Veto thermodynamique contre les duplicateurs perpétuels de seconde espèce

O. Costa de Beauregard

Institut Henri Poincaré
11 rue P. et M. Curie, 75005 Paris

RESUME. Il a été récemment argué que, dans le processus standard *codage, traitement, décodage* (et effacement) d'une information, la dernière étape seule est *nécessairement* irréversible –ce qui est parfaitement vrai. Il en va de même avec le cycle de Carnot : *chaleur tirée d'une source chaude*, *production de travail*, *décharge de chaleur dans le condenseur*. Ce qu'oublient les auteurs en question est la *nécessité* de soutirer au départ de la néguentropie pour le codage (éventuellement la duplication) d'un "message" ; les modèles qu'ils proposent sont alors des schémas de "duplicateurs perpétuels de seconde espèce", ce que Brillouin refuse explicitement. Brillouin explique la génération d'ordre, ou d'information, en termes de néguentropie puisée en amont, alors que Prigogine l'explique en termes d'entropie déchargée vers l'aval. En unissant les deux points de vue on obtient un schéma d'ordinateur exactement parallèle au schéma de la machine de Carnot, la différence étant que l'un produit de l'information et l'autre du travail. J'achève ma réfutation de la partie discutable de la thèse incriminée en montrant, sur les modèles proposés, qu'à *chaque* bit d'information *codé* à une température T_c, *traité* à une température T, *décodé* à une température T_f ($T_c \geq T \geq T_f$) *doit nécessairement être associé un quantum d'énergie* valant au moins $(\ln 2)\, k$ fois T_c, T ou T_f ; *faute de cela, l'information se perdrait dans le bruit.*

ABSTRACT. It has been argued recently that, in the standard procedure of coding, treating, decoding *(and erasing) an information, only the last step is* necessarily *irreversible, –which is quite true. The same holds with the Carnot cycle : borrowing heat from a firebox, generating work, dumping heat in the condenser. What the authors I refer to do forget is the* necessity *of borrowing negentropy for the process of coding (or possibly duplicating) a "message" ; the models they propose are thus schemes of "perpetual duplicators of the* *second kind", a possibility explicitly denied by Brillouin. Brillouin explains the generation of order, or information, by tapping negentropy from upstream, while Prigogine explains it by dumping entropy downstream. Joining both ideas yields a scheme of the computer paralleling exactly that of Carnot's heat engine, the differences being that the one yields information and the other, work. My refutation of part of the thesis of the authors I refer to consist of showing that to* each *bit of information* coded *at a temperature* T_c, treated *at a temperature* T, decoded *at a temperature* T_f ($T_c \geq T \geq T_f$) must *be attached* an energy quantum *at least* ($\ln 2$) k *times equal to* T_c, T *or* T_f. Otherwise, the information would be drowned in the thermal noise.

1. Introduction

Récemment plusieurs auteurs [1][2][3][4] ont argué que, lors du processus général *codage, traitement, décodage* d'une information, la seule irréversibilité physique irrécupérable est celle survenant à la dernière étape : effacement de l'enregistrement ou "décodage" au sens matériel du terme. Cela n'a rien de surprenant si l'on se souvient qu'il en va de même dans le cycle de Carnot : la seule énergie *en principe* irrécupérable est celle évacuée vers la source froide, ou condenseur. Si l'on stocke les enregistrements au lieu de les "dégrader", ils deviennent vite encombrants, et "l'on ne peut plus s'y retrouver". C'est alors, dit Bennett [4], l'entropie de la mémoire, plutôt que celle de l'environnement, qui augmente ...

Le phénomène étudié par les auteurs que je cite correspond exactement à celui qui est au coeur des analyses de Prigogine [5] : production d'ordre par décharge d'entropie dans l'environnement. Mais souvenons-nous que Brillouin [6] analyse la production d'ordre en termes de néguentropie soutirée de sources existant en amont, analogues aux sources chaudes de Carnot. Faute de prendre en considération cette partie de la phénoménologie totale, les auteurs [1][2][3][4] aboutissent à l'incroyable conclusion que la duplication, ou la multiplication, d'un document codé serait possible tout à fait gratuitement, sans qu'il soit besoin de payer en néguentropie tout ce surcroit d'information exprimée (ou imprimée, comme on voudra). *C'est ce que dénie explicitement Brillouin* [6].

