point col; T : croisement normal (d'après Koenderink, 1979, p. 214).*

existe deux directions principales. Les trajectoires de ces deux champs de directions (dites lignes asymptotiques) admettent pour enveloppe les lignes de points paraboliques.

Koenderink montre que la famille des CA de T permet de déterminer le *type* de la métrique en chaque point (et donc en particulier les propriétés de convexité de T)[5]. Pour cela, il analyse avec soin les composantes de CA introduites par les déplacements du point de vue p sur T et il en explicite la structure à partir de *l'application de Gauss* de T, c'est-à-dire de l'application $G : T \rightarrow S^2$ qui à $x \in T$ associe le vecteur normal unitaire (externe, on suppose T orientable) n(x) à T en x (S^2 est la sphère unité de $\mathbb{R}^3$)[6]. Les lignes paraboliques de T correspondent aux plis P_G de G. Si x est un point du générateur Γ d'un CA de T, n(x) est normal à la direction δ (on suppose p à l'infini pour simplifier). L'image de Γ par G est donc incluse *dans un grand cercle* Γ_G de S^2. Lorsque l'on bouge δ, Γ_G se déplace et, en étudiant les transformations de sa position par rapport à P_G, on peut reconstruire qualitativement la géométrie proto-riemanienne de T. Cela résout le problème du contour.

A partir de cet acquis, Koenderink développe alors l'argument suivant, qui nous paraît fondamental. La déformation, par transformation des positions relatives de p et de T, des CA de T — qui ont une réalité perceptive bien établie — permet de reconstruire la géométrie *intrinsèque* (objective) de T. Elle permet donc de *prédire — d'anticiper sur —* ces déformations. Celles-ci, parce que prédictibles, peuvent être interprétées

5. Ce niveau plus fort que le différentiable et plus faible que le métrique ne semble pas avoir été très étudié mathématiquement. Il est en quelque sorte encore qualitatif et déjà proto-métrique, bien que sans notion de distance et de géodésiques.

6. L'application de Gauss est évidemment couramment utilisée dans les modèles de vision computationnelle puisqu'elle représente le champ des orientations locales d'une surface (cf. plus haut). Mais en général on n'utilise pas sa relation avec les CA.

comme d'origine *proprioceptive,* ce qui explique *l'invariance* objective de l'objet malgré la grande variation subjective de l'*input* visuel. « Our geometrical theory enables us to understand the structure of the observer's internal models of external bodies » (Koenderink, 1976, p. 59).

Dans un travail plus récent, Koenderink aborde la généralisation de la théorie de Marr. Sa première idée est d'abord, étant donné un pattern d'intensité 2D $I(x, y)$, d'en représenter la morphogenèse en l'incluant dans une déformation $F = I_t$ conduisant de $I = I_1$ à un pattern I_0 trivial. La déformation inverse $I_0 \rightarrow I_1$ est donc un chemin de genèse de I. Koenderink choisit alors pour déformation une solution d'une *équation de diffusion* (type équation de la chaleur) $\frac{\partial F}{\partial t} = \Delta F$. La raison en est qu'une telle solution équivaut à lisser I par convolution avec une gaussienne (dépendant de t). On reprend donc l'algorithme de Marr mais en lui donnant un nouvel éclairage : « Gaussian blurring is the only sensible way to embed a primal image into a one-parameter family » (Koenderink, 1979, p. 365). L'auteur étudie ensuite la structure locale de F en termes

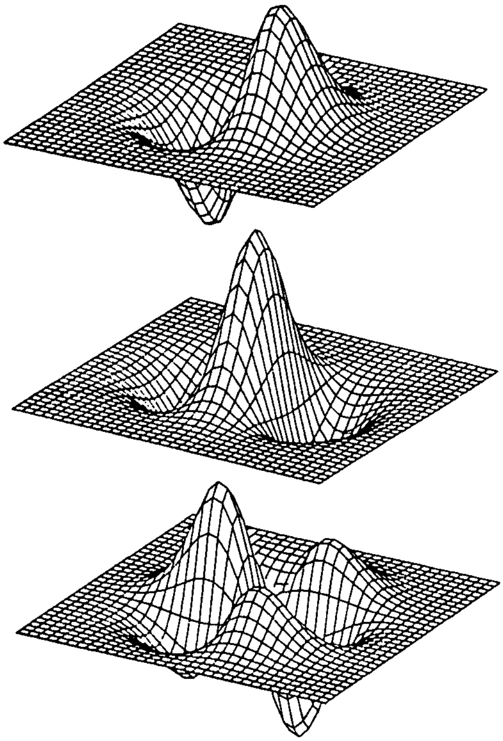


Figure 9. *Profils récepteurs permettant selon Koenderink d'effectuer des calculs de jets (d'après Koenderink, 1987, p. 371).*

