Sensors and Materials, Vol. 13, No. 1 (2001) 001–012
MYU Tokyo

S & M 0425

The Diode Hall Effect and Its Sensor Applications - An Overview

Chavdar Roumenin*, Konstantin Dimitrov, Dimitar Nikolov and Avgust Ivanov

Institute of Control and System Research, Bulgarian Academy of Sciences, 1113 Sofia, P.O. Box 79, Bulgaria

(Received February 14, 2000; accepted December 22, 2000)

Key words: diode Hall effect, Hall voltage, Hall current mode, multisensors for temperature and magnetic fields, multidimensional sensing

This first comprehensive review of the diode Hall effect – the Hall phenomenon in devices with bipolar nonequilibrium conductivity (diodes and transistors) – considers the special features of this unusual fundamental peculiarity, which is regarded as uncommon for this class of structure. Depending on the injection level and the geometry, the Hall effect (voltage or current) exists and has been experimentally confirmed. The conditions for the appearance of the diode Hall effect are presented in detail along with the applicability of the Hall current mode of operation for metrology. The most attractive applications of the diode Hall effect are in the creation of functional multisensors for magnetic fields and temperature, and for low-field and vector magnetometers.

1. Introduction

The development of microsensors for magnetic fields involves not only the design of new devices, optimization of operation modes, application of sophisticated technological steps, and 2D and 3D simulations of their performance. Rather, the most innovative breakthroughs have been made where the magnetosensitive device implements a newly discovered effect or phenomenon, which dramatically improves its characteristics and behavior. Such is the case with the quantum Hall effect, which is the basis for the function of the most precise resistor standard, and whose application in the metrology of electrical quantities is extremely sensitive.

*Corresponding author, e-mail address: roumenin@iusi.bas.bg

When the diode Hall effect was first observed experimentally,[1,2] it was in essence the unanticipated answer to the question of whether the Hall voltage V_H (the dominant manifestation of the classical Hall effect) is generated in diode and transistor structures. There is a paradox in the point of view held for decades in that if the conductivity in a semiconductor is bipolar (equilibrium or nonequilibrium) with approximate equality of electrons and holes, then according to Hall's theory, the voltage V_H is expected to be absent. The reason for this is the simultaneous Lorentz deflection of both types of carriers to the same surface.[3] Analysis of the Hall effect in different types of diodes and transistors, however, has established that this assumption is not valid in the general case. Depending on the injection level and the geometry, the voltage V_H exists and can reach up to 15–25 mV; this has been experimentally confirmed.[1] Initially, this new fundamental property was used to elucidate the synergic action of multiple sensor mechanisms in magnetotransistors.[4–6] Recently, the diode Hall effect has found a very promising application in different silicon microdevices for measuring magnetic fields with amperometric output – functional multisensors for simultaneous and independent registration of magnetic field and temperature, and 2D and 3D magnetometers.

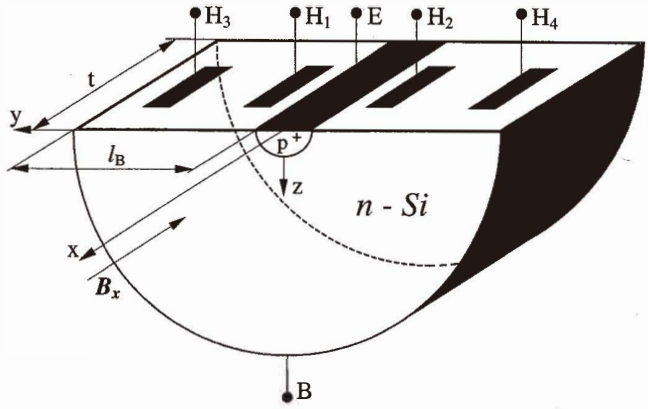
In this overview, the results of the investigations on the diode Hall effect are summarized and described for the first time, and new transducers based on this effect are presented.

2. Conditions for Appearance of the Diode Hall Effect

The main concept of the Hall effect in diodes and transistors can be illustrated with the semicircular p^+-n abrupt diode shown in Fig. 1(a). The selected geometry is not of particular significance because the result can be generalized for the rectangular structure shown in Fig. 1(b) by conformal mapping theory.[7] Because of the relatively high doping level, any voltage can be considered negligible in the p^+ - emitter region, $p^+ \gg n_0 \approx N_D$, where n_0 represents the equilibrium concentration of the majority carriers in the low-doped base region. The presence of low and moderate injection levels $\Delta p / n_0 < 1$ is also assumed.