En faisant la synthèse des points de vue de Brillouin et de Prigogine on obtient un schéma de l'ordinateur exactement parallèle à celui de la machine thermique de Carnot, que je présenterai dans la section 2. La différence est que la "machine à feu" idéale fonctionne à bilan de néguentropie nul et distribue du travail, tandis que l'ordinateur idéal fonctionne à bilan d'énergie nul et distribue de l'information. "De la machine à feu à l'ordinateur" pourrait titrer un chapître de l'histoire sociale des 19è et 20è siècles... Cette section 2 contiendra donc à la fois une approbation de l'analyse que font les auteurs [1][2][3][4] du rôle du "condenseur" et une critique de l'oubli qu'ils commettent du rôle nécessaire du "foyer" de Brillouin.

Les sections 3, 4 et 5 réfuteront les modèles –les "expériences pensées"– par lesquels ces auteurs disent illustrer la duplication (ou le traitement de l'information) à un coût de néguentropie rigoureusement nul. Je montrerai pour cela qu'à *chaque bit* d'information *codé* à la température T_c, *traité* à la température T, *décodé* à la température T_f ($T_c > T > T_f$) correspond *nécessairement* (au moins) *un quantum thermique d'énergie encapsulé, transporté, dégradé*, valant $(\ln 2)\, k$ fois T_c, T, T_f, respectivement.

Faute de cela, l'expression, ou l'impression (comme on voudra) du document ne résisterait pas à l'agitation thermique ; elle serait "noyée dans le bruit". L'ordinateur ne serait qu'un presse-purée ...

A l'égard de Brillouin les auteurs que je critique ne commettent pas seulement l'oubli que je viens de mentionner, mais aussi un contresens. Ils croient qu'au codage selon Brillouin est associée une dégradation. C'est exactement le contraire qui est vrai : lors du codage, le quantum thermique $(\ln 2)kT_c$ n'est pas "perdu", mais bel et bien "sauvé" par le fait de son encapsulage. Sans cela il s'écoulerait sans profit aucun, "naturellement", de la source chaude vers l'ambiance plus froide.

La vérité est donc que, comme la machine de Carnot, l'ordinateur fonctionne en dérivation sur la cascade universelle de la néguentropie ; l'une et l'autre parasitent, en quelque sorte, cette cascade qui, à elle seule, ne ferait que laisser la chaleur couler de la source chaude à la source froide.

Impossibilité du duplicateur perpétuel de seconde espèce : voilà qui suffit à fonder les équations fondamentales de l'ordinateur ; c'est le *Principe de Carnot généralisé*, ou *Principe de Brillouin*, que je vais ici "défendre et illustrer".

2. La machine à feu et l'ordinateur

Rappelons que la "machine à feu" de Carnot emprunte une quantité de chaleur dQ_c à une source chaude de température T_c et rejette une quantité de chaleur $dQ_f < dQ_c$ à une source froide de température $T_f < T_c$, libérant la différence

$$dW = dQ_h - dQ_c > 0 \tag{1}$$

sous forme de travail. Le rendement maximum est obtenu si

$$-T_c^{-1}dQ_c + T_f^{-1}dQ_f = 0 \tag{2}$$

ou

$$dN_c - dN_f = 0, \tag{3}$$

$dN \equiv T^{-1}dQ$ notant une néguentropie intérieure à la machine ; la machine idéale opère ainsi à bilan de néguentropie nul.

Un point très important est que *la seule perte irréparable (par principe) survient à la dernière étape : décharge de la néguentropie* $-T_f^{-1}dQ_f$ *dans le condenseur*, conjointement à une énergie de basse qualité dQ_f. Le contenu des poubelles n'est pas tout à fait sans valeur ...

Voyons maintenant l'ordinateur, en combinant les idées de Brillouin [6] et celles de Prigogine [5]. L'ordinateur idéal opère à bilan d'énergie nul

$$dW = dQ_c - dQ_f = 0 \tag{4}$$

mais il fournit une information

$$(\ln 2)kdI = dN_c - dN_f > 0 \tag{5}$$

ou

$$(\ln 2)kdI = -T_c^{-1}dQ_c + T_f^{-1}dQ_f > 0, \tag{6}$$

différence entre la néguentropie empruntée en amont et celle abandonnée à l'aval, lors du *codage* et du *décodage*, respectivement. Ceci définit le rendement optimal d'un ordinateur fonctionnant entre les températures T_c du "foyer de Brillouin" et T_f du "condenseur de Prigogine". Comme exemple on a le laser, une sorte de "photomultiplicateur", où le "gain" est réparti entre beaucoup de "petits porteurs".