de jets. Pour cela, il reprend l'idée directrice de Marr d'une analyse multirésolution (locale et multiéchelle) *par des convolutions avec des profils bien choisis* f_n *(x, t) de champs récepteurs.* Ces f_n (x, t) sont, comme chez Marr, des dérivées partielles (d'ordre n) de gaussiennes (cf. figure 9).

Koenderink explique alors l'importance du concept de jet pour la vision computationnelle. Dans F, l'information morphologique est continuement distribuée. Elle est *multilocale.* Dans les jets j^kF elle devient au contraire *ponctuelle* et traitable par des processeurs ponctuels.

> « Routines accessing a single location may aptly be called point processors, those accessing multiple location *array processors.* The difference is crucial in the sense that point processors need no geometrical expertise at all, whereas array processors do » (Koenderink, 1987, p. 370).

Les profils de champs récepteurs fournissent une implémentation des détecteurs de données différentielles. A partir d'eux, on peut construire *des processeurs de jets* qui sont des *détecteurs de traits morphologiquement significatifs.* « The order of the jets in the representation determines the " features " (the geometrical properties) that can be computed by a point processor » *(ibid.,* p. 370). Certaines hypercolonnes corticales seraient des champs de tels détecteurs :

> « the modules (like " cortical columns " in the physiological domain or " records " of raw data in the syntactic domain) of the sensorium are local approximation (N^{th} order jets) of the retinal illuminance that can be adressed as *a single datum* by the point processors. »

Les jets sont des K-uples de *nombres* possédant « a semantic content in terms of certain visual routines ».

> « That looking at a retinal illuminance distribution through a receptive field profile (or even through several layers of them !) is equivalent to looking at certain partial derivative of a blurred pattern is a new insight that immediately leads to useful interpretation in terms of differential geometry » *(ibid.,* p. 374).

6. *Éléments d'une théorie morphodynamique intégrée*

Dans ce qui précède, nous nous sommes focalisés sur *un* point qui nous paraissait névralgique. Nous aimerions maintenant brièvement faire le lien avec d'autres recherches qui sont susceptibles de conduire à une théorie intégrée.

6.1. La nature d'une optique morphologique.

Une thèse réaliste sur l'information morphologique n'est évidemment tenable que si l'on peut montrer que les CA (c'est-à-dire des singularités d'applications différentiables) peuvent effectivement être encodés dans le signal lumineux, c'est-à-dire dans des solutions des équations de Maxwell. Le problème est loin d'être trivial mathématiquement. Il est résolu (depuis peu de temps seulement) en ce qui concerne un cas plus simple que celui des CA, à savoir celui des caustiques. Les caustiques sont les enveloppes de rayons lumineux qui apparaissent lorsque des faisceaux lumineux sont soumis à des contraintes dioptriques de convergence (de focalisation). Ce sont des singularités (des lieux critiques d'applications) faciles à décrire géométriquement. Elles dominent les images optiques et sont phénoménologiquement structurantes. Comment sont-elles encodées dans le signal optique ? Comment leur information, typiquement de nature morphologique, peut-elle être, comme dirait Marr, véhiculée par les photons ? Une réponse peut être donnée dans le cadre de l'approximation géométrique de l'équation des ondes. On montre (c'est très technique : théorie des intégrales oscillantes) qu'à chaque singularité générique de caustique (pli, cusp, ombilic, etc.) est associée une intégrale oscillante typique qui est une structure ondulatoire fine construite sur l'infrastructure géométrique de la singularité (pour des précisions, cf. Arnold *et al.*, 1986; pour une introduction, cf. Petitot, 1986b, 1989g ainsi que leurs bibliographies). Il s'agit là d'un exemple, en tous points remarquable et entièrement mathématisé, d'émergence qualitative de formes perceptivement significatives à partir de la physique fondamentale.