A typical feature of the device in Fig. 1(a) is that the base width l_B exceeds the diffusion length of the minority carriers L_p (in our case the holes), $l_B \geq L_p$, i.e., the structure represents a long diode. This means that during the diffusion of the minority carriers in the directions $y > 0$ and $z > 0$, they completely recombine in the volume before reaching the base contact B, $y = l_B$, i.e., $\Delta p|y = l_B = 0$. The distribution of the injecting holes $\Delta p(y)$ and the corresponding current I_p decrease exponentially from the distance y, while the total diode current I_{EB} remains constant, $I_{EB} = I_p(y) + I_n(y) = \text{const}$. Therefore, with an increase in the distance $y > 0$, the current from the majority carriers I_n increases far from the emitter $I_{EB} \approx I_n$. Essentially, the component $I_n \sim v_d \cdot q \cdot n_0$ is the recombination current from the base contact B to neutralize the charge of the injected holes. In the presence of the magnetic field $\boldsymbol{B}$, the current I_n generates the Hall voltage V_H in the diode. The voltages $V_{H1,2}$ between the respective contacts H_1 and H_2, and $V_{H3,4}$ between H_3 and H_4 in Fig. 1(a) are defined by the expressions:

(a)



(b)

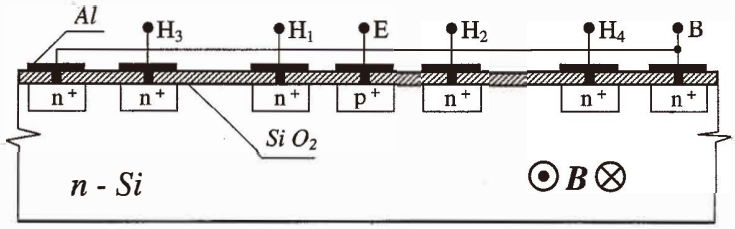


Fig. 1. (a). Semicircular p^+-n abrupt diode to illustrate the diode Hall effect. (b). Rectangular geometry of a long p^+-n structure for observation of the diode Hall effect.

$$V_{H1,2} = \int_{H1}^{H2} E_H dl; V_{H3,4} = \int_{H3}^{H4} E_H dl, \tag{1}$$

respectively, where dl is an infinitely small segment of the respective distances $\overline{H_1H_2}$ and $\overline{H_3H_4}$, $E_H = R_H(y) \cdot j \times \boldsymbol{B}$ is the local Hall field, j is the total current density $j_{EB} = j_p(y) + j_n(y)$, and $R_H(y)$ is the Hall coefficient. Since the current j is perpendicular to dl and to $\boldsymbol{B}$, we can write:

$$V_{H1,2} \cong R_H(y') \cdot B \cdot l / t \cdot \int_{H1}^{H2} j \cdot t \cdot dl; V_{H3,4} \cong R_H(y'') \cdot B \cdot l / t \cdot \int_{H3}^{H4} j \cdot t \cdot dl, \tag{2}$$

where $y' < y''$ (y' and y'' are variable distances on y axis), or

$$V_{H1,2} \cong R'_H \cdot B \cdot I_{EB} / t; V_{H3,4} \cong R''_H \cdot B \cdot I_{EB} / t, \quad (3)$$

in which t is the thickness of the diode from Fig. 1(a) in the x-direction of the field B_x.

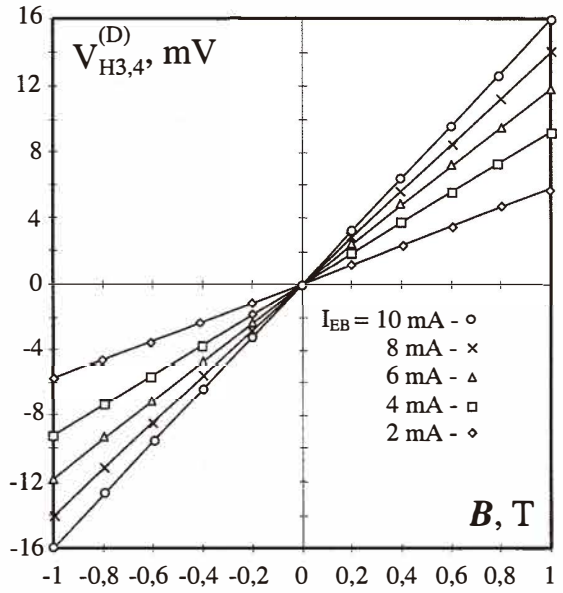
The concentration of the minority carriers Δp is greater near the emitter E, i.e., $\Delta p(y') > \Delta p(y'')$, hence the corresponding Hall coefficient $R'_H < R''_H$ and consequently, the inequality $V_{H1,2} < V_{H3,4}$ is valid. Clearly the Hall voltage V_H is minimal next to the depletion layer of the p⁺- n junction. The short-circuiting effects of the p⁺ contact on V_H are virtually eliminated by the relatively high resistance of the depletion layer of the p⁺ - n junction at low and moderate injection levels.[1] In summary, the diode Hall voltage at a maximal value far from the emitter E, $V_H \cong I_n \cdot B / q \cdot n_0 \cdot t$, is bound to decrease as we approach the p⁺- n junction and the base ohmic contact B. An increase in the total current I_{EB}, with respect to the injection level, leads to saturation of the voltage V_H.