Comme y insistent justement Bennett [3][4], Landauer [1][2], et d'autres qu'ils citent, *la seule perte irréparable survient à la dernière étape*, le décodage, où l'information exprimée est effacée pour faire place à une nouvelle opération. Avec elle, une énergie dQ_c (idéalement égale à celle investie au départ) est "dégradée" à la température T_f.

Cependant, la “fixation” que font Bennett, Landauer, et d’autres qu’ils citent, sur les problèmes du condenseur les aveuglent quant à la présence de la “source chaude” de Brillouin, *absolument nécessaire au codage.*

Ce parallélisme formel entre la machine à feu et l’ordinateur illustre à sa manière la transition entre une technologie –et– culture à dominante énergétique et une autre à dominante informatique.

3. Critique d’un duplicateur “idéal” proposé par Landauer

Les Figures 1a à 1d illustrent un procédé de “duplication totalement gratuite” proposé par Landauer [2]. *Etape a* : le “système bistable” formé de deux puits de potentiel appariés contenant une seule particule affiche un bit d’information. *Etape b* : ayant abaissé la barrière, Landauer couple la particule à une autre mise au fond d’un potentiel simple, au moyen d’un ressort détendu. *Etape c* : la barrière initiale est relevée. *Etape d* : une nouvelle barrière est élevée au milieu du puits simple. C’est fini. La copie est conforme. Le bit initial est devenu deux bits ... Et rien n’oblige à s’arrêter là. Voilà le prototype d’un “duplicateur perpétuel de seconde espèce”... et (pourquoi pas) la preuve que $1 = 2 = 3 = \; ...n = \; .....$

Au temps de Huygens et d’Euler, où des structures métastables de ce type pouvaient être admises comme indestructibles, et où le concept d’information n’avait pas encore été dégagé du calcul des probabilités de Pascal et de Fermat, l’argument aurait pu passer. Depuis Maxwell, il faut penser au désordre thermique.

Si ce modèle est supposé fonctionner dans une ambiance isotherme à T^0K, la particule a une énergie cinétique moyenne $(3/2)kT$, et l’on ne peut pas abaisser totalement la barrière, sous peine de la laisser s’échapper “Dieu sait où”. Donc le double puits de potentiel *doit piéger à la fois* 1 bit d’information et $(3/2)kT$ unités d’énergie (au minimum). De plus, le ressort vibre, et possède une énergie $(1/2)kT$ s’il a un degré de liberté. “Le problème” est alors qu’il doit être assez raide pour assurer un positionnement précis, et assez souple pour ne pas violenter la barrière de potentiel. Un problème requérant une discussion attentive ...

Pour dire la chose autrement, *c’est manifestement Landauer en personne* qui, à l’étape *b*, a inséré le second bit, sous la forme d’une “néguentropie structurelle”. Comme $1 + 1 = 2$, l’arithmétique est sauve ...

En conclusion, *ce schème de duplicateur perpétuel n'est pas viable.*

4. Critique du télégraphe lent à puits de potentiel jumeaux

La Figure 2 illustre quelque chose d'analogue à ces télésièges biplaces qu'on voit dans certaines stations de ski. Un "mur de potentiel" sépare les deux conduits ; et des "crans de potentiel en mouvement" séparent chaque paire de "godets" de la précédente et de la suivante. Chaque bit d'information codé, transporté et décodé l'est sous la forme d'une particule occupant l'un ou l'autre des godets. Le point est que, si le dispositif est supposé fonctionner dans un environnement monotherme à température T, les murs de potentiel doivent avoir au moins la hauteur $(3/2)kT$. Donc à *chaque* bit *codé, transporté* et *décodé doit être biunivoquement associé* à un quantum d'énergie $(3/2)kT$ *encapsulé, transporté* et *libéré.* A l'arrivée du télésiège, l'occupant se perd dans la foule et le quantum d'énergie avec lui.

Une question : pourquoi apparaît-il le facteur 3/2 au lieu du facteur $\ln 2$ proposé par la théorie générale ? Parce qu'un raisonnement "concret" n'a pas la précision d'un raisonnement abstrait. Essayons donc de raffiner.

