SOLID STATE CHEMICAL SENSORS

Edited by JIŘÍ JANATA

ROBERT J. HUBER

SOLID STATE CHEMICAL SENSORS

Edited by JIŘÍ JANATA

Department of Bioengineering
The University of Utah
Salt Lake City, Utah

ROBERT J. HUBER

Department of Electrical Engineering
The University of Utah
Salt Lake City, Utah

1985

ACADEMIC PRESS, INC.
(Harcourt Brace Jovanovich, Publishers)
Orlando San Diego New York London
Toronto Montreal Sydney Tokyo

ACADEMIC PRESS, INC.
Orlando, Florida 32887

United Kingdom Edition published by
ACADEMIC PRESS INC. (LONDON) LTD.
24–28 Oval Road, London NW1 7DX

Library of Congress Cataloging in Publication Data
Main entry under title:

Solid state chemical sensors.

Includes index.
1. Chemical detectors. 2. Solid state chemistry.
I. Janata, Jiří. II. Huber, Robert J.
TP159.C46S65 1985 681'.2 84-20374
ISBN 0-12-380210-5 (alk. paper)

PRINTED IN THE UNITED STATES OF AMERICA

85 86 87 88 9 8 7 6 5 4 3 2 1

Contributors

Robert J. Huber
Jiří Janata
Ingemar Lundström
Christer Svensson
Jay N. Zemel

List of Contributors

Numbers in parentheses indicate the pages on which the authors' contributions begin.

ROBERT J. HUBER (119), Department of Electrical Engineering, The University of Utah, Salt Lake City, Utah 84112

JIŘÍ JANATA (65), Department of Bioengineering, The University of Utah, Salt Lake City, Utah 84112

INGEMAR LUNDSTRÖM (1), Laboratory of Applied Physics, Department of Physics and Measurement Technology, Linköping Institute of Technology, S-581 83 Linköping, Sweden

CHRISTER SVENSSON (1), Laboratory of Applied Physics, Department of Physics and Measurement Technology, Linköping Institute of Technology, S-581 83 Linköping, Sweden

JAY N. ZEMEL (163), Center for Chemical Electronics, Department of Electrical Engineering, University of Pennsylvania, Philadelphia, Pennsylvania 19104

Preface

Detection of chemical species using solid state circuitry is a relatively new field that is generating a great deal of interest, both in academia and in industry. Several conference proceedings have been published on this subject, and review articles have been written that deal with particular subsets of solid state sensors. This volume is our attempt to review the basic chemical and physical principles—and problems—involved in the construction and operation of some of these devices.

A major portion of the book is devoted to explanation of the basic mechanism of operation and the many actual and potential applications of field effect transistors for gas and solution sensing. A chapter describing the basics of device fabrication is included so that the nonspecialist reader may gain an appreciation of the complexity of semiconductor fabrication methods. The chapter on piezoelectric and pyroelectric chemical sensors outlines early work in the development of new techniques of chemical detection. Chemical sensing covers a vast territory, and it was necessary to omit many areas of important research, e.g., high-temperature surface conductivity sensors, chemiresistors, etc. Although this volume is not intended to be used as a textbook, some of the material is suitable for inclusion in graduate courses.

This emerging technology is dependent on the research efforts of two groups of scientists, electrical engineers and chemists, because solid state chemical sensors are hybrid devices that employ the principles of both these fields. There are many similarities in the laws that describe the seemingly quite different phenomena in these two disciplines. Our major goal in this book, then, is to demonstrate this coincidence in the expressions of these phenomena and to use it to assist electrical engineers in understanding the chemistry involved and to educate chemists in solid state science. We hope that our efforts will help accelerate progress in the exciting new field of solid state chemical sensors.