The structure in Fig. 1(a) was tested experimentally with a p⁺- n diode, the cross-section of which is shown in Fig. 1(b) (to elucidate our conclusions). The action of the silicon device in Fig. 1(b) is practically the same as that in Fig. 1(a).

The dependencies $V^{(D)}_{H3,4}(\boldsymbol{B})$ and $V^{(D)}_H(I_{EB})$ are shown in Fig. 2. In addition to the linear dependence of $V^{(D)}_H(\boldsymbol{B})$ in Fig. 2(a), the values of the voltage $V^{(D)}_{H1,2}$ are lower than those of $V^{(D)}_{H3,4}$. With an increase in the injection level, the growth rate of both voltages decreases, as shown in Fig. 2(b). It has been established that when the Lorentz force deflects the majority carriers, which are on the right-hand side of the emitter E, toward the Si - SiO_2 interface, the voltages $V^{(D)}_{EH2}$ and $V^{(D)}_{EH4}$ increase, while the corresponding voltages on the left-hand side of the emitter, $V^{(D)}_{EH1}$ and $V^{(D)}_{EH3}$, decrease (after compensation for the magnetoresistance MR ~ B^2). The magnetodiode effect is not observed based on the dependence $I_{EB}(V_{EB})$ in the structure shown in Fig. 1(b). The diode Hall effect is also manifested in transistor structures similar to the diode shown in Fig. 1(b) in which the Hall terminals H_1 and H_2 are replaced with p⁺- collectors. At the same time, the voltage $V^{(D)}_{H3,4}$ at current I_{EB} = 10 mA and B = 1T reaches 20 mV.[2]

The experimental data confirm the conclusion that the Hall voltage is generated by the majority carriers in the base region of the diode and transistor structures. Both the higher values of $V^{(D)}_{H3,4}(\boldsymbol{B})$ in comparison with $V^{(D)}_{H1,2}(\boldsymbol{B})$ and the reduced growth rate of the voltage $V^{(D)}_H$ with an increase of the current I_{EB} (as shown in Fig. 2) are primarily due to the reduction of the Hall coefficient $R_H(\Delta p)$ because of the high minority carrier concentration Δp.[8,9]

(a)



(b)

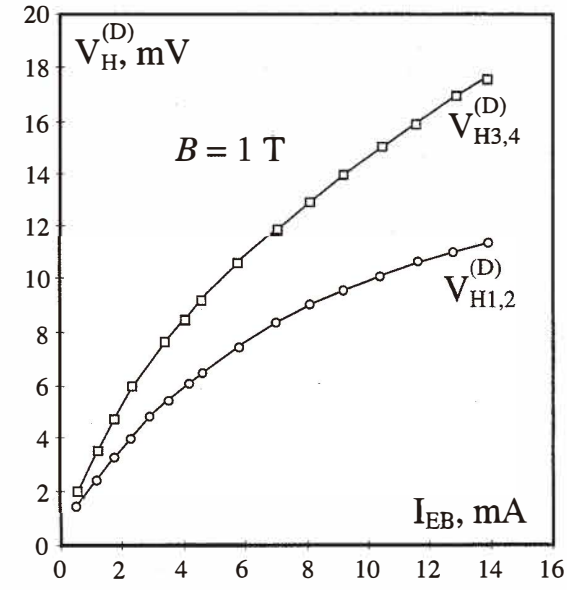


Fig. 2. (a). Dependence of the diode Hall voltages $V^{(D)}_{H3,4}(\boldsymbol{B})$ as a function of the magnetic field with the current I_{EB} as a variable. (b). Dependence of the diode Hall voltages $V^{(D)}_{H1,2}$ and $V^{(D)}_{H3,4}$ as a function of the supply current I_{EB}.