Contrairement aux idées reçues, il existe donc bien une optique morphologique et il est par conséquent légitime de faire l'hypothèse que l'information géométrique est non seulement géométriquement objective, mais également *physiquement* objective.

6.2. Le niveau 2-1/2 D, le niveau 3 D et la Structure conceptuelle.

Le niveau 3 D objective le niveau morphologique 2-1/2 D. Ses algorithmes commencent à être bien compris (on a, par exemple, étudié en détail le nombre minimal de projections planes et de CA dont on a besoin pour reconstruire de façon non ambiguë une forme tridimensionnelle : cf. entre autres Hoffman-Bennett, 1986). Mais beaucoup de ses constituants s'enracinent dans le niveau 2-1/2 D. Par exemple, la décomposition (relativement canonique) d'un objet en parties s'opère essentiellement sur des bases morphologiques (les lignes de décomposition sont des lignes de forte courbure, etc.). Contrairement à ce que l'on croit habituellement, elle n'est pas descendante (reconnaissance d'occurrences de modèles de parties prototypiques stockées dans une mémoire à long terme) mais ascendante. Ainsi que l'affirment Hoffman et Richards :

> « the visual system decomposes shapes into parts [...] using a rule defining part boundaries rather than part shapes, [...] the rule exploits a uniformity of nature — transversality, and [...] parts with their description and spatial relations provide a first index into a memory of shapes » (Hoffman-Richards, 1984, p. 65).

Ce n'est que postérieurement à cette décomposition morphologique qu'interviennent les segmentations en constituants géométriquement typiques reposant sur un vocabulaire fini de primitives (cf., par exemple, Biederman, 1987) et que les niveaux supérieurs de représentation et d'organisation hiérarchisée de l'information visuelle deviennent de format similaire à, et compatibles avec, ceux de l'information non visuelle.

De même les phénomènes de *catégorisation* proviennent essentiellement de la forte non-linéarité du contrôle des formes par des paramètres de déformation (cf. Petitot, 1989e). A l'intérieur des catégories, les formes sont stables par rapport à la variation du contrôle. Les frontières des catégories sont, au contraire, des lieux critiques à la traversée desquels les formes deviennent structurellement instables par rapport au contrôle et, donc, changent de type qualitatif. De façon générale, ainsi qu'y insiste R. Jackendoff, énormément de traits qui servent à catégoriser les objets *sont morphologiques et non pas sémantiques*. Le niveau 3D est celui où le langage se branche sur la vision à travers la structure conceptuelle et le langage en hérite *de fortes composantes morphologiques* (au sens adopté ici, non linguistique, de morphologie).

6.3. Vision et langage.

Dans un certain nombre de travaux (en particulier, Petitot, 1979, 1982, 1985, 1989a, c, f) nous avons développé l'idée maîtresse de Thom selon laquelle les relations actantielles entre les actants spatio-temporels d'une scène visuelle étaient morphodynamiquement — et non pas seulement symboliquement — descriptibles.

Nous avons montré comment cette idée permettait de fonder et de développer mathématiquement ce que l'on appelle *l'hypothèse localiste* en linguistique et d'en déduire une théorie actantielle (casuelle), une théorie de l'aspectualité et une théorie de l'agentialité. Nous avons, enfin, analysé le rapport qu'une telle schématisation morphodynamique entretient avec certains des courants fondamentaux de la linguistique cognitive actuelle (Langacker, Talmy, Jackendoff).

Enfin nous avons montré comment une telle théorie de la syntaxe actantielle permettait de répondre aux objections de principe élevées par J. Fodor et Z. Pylyshyn contre le connexionnisme (cf. Petitot, 1989f, i).

III. — LA MORPHODYNAMIQUE VISUELLE COMME RÉPONSE AUX PROBLÈMES DE L'ÉCOLOGISME ET DE L'INTENTIONALITÉ

Sur le plan *épistémologique,* nous considérons que l'existence d'un niveau de réalité morphologique assurant la médiation entre le physique et le symbolique est d'une grande importance dans la mesure où elle permet de résoudre un certain nombre de problèmes cruciaux qui resteraient autrement aporétiques. Donnons brièvement, pour conclure, quelques indications à propos de deux d'entre eux.

