

Les premiers transistors : source Wikipedia.

Le transistor : innovation majeure de l'électronique du XX^{ème} siècle

Marc Leconte

Membre émérite de la SEE

Les premiers pas de l'électronique

Le XIX^{ème} siècle avait été marqué par la découverte des lois de l'électricité et du magnétisme dont les premières applications concernaient la production et la distribution de l'électricité et à la fin du siècle la télégraphie sans fil. Mais cette dernière application utilisait encore avant 1900 des moyens rudi-

Le 1^{er} juillet 1948 le transistor entrait dans l'histoire, mais comment le qualifier ? S'agissait-il d'une innovation de produit ou de procédé, autrement dit était-ce une révolution physique ou technique ? La réponse n'allait pas de soi et dans ce qui suit nous allons exposer ce qui a conduit à remplacer les tubes à vide qui marquaient les débuts de l'électronique par un composant à semi-conducteur.

mentaires pour la détection et l'émission des ondes. Le véritable début de l'électronique datait au fond de la découverte de la radio ou plus exactement de la télégraphie sans fil, qualification qui restera longtemps pour désigner les communications radiodiffusées. Ceci s'inscrit dans le contexte théorique de la découverte de l'électron dans la dernière décennie du XIX^{ème} siècle et de la théorie de l'électromagnétisme formalisée par Maxwell.

Ensuite, en 1888, Heinrich Hertz mit en évidence les ondes électromagnétiques et par la suite les chercheurs tentèrent de développer des appareils capables de les émettre et de les détecter. Une première approche reposait sur les techniques électriques comme les éclateurs à étincelles et les tubes à limaille et une autre approche était fondée sur des composants qui régulaient et pouvaient amplifier des courants faibles ; ainsi naquit

l'électronique. Ces premiers composants furent les tubes à vide.

Les tubes à vide

De nombreuses expériences effectuées par les physiciens avaient mis en évidence l'existence des électrons et les phénomènes thermoioniques. Edison avait fait l'observation en étudiant la fabrication des ampoules électriques d'un dépôt de carbone ou de métal sur les parois de l'ampoule. Edison chercha à déterminer les causes de ces dépôts et disposa à l'intérieur une plaque métallique en face du filament. Il constata qu'il s'établissait un courant entre le filament chauffé et la plaque métallique et décrit le phénomène. Il avait virtuellement découvert un redresseur car le flux d'électron ne pouvait aller que dans un seul sens mais étant opposé à l'utilisation du courant alternatif, il n'exploita pas cette découverte dont l'utilité lui échappait. Dès 1883 un chercheur, J. A. Fleming, qui poursuivait une activité d'universitaire et d'industriel étudiait par intermittence le phénomène décrit par Edison alors qu'il travaillait au sein de sa filiale anglaise. C'est plus tard, en 1903, qu'il s'intéressa au phénomène de manière active alors qu'il œuvrait au sein de la compagnie de Marconi autour de la recherche d'appareils de détection des ondes. Fleming découvrit que l'effet Edison permettait effectivement d'obtenir un effet de détection en ne laissant passer le courant que dans un seul sens. Il déposa pour cela un brevet en 1904, brevet qui fera date plus tard car il marquait le début du développement des tubes à vide dits à l'époque thermoioniques mais plus précisément le début de l'électricité des courants faibles autrement appelée électronique.

La triode

La diode à vide appelée à l'époque valve thermoionique était plus facile d'utilisation et plus stable que la galène mais il était nécessaire de porter le filament à incandescence par une source de courant séparée et cet inconvénient n'était pas compensé par une détection plus sensible. Cela explique que le tube de Fleming ne fut pas utilisé immédiatement dans les radiocommunications. Une première amélioration vint des Etats-Unis