3. Features of the Diode Hall Effect

The principle features of the diode Hall effect are related to the bipolar conductivity substantially reducing the voltage $V_H^{(D)}$, which has been confirmed experimentally. On the other hand, however the equivalent alternative to the Hall voltage – the Hall current I_H to a first approximation – must be commensurate with the I_H of the classical Hall element in the current mode of operation. The main reason for this is that the current, because of its circulatory nature, flows first through the internal resistance of the Hall plate which is reduced by the injection (the diode Hall effect) and then through another internal resistance which is several times higher (the classical Hall effect). The current $I_H^{(D)}$ is determined by the relation:

$$I_H^{(D)}(\boldsymbol{B}) = V_H^{(D)}(\boldsymbol{B})/(R^* + r), \tag{4}$$

where R* is the distributed bulk resistance of the base in the region between the Hall contacts, and r is the resistance of the measurement resistor, $r \ll R^*$.[10,11]

The magnitude of the Hall current $I_H^{(D)}$ is determined simultaneously by the values of the diode Hall voltage $V_H^{(D)}$, and the resistance R*. The analysis shows that for a given p^+ - n device structure with an increase of the injection (current I_{EB}), in first approximation at $l_B > L_p$, the resistance R* decreases, $R^* \sim l/I_{EB}$.[12] As a result, at the moderate injection level, the reduction of the effective resistance R* substantially exceeds the increase in the voltage $V_H^{(D)}(I_{EB})$. The amperometric (current) output of the bipolar sensor based on the diode Hall effect is of practical value for metrological applications.

The current $I_H^{(D)}$ is detected through a voltage drop V_r on a low-ohm resistor r placed between the Hall terminals H_1 and H_2. Note that the output resistance R_{out} of a silicon diode Hall sensor in amperometric mode is about one degree of magnitude lower than that of the classical Hall element with current output and that R_{out} reaches 100–200 Ω. Therefore the resistor r has a value of $r = 10$ Ω, which is preferable.[10,11] State-of-the-art measuring instruments are capable of detecting the relatively low values of the voltage V_r with high accuracy.

The low output resistance R_{out} of the amperometric diode Hall sensor contributes to the increase in noise immunity in transmitting the signal V_r at a distance. The internal noise of the sensor itself is reduced in comparison with that of the classical Hall element, and the signal-to-noise ratio is improved.

4. Sensor Applications of the Diode Hall Effect

4.1 *Linear multisensors for magnetic field and temperature*

There are several device structures for parallel-field Hall microsensors, as shown in Figs. 3(a) to 3(e). The performance of these novel elements reflects the classical Hall effect, as explained in refs. 8 and 9. By short-circuiting the Hall voltage $V_H(\boldsymbol{B})$, the information signal for the field $\boldsymbol{B}$ becomes the Hall current $I_{out}(\boldsymbol{B}) \equiv I_{H1,2}(\boldsymbol{B})$ (Hall current mode of operation).

a) The first question that arises is whether sensors utilizing the diode Hall effect can be used in the parallel-field Hall microdevices as shown in Figs. 3(a) to 3(e). We consider that they can be, and the new type of transducer is shown in Figs. 3(a′) to 3(e′). The transport process to the left and to the right of the emitter E, for example in Fig. 3(d′), turns from diffusion-drift into drift. Hence, far from the emitter E, a parallel-field Hall effect is generated in the magnetic field $\boldsymbol{B}_x$ on the upper surface of the chip. This is related to the base current from the majority carriers, in our case, the electrons.

The direction of the supply current I_{EB} in the microsensors from Figs. 3(a′) and 3(e′) clearly cannot be reversed in the same way as in the Hall elements from Figs. 3(a) to 3(e). Another peculiarity of these devices is that the output Hall terminals H_1 and H_2 must be placed near the base contacts B_i, where the base current from the majority carriers is dominant and the Hall effect generated by it on the chip surface containing the contacts is maximized,[1,2,8,9] for example transducers (b′) and (e′). In some of the new sensors, the adjustment of the current offset to zero is easily achieved by the trimmer r*, as in Figs. 3(a′) and 3(d′).