Contents

1

Gas-Sensitive Metal Gate Semiconductor Devices

INGEMAR LUNDSTRÖM AND
CHRISTER SVENSSON

LABORATORY OF APPLIED PHYSICS
DEPARTMENT OF PHYSICS AND MEASUREMENT TECHNOLOGY
LINKÖPING INSTITUTE OF TECHNOLOGY
LINKÖPING, SWEDEN

ISBN 0-12-380210-5

I. Introduction

The first descriptions of a hydrogen-sensitive metal–oxide–semiconductor (MOS) field effect transistor were published in 1975 (Lundström *et al.*, 1975a,b). This device represents—to our knowledge—the first application of a chemically active metal gate, namely palladium, in an active semiconductor device. Since then a number of papers have been published on the subject. Most studies on the Pd-MOS devices have been made in Sweden, but similar devices have also been studied by other groups, as reviewed in the following. This chapter deals mainly with Pd–SiO_2–Si structures, although other devices, such as Pd–semiconductor Schottky barriers, are also considered. Our main purpose is to give a simple physical description of semiconductor devices with catalytic metal gates. In addition, we indicate the present level of understanding of these devices, their drawbacks, and their promise. Special attention is paid to the behavior of hydrogen in the Pd–SiO_2 system. Not only the wanted signal but also some hysteresis and long-term drift phenomena are due to the properties of this system. A description of some applications of hydrogen-sensitive transistors is given, e.g., smoke detection and biochemical reaction monitoring.

II. MOS Device Physics

A. Introduction to the Semiconductor Surface

Semiconductors in general contain relatively few free charge carriers. This facilitates control of the concentration and behavior of these charge carriers by external means. Many semiconductor phenomena have therefore became very attractive for technical applications. Furthermore, one semiconductor, silicon, has excellent stability, as has its oxide, silicon dioxide. The silicon–silicon dioxide system has therefore made semiconductor technology perfectly suited for industrial products. This technology has already led to an industrial revolution in electronics and information science, and it is natural to seek even more applications for it—for example, in chemical sensors.

The most important property of a semiconductor is its concentration of charge carriers. In an absolutely clean semiconductor there are equal amounts of negative free electrons and positive free holes, created by the excitation of valence electrons from the valance states in the crystal to the first band of excited states, called the conduction band (see Fig. 1). By such thermal excitation free electrons are formed in the conduction band, leaving unoccupied valence states. These unoccupied states are also considered as free charge carriers, positively charged (as the crystal was neutral when the state was filled), and are called holes. The concentration of these two carriers is controlled by the mass action law applied to the reaction $h^+ + e^- \rightleftarrows$ crystal:

$$np = n_i^2 \tag{1}$$

where n and p are the electron and hole concentrations, respectively, and n_i is a constant. For the clean semiconductor $n = p$, because it is electrically neutral; thus both n and p are equal to n_i, the intrinsic carrier concentration.

The intrinsic carrier concentration is very small in silicon, about 10^{16} m^{-3} at room temperature. However, n or p may be increased by doping. Consider a semiconductor doped by N_D positive ions per unit volume: the ions are assumed completely dissociated. We then have two equations to fulfill, Eq. (1) and the electrical neutrality condition

$$n = p + N_D \tag{2}$$

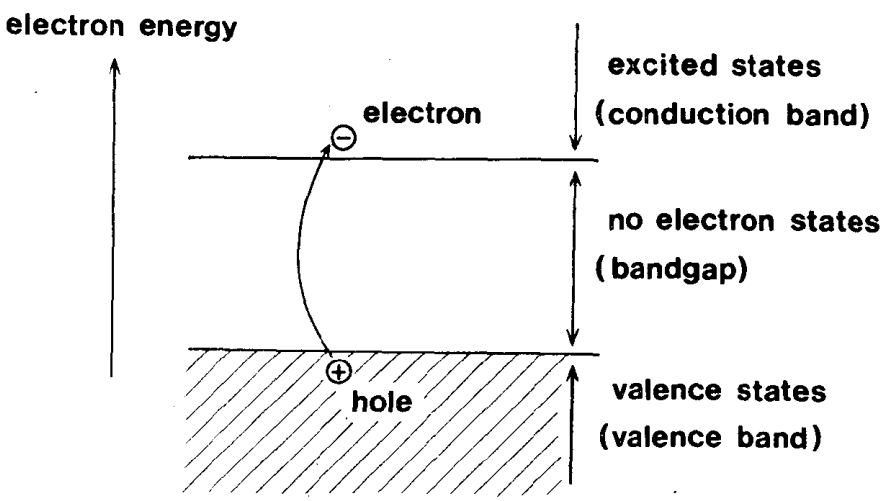

Fig. 1. Energy bands characterizing a semiconductor. Free electrons occur at the bottom of the conduction band. A hole corresponds to the lack of a valence electron and behaves like a mobile positive charge.