avec les travaux d'un chercheur américain, Lee de Forest (figure 1), qui imagina une modification décisive en interposant une électrode appelée grille entre la cathode et l'anode. En faisant varier le potentiel de la grille il était possible d'agir sur le courant filament plaque et d'en faire varier l'intensité. Lee de Forest appela cette nouvelle lampe l'audion mais l'utilisa d'abord comme un détecteur plus perfectionné que la diode de Fleming. Le progrès déterminant fut de constater que ce dispositif pouvait constituer un amplificateur car une infime variation de potentiel de la grille produisait une variation importante et proportionnelle du courant dans le circuit anodique. Cette possibilité de contrôler un courant important à partir de signaux plus faibles permit l'essor de l'électronique. L'audion ou, comme ce tube fut appelé plus tard, la triode, resta pendant plusieurs années dans les laboratoires sans utilisation particulière. Après 1910, Irving Langmuir qui travaillait sur les lampes d'éclairage en atmosphère gazeuse perfectionna la triode en montrant que le fonctionnement

était mieux assuré lorsque le vide parfait régnait dans le tube. Peu à peu, les sociétés de téléphone utilisèrent la triode comme amplificateur redresseur sur les lignes pour les communications à grande distance et elle fut mise en service pour le téléphone en 1915 et employée en radiotélégraphie à peu près en même temps. La guerre apporta des améliorations dans son fonctionnement avec notamment la cathode à oxyde qui permettait de prolonger la durée des tubes de manière importante. Le groupe du général Ferrié en France fut informé très tôt du développement de la triode et Henri Abraham fut chargé de la production industrielle qui débuta à la fin de 1914. D'autres chercheurs comme Camille Gutton et Lucien Levy étudièrent des montages pour expérimenter les applications possibles de la triode et, dès 1915, les triodes équipaient des amplificateurs basse et haute fréquence. C'est dans ce cadre qu'apparut la célèbre lampe TM (télégraphie militaire) produite à des centaines de milliers d'exemplaire durant la guerre. A partir des années 1920, la radiodiffusion et



Figure 1 : Lee de Forest avec ses tubes expérimentaux. Source <https://www.radioworld.com>.

“ Les pionniers de la radio avaient effectué leurs premières expériences avec des ondes de quelques centimètres mais ensuite la pratique montra l’intérêt d’augmenter la longueur d’onde de plusieurs dizaines à plusieurs centaines de mètres. ”

- la commercialisation dans le grand public des récepteurs radio déterminèrent l’essor extraordinaire de l’industrie des tubes.

Généralisation des tubes

Pour améliorer la triode, des modifications importantes furent apportées, en particulier dans la structure du tube en ajoutant une ou plusieurs électrodes. La première réalisation consista en 1916 à introduire une grille supplémentaire positive entre la cathode et la grille de commande. Cette électrode accélérât les électrons produits par la cathode et permettait de réduire la tension anodique. Par la suite, on découvrit que la capacité entre la grille de commande et l’anode pouvait créer un couplage parasite entre les circuits, problème qui fut résolu par une grille supplémentaire chargée positivement entre la grille de commande et l’anode car il se formait un écran électrostatique entre les deux circuits. Ce tube, appelé tétrode, avait un coefficient d’amplification accru mais il avait d’autres inconvénients car les électrodes positives près de la plaque favorisaient l’extraction d’électrons secondaires arrachés à l’anode et une solution fut apportée avec l’introduction d’une troisième grille dite d’arrêt réunie à la cathode entre la grille écran et l’anode. Ce nouveau tube fut appelé pentode. En 1928, la pentode devenait un tube extrêmement répandu dans les récepteurs radio et d’autres modèles suivirent, double diode, double triode, double pentode ou diode pentode. Ces derniers tubes concernaient les tubes de réception à faible puissance. Pour l’émission radio à grande puissance qui jadis était assurée par des alternateurs, on continua d’utiliser des triodes en augmentant progressivement la puissance. L’invention de la soudure verre métal en 1921 permit de réaliser des tubes de plus

en plus puissants qui permettaient d’arriver jusqu’à 500 kilowatts.