b) The second basic question is whether the diodes from Figs. 3(a′) to 3(e′) possess the linear thermometric characteristics of the emitter-base voltage $V_{EB}(T)$ at constant supply current I_{EB}. We consider that they do. The linear dependence of the junction forward voltage drop $V_{EB}(T)$ of a diode at a constant current (I_{EB} = const) due to the temperature dependence of the semiconductor band gap is well known.[3,9,13] In this case the voltage in the spatial charge region of the emitter p^+ - n junction decreases linearly as the parameter T increases:

$$V_{p-n} = \varphi_0 - \left[\left(k_B T / q\right)\ln\left(I_{s\infty} / I_{EB}\right)\right] = \varphi_0 - \left[T / 11600 \ln\left(I_{S\infty} / I_{EB}\right)\right], \tag{5}$$

where $\varphi_0 = \Delta E_g / q$ is the band gap in Si, $I_{S\infty}$ is the saturation current at $T \rightarrow \infty$, $I_{EB} \ll I_{S\infty}$ and k_B is the Boltzmann constant. The other notations are conventional. According to eq. (5), the coefficient of thermosensitivity of the p-n junction $K_{p-n} \equiv \partial V_{p-n} / \partial T$ is negative, i.e., $K_{p-n} = -(k_B / q)\ln(I_{S\infty} / I_{EB})$. The voltage drop V_B for the resistance R_B of the base region can be determined by the dependence $V_B = I_{EB} \cdot R_B$, as $R_B \sim \rho \sim 1/q \cdot n_0 \cdot \mu_n$, where the notations are conventional. In fact, the resistance R_B increases linearly with temperature T because of decreasing carrier mobility $\mu_n(T)$ and hence the temperature coefficient K_B is positive.[3,9,12] Therefore, by adding the two voltages V_{p-n} and V_B in series, the temperature signal is reduced, i.e., the resultant coefficient $K_{EB} \equiv \partial V_{EB} / \partial T$ depends on the distance $l_{EB1} = l_{EB2}$ in

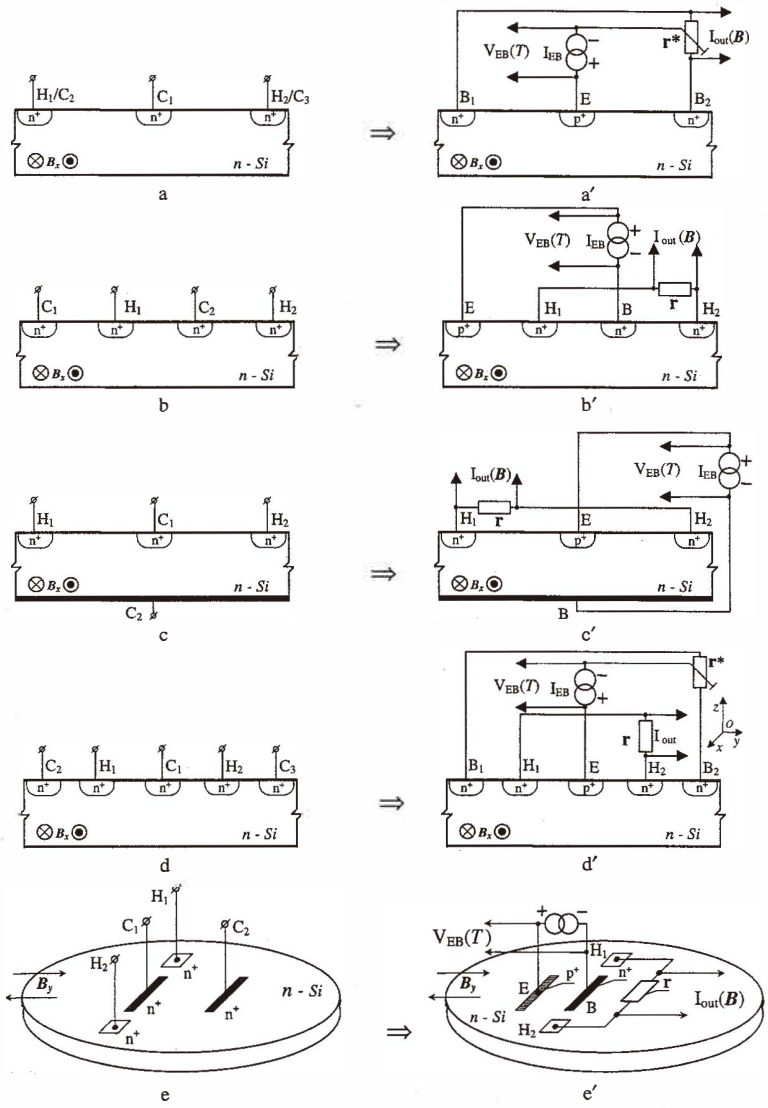


Fig. 3. Classification of the new silicon multisensors (a′) to (e′) based on the diode Hall effect and the related parallel-field Hall microdevices (a) to (e). The trimmer r* is for offset regulation of the amperometric output, H_1 and H_2 represent Hall terminals, and C_1 and C_2 represent current contacts.

example Fig. 3(d′), and current I_{EB} is smaller than the standard value $K_{p\text{-}n} \approx -2$ mV/°C for the silicon p - n junction. This is why, in our case, the thermometric sensor output represents the linear voltage $V_{EB}(T)$.

