Lightning and NEMP Transient Protection with Metal Oxide Varistors

François D. Martzloff
General Electric Company


Significance:
Part 6: Tutorials, textbooks and reviews
Part 7: Mitigation techniques

After a brief review of varistor fundamentals, including a discussion of the infamous "speed of response" issue, the report describes several unconventional packaging implementations of varistor material (usually a flat disc between metallized electrodes), such as molded axial package, bulkhead connectors, coaxial mounting, reconstituted material, and varistor substrates with surface conduction between electrodes or bulk conduction to a ground plane.
Protection of electronic components against the transient overvoltages produced by lightning or nuclear electromagnetic pulses (NEMP) involves several techniques, starting with hardening and shielding the circuits and, finally, with protecting individual components. Metal oxide varistors, in their conventional form or in unconventional packages described in this report, can offer protection at various stages of the circuit, in bulkhead connectors, in printed circuit cards, or in substrate form.

In this report, a brief overview is given on varistor fundamentals, including the limitations brought about by leads in NEMP protection. Several examples of novel, unconventional approaches are then described. These approaches have been proposed and reduced to practice as far as basic feasibility is concerned, but the details have not yet been developed.
FUNDAMENTALS ON VARISTORS
The term varistor is derived from its function as a variable resistor. It is also called a voltage-dependent resistor, but that description misleadingly implies that the voltage is the independent parameter in surge protection. Two very different devices have been successfully developed as varistors: silicon carbide discs have been used for years in the surge arrester industry, and, more recently, metal oxide varistor technology has come of age.

Structure and Packaging
Composition of the varistor consists primarily of zinc oxide with small additions of selected metal oxides. These materials are sintered at high temperature to produce a polycrystalline ceramic body. The structure of the body is a matrix of conductive zinc oxide grains separated by a highly resistive intergranular boundary (Figure 1).

Figure 1. Photomicrograph of a polished and etched section of a metal oxide varistor.

Since electrical conduction occurs between zinc oxide grains distributed throughout the bulk of the device, a varistor is inherently more rugged than its single junction counterparts, such as zener diodes. In the varistor, energy is absorbed throughout the body of the device with the resultant heat spread evenly through its volume. Electrical properties are controlled mainly by the physical dimensions of the varistor body, which is generally sintered in the shape of a disc. Energy rating is determined by volume, voltage rating by thickness, and current capability by area.

In conventional applications, a varistor device consists of a disc of the material, with electrodes applied by silk screen, metal spray, etc., to opposite, parallel faces. A connection is made to the electrodes for feeding current in and out of the varistor by attachment of the device leads to the disc, and thence to the circuit. The attachment can take many forms (Figure 2), but all involve the insertion of some lead length, a potential problem that will be discussed later.

Electrical Parameters
Varistor V-I Characteristics
Varistor electrical characteristics are conveniently displayed using a log-log format in order to show the wide range of the V-I curve. The log format is also clearer than a linear representation which tends to exaggerate the nonlinearity in proportion to the current scale chosen. A typical V-I characteristic curve is shown in Figure 3. In order to illustrate three distinct regions of electrical operation, this plot shows a wider range of current than is normally provided on varistor data sheets.

In the leakage region (Figure 3), the predominant parameter is the $R_{on}$ resistance (Figure 4), shown as the $10^9$ ohm line in the figure. As the current density increases, normal varistor operation occurs, with $R_s$ being the predominant parameter. Finally, in the upturn region, where $R_s$ has become very small compared to $R_{on}$, the latter dominates. Two other parameters, the capacitance $C$ (reflecting a high effective dielectric constant of the material) and the lead inductance $L$ are significant at high frequencies.

The leakage region is of little interest to the user; in the normal operating region, the flatness of the characteristic is a criterion of clamping effectiveness. Departure from this flat characteristic in the upturn region could be viewed as undesirable; however, this upturn is useful in providing
Figure 2. Typical lead configurations.

Figure 3. Typical varistor V-I curve plotted on log-log scale.

Figure 4. Varistor equivalent circuit model.
current sharing across the section of the device at high currents.