La marche vers les ondes courtes

Les pionniers de la radio avaient effectué leurs premières expériences avec des ondes de quelques centimètres mais ensuite la pratique montra l’intérêt d’augmenter la longueur d’onde de plusieurs dizaines à plusieurs centaines de mètres. A partir de 1915, de nombreux émetteurs transocéaniques fonctionnaient avec des ondes très longues jusqu’à 20 km. Cela nécessitait des puissances et des antennes énormes donc très coûteuses avec des rendements médiocres. Cependant, les essais organisés par l’union internationale des radioamateurs dans les années 20 révélèrent les possibilités des ondes courtes impliquant une inflexion vers des longueurs d’onde de plus en plus faibles. Les premières études des systèmes de radiodétection qu’on appellera durant la guerre, radars, se développèrent dans les années 30 avec l’utilisation de tubes spéciaux comme le klystron et le magnétron qui permettaient d’obtenir des puissances d’émission élevées dans les bandes métriques puis, au cours de la guerre, dans les bandes centimétriques pour les radars embarqués. Avec ces derniers, le besoin de miniaturisation des composants électroniques se révéla indispensable pendant et après la guerre. Dans le domaine professionnel, les premiers calculateurs expérimentés durant la guerre, tel l’ENIAC, pesaient 30 tonnes, occupaient 160 m², consommaient 175 kW pour alimenter leurs quelques 18 000 tubes et tombaient souvent en panne ! Tout cela imposait une révision radicale des composants qui arriva avec l’entrée dans le système technique de l’électronique des composants à état solide.

Les semi-conducteurs

Avant le transistor

Les recherches sur la conductibilité des solides commencèrent au XIX^{ème} siècle avec l’étude des propriétés de l’électricité et continuèrent tout au long de ce siècle et du suivant. Les premiers travaux consistaient à décrire les caractéristiques de ce que nous nommons aujourd’hui les semi-conducteurs dont la résistivité, contrairement au cas des métaux, diminuait avec la température. La première observation remontait à Faraday qui découvrit que le sulfate d’argent possédait une résistance décroissante quand la température s’élevait. En 1839, Edmond. Becquerel décrivit l’effet d’une tension générée par une jonction entre un semi-conducteur et un électrolyte lorsque celui-ci est illuminé, ce qui pouvait être considéré comme le premier exemple observé de l’effet photovoltaïque. Vers la fin du siècle, les chercheurs avaient mis en évidence les quatre propriétés fondamentales des semi-conducteurs : résistance à coefficient négatif, photoconductivité, photoélectricité et effet redresseur. Avec les premiers développements de la radio, des chercheurs américains, de 1904 à 1910, établirent que le contact d’une pointe métallique avec un morceau de galène constituait un détecteur, phénomène qui a été utilisé dans les postes à galène. L’invention de la télégraphie sans fil joua un rôle de stimulation et conduisit à une étude approfondie de nombreux minéraux naturels comme détecteurs de rayonnement électromagnétique. Dans les années 20, les redresseurs à état solide commençaient à être utilisés pour transformer le courant alternatif en courant continu et également comme détecteurs radio. Avec la diffusion de la radioélectricité, de plus en plus de radioélectriciens souhaitaient avoir à disposition un dispositif amplificateur sans cathode chaude et d’un meilleur rendement.

Théorie de la conduction

Le développement, dans les années 20, de la théorie quantique fournit les premiers outils théoriques pour la compréhension

