Zinc oxide (ZnO) varistor, which is a kind of polycrystalline semiconductor ceramic composed of multiple metal oxides and sintered using conventional ceramic technology, is a voltage-dependent switching device, which exhibits highly nonohmic current-voltage characteristics above the breakdown voltage. Basic information on ZnO varistors, including the fabrication, microstructure, and typical electrical parameters, is introduced. The history and applications of ZnO varistors are also presented. The panorama of alternative varistor ceramics for Bi<sub>2</sub>O<sub>3</sub>-based ZnO varistors is mapped out. Especially, the ceramic-polymer composite varistors with lower breakdown voltage, incorporating varistor particles such as semiconducting particles, a combination of metal and semiconducting particles, and ZnO microvaristors, in a polymeric matrix are reported. Now, varistors are available that can protect circuits over a very wide range of voltages, from a few volts for low voltage varistors in semiconductor circuits to 1000 kV AC and ±1100 kV DC in electrical power transmission and distribution networks. Correspondingly, they also can handle an enormous range of energies from a few joules to many megajoules.

### 1.1 ZnO Varistors

A varistor is an electronic component with a "diode-like" nonlinear current–voltage characteristic, which is a portmanteau of variable resistor [1]. Functionally, varistors are equivalent to a back-to-back Zener diode and are typically used in parallel with circuits to protect them against excessive transient voltages in such a way that, when triggered, they will shunt the current created by the high voltage away from sensitive components.

The most common type of varistor is the metal oxide varistor (MOV), which contains a ceramic mass of ZnO grains, in a matrix of other metal oxides, such as small amounts of bismuth, cobalt, and manganese, sandwiched between two metal electrodes. The boundary between each grain and its neighbor controls the current according to the applied voltage, and allows current to flow in two directions. The mass of randomly oriented grains is electrically equivalent to a network of back-to-back diode pairs, each pair in parallel with many other pairs. A varistor's function is to conduct significantly increased current when voltage is excessive. Only nonohmic variable resistors are usually called varistors [1].

1

In normal use, a varistor is subject to a voltage below its characteristic breakdown voltage and passes only a tiny leakage current. When the voltage exceeds the breakdown voltage, the varistor becomes highly conducting and draws a large current through it, usually to ground. When the voltage returns to normal, the varistor returns to its high-resistance state [2]. The result of this behavior is a highly nonlinear current-voltage characteristic [3-5], in which the MOV has a high resistance at low voltages and a low resistance at high voltages; usually, the resistivity of a ZnO varistor is more than  $10^{10} \Omega$  cm below the breakdown voltage, whereas it is less than several ohm-centimeters above the breakdown voltage [6]. The switch is reversible with little or no hysteresis although it can degrade under electrical loading [2]. A varistor remains nonconductive as a shunt-mode device during normal operation when the voltage across it remains well below its "clamping voltage"; thus varistors are typically used for suppressing line voltage surges. However, a varistor may not be able to successfully limit a very large surge from an event such as a lightning strike where the energy involved is many orders of magnitude greater than it can handle. Follow-through current resulting from a strike may generate excessive current that completely destroys the varistor [1].

ZnO varistors are voltage-dependent switching devices, which exhibit highly nonohmic current–voltage (I-V) characteristics above the breakdown voltage. The nonohmic I-V characteristics are usually expressed logarithmically, as shown in Figure 1.1 [6]. The degree of nonohmic property is usually expressed by a nonlinear coefficient  $\alpha$  defined by the following equation:

$$\alpha = \frac{V}{I} \frac{\mathrm{d}V}{\mathrm{d}I} \tag{1.1}$$

Empirically, the following simple equation is used:

$$I = \left(\frac{V}{C}\right)^{\alpha} \tag{1.2}$$



Figure 1.1 I-V characteristics of a typical ZnO varistor. Source: Adapted from Eda [6].





where *C* is a constant. Typical  $\alpha$  values of ZnO varistors are from 30 to 100; therefore, the current varies by orders of magnitude with only small changes in voltage. A more accurate measure of the nonlinearity is the dynamic conductance, the differential of the characteristic, at each voltage [2]. On the contrary,  $\alpha$  values of conventional varistors such as SiC varistors do not exceed 10.