1. L'objectivité écologique

Dans un important article, J. Fodor et Z. Pylyshyn ont ruiné théoriquement les thèses écologistes. Ils partent de l'hypothèse classique : parce que cognitive, la perception doit nécessairement être un processus computationnel symbolique et inférentiel. Ils cherchent alors à invalider la thèse gibsonienne selon laquelle la perception est à même d'extraire de l'environnement des invariants possédant un contenu objectif. Pour cela ils dégagent, avec une acuité remarquable, les inconsistances de la théorie écologique. Selon eux, la principale consiste à fonder toute la théorie sur l'existence d'une information objective (mais non physique) qui serait présente dans le médium lumineux (discontinuités, déformations, formes, textures, réflectances, etc. des surfaces visibles), alors qu'on reste dans l'impossibilité de la définir. Que peut être, en effet, cette énigmatique « information in the light » (Fodor-Pylyshyn, 1981, p. 143) ? Pour les auteurs, en vertu du dualisme physique/symbolique, l'information est soit physique, soit symbolique. Si donc elle n'est pas à proprement parler physique, mais « écologique », elle doit nécessairement être symbolique. Bref, Gibson introduit une objectivité écologique introuvable. Il critique d'un côté la physique physicaliste et de l'autre la psychologie mentaliste. Il ne fournit toutefois pas d'alternative. D'où un cercle vicieux. « What we need, of course, is some criterion for being ecological *other than perceptibility*. This however, Gibson fails to provide » (p. 146). Il faudrait une optique écologique différente de l'optique physique, capable de caractériser ce qui est *phénoménologiquement significatif*. Or, celle-ci demeure, selon les auteurs, inaccessible.

Le syllogisme est au fond le suivant. La seule extraction directe d'invariants ne peut être que celle effectuée par la transduction. Les transducteurs ne peuvent être sensibles qu'aux propriétés physiques du signal lumineux car leur fonctionnement est régi par des lois et les seules lois existantes sont les lois physiques. Il ne saurait donc exister de transducteurs (même compilés, c'est-à-dire opérant modulairement jusqu'à des niveaux post-rétiniens) qui extraient du signal des propriétés écologiques non physiques. Fodor et Pylyshyn dénoncent alors ce qu'ils considèrent être une *subreption* chez Gibson. Pour Gibson, il existe *de l'information contenue dans la lumière*. Mais, selon les auteurs, le concept d'information est *relationnel*. *La lumière contient de l'information sur l'environnement*, et « contenir de l'information sur » signifie « être corrélé avec ». Les propriétés de l'environnement sont donc *inférées* à partir de la structure du signal lumineux sur la base de la connaissance que possède le système perceptif sur ces corrélations. En remplaçant « contenir de l'information sur » par « information contenue dans », Gibson aurait subrepticement *réifié* le concept relationnel d'information. Il l'aurait traité « as a thing, rather than a relation » (p. 167). Une information *ne peut pas* affecter un système perceptuel. Seules des propriétés physiques le peuvent. Elles peuvent alors certes être « informatives sur quelque chose », mais seulement au moyen d'inférences. Car la *corrélation* elle-même qu'est l'information *ne peut pas* être un état d'un récepteur. Le problème est : « how (by what mental process) does the organism get from the detection of an informative property of the medium to the perception of a correlated property of the environment ? » Et la réponse est : par inférences. « X contient de l'information sur Y » est une relation *sémantique* et dépend donc de la façon dont X est mentalement représenté comme une prémisse d'inférences de X vers Y.

On voit que toute cette discussion (poussée beaucoup plus loin par les auteurs) repose sur le double préjugé que la réalité physique ne possède aucune propriété émergente et que ce qui est significatif doit nécessairement s'abstraire en sémantique et être produit par une intentionalité (la façon dont les représentations mentales dénotent). Par conséquent, il ne saurait exister dans l'environnement de structures intrinsèquement significatives encodables dans le signal lumineux.

L'existence d'une information morphologique géométriquement, *phéno*-physiquement et optiquement objective dément ce préjugé et permet de fonder *un écologisme morphodynamique*. Gibson était dans le vrai avec son concept d'extraction d'invariants. Mais Fodor et Pylyshyn sont également dans le vrai en dénonçant chez lui un cercle vicieux. Il est effectivement vrai que « what we need is some criterion of being ecological *other than perceptibility* ». Mais ce critère, *c'est précisément le critère morphologique*. L'information morphologique *n'est pas sémantique*. Non relationnelle, elle est pourtant intrinsèquement significative. Elle peut affecter les systèmes sensoriels et perceptuels. A la suite de Thom, il faut méditer profondément sur ce statut *« sémio-physique »* des discontinuités qualitatives.