de la conductibilité dans les solides. L'étude du mouvement d'un électron dans un cristal montrait que, en raison de la périodicité du réseau cristallin, l'énergie des électrons devait être comprise dans certaines bandes dont les limites correspondaient à des longueurs d'onde déterminées par les zones de Brillouin, bases de la physique moderne des solides. La théorie des bandes, découverte en 1927 et développée par Bloch et Brillouin, donna la clé de la distinction entre isolants et conducteurs, entre les électrons liés et les électrons semi-libres. En 1931, l'idée fondamentale qui émergeait était qu'un corps n'ayant que des bandes complètes était un isolant et au contraire un corps qui possédait une bande incomplètement remplie était un conducteur ; cette bande semi-pleine était sa bande de conduction. Un troisième cas pouvait se présenter : celui des semi-conducteurs qui étaient des isolants à basse température et avaient la particularité d'avoir une bande interdite étroite au-dessus de la dernière bande pleine qui constitue la bande de valence. Ainsi, lorsque la température s'élevait, l'agitation thermique du réseau pouvait faire passer un électron de la bande de valence vers la bande de conduction et cet électron laissait une place libre équivalente à un trou dans la bande de valence du réseau cristallin comme si une particule positive était apparue au point correspondant. La croissance très rapide du nombre de paires électron-trou avec la température déterminait la propriété caractéristique des semi-conducteurs qui était la diminution de leur résistance électrique avec l'élévation de la température. Les semi-conducteurs intrinsèques les plus utilisés étaient des éléments de la quatrième colonne du tableau de Mendeleïev comme le germanium ou le silicium. Il existait également des semi-conducteurs extrinsèques qui devaient leurs propriétés à la présence d'impuretés. En introduisant un élément du cinquième groupe, tel que l'antimoine ou l'indium, dans un cristal de silicium on introduisait un électron supplémentaire, celui-ci faiblement lié pouvait passer dans la bande de conduction ce qui avait pour conséquence de produire une conductibilité par excès d'électrons appelé pour cela de type n et si au contraire on introduisait un élément

du troisième groupe tel que le bore on laissait une place vacante pour un électron. Dès qu'un électron d'un atome de silicium venait s'y fixer il y avait création d'un trou dans la bande de valence du cristal qui possédait alors une conductibilité, par défauts d'électrons, appelé de type p. Les éléments introduits de la sorte étaient appelés des impuretés et l'opération appelée dopage. Dans les deux cas précités, les atomes de l'impureté introduisaient des niveaux supplémentaires situés dans les bandes interdites. Dans les semi-conducteurs de type n, il s'agissait de niveaux donneurs d'électrons qui étaient situés au voisinage de la bande de conduction et, dans les semi-conducteurs de type p, de niveaux accepteurs au voisinage de la bande de valence. Par ces opérations, la conduction des cristaux pouvaient dépendre de la présence d'impuretés. Les opérations de dopage jouaient, dans la technologie des semi-conducteurs, un rôle analogue à celui de la technique du vide pour les tubes électroniques.

L'invention du transistor

Le 1^{er} juillet 1948 le New-York times annonçait dans sa page radio :

« Un dispositif appelé transistor qui peut faire fonctionner un système là où un tube à vide est habituellement employé a été présenté pour la première fois hier au laboratoire Bell où il a été inventé. La démonstration était réalisée à l'aide d'un récepteur radio qui peut effectivement fonctionner sans aucun tube conventionnel. Le dispositif peut de la même façon être intégré à un système téléphonique ou à un poste de télévision ; dans chaque cas le transistor est utilisé comme amplificateur mais on prétend qu'il peut également opérer comme oscillateur afin de générer des ondes radios. Il possède la forme d'un petit cylindre d'environ un centimètre et ne contient ni vide, ni grille, ni revêtement susceptibles de le protéger de l'air ambiant. Son action est instantanée et ne nécessite aucun délai de chauffe et n'entraîne aucun dégagement de chaleur contrairement au tube à vide. Les parties actives du dispositif consistent simplement en deux films métalliques qui s'enfoncent dans un petit morceau de matériau semi-conducteur soudé à une base métallique, la substance se trouvant sur cette base amplifie le courant auquel un des deux fils est soumis l'autre fil restituant le signal amplifié ».

**The Transistor,
A Semi-Conductor Triode**
J. BARDEEN AND W. H. BRATTAIN
Bell Telephone Laboratories, Murray Hill, New Jersey
June 25, 1948

A THREE-ELEMENT electronic device which utilizes a newly discovered principle involving a semi-conductor as the basic element is described. It may be employed as an amplifier, oscillator, and for other purposes for which vacuum tubes are ordinarily used. The device consists of three electrodes placed on a block of germanium¹ as shown schematically in Fig. 1. Two, called the emitter and collector, are of the point-contact rectifier type and are placed in close proximity (separation $\sim .005$ to $.025$ cm) on the upper surface. The third is a large area low resistance contact on the base.