Le facteur 3 se réfère aux 3 dimensions de l'espace euclidien ; il serait n si l'on raisonnait avec n degrés de liberté. Pour descendre au niveau fondamental, il faut faire $n = 1$.

Alors, le dénominateur 2 est gênant, parce que $1/2 < \ln 2$. Mais, si l'on veut coder proprement une alternative, *il faut que les deux états soient occupés.* Prenons alors deux boules, une blanche signifiant 0, une noire signifiant 1, et codons sous la forme $(01) = 1$ *ou* $(10) = 2$. Alors, chaque paire de godets transportent 1 bit transporte aussi kT unités d'énergie.

Quid enfin du facteur $\ln 2$ manquant ? Pour coder convenablement une *alternative* le bit, et non le "neperbit", est la bonne unité. D'où le facteur $\ln 2$.

Et si quelqu'un venait dire : "en mécanique quantique il y a gel des degrés de liberté à basse température, où l'on pourrait se contenter de barrières de potentiel moins hautes", la réponse serait : "n'oublions pas l'effet tunnel" ...

Les auteurs que je discute arguent de ce qu'*un ordinateur est exactement réversible en ce qui concerne l'information exprimée* (ou, comme l'appelle Bennett [4], l'*information algorithmique*). Mais *il ne suit pas* de là qu'il soit réversible en ce qui concerne l'information thermodynamique sous-jacente. *Cette seconde irréversibilité n'est convenablement préservée que si à chaque bit traité est biunivoquement associé un quantum d'énergie* $(\ln 2)kT$.

Si la file des godets est très longue, des échanges entre deux godets appariés finiront par avoir lieu, soit par dessus, soit à travers la barrière de potentiel. C'est pourquoi Brillouin dit que (par exemple) le temps de transmission d'un télégramme doit être inférieur au temps de relaxation thermique du télégraphe.

5. Critique du modèle balistique ou "des sphères dures"

Discuté par Sinai [7] le gaz de sphères dures est un "système chaotique" intrinsèquement réversible ; s'il est affligé d'une incertitude initiale arbitrairement petite, sa prédiction, ou sa rétrodiction, divergent exponentiellement, l'une comme l'autre. Ceci vaut au zéro absolu, une clause à exclure s'agissant d'un ordinateur.

A $T^0 K$ chaque sphère est affectée d'une énergie thermique $(3/2)kT$ –une propriété mise à profit par Jean Perrin dans une détermination fameuse du nombre d'Avogadro.

Ce modèle en vaut bien un autre pour discuter codage, traitement, décodage de l'information [1]. Mais, naturellement, toute sphère recevant ou transmettant 1 bit d'information *doit impérativement* avoir alors une énergie cinétique au moins égale à $(3/2)kT$ ... comme au jeu de billard.

Tout joueur de billard sait que les boules ralentissent et finissent par s'arrêter, et qu'il n'a donc qu'un temps limité pour maîtriser ses "effets". *Mutatis mutandis*, cela vaut aussi pour un télégraphe.

6. Conclusions : défense et illustration du "principe de Brillouin"

On peut trouver chez Maxwell, Boltzmann, Gibbs des citations montrant qu'ils assimilaient plus ou moins consciemment les deux concepts *désordre* et *connaissance incomplète.* La même idée est implicite dans le "coarse graining", même (et surtout) s'il n'est pas arbitraire, mais inhérent à la manière même de poser un problème, de choisir ce qu'on décide de *connaître* et de *contrôler*, ou *de ne pas connaître* et *ne pas* *contrôler*. On peut lire à ce sujet Jaynes [8], et méditer un fameux exemple de Bertrand [9] en calcul des probabilités. Chez les philosophes, Bergson [10] a fait le même rapprochement.

Un corollaire de la précédente assimilation est la reconnaissance que *l'information est un concept à deux faces : acquisition de connaissance*, symbolisée par $N \to (\ln 2)kI$; et *pouvoir d'organisation* symbolisé par $(\ln 2)kI \to N$. La cybernétique a ainsi redécouvert ce qu'avait discuté (de façon moins formalisée) Aristote.

Faut-il ou non assimiler l'information codée (information algorithmique de Bennett [4]) à l'information thermodynamique, et poser le *Principe de Brillouin* qu'au sein d'un système isolé la grandeur

$$N + (\ln 2)kI$$

ne peut pas croître ? C'est un prolongement assez naturel des intuitions de Maxwell, Boltzmann et Gibbs, et cependant un très audacieux postulat qui, comme d'autres du même genre (principe d'inertie, principe de Carnot) ne peut que procéder de la méditation d'exemples, et n'être justifié que par ses conséquences.