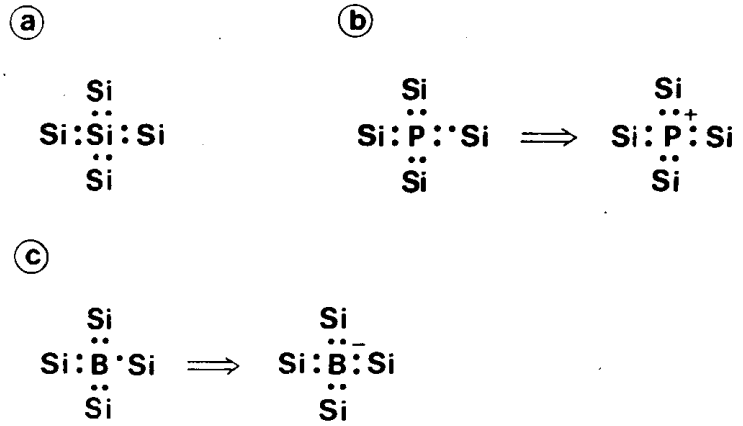

Fig. 2. Chemical bonding structure of silicon (a) compared to silicon doped with phosphorus (b) and boron (c). In the case of boron one electron is taken from the surrounding silicon atoms (giving rise to a hole) to complete the sp^3 hybrid bonds. In the case of phosphorus one electron does not fit into the sp^3 bonds and is therefore free (donated to the conduction band).

Normally, N_D is chosen much larger than n_i; N_D determines n, thus $n = N_D$. We have formed an n-type semiconductor, dominated by electron conduction. Note that the ions are fixed in the crystal. In silicon these ions may be, for example, phosphorus, which forms ions at silicon lattice positions, as shown in Fig. 2b. In the same way we may form a p-type semiconductor by doping with negatively charged ions, for example, boron (Fig. 2c).

The electron concentration can also be described in terms of the chemical potential or Fermi energy ϕ_F

$$n = n_i \exp\left(\frac{q\phi_F}{kT}\right) \tag{3a}$$

where q is the electron charge. As holes are just lack of electrons they are related to the same chemical potential or Fermi energy; thus from Eq. (1):

$$p = n_i \exp\left(\frac{-q\phi_F}{kT}\right) \tag{3b}$$

Energy bands and the Fermi energy are often represented in a "band diagram," shown in Fig. 3 as electron energy versus some space parameter (x axis). The band of valence states, the valence band, is shown normally filled with electrons in a simple covalent material. Holes will occur at the top of the valence band. The band of conduction states (first excited states), the conduction band, is normally empty. Free electrons occur at the bottom of this band. The Fermi energy is represented in this

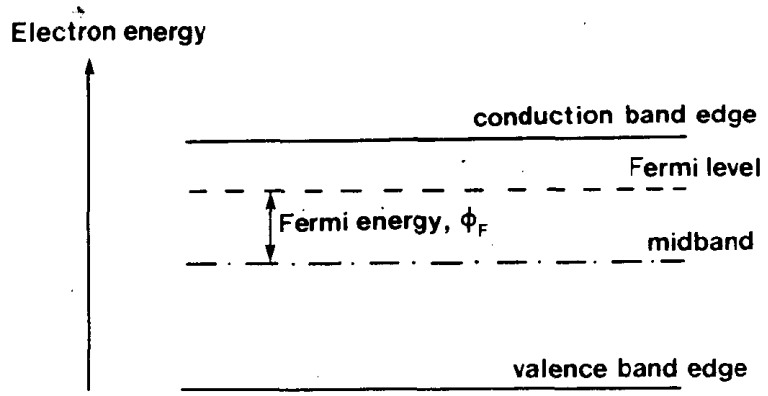

Fig. 3. Simplified energy band diagram for a semiconductor. The Fermi level and the Fermi energy are defined. (The example is for an *n*-type semiconductor.)

diagram as a dotted line. It is easy to remember that more electrons occur if this line is close to the conduction band (making ϕ_F large), and vice versa for holes.