The germanium is prepared in the same way as that used for high back-voltage rectifiers.² In this form it is an N-type or excess semi-conductor with a resistivity of the order of 10 ohm cm. In the original studies, the upper surface was subjected to an additional anodic oxidation in a glycol borate solution³ after it had been ground and etched in the usual way. The oxide is washed off and plays no direct role. It has since been found that other surface treatments are equally effective. Both tungsten and phosphor bronze points have been used. The collector point may be electrically formed by passing large currents in the reverse direction.

Each point, when connected separately with the base electrode, has characteristics similar to those of the high

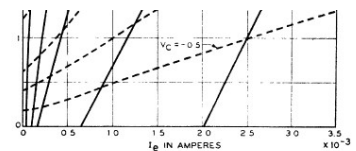


FIG. 2. d.c. characteristics of an experimental semi-conductor triode. The currents and voltages are as indicated in Fig. 1.

back-voltage rectifier. Of critical importance for the operation of the device is the nature of the current in the forward direction. We believe, for reasons discussed in detail in the accompanying letter,⁴ that there is a thin layer next to the surface of P-type (defect) conductivity. As a result, the current in the forward direction with respect to the block is composed in large part of holes, i.e., of carriers of sign opposite to those normally in excess in the body of the block.

When the two point contacts are placed close together on the surface and d.c. bias potentials are applied, there is a mutual influence which makes it possible to use the device to amplify a.c. signals. A circuit by which this may be accomplished is shown in Fig. 1. There is a small forward (positive) bias on the emitter, which causes a current of a few milliamperes to flow into the surface. A reverse (negative) bias is applied to the collector, large enough to make the collector current of the same order or greater than the emitter current. The sign of the collector bias is such as to attract the holes which flow from the emitter so that a large part of the emitter current flows to and enters the collector. While the collector has a high impedance for flow of electrons into the semi-conductor, there is little impediment to the flow of holes into the point. If now the emitter current is varied by a signal voltage, there will be a corresponding variation in collector current. It has been found that the flow of holes from the emitter into the collector may alter the normal current flow from the base to the collector in such a way that the change in collector

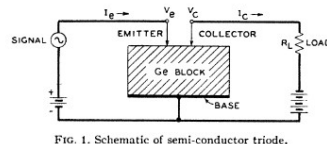


FIG. 1. Schematic of semi-conductor triode.

Figure 2 : Extrait de la Physical Review qui détaille pour la première fois le transistor à pointe.

Patented Oct. 3, 1950

2,524,035

UNITED STATES PATENT OFFICE

2,524,035

THREE-ELECTRODE CIRCUIT ELEMENT UTILIZING SEMICONDUCTIVE MATERIALS

John Bardeen, Sumit, and Walter H. Brattain, Morristown, N. J., assignors to Bell Telephone Laboratories, Incorporated, New York, N. Y., a corporation of New York

Application June 17, 1948, Serial No. 33,466

40 Claims. (CL 179-171)

Figure 3 : Premier brevet du transistor dont la date d'application est voisine de l'article de la Physical Review.

●●● Cette annonce suivait de près de 6 mois la découverte effective du transistor par trois chercheurs des Bell Labs. Comme nous l'avons indiqué plus haut, le transistor s'inscrivait dans un mouvement de recherche sur les semi-conducteurs qui remontait à plusieurs décennies. Avant la première guerre

mondiale, les Bell Labs avaient constitué un groupe de chercheurs pour l'étude du courant dans les solides. William Shockley était l'un des principaux théoriciens de ces recherches quand il intégra les Bell Labs en 1936. Durant la guerre, le laboratoire fut impliqué dans le développement du radar