The I-V characteristics of ZnO varistors are classified into three regions, as shown in Figure 1.2 [6]. In region I (low electrical field region, or pre-breakdown region), below the breakdown voltage (typically a voltage at 1  $\mu$ A cm<sup>-2</sup>), the nonohmic property is not so prominent and can be regarded as ohmic, and the leakage current is highly temperature dependent. In region II (medium electrical field region, nonlinear region, or breakdown region), between the breakdown voltage and a voltage at a current of about 100 A cm<sup>-2</sup>, the nonohmic property is very prominent and almost independent of temperature. In region III (high electrical field region or upturn region), above 100 A cm<sup>-2</sup>, the nonohmic property gradually decays, and the varistor is again ohmic. These three regions in engineering applications are also called as the low current region, medium current region, and high current region, respectively.

ZnO varistors are characterized by the magnitude of the  $\alpha$  values and the width of the range where highly nonohmic property is exhibited. In contrast to the pre-breakdown region, the nonlinear region and upturn region are little affected by temperature. The I-V characteristics below 100 mA cm<sup>-2</sup> are usually measured using a DC electric source, whereas those above 1 A cm<sup>-2</sup> are measured by an impulse current source to avoid heat generation and thermal breakdown. The waveform of the impulse current is 8/20 µs with an 8 µs rise time and 20 µs decay time up to one-half the peak value [6], which is used as a standard impulse current to test ZnO varistors. The I-V characteristics measured by the impulse currents show voltages higher than those measured using DC. The discrepancy is usually 10–20%, as shown by arrows in Figure 1.1 [6]. This discrepancy is caused by the delay in electrical response in the ZnO varistors. The response delay is speculated to be caused by electron trapping and hole creation at the grain boundaries.

# 1.2 Fabrication of ZnO Varistors

ZnO has a wurtzite structure in which the oxygen atoms are arranged in a hexagonal close-packed type of lattice with zinc atoms occupying half the

tetrahedral sites, as shown in Figure 1.2; Zn and O atoms are tetrahedrally coordinated to each other and are equivalent in position. The ZnO structure is thus relatively open with all the octahedral and half the tetrahedral sites empty. It is, therefore, relatively easy to incorporate external dopants into the ZnO lattice. The open structure also has a bearing on the nature of defects and the mechanism of diffusion, and the most common defect in ZnO is the metal in the open interstitial sites [8]. Pure ZnO without any impurities or dopants is a nonstoichiometric n-type semiconductor with linear or ohmic behavior, and with a wide band gap (3.437 eV at 2 K) [8, 9].

ZnO varistors are semiconducting ceramics fabricated by sintering of ZnO powders with small amounts of various metal oxide additives such as  $Bi_2O_3$ , CoO, MnO, and  $Sb_2O_3$ . The nonohmic property comes from grain boundaries between semiconducting ZnO grains. These oxides are added to control one or more of the properties such as the electrical characteristics (breakdown voltage, non-linearity, and leakage current), grain growth, or the sintering process [10–18].  $Bi_2O_3$  is the most essential component for forming the nonohmic behavior, and addition of CoO and MnO enhances the nonlinear properties [11, 12, 19–22]. A very low concentration of  $Al_2O_3$  increases the ZnO grain conductivity while  $Sb_2O_3$  controls the ZnO grain growth [23–32]. Combination of additives such as  $Cr_2O_3$ , MnO,  $Bi_2O_3$ , and CoO produces greater nonlinearity than a single dopant [10]. During high temperature sintering different phases are formed and the microstructure of a ZnO varistor comprises conductive ZnO grains surrounded by electrically insulating grain boundary regions.

Dopants play at least three major roles in forming varistors by affecting grain growth during sintering, the dewetting characteristics of the liquid phase during cooling, and the electronic defect states that control the overall varistor characteristics [2]. In order to obtain high performance ZnO varistors, the compositions, impurities, mixing methods, particle sizes, and sintering conditions, such as maximum temperature and holding time, temperature rising/lowering rates, and oxygen and  $Bi_2O_3$  partial pressures, should be controlled precisely. This will be discussed in detail in Chapter 3.