Let us now consider a semiconductor surface. Figure 4a shows a band diagram of a *p*-type semiconductor with a surface that does not disturb the interior of the semiconductor. We have total electrical neutrality and constant potentials. Assuming that the surface contains a positive charge,

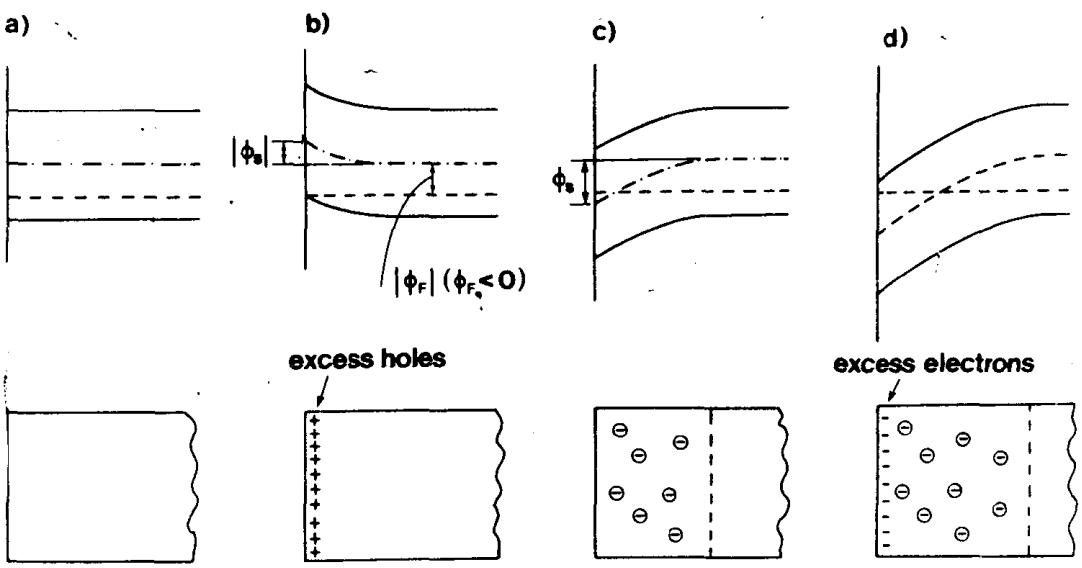

Fig. 4. Schematic illustration of the possible charge states of a *p*-type semiconductor and the corresponding energy band diagrams: (a) neutral, (b) accumulation, (c) depletion, (d) inversion.

the band diagram will be as shown in Fig. 4b. This situation occurs, for example, if we have a negative potential or negative charge somewhere outside the semiconductor surface. The positive charge in the surface consists of an extra hole concentration. Because the surface is no longer neutral, there is also a variable electric potential ϕ. In Fig. 4 this potential is shown as a corresponding change in electron energy, that is, as a bending of the conduction and valence band edges. The hole concentration is, of course, associated with the potential (note that $\phi_F < 0$ in a p-type material):

$$p = n_i \exp \frac{q}{kT}(-\phi - \phi_F) \tag{4a}$$

where ϕ becomes more negative closer to the surface. The change in semiconductor potential is largest at the surface. The value of the potential at the surface is called the surface potential ϕ_S. The situation discussed is called the accumulation regime, because holes accumulate at the surface.

In the case of a positive external charge or potential the semiconductor becomes negatively charged. The first effect of a positive external charge is to push the free holes away from the surface, leaving negative ions behind (Fig. 4c). Again, the hole concentration follows Eq. (4a). As ϕ is now positive the hole concentration is very small and we have a region in the semiconductor that is depleted of holes, the depletion region. The width of this region depends on doping, that is, on the concentration of negative ions, and is of the order of 0.1–1 μm in most cases. In contrast, the accumulation region, described above, is much narrower. The situation depicted in Fig. 4c is called the depletion regime. Note also that the electron concentration is controlled by an equation similar to Eq. (4a):

$$n = n_i \exp \frac{q}{kT}(\phi + \phi_F) \tag{4b}$$

Thus, if the potential ϕ changes enough the electron concentration may be of importance. Specifically, for $\phi_S = 2|\phi_F|$, the electron concentration at the surface is equal to the bulk hole concentration. This is the point of onset of the next regime, the inversion regime, which occurs for a higher external positive potential or charge. The inversion regime is demonstrated in Fig. 4d. The surface has been inverted from p-type to n-type, and a channel of electrons has formed at the surface. This channel is the basis of the MOS transistor discussed below. After reaching the inversion regime, all new charges are accumulated in the channel; thus the depletion width remains constant.