qui bénéficia des travaux sur le germanium et le silicium pour les détecteurs hyperfréquences. Après la guerre, les recherches théoriques fut reprises avec l'arrivée de John Bardeen. L'étude de la résistance d'un film de semi-conducteur soumis à un champ électrique fut la première tentative de commande d'un courant à l'aide d'un champ pour reproduire un effet d'amplification comme dans une triode. L'effet fut observé mais ne correspondait pas à la théorie. J. Bardeen élabora une nouvelle théorie du comportement des électrons et des trous qui étaient sur la surface du solide, théorie que Walter Brattain expérimenta en explorant une zone voisine d'un contact ponctuel métal semi-conducteur avec une deuxième pointe métallique. Il avait ainsi découvert expérimentalement la structure du transistor à pointe. L'amplification n'utilisait pas un champ électrique mais l'effet transistor était décrit comme l'interaction entre deux jonctions voisines. Cette découverte intervint en décembre 1947, mais elle fut présentée au monde scientifique, en juin 1948, par diverses annonces dont un article dans la Physical Review (figure 2) et un brevet (figure 3). Le transistor à pointe constitué par du germanium était encore rudimentaire et fragile ; il ne connut pas de succès commercial et allait être supplanté par le transistor à jonction.

Une nouvelle industrie

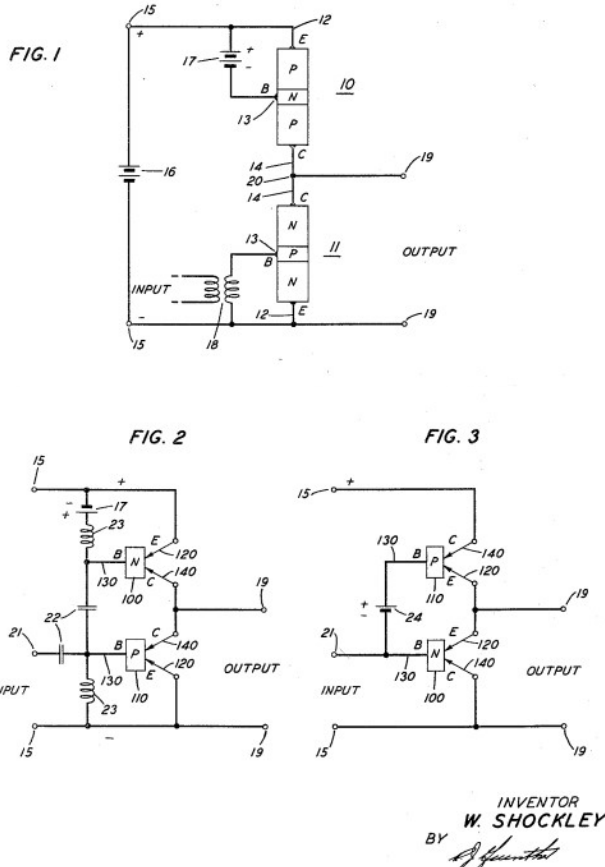
En juillet 1949, W. Shockley, dans un article dans lequel il exposait la théorie de la jonction n-p proposait un autre type de transistor formé de deux jonctions très rapprochées polarisées en sens contraire dans un même cristal de semi-conducteur. Il avait inventé le transistor à jonction qui marquait le début de l'électronique moderne et un brevet était déposé en 1951 (figure 4).

Très vite, les productions de masse du transistor à jonction commencèrent dès 1951 grâce à des commandes militaires du DoD à des grands groupes des radiocommunication comme Western Electric, General Electric et Raytheon. L'armée américaine fut l'acteur majeur de la diffusion du transistor pour remplacer les tubes à vide. En octobre 1952, le département de la défense commanda 5000 unités par semaine par contractant à

Jan. 19, 1954

W. SHOCKLEY
TRANSISTOR AMPLIFIER
Filed Sept. 13, 1951

2,666,818



INVENTOR
W. SHOCKLEY
BY [Signature]

Figure 4 : Premier brevet du transistor, déposé en 1951, à une date voisine de la parution de l'article dans Physical Review.



Figure 5 : John Bardeen
(Source : <https://losmundosdebrana.com>).



Figure 6 : Walter Brattain
(Source : Wikipedia).