2. L'intentionalité

Un autre problème de base qu'une morphodynamique permet de résoudre sur le plan des principes est celui *de l'intentionalité* (cf. Petitot, 1984, 1986a, 1989b). On considère en général comme une évidence que l'intentionalité (la directionalité vers le monde externe) des représentations mentales est un fait sémantique. Selon nous, une telle approche, bien que traditionnelle, demeure irrémédiablement insuffisante. L'intentionalité est d'origine perceptive et les contenus sémantiques en héritent à travers la fondation de la structure conceptuelle dans le niveau 3D (au sens de Jackendoff-Marr). « Le problème des problèmes », comme dirait Husserl, est donc celui de l'intentionalité visuelle.

Or ce problème se trouve recevoir au niveau morphologique une réponse fort proche philosophiquement (mais évidemment fort éloignée mathématiquement) de celle qu'avait conçue Husserl. L'intentionalité visuelle se ramène essentiellement au passage des esquisses perceptives 2D à un objet identitaire 3D. Ce sont donc :

(i) le saut dimensionnel 2D $\longrightarrow$ 3D ;

(ii) le principe de cohérence (le principe d'identité) que constitue l'objet pour la famille (l'espace fonctionnel) de ses esquisses,

qui en définissent le concept. Or nous avons vu que ce problème fondamental peut être désormais considéré comme résolu. L'intentionalité sémantique en perd du coup ses aspects aporétiques.

Cela montre bien toute l'importance de cette médiation morphologique entre le physique et le symbolique que nous avons tenté ici d'expliciter sur un exemple précis.

Jean PETITOT,
École des Hautes Études en Sciences Sociales.

BIBLIOGRAPHIE

AMIT (D.), 1989, *Modeling Brain Function*, Cambridge, Cambridge University Press.

ANDLER (D.), 1987, « Progrès en situation d'incertitude », *Le Débat, 47*, p. 5-25.

ARNOLD (V.), VARCHENKO (V.), GOUSSEIN-ZADE (S.), 1986, *Singularités des applications différentiables*, Moscou, Éditions Mir.

BALLARD (D. H.), BROWN (C. M.), 1982, *Computer Vision*, Englewood Cliffs, N.J., Prentice Hall.

BARROW (H. G.), TENENBAUM (J. M.), 1978, « Recovering Intrinsic Scene Characteristics from Image », in *Computer Vision Systems*, A. R. HANSON, E. M. RISEMAN eds, New York, Academic Press.

BIEDERMAN (I.), 1987, « Recognition-by-Components : A Theory of Human Image Understanding », *Psychological Review*, 94, *2*, p. 115-147.

BRADY (M.), 1982, « Computational Approaches to Image Understanding », *Computing Surveys*, 14, *1*, p. 3-71.

BRANDT (P.-A.), 1986, *La Charpente modale du Sens*, thèse de doctorat d'État, Université de Paris III.

BUSER (P.), IMBERT (M.), 1987, *Vision*, Paris, Hermann.

CHURCHLAND (P. M.), 1984, *Matter and Consciousness*, Cambridge, MA, MIT Press.

DESCLÉS (J.-P.), 1986, « Représentation des connaissances, archétypes cognitifs, schèmes conceptuels, schémas grammaticaux », *Actes sémiotiques*, VII, *69/70*.

FELDMAN (J. A.), 1985, « Four Frames Suffice : A Provisional Model of Vision and Space », *The Behavorial and Brain Sciences, 8*, p. 265-289.

FODOR (J. A.), PYLYSHYN (Z. W.), 1981, « How Direct Is Visual Perception ? Some Reflections on Gibson's " Ecological Approach " », *Cognition*, 9, p. 139-196.

FODOR (J. A.), 1984, *The Modularity of Mind*, Cambridge, MA, MIT Press.

GIBSON (J. J.), 1979, *The Ecological Approach to Visual Perception*, Boston, Houghton-Mifflin.

GRIMSON (W. E. L.), HILDRETH (E. C.), 1985, « Comments on Haralick 1984 », *IEEE, Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-7, p. 121-126.

HARALICK (R. M.), 1984, « Digital Step Edges from Zero Crossings of Second Directional Curvature », *IEEE, Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-6, p. 58-68.