Examinons la question sur un exemple. Un domino du type $(0 \mid 1)$ exhibe un bit d'information. Comment va-t-on le "dupliquer" ? Par exemple, en perçant à travers le point noir, et en le fendant dans son épaisseur, ce qui peut être assez facile s'il est du type "sandwich". Quel sera le coût en néguentropie ? Au lieu de percer un domino, imaginons qu'on perce à travers un bloc-notes au moyen d'un emporte-pièces. L'énergie à dépenser pour percer g feuilles vaudra g fois l'énergie nécessaire pour percer une feuille ; cela fait, on peut considérer la diffusion de cette information comme gratuite ...

Le principe de Brillouin énonce que *pour imprimer* 1 *bit d'information codée à la température ambiante* T^0K *il faut dépenser au moins* $(\ln 2)kT$ *unités d'énergie.* Il justifie cela par des exemples, et par de nombreuses conséquences. Le présent article est proposé à l'appui du Principe de Brillouin.

Une fois décrété qualitativement, comme le Principe de Carnot, sous une forme ou une autre (par exemple, impossibilité du duplicateur perpétuel) le Principe de Brillouin se formalise au moyen des *quelques formules universelles* suivantes : *Entropie de Clausius*

$$dS = T^{-1}dQ \ ;$$

entropie de Boltzmann

$$S = k \ln g \quad \text{ou} \quad N = k \ln(1/g) \ ;$$

information

$$I = log_2(1/g) \ ;$$

et, bien sûr, $log_e g = log_e 2 \cdot log_2 g$.

Faut-il rappeler que ce modèle est tout à fait congru à celui de Prigogine traitant de la génération cosmologique du hadron, de l'atome, de la molécule, etc. à partir du "Big Bang", par refroidissement de l'Univers et copie de "gabarits virtuels", ou "modèles mathématiques formels" ? C'est bien souvent que la physique a recouru au virtuel, depuis les "images virtuelles" de l'optique classique, et les énergies potentielles négatives ...

J'ai gardé pour la fin une remarque de Brillouin tout à fait appropriée à cette discussion : le quantum d'énergie $(\ln 2)kT$ nécessaire pour *imprimer* à la température T un bit sous la forme d'une structure résistant à la dégradation thermique *est si petit qu'on risque de l'oublier ...*

En terminant je tiens à remercier R. Landauer pour la très intéressante, quoique vive, correspondance que nous avons échangée (et qui du reste se poursuit). Merci aussi à lui de sa courtoisie de m'avoir envoyé deux de ses calques pour reproduction dans un article. De la discussion jaillira peut-être la lumière ...

Post scriptum (4 Mars 1989)

La discussion s'est poursuivie en effet, aboutissant pour le moment à deux articles contradictoires [11][12] sous presse à *Foundations of Physics* (livraison de Juin 1989). J'avais demandé à Georges Lochak de différer la publication du présent article afin de rendre possibles quelques ajustements ; voilà qui est fait.

Encore une ultime remarque. Certains feront observer que le codage d'un traitement informatique ne se fait pas "évidemment" à partir d'une source chaude. C'est vrai : on tape sur des touches, ou autre opération analogue. Le *hic* est que *c'est par cette opération qu'on injecte* dans l'ordinateur *l'énergie* et la néguentropie nécessaires au "traitement de l'information" ; celles-ci doivent être au moins égales aux deux sommes d'énergie et d'information délivrées par la machine –y compris celles finalement dégradées.

Bien entendu, cette énergie et cette néguentropie sont soutirées à la cascade néguentropique universelle. Si le codage est fait au sein d'une ambiance à la température T^0K, les valeurs mentionnées dans le texte doivent être satisfaites.