The capacitance may be objectionable at high frequencies. Adding a low capacitance avalanche diode in series is a method of reducing the effective capacitance. The unconventional examples described in the next section also show means of reducing capacitive effects.

The lead inductance, already mentioned, will be discussed at greater length in the following subsections.

The V-I characteristic, then, is the basic application design tool for selecting a device to perform a protective function. For a successful application, however, other factors, which are discussed in detail in the information available from manufacturers, must also be taken into consideration. Some of these factors are:

- Selection of the appropriate nominal voltage for the line voltage of the application
- Selection of energy-handling capability (including source impedance of the transient, waveshape, and number of occurrences)
- Heat dissipation
- Proper installation in the circuit (configuration, lead length)

"Overshoot" — A Lead Effect

Enough instances of poor installation practices have been observed, and enough questions have been raised on alleged "overshoot," that a brief discussion of lead effects is in order. To illustrate the effect of lead length on the overshoot, two measurement arrangements were used. As shown in Figures 5a and 5b, respectively, 0.5 cm² and 22 cm² of area were enclosed by the leads of the varistor and of the voltage probe.

The corresponding voltage measurements are shown in the oscillograms of Figures 6a and 6b. With a slow current front of 8 μs, there is little difference in the voltages occurring with a small or large loop area, even with a peak current of 2.7 kA. With the steep front of 0.5 μs, the peak voltage recorded with the large loop is nearly twice the voltage of the small loop. Note (Figure 6b) that at the current peak, $L \frac{di}{dt} = 0$, and the two voltage readings are equal; before the peak, $L \frac{di}{dt}$ is positive, and after, it is negative.

Hence, when one is making measurements as well as when one is designing a circuit for a protection scheme, it is essential to be alert to the effects of lead length (or, more accurately, of loop area) for connecting the varistors. This warning is especially important when the currents are in excess of a few amperes with rise times of less than one microsecond.

**Speed of Response**

The varistor action depends upon a conduction mechanism similar to that of other semiconductor devices. For this reason, conduction occurs very rapidly, with no time lag apparent, even into the nanosecond range. Figure 7 shows a composite photograph of two voltage traces with and without a varistor inserted in a very low inductance impulse generator. The second trace (which is not synchronized with the first, but merely superimposed on the oscilloscope screen) shows that the voltage clamping effect of the varistor occurs in less than one nanosecond.

In the conventional lead-mounted devices, the inductance of the leads would completely mask the fast action of the varistor. Therefore, the test circuit for Figure 7 required insertion of a small piece of varistor material in a coaxial line to demonstrate the intrinsic varistor response.

Tests made on lead-mounted devices, even with careful attention to minimizing lead length, show that the voltages induced in the loop formed by the leads contribute a substantial part of the voltage appearing across the terminals of a varistor at high current and fast current rise. In typical power applications where the threat is lightning or switching surges, the impressed current rise time is sel-

![Diagram of measurement circuit for varistors](attachment:image.png)
dom short enough for this effect to become significant. Competitive technology will make claims that an “overshoot” occurs, but that has been shown to be primarily a lead effect as discussed in the preceding subsection.

Voltage rate-of-rise is not the best term to use when discussing the response of a varistor to a fast impulse, unlike spark gaps where a finite time is involved in switching from a non-conducting to a conducting state. The response of the varistor to the transient current delivered by a circuit is the appropriate characteristic to consider.

Applications involving NEMP protection will be fast-rising voltages and current surges, so that the packaging will become very significant, and a departure will be required from the conventional packages currently found in commercial devices. In the next section, these unconventional applications will be briefly presented.