HOFFMAN (D. D.), 1983, *Representing Shapes for Visual Recognition*, Doctoral Dissertation, MIT.

HOFFMAN (D. D.), RICHARDS (W. A.), 1984, « Parts of Recognition », *Cognition*, 18, p. 65-96.

HOFFMAN (D. D.), BENNETT (B. M.), 1986, « The Computation of Structure from Fixed-Axis Motion : Rigid Structures », *Biological Cybernetics*, 54, p. 71-83.

IKEUCHI (K.), « Shape from Regular Patterns », *Artificial Intelligence*, 22, p. 49-75.

JACKENDOFF (R.), 1983, *Semantics and Cognition*, Cambridge, MA, MIT Press.

JACKENDOFF (R.), 1987, *Consciousness and the Computational Mind*, Cambridge, MA, MIT Press.

KITCHER (P.), 1988, « Marr's Computational Theory of Vision », *Philosophy of Science*, 55, p. 1-24.

KOENDERINK (J.J.), VAN DOORN (A.J.), 1976, « The Singularities of the Visual Mapping », *Biological Cybernetics*, 25, p. 51-59.

KOENDERINK (J.J.), VAN DOORN (A.J.), 1979, « The Internal Representation of Solid Shape with Respect to Vision », *Biological Cybernetics*, 32, p. 211-216.

KOENDERINK (J.J.), VAN DOORN (A.J.), 1986, « Dynamic Shape », *Biological Cybernetics*, 53, p. 383-396.

KOENDERINK (J.J.), VAN DOORN (A.J.), 1987, « Representation of Local Geometry in the Visual System », *Biological Cybernetics*, 55, p. 367-375.

KOSSLYN (S. M.), 1980, *Image and Mind*, Cambridge, MA, Harvard University Press.

LANGACKER (R.), 1987, *Foundations of Cognitive Grammar*, Stanford University Press.

LE DÉBAT, 1987, « Une nouvelle science de l'esprit », *Le Débat*, *47*.

LTC, 1989, *Logos et Théorie des catastrophes*, Colloque de Cerisy à partir de l'œuvre de René Thom, Jean PETITOT, éd., Genève, Éditions Patiño.

MARR (D.), 1982, *Vision*, San Francisco, Freeman.

MEYER (Y.), 1989, « Ondelettes, filtres miroirs en quadrature et traitement numérique de l'image », *Gazette des mathématiciens*, 40, p. 31-42.

MINGOLLA (E.), TODD (J.T.), 1986, « Perception of Solid Shape from Shading », *Biological Cybernetics*, 53, *3*, p. 137-151.

OUELLET (P.), 1987, « Une physique du sens », *Critique*, *481/482*, p. 577-597.

P.D.P., 1986, *Parallel Distributed Processing*, David E. RUMELHART, James L. MCCLELLAND eds, Cambridge, MIT Press.

PETITOT (J.), 1977, « Topologie du carré sémiotique », *Études littéraires*, p. 347-428, Québec, Université de Laval.

PETITOT (J.), 1979, « Hypothèse localiste et Théorie des catastrophes », in *Théories du langage, théories de l'apprentissage*, M. PIATTELLI, éd., Paris, Le Seuil.

PETITOT (J.), 1982, *Pour un schématisme de la structure*, thèse de doctorat d'État, Paris, E.H.E.S.S.

PETITOT (J.), 1983, « Théorie des catastrophes et structures sémio-narratives », *Actes sémiotiques*, V, *47/48*, p. 5-37.

PETITOT (J.), 1984, « La lacune du contour », *Anàlise*, 1, *1*, p. 101-140, Lisbonne.

PETITOT (J.), 1985, *Morphogenèse du Sens*, Paris, Presses Universitaires de France.

PETITOT (J.), 1986a, « Le " morphological turn " de la phénoménologie », *Document du C.A.M.S.*, Paris, E.H.E.S.S.

PETITOT (J.), 1986b, « Épistémologie des phénomènes critiques », *Document du C.A.M.S.*, Paris, E.H.E.S.S.

PETITOT (J.), 1988, « Approche morphodynamique de la formule canonique du mythe », *L'Homme*, 106-107, XVIII *(2-3)*, p. 24-50.

PETITOT (J.), 1989a, « Éléments de dynamique modale », *Poetica et Analytica*, 6, p. 44-79, Université d'Aarhus.