#### 1.2.1 Preparation of Raw Materials

ZnO varistors are commonly produced by conventional ceramic technology. The preparation process of ZnO varistors is basically the same as for general ceramics, whose primary processes include raw materials preparation, mixing, stoving and sieving, pre-sintering, smashing, prilling, molding, sintering, silvering, and testing. The preparation begins with weighing, mixing, and milling of oxide powders in ball mills. After the addition of organic binder substances, the aqueous slurry is spray-dried to produce a granulate. After sieving appropriate fractions to extract dust and large agglomerated particles, the material is transferred to hydraulic pressing machines to mold the "green" disks [33]. In the whole process, raw materials preparation and sintering are the two key steps that determine the electrical properties of ZnO varistors.

The preparation of material powders of ZnO varistors actually involves two steps: the preparation of ZnO powder and the mixing of ZnO powder with other additives. Usually, both steps are often combined as an inseparable one, so as to achieve a one-step preparation of composite raw powders of the ZnO varistor ceramics.

The preparation of pure ZnO powder is divided into three traditional methods:

- *Direct method:* Zinc powder is added to reducing agents such as coke and coal, and heated at a high temperature to form reduced Zn steam, and then oxidized to ZnO by air.
- *Indirect method:* the zinc ingots are melted into the hot crucible to evaporate, so as to form the Zn steam, and then oxidized to ZnO by air.
- Wet method: The ZnSO<sub>4</sub> or Zn salt solution such as ZnCl<sub>2</sub> reacts with the Na<sub>2</sub>CO<sub>3</sub> solution to crystallize the ZnCO<sub>3</sub> precipitation, and ZnO is generated by washing and filtering and heating to decompose the sediment.

In order to prepare high quality ZnO powder, some liquid- and gas-phase methods have been also used for the preparation of ZnO powders alone.

There are many ways of preparing ZnO varistor ceramic composite powders. According to the material states, the methods can be summarized as the solid phase, liquid phase, and gas phase one.

- Solid-phase method: it applies for the uniform mixing of ZnO solid powder and solid powder of other additives by mechanical milling. This is the conventional method used in the industry. However, it is difficult to obtain powders with uniform particle size distribution through this process.
- *Liquid-phase method:* it is developed from the wet method in the pure ZnO powder preparation. In liquid-phase synthesis, as components are fully dispersed in the liquid phase the content of each component can be precisely controlled. The solid-phase grain size produced is small with narrow size distribution, and particles of spherical or nearly spherical shape are obtained. The liquid-phase method is now becoming more and more common in the ZnO varistor ceramics industry, and it includes precipitation calcination, sol-gel, spray pyrolysis, hydrothermal method, and so on.
- Gas phase method: this method includes chemical vapor oxidation, laser heating, and so on. But the gas-phase method is of high cost, and it is difficult to achieve mass production using this.

Combustion synthesis technique was also used to successfully produce pure and doped pure crystalline ZnO varistor powders with good compositional control. The combustion synthesis route enables synthesis at low temperatures and the products obtained are in a finely divided state with large surface areas. Combustion synthesis offers as added advantages, the simplicity of experimental setup, the surprisingly short time between the preparation of reactants and the availability of the final product, the savings in external energy consumption, and the equally important potential of simplifying the processing prior to forming, providing a simple alternative to other more elaborate techniques [34].

#### Sintering of ZnO Varistors 1.2.2

ZnO varistor is fabricated by the traditional ceramic sintering technology, and the temperature control of the sintering process has a critical influence on the performances of the final products. The sintering procedure includes warm-up,

constant-temperature, and cooling periods. The pressed disks are sintered in an electric furnace at 1150-1350 °C for one to five hours in air, and the temperature is increased or decreased at 50-200 °C h<sup>-1</sup>. During the warm-up and constant-temperature periods several chemical phase transformations take place, which result in a complete rearrangement of the microscopic particles and the creation of a dense polycrystalline matrix of ZnO grains and other phases, which are incorporated therein. The homogeneity of the ceramic microstructure strongly depends on the initial distribution of the reacting particles in the green body, and inhomogeneously distributed additives can result in islands of exaggerated or hampered ZnO grain growth or porosity, while such defects lead to bad electrical performance [33]. Finally, electrodes are prepared on both surfaces. The best electrodes are painted In–Ga alloys or evaporated Al films. Conventional silver electrodes painted and fired at 500–800 °C are also used [6].