Consequently the superposition of the above two phenomena in the same zone or the functional integration in the same silicon structure of a parallel-field Hall microsensor and a p^+ - n thermodiode creates a new type of linear multisensor for magnetic fields and temperature. In these multisensors, the magnetosensitive output $I_{out}(\boldsymbol{B})$ is amperometric, and the thermometric one is the voltage $V_{EB}(T)$. For illustration, in Figs. 4(a) and 4(b) the output characteristics of the multisensor shown in Fig. 3(d′) are given. The characteristics $I_{out}(\boldsymbol{B})$ are odd and linear. The nonlinearity (NL) is about 1% in the range $-1.2\text{T} \le B \le 1.2\text{T}$. We consider that the absolute magnetosensensitivity $S_A^{(I)} \equiv I_{out}(B) / \Delta B$ of the new multisensors is completely sufficient for a wide range of applications and is comparable to the parallel-field Hall sensors with current output. The new devices are capable of independently detecting two nonelectrical measurands simultaneously – a feature which is absent in the well-known and now classical Hall elements and p^+ - n thermometers.

It is preferable to use the p^+- n diode structures instead of the n^+- p ones because of the threefold greater mobility of the electrons μ_n with respect to μ_p of the holes[9,12] (the diode Hall effect is generated by the majority carriers in the base, i.e., by the electrons).

The metrological qualities of the suggested sensors are closely related to the cross-sensitivity effects, mainly due to the influence of the magnetic field $\boldsymbol{B}$ on the thermometric signal $V_{EB}(T)$. The magnetoresistance, magnetoconcentration, magnetodiode effect, and emitter - injection modulation have a disturbing influence on the voltage $V_{EB}(T)$.[4,9] It has been experimentally established that in optimized device structures the negative influence of these galvanomagnetic phenomena at $B < 1$T and $T = 300$ K on the voltage $V_{EB}(T)$ does not exceed 1%.[10,11]

The magnetosensitive signal quality may be considerably improved in the new multisensors. In practice, the magnetic field $\boldsymbol{B}$ does not influence the temperature output $V_{EB}(T)$, but the output $I_{out}(\boldsymbol{B})$ is thermo-dependent with a temperature coefficient T.C. = –0.4%/°C. Since the signal $V_{EB}(T)$ adequately reflects the temperature of the chip, it contributes to eliminating the temperature dependence of magnetosensetivity in a wide range of ΔT. For this purpose the voltage $V_{EB}(T)$ is utilized to control the gain coefficient of the operational amplifier for the signal $V_r(\boldsymbol{B})$.[9,14]

4.2 *Low-field and vector magnetometry*

For the purpose of low-field magnetometry it is of interest to enhance the magnetosensetivity of the microsensors shown in Figs. 3(a′) to 3(e′) by operating in a cryogenic environment that increases the mobility of the carriers.[11,12] For this reason, the devices have been investigated experimentally at the boiling point of liquid nitrogen $T = 77$ K. It is established that in the linear interval $-0.2\text{ T} \le \text{B} \le 0.2\text{ T}$ at $T = 77$ K the sensitivity of the current output increases to more than fourfold that at $T = 300$ K. The repeated cycles of cooling to $T = 77$ K do not lead to the degradation of the sensor characteristics for any of the transducers. Therefore the cryogenic environment extends the applicability of the new devices.