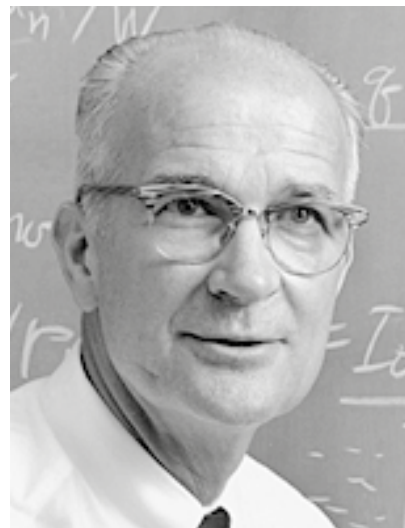


Figure 7 : William Shockley
(Source : <https://perspective.usherbrooke.ca>).

Raytheon, General Electric, RCA et Sylvania. L'industrie des tubes devait alors se convertir progressivement au transistor et, comme souvent, ce sont les commandes militaires qui permirent de financer le passage des tubes aux transistors dans le domaine civil. En quelques années, les postes à transistors devaient peu à peu envahir le marché de l'électronique. En 1955, quand Bertrand Daugny tenta de persuader Marcel Dassault de financer un département d'électronique, il lui démontra l'avenir de l'électronique en construisant un poste de radio miniature à transistors appelé Mirage par esprit de famille. Le département électronique devait devenir Dassault Electronique et le transistor devait devenir le nom des postes radio grand public. Le premier grand succès du département électronique des avions Dassault était un autodirecteur de missile complètement transistorisé et développé à la fin des années 60.

Epilogue

Les tubes feront encore de la résistance, en particulier dans les émetteurs haute et hyperfréquence pour la radiodiffusion et les radars mais, à l'orée des années 2000, les klystrons et magnétrons tendaient à disparaître sauf dans les fours domestiques. De manière équivalente au passage dans le système de l'éclairage du gaz à l'électricité, l'électronique est passée des tubes à vide au transistor en l'espace de quelques an-

nées. La conséquence la plus visible était la miniaturisation de plus en plus poussée des circuits électroniques ; le transistor sera presque dématérialisé en tant que composant en étant intégré dans des circuits imprimés dans les décennies postérieures à 1960. Mais ses fonctions, soit comme amplificateur, soit comme commutateur sont restées les mêmes. La découverte du transistor peut donc, au même titre que l'invention de la lampe d'Edison ou le tube de Lee de Forest, être considérée comme une révolution technique à part entière qui a donc 75 ans et qui a forgé le monde technique dans lequel nous vivons aujourd'hui. L'importance de cette découverte fut d'ailleurs concrétisée par le prix Nobel de physique attribué en 1956 à W. Shockley, J. Bardeen et W. Brattain (figures 5, 6 et 7) pour leurs recherches sur les semi-conducteurs et l'effet transistor. ■

Références

- DAUMAS, Maurice. L'histoire générale des techniques. 1960. Collection quadriga
- TATON, R. XX^e siècle, t. V, vol. de l'histoire générale des sciences. 1961. PUF
- MIZRAHI, Jean. L'échiquier de l'électronique : une géopolitique des technologies de l'informatique. Fenix, 1986.
- MORRIS, Peter Robin. A history of the world semiconductor industry. IET, 1990.

L'auteur

Marc Leconte est ancien secrétaire du club RSSR de la SEE (radars, sonars et systèmes radioélectriques), membre du



comité de rédaction de la REE, membre émérite SEE et médaillé Ampère. Au sein de Dassault Electronique, il a passé une quinzaine d'années à l'étude, au développement et aux essais en vol du radar RDI du Mirage 2000. Ensuite pendant trois ans, il a participé à l'étude d'un démonstrateur laser franco-britannique CLARA. A partir de 1995, il a élargi son activité aux domaines des études concurrentielles et stratégiques dans les domaines des radars aéroportés et de la guerre électronique. Il a exercé les mêmes activités dans la division aéronautique de Thales après la fusion de Dassault Electronique et de Thomson-CSF. A partir des années 90 et en parallèle, il s'est intéressé à l'histoire des sciences et des techniques et a publié plusieurs articles s'y rapportant.