The solid-state-based ceramic processing route still remains the preferred method of manufacturing because of the simplicity, cost, and availability of the metal oxide additives; this process has considerable disadvantages especially for many modern commercial applications. A major disadvantage of this route is the difficulty in obtaining additive homogeneity at a microscopic level, which is especially important for the manufacture of miniaturized electronic equipment. Processing methods and additive homogeneity are the critical parameters to produce a varistor material with favorable electrical characteristics. Inhomogeneous microstructures could cause degradation of varistors during electrical operation. Therefore, careful control of the microstructure is required to produce a varistor of desired characteristics [10].

Chemical processing, such as sol–gel, solution, precipitation, and microemulsion techniques, facilitates homogeneous doping at the molecular level. Among those various methods reported, sol–gel processing is found to be a promising one to obtain a miniature device with higher breakdown voltage, low leakage current, and reasonably good nonlinear characteristics [10]. However, controlling the grain growth at high sintering temperatures still remains a challenge for varistor ceramics, and the electrical properties could be improved if the materials could be sintered to full density with sub-micrometer or nanometer grain size. Researchers have proposed seed law method, stacking method, hot pressing method, microwave sintering method, and other technologies for the sintering of ZnO varistor ceramics [6]. Indeed, a significant enhancement in electrical properties has been achieved by novel sintering procedures such as step sintering, spark plasma sintering, or using microwave irradiation [10]. This will be introduced in Chapter 9.

# 1.3 Microstructure

Conventional varistors are usually prepared by mixing  $0.2-1 \ \mu m$  sized ZnO powders with oxide additives. Although the microstructures of varistors exhibit considerable variation from one manufacturer to another, they all exhibit the characteristics of a typical ceramic prepared by liquid-phase sintering. ZnO



Figure 1.3 SEM photomicrograph of ZnO varistor microstructure.

varistors are polycrystalline materials composed of semiconducting ZnO grains with their attendant grain boundaries as shown in Figure 1.3. Moreover, the microstructure contains particles of one or more types of spinel, and a lot of remaining pores can be primarily found in the grain boundaries. Depending on the composition, twins within the ZnO grains and a small amount of a pyrochlore phase can also form. The ZnO–Bi<sub>2</sub>O<sub>3</sub>–Sb<sub>2</sub>O<sub>3</sub> system forms a pyrochlore phase Zn<sub>2</sub>Bi<sub>3</sub>Sb<sub>3</sub>O<sub>14</sub> above 650 °C. With ZnO, pyrochlore further reacts to form a spinel (Zn<sub>7</sub>Sb<sub>2</sub>O<sub>12</sub>) [12, 35].

The grain boundary has a relatively thick  $(0.1-1 \text{ pm}) \text{ Bi}_2 \text{O}_3$ -rich intergranular layer, which becomes thinner (10-1000 Å) as it approaches the contact points of the particles [6]. The sizes of ZnO grains are usually 5–20 µm and depend on the material composition, temperature, and time of sintering. The resistivity of a ZnO grain is  $0.1-1 \Omega$  cm. The grain boundaries are highly resistive and show nonohmic property, and the corresponding breakdown voltage is around 200–400 V mm<sup>-1</sup> for conventional varistors, and the breakdown voltage per grain boundary is about 3 V. The breakdown voltage of the sintered body is proportional to the number of grain boundaries between the two electrodes. This indicates that the breakdown voltage is proportional to the inverse of the ZnO grain size [6].

The characteristics of ZnO varistor ceramics are closely related to their microstructure, which is characterized by the following parameters: ZnO grain size and the grain size distribution, grain boundaries, secondary phases such as the  $Bi_2O_3$ -rich phase,  $Zn_7Sb_2O_{12}$  spinel-type phase, and  $Bi_3Zn_2Sb_3O_{14}$  pyrochlore-type phase, the distribution of secondary phases along the grain boundaries, and the size and distribution of pores [36].