B. MOS Devices

MOS devices are normally based on the semiconductor silicon. If single-crystal silicon is oxidized in oxygen or water vapor at a high temperature, a high-quality silicon dioxide film is formed on the surface. The discovery of this phenomenon is the basis of the evolution of modern integrated circuit technology and the MOS technology. An MOS device is one based on the combination metal–oxide–silicon.

MOS devices can be made in two forms, as MOS capacitors or MOS transistors (Fig. 5). The devices can be made out of *p*-type silicon doped with trivalent impurities, such as boron. The impurities exist as negative ions in the crystal, and they will be compensated by an equal concentration of free holes (positive). The two concentrations are of the order of 10^{21} m^{-3}. The *p*-type silicon wafer is oxidized to an oxide thickness of the order of 100 nm and covered by a metal dot—for example, aluminum of about the same thickness. In the case of the transistor we have also formed two *n*-doped contacts in the silicon crystal. These areas are doped with, for example, phosphorus, and contain free electrons. The distance between the two regions is of the order of 10 μm.

For a negative potential on the metal electrode (or gate) the silicon surface will be charged positively through a higher concentration of holes. The surface will have metallic conductivity because of the high concentration of free charges, and the MOS capacitor will act like a normal plate capacitor with the oxide as a dielectric, having a capacitance equal to C_0 (Fig. 6). A positive potential on the gate pushes the holes away from the silicon surface, leaving only the negatively charged doping atoms near the surface, and these now form the negative charge on the silicon side of the capacitor. The capacitance of the MOS structure will be lower than before, because it is acting as a plate capacitor with a large dielectric thick-

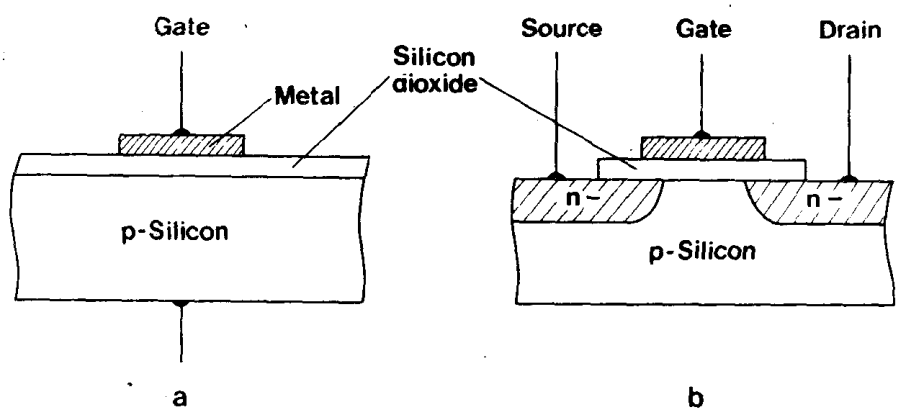

Fig. 5. The basic MOS capacitor (a) and MOS transistor (b).

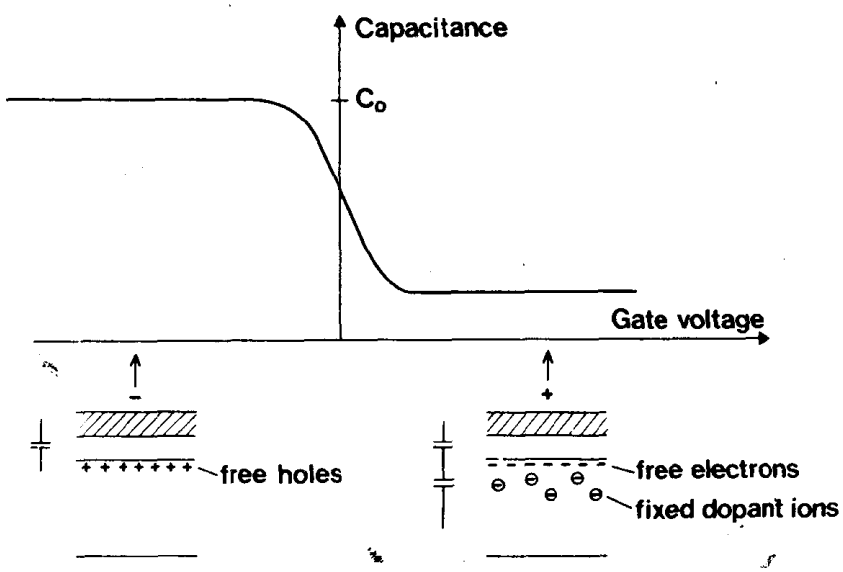

Fig. 6. The capacitance–voltage curve for an MOS capacitor. Schematic pictures of the charge distribution in the capacitor for negative and positive voltages are also shown.