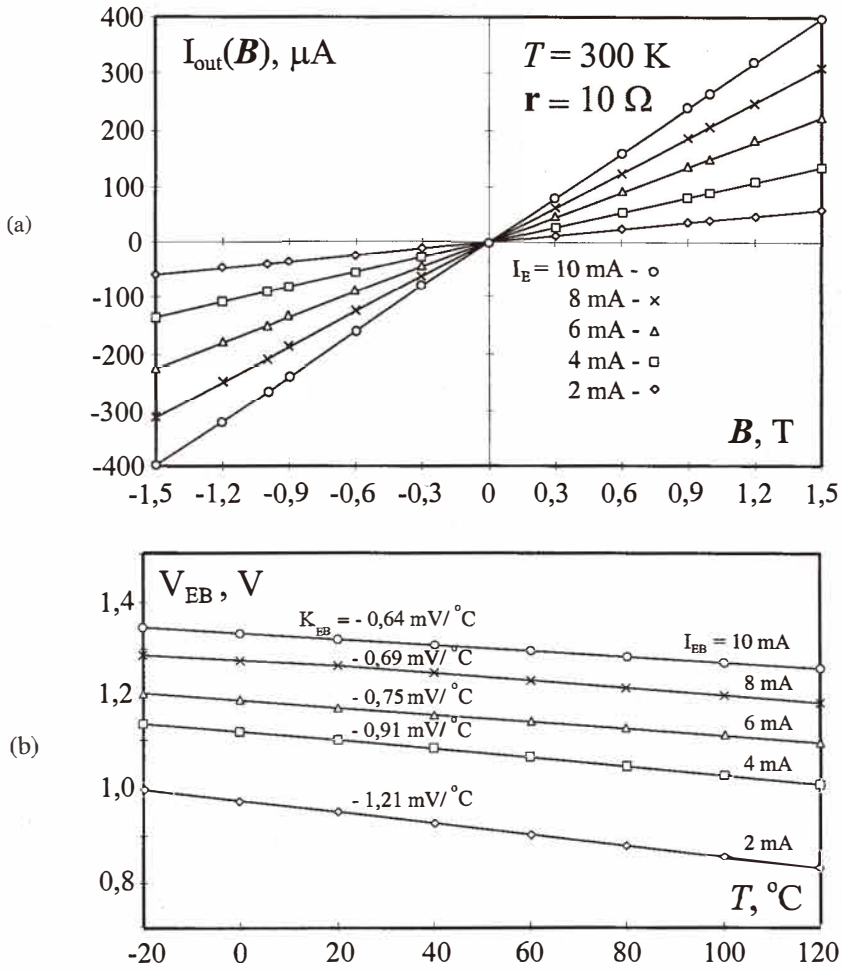


Fig. 4. (a). Output magnetosensitive characteristics at $T = 300$ K for the multisensor based on the diode Hall effect with supply current I_{EB} as a variable. (b) Temperature dependence of the bias V_{EB} with current I_{EB} as a variable at a field $\boldsymbol{B} = 0$.

On the basis of the symmetry of some of the parallel-field device structures of the microsensors in Figs. 3(a′), 3(c′) and 3(d′) with respect to the emitter p^+- n junction it is possible to design 2D vector magnetometers with linear output for simultaneous detection of the in-plane $\boldsymbol{B}_x$ and $\boldsymbol{B}_y$ components as in Fig. 5.

These devices consist of two identical elements with a common emitter E and are oriented with respect to one another at 90°. Their action is related to the nonlinear