### 1.4 Typical Parameters of ZnO Varistors

The nonlinear voltage–current (V-I) characteristics of ZnO varistors can be approximately expressed by Eq. (1.2). In Figure 1.4 [37], the V-I nonlinear curve is shown, where the bi-logarithm coordinates are used, and the voltage and current are represented by the voltage gradient and current density, respectively.



Figure 1.4 Nonlinear voltage-current (V-I) characteristics of ZnO varistors.

The nonlinear V-I characteristics of ZnO varistors is mainly indicated by the following parameters, some of which only apply to the high voltage electrical applications [37].

- *Nonlinear coefficient*  $\alpha$ : it represents the reciprocal of the slope of the straight part in the *V*–*I* curve, and generally refers to the corresponding value of the breakdown region. The larger the value is, the better the performance.
- *Varistor voltage*  $U_{nmA}$ : it refers to the voltage of the varistor at the turning point of the curve, at a specific value of current (between 0.1 and several milliamperes, usually 1 mA), also called the reference voltage. For the convenience of performance comparison, it is usually represented by the voltage gradient  $E_{nmA}$ :

$$E_{n\mathrm{mA}} = \frac{U_{n\mathrm{mA}}}{d} \tag{1.3}$$

where *d* is the thickness of the varistor disk.

- *Chargeability (or voltage undertaking ratio) S:* it is defined by the ratio of the crest value of the continuous operating voltage to the varistor voltage  $U_{nmA}$ ; it represents the load voltage intensity at continuous operation condition.
- *Leakage current*  $I_L$ : refers to the current flowing through the varistor at the set temperature and voltage, which is less than the varistor voltage, and usually takes 0.75  $U_{1\text{mA}}$  as the test condition.
- *Residual voltage*  $U_{NkA}$ : it refers to the impulse voltage occurring at the set impulse current of N kA. In high voltage applications, it is usually represented by impulse current residual voltage, lightning current residual voltage, and steep-wave current residual voltage.
- *Residual voltage ratio K:* it refers to the ratio of the residual voltage to the varistor voltage, namely,  $U_{NkA}/U_{nmA}$ .
- *Discharge capacity:* it refers to the maximum impulse current value allowed to flow through the varistor in the set conditions, including a 2 ms square wave and a  $4/10 \mu$ s impulse current.

Energy absorption capability (or energy handling capability) is the second most important property of ZnO varistors next to nonlinearity. ZnO varistors as the core elements of surge arresters in high voltage systems or surge protection devices (SPDs) in low voltage systems are required to absorb substantial amounts of energy resulting from temporary overvoltages, switching surges, or lightning impulses. Therefore, the energy absorption capabilities of ZnO varistors are crucial for the integrity of equipment and systems. However, it has been observed in experiments that differences in barrier voltages, grain sizes, and grain boundary characteristics inside the same ZnO varistor and among different ZnO varistors cause nonuniformity in the microstructurally electrical and thermophysical characteristics of ZnO varistors, and finally result in the differences in their current handling capabilities, which is also called as the energy absorption capability and has a direct relation to the failure modes. Failure modes include electrical puncture, physical cracking, and thermal runaway, which happen under different currents. The energy absorption capability can be divided into thermal energy absorption capability and impulse energy absorption capability. For the impulse energy absorption capability single impulse stress, multiple impulse stress (without sufficient cooling between the impulses), and repeated impulse stress (with sufficient cooling between the stresses) have to be considered. Thermal energy absorption capability, on the other hand, can only be considered for complete metal oxide arresters (MOAs) or SPDs, as it is mainly affected by the heat dissipation capability of the overall arrester design, besides the electrical properties of the MOVs.

Another important electrical characteristic of ZnO varistors is the dielectric property. Below the breakdown voltage, ZnO varistors are highly capacitive. The dielectric constant of ZnO is 8.5, whereas an apparent dielectric constant of a ZnO varistor is typically 1000. The dielectric properties are mainly caused by thin depletion layers (~1000 Å) at the grain boundaries [6].