ness: the oxide thickness plus the thickness of the depleted layer in the silicon. If we increase the gate voltage further the capacitance will continue to decrease until it reaches a point at which the electron concentration starts to increase at the silicon surface. In other words, the electric potential in the surface has lowered the hole concentration so much that the electron concentration increases according to the mass action law, forming an inversion layer. Figure 6 shows the total capacitance versus voltage curve that is used for analyzing the MOS capacitor.

In the case of a transistor we observe the conductance between the two *n*-type electrodes, the source and the drain, when the gate voltage is changed. With no gate voltage applied the conductance is almost zero. The two electrodes act as two opposite *pn* junctions or diodes, one diode always operating in its reverse direction. This is true in all situations except when an inversion layer is formed. An inversion layer forms an *n*-type channel between the two *n*-type contacts, and this channel provides a conductive path between the source and the drain. Figure 7 shows the drain–source conductance versus the gate voltage at small drain voltages. The conductance is zero until we reach the threshold voltage V_T at which the inversion layer starts to form. Then the conductance increases linearly with the gate voltage.

The gate capacitance stores a channel charge Q given by

$$Q = C_g(V_G - V_T) \tag{5}$$

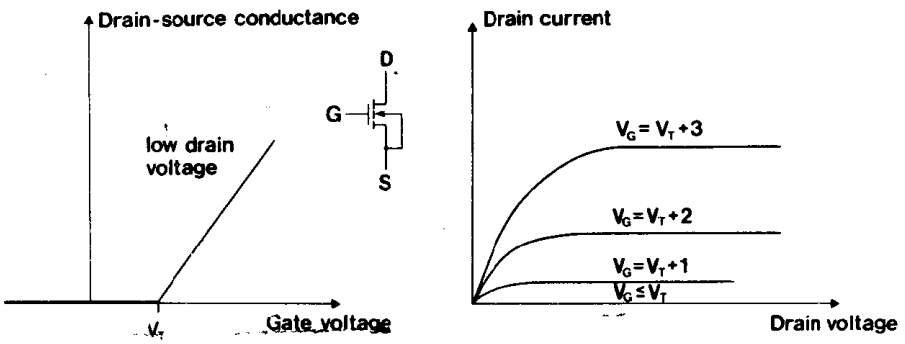

Fig. 7. Characteristic curves for an MOS transistor. The symbol of the *n*-channel MOS transistor is also shown. The arrow indicates the *p*-type substrate, which is normally connected to the source.

where C_g is the gate capacitance (C_0 times gate area), V_g is the gate voltage, and V_T is the gate voltage at which the channel starts to form, called the threshold voltage (Q is actually negative for an *n*-channel device). The gate charge passes the channel in a time t, given by the channel length L divided by the electron velocity v. The velocity is given by the electron mobility μ times the electric field along the channel, V_D/L

$$t = \frac{L}{v} = \frac{L^2}{\mu V_D} \tag{6}$$

From this we can calculate the transistor current I_D, given by the amount of charge that passes the channel per unit time

$$I_D = \frac{Q}{t} = \left(\frac{\mu C_g}{L^2}\right)(V_G - V_T)V_D \tag{7}$$

This formula is, in fact, correct for small drain voltages. At larger drain voltages, the potential varies along the channel and the induced charge $q(x)$ in the channel (per unit length) is a function of distance along the channel

$$q(x) = \frac{C_g}{L}[V_G - V_T - V(x)] \tag{8}$$

Furthermore, the electric field along the channel is $-dV/dx$ and the current is

$$I_D = q(x)\mu \frac{dV}{dx}$$

I_D must be constant along the channel, and by integrating the expression above with $q(x)$ given by Eq. (8) from $x = 0$ to $x = L$ (or from 0 to $V = V_D$)