Références

[1] R. Landauer, Foundations of Physics, 16, 551 (1986).
[2] R. Landauer, *Computation, Measurement, Communication and energy dissipation in* Signal Processing, S. Haykin ed., Prentice Hall (sous presse).
[3] C.H. Bennett, Int. Journ. Theor. Phys. 21, 905 (1982).
[4] C.H. Bennett, Scientific American 255, 108 (1987).
[5] G. Nicolis et I. Prigogine, *Self Organization in Non equilibrium Systems* (Wiley, New York, 1977).
[6] L. Brillouin, *Science and Information Theory*, 2ème édition, Academic Press, New York, 1962. Voir notamment Chapitres 12 et 18.
[7] Y.G. Sinai, *Introduction to ergodic theory*, Math. Notes, Princeton University Press, 1976.
[8] E.T. Jaynes, Papers on Probability, Statistics and Statistical Physics, R.D. Rosekranz, ed., D. Reidel, Dordrecht, 1983 ; voir pp. 85-86.
[9] J. Bertrand, *Calcul des Probabilités*, Gauthier-Villars, Paris, 1888, pp. 4-5.
[10] H. Bergson, *L'Evolution Créatrice*, Chapitre 3.
[11] O. Costa de Beauregard, The computer and the heat engine, Found. Phys. **19**, no. 6, (sous presse).
[12] R. Landauer, Response to "The computer and the heat engine", ibid.

(Manuscrit reçu le 9 février 1988)

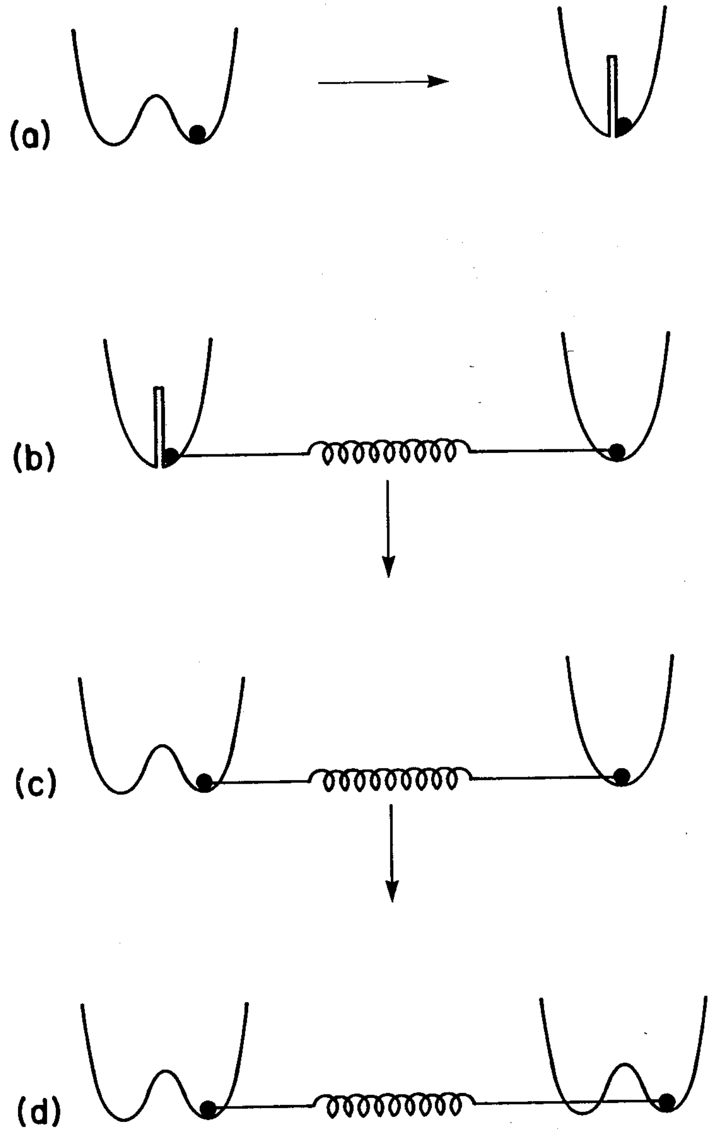

Figure 1.

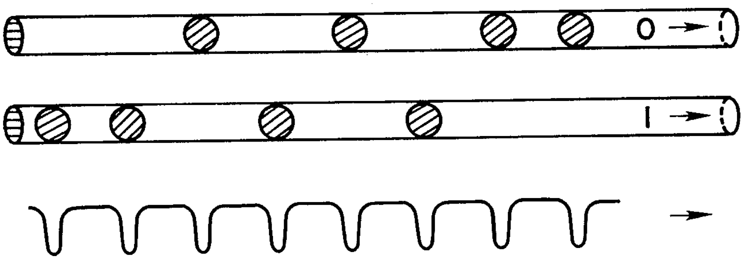

Figure 2.