PETITOT (J.), 1989b, « Structuralisme et Phénoménologie », *LTC*, 1989, p. 345-376.

PETITOT (J.), 1989c, « On the Linguistic Import of Catastrophe Theory », *Semiotica*, 74, *3/4*, p. 179-209.

PETITOT (J.), 1989d, « Catastrophe Theory and Semio-Narrative Structures », in *Paris School of Semiotics*, P. PERRON, F. COLLINS, eds, Amsterdam, John Benjamins, p. 177-212.

PETITOT (J.), 1989e, « Morphodynamics and the Categorical Perception of Phonological Units », *Theoretical Linguistics*, 15, *1/2*, p. 25-71.

PETITOT (J.), 1989f, « Hypothèse localiste, Modèles morphodynamiques et théories cognitives : remarques sur une note de 1975 », *Semiotica*, 77, *1/3*, p. 65-119.

PETITOT (J.), 1989g, « Forme », *Encyclopaedia Universalis*, XI, p. 712-728, Paris.

PETITOT (J.), 1989h, « La modélisation : formalisation ou mathématisation ? L'exemple de l'approche morphodynamique dans les sciences du langage », in *Perspectives méthodologiques et épistémologiques dans les sciences du langage*, M.J. REICHLER-BÉGULIN, éd., Bern, Peter Lang, p. 205-220.

PETITOT (J.), 1989i, « Why Connectionism Is Such a Good Thing ? », in *Workshop Connectionism and Language*, San Marino, Università degli Studi.

PINKER (S.), 1984, « Visual Cognition : An Introduction », *Cognition*, 18, p. 1-63.

PINKER (S.), ed., 1984, *Visual Cognition, Cognition*, 18, Cambridge, MA, MIT Press.

POGGIO (T.), 1984, « Vision by Man and Machine », *Scientific American*, 250, *4*, p. 68-78.

PRÉFACES, 1988, « Un tournant cognitif dans les sciences humaines », *Préfaces*, *10*, p. 67-105.

PROUST (J.), 1987, « L'intelligence artificielle comme philosophie », *Le Débat*, 47, p. 88-102.

PYLYSHYN (Z.), 1986, *Computation and Cognition*, Cambridge, MA, MIT Press.

RICHTER (J.), ULLMAN (S.), 1986, « Non-Linearities in Cortical Simple Cells and the Possible Detection of Zero Crossings », *Biological Cybernetics*, 53, *3*, p. 195-202.

SHEPARD (R. N.), COOPER (L. A.), 1982, *Mental Images and their Transformations*, Cambridge, MA, MIT Press.

SMOLENSKY (P.), 1988, « On the Proper Treatment of Connectionism », *The Behavioral and Brain Sciences*, 11, p. 1-74.

STILLINGS (N. A.), *et al.*, 1987, *Cognitive Science. An Introduction*, Cambridge, MA, MIT Press.

TALMY (L.), 1978, « Relation of Grammar to Cognition », in *Proceedings of TINLAP-2*, D. WALTZ, ed., Urbana, University of Illinois.

TALMY (L.), 1983, « How Language Structures Space », in *Spatial Orientation : Theory, Research and Application,* H. PICK, L. ACREDOLO, eds, New York, Plenum Press.

TALMY (L.), 1985, « Force Dynamics in Language and Thought », *Parasession on Causatives and Agentivity,* Chicago Linguistic Society, 21st. Regional Meeting.

THOM (R.), 1972, *Stabilité structurelle et Morphogenèse,* New York, Benjamin, Paris, Édiscience.

THOM (R.), 1978, « Morphogenèse et Imaginaire », *Circé, 8-9,* Paris, Éditions Lettres Modernes.

THOM (R.), 1980, *Modèles mathématiques de la Morphogenèse,* 2e éd., Paris, Christian Bourgois.

THOM (R.), 1988, *Esquisse d'une Sémiophysique,* Paris, Inter-Éditions.

ULLMAN (S.), 1979, *The Interpretation of Visual Motion,* Cambridge, MA, MIT Press.

ULLMAN (S.), 1984, « Visual routines », *Cognition,* 18, p. 97-159.

WILDGEN (W.), 1982, *Catastrophe Theoretic Semantics,* Amsterdam, Benjamins.

ZEEMAN (Ch.), 1977, *Catastrophe Theory,* Massachusetts, Addison-Wesley.