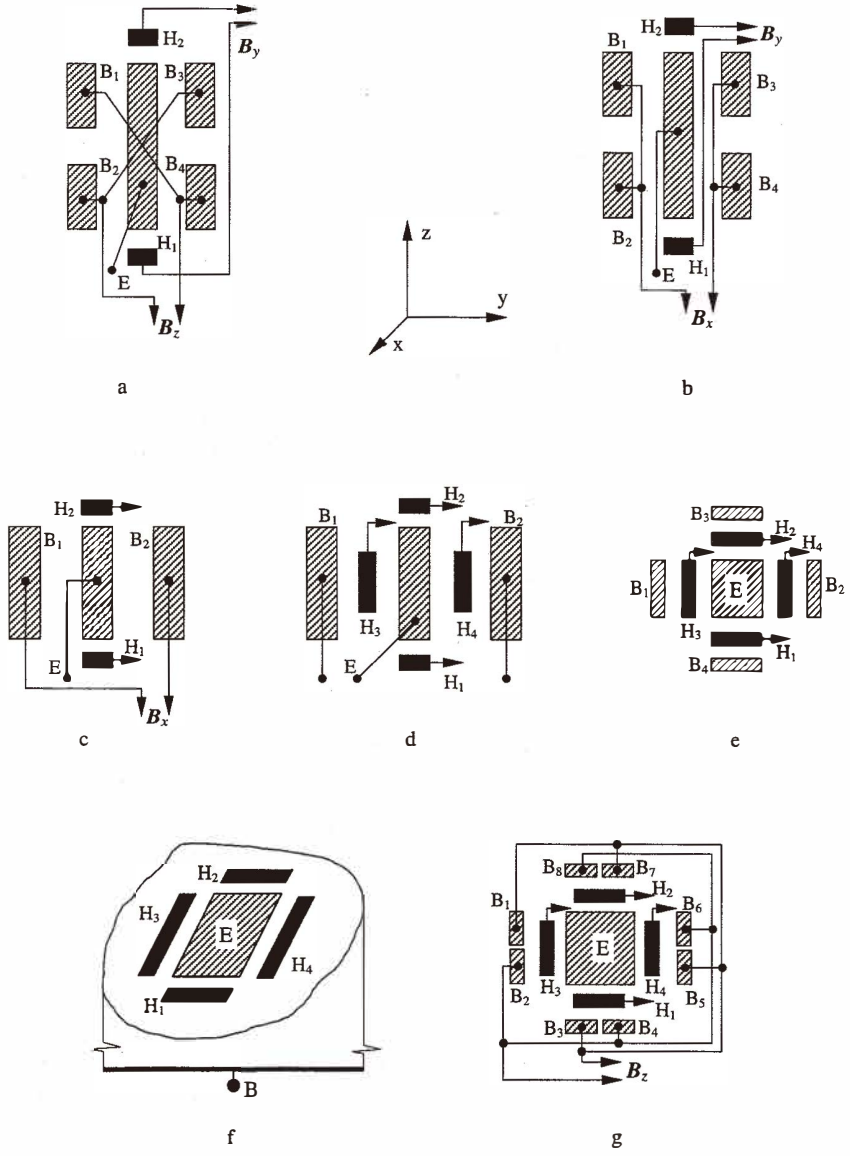


Fig. 5. 2D and 3D vector magnetometers based on the diode Hall effect. The Hall terminals H_1 and H_2, and H_3 and H_4 measure, respectively, the $\boldsymbol{B}_y$ and $\boldsymbol{B}_x$ components of the magnetic field $\boldsymbol{B}$.

trajectory of the carriers (the existence of two components of the velocity v_x and v_y) and is analogous to vector magnetometers using the classical Hall effect, as explained in refs. 7, 9 and 15. An integrated 3D vector magnetometer for measuring the three components $\boldsymbol{B}_x$, $\boldsymbol{B}_y$ and $\boldsymbol{B}_z$ is shown in Fig. 5(g). Its distinctive feature is the splitting of the base contacts into two equal parts and their cross-coupling. This solution allows the detection of the $\boldsymbol{B}_z$ component. For the other two components, $\boldsymbol{B}_x$ and $\boldsymbol{B}_y$, the situation is as in the 2D magnetometers. The cross-sensitivity in the described magnetometers based on the diode Hall effect is generally overcome through the device symmetry. All are fabricated by either bipolar or CMOS technologies. The new vector magnetometers also measure the temperature T of the environment by the voltage $V_{EB}(T)$.[10,11]

5. Conclusions

The diode Hall effect in devices with bipolar conductivity has evolved from a purely novel to a practical phenomenon which has contributed to the understanding of the action of microsensors and has already found practical application in multiple microtransducers. In the authors' view, some galvanomagnetic characteristics that to date have been difficult to explain in diodes and transistors can be consistently interpreted with this effect. It is an impressive fact that almost all technical solutions based on the classical Hall effect can be successfully used in bipolar sensors – analogs with amperometric output. Consequently, the diode Hall effect is a strategic technology for the design of functionally integrated microsensors for more than one measurand.

References

1. Ch. S. Roumenin: Compt. Rendus ABS **38** (1985) 1501.
2. Ch. S. Roumenin: Compt. Rendus ABS **38** (1985) 579.
3. S. Middelhoek and S. A. Audet: Silicon Sensors (Acad. Press, London, 1989) Chap. 5.
4. A. W. Vinal and N. Masnari: IEEE Trans. Electron Devices ED–**31** (1984) 1486.
5. A. Nathan, K. Maenaka, W. Allegretto, H. P. Baltes and T. Nakamura: IEEE Trans. Electron Devices ED–**36** (1989) 108.
6. Ch. S. Roumenin: Sensors and Actuators A **24** (1990) 83.
7. R. F. Wick: J. Appl. Phys. **25** (1954) 741.
8. Ch. S. Roumenin: Sensors and Actuators A **30** (1992) 77.
9. Ch. S. Roumenin: Solid State Magnetic Sensors (Elsevier, Amsterdam, 1994) Chap. 4.
10. Ch. S. Roumenin, A. Ivanov and P. Nikolova: Sensors and Actuators A **69** (1998) 16.
11. Ch. S. Roumenin, A. Ivanov and P. Nikolova: Sensors and Actuators **72** (1999) 27.
12. S. M. Sze: Physics of Semiconductor Devices (J. Wiley & Sons, New York, 1981).
13. G. C. M. Meijer: Sensors and Actuators A **10** (1986) 103.
14. K. Maenaka and M. Maeda: Sensors and Materials **5** (1994) 265.
15. Lj. Ristic and M. Paranjape: Sensors and Materials **5** (1994) 301.