### 1.5 History of ZnO Varistors

The ZnO varistors were first developed by Matsuoka and his research group at Matsushita Electric (Japan) in 1969 [38] and were commercialized under the trade name ZNR in the following year [6]. In 1979, they introduced many of the essential features of ZnO varistors as we know them today in detail [3], which included making ZnO semiconducting by adding substitutional ions, densification by liquid-phase sintering with a  $Bi_2O_3$ -rich liquid phase, segregation of large ions to the grain boundaries, and the main formulas with high nonlinearity by doping manganese and cobalt. Although work on the electrical properties of ZnO ceramics apparently had been underway in Russia in the early 1950s [39, 40], it was Matsuoka's work [3] that captured attention, and, shortly thereafter, a joint development effort was undertaken by GE and Matsushita [41]; one of Matsuoka's contributions was the discovery of the role of aluminum [42]. It was in the period following this that extensive literature was documented and the scientific basis for varistors as well as many of the

key technological developments was established. Concurrently, the superiority of ZnO varistors over SiC varistors for many applications was established, and expertise in utilizing ZnO varistors spread geographically. More companies began to manufacture varistors, and alternative sets of dopants were introduced, but the essential features introduced by Matsuoka remained [2]. The one exception was the use of  $Pr_2O_3$  in the place of  $Bi_2O_3$  [43, 44].

The performance parameters of ZnO are closely related to the raw formulas. By changing the percentage and compositions of the additives, the performance of ZnO varistors changes correspondingly. According to the additives, industrial ZnO varistors are mainly divided into two series:

- The ZnO-Bi<sub>2</sub>O<sub>3</sub> series, or Bi-series, mainly includes additives such as Bi<sub>2</sub>O<sub>3</sub>, Sb<sub>2</sub>O<sub>3</sub>, Co<sub>2</sub>O<sub>3</sub>, MnO<sub>2</sub>, and Cr<sub>2</sub>O<sub>3</sub> [3];
- The ZnO–Pr<sub>2</sub>O<sub>3</sub>/Pr<sub>6</sub>O<sub>11</sub> series, or Pr-series, mainly includes additives such as Pr<sub>2</sub>O<sub>3</sub>/Pr<sub>6</sub>O<sub>11</sub>, Co<sub>2</sub>O<sub>3</sub>, and CaO [43, 44].

Many efficient additives were discovered. Eda [6] and Gupta [8] separately summarized and classified the additives into four types according to their effects:

- Additives such as Bi<sub>2</sub>O<sub>3</sub>, Pr<sub>2</sub>O<sub>3</sub>, BaO, SrO, and PbO, which make ZnO form a grain boundary structure, mainly promote liquid-phase sintering to form grain boundaries and make the materials nonlinear.
- Additives such as Co<sub>2</sub>O<sub>3</sub>, MnO<sub>2</sub>, Sb<sub>2</sub>O<sub>3</sub>, Al<sub>2</sub>O<sub>3</sub>, and Ga<sub>2</sub>O<sub>3</sub> improve the nonlinearity of the ZnO varistor. Part of these additives solidly solutes in ZnO grains to form impurity traps to provide carriers as donors, while the others form traps to promote the barrier height as acceptors.
- Additives such as Sb<sub>2</sub>O<sub>3</sub>, SiO<sub>2</sub>, and TiO<sub>2</sub> affect the growth of ZnO grains by promotion or inhibition; the size of ZnO grains can be controlled, and the voltage gradient is tailored.
- Additives such as Sb<sub>2</sub>O<sub>3</sub>, SiO<sub>2</sub>, NiO, and frit mainly improve the stability of ZnO varistors.

In the first decade after the Matsuoka's invention, various additives to improve the electrical characteristics were discovered and the processing conditions were optimized. This will be introduced in Chapter 3. In the next decade, the microstructures and the physical properties of the grain boundaries were gradually identified, and applications were rapidly found in protecting electrical equipment and electronic components such as transistors and integrated circuits (ICs) against voltage surges [6].

Many efforts have been made in the past several decades focusing on the conduction mechanism in the ZnO varistor; among the numerous conduction models proposed by different research groups, the one presented by Pike [45, 46] and further developed by Blatter and Greuter [47–49] has been widely recognized and may meet most of the experimental phenomena. Essentially, it is believed that the nonlinear V–I characteristic of the ZnO varistor is caused by the charge carriers transporting across the double Schottky barrier (DSB) formed at the grain boundary. The conduction mechanism of the ZnO varistor will be introduced in Chapter 2.