**Failure Modes**

An electrical component is subject to failure either because its capability was exceeded by the applied stress or because some latent defect in the component went by unnoticed in the quality control processes. While this situation is well recognized for ordinary components, a surge protective device, which is no exception to these limitations, tends to be expected to perform miracles, or to at least fail graciously in a “fail-safe” mode. The term “fail-safe,” however, may mean different failure modes to different users and, therefore, should not be used. To some users, fail-safe means that the protected hardware must never be exposed to an overvoltage, so that failure of the protective device must be in the fail-short mode, even if it puts the system out of operation. To other users, fail-safe means that the function must be maintained, even if the hardware is left temporarily unprotected, so that failure of the protective device must be in the open-circuit mode. It is more accurate and less misleading to describe failure modes as “fail-short” or “fail-open,” as the case may be.(4)

Metal oxide varistors, when subjected to excessive energy (high surge current), will fail in the “fail-short” mode. If the varistor is connected to a power system with large short-circuit capability, the fault current may produce melting of a soldered
connection or shattering of the ceramic, ultimately producing a "fail-open" condition.

UNCONVENTIONAL APPLICATIONS

For the protection of fast rise transients occurring with NEMP, it is apparent that lead-mounted devices are likely to be unsatisfactory. To still benefit from the inherent advantages of metal oxide varistors, novel packaging is required. In this section, several examples are given: bulkhead connectors and coaxial connectors, each with conventional sintered blocks or with reconstituted material. Both have been shown to be feasible, although the complete details have not yet been fully developed.

Two further applications are also described, which are still in the conceptual stage, but show promise of useful application in the integrated protection of circuits: as thick film layers between conductor runs, and as substrates.

Bulkhead Connectors

The inductance of the leads whose effects have been discussed can be essentially eliminated by construction in a coaxial configuration. This type of packaging would be especially effective and convenient in conjunction with connectors for cable entry into shielded boxes.

Figure 8 shows a configuration for a single wire, and Figure 9 for multiple wires or pairs. In both figures, a varistor bead is inserted between the conductor(s) and an opening of the bulkhead plate, which is part of the grounded shell. In this manner, any impinging surge current flowing in the conductor is diverted to the grounded chassis of the enclosure, with minimum inductance inserted.

In principle, this construction is quite simple. The details of the bonding between conductors and varistor will require development, which should be within the skills of connector manufacturing.

A possible objection to this construction would be that the whole volume of the varistor bead acts as the dielectric of a capacitor inserted between the conductor and its sheath. With an effective dielectric constant of 1000, and the small radial distance required for low-voltage clamping, the resulting capacitance might be excessive for high-frequency applications. We note, however, that some advanced varistor materials have been fabricated with lower dielectric constants.

An alternate construction is shown in Figure 10, where the varistor disc has two concentric electrodes that are silk-screened on the side of the disc, rather than on the inside and outside surfaces, as in Figures 8 and 9. In this manner, the disc can have a substantial distance between the inside and outside surfaces, reducing the resulting capacitance. The clamping voltage is then determined by the distance between the edges of the two electrodes. This arrangement offers the advantages of lower
capacitance and greater accuracy in the distance between electrodes, albeit at the cost of reducing the energy handling capability since only a thin surface layer of the varistor disc will conduct the surge current.

Coaxial Mounting

A varistor can also be inserted in a shielded cable by a coaxial "Tee" configuration. In fact, that configuration is required for making valid measurements of the varistor performance in the sub-microsecond time range.\(^{(3)}\)

Here again, the capacitance of the varistor might be objectionable at high frequencies. A possible remedy to this situation is shown in the configuration of Figure 11, where the capacitance of the varistor is series-tuned to the inductance of the clamping spring. This achieves high impedance to the UHF signal frequency, thus low loading, but offers a lower impedance and good suppression at the lower frequencies associated with lightning surge currents.\(^{(6)}\) This would not, of course, be suitable for NEMP protection.

Reconstituted Material

In the two preceding examples, the description of the varistor simply stated "varistor disc," with the assumption that the material had been prepared in the conventional sequence of pressing a powder into shape and sintering the piece in its final form. There is another method of producing varistor components, especially attractive for parts shaped in a form different than conventional mass-produced discs.

In this method, the varistor material, obtained by the usual sintering, is ground into particles of controlled dimension, mixed with a suitable binder in proper proportions, and molded into the finished part. By this method, it is possible to embed electrodes and conductors directly into the varistor material, as well as to fill a cavity such as a conductor shell.\(^{(7)}\)

Figure 12 shows the structure of a portion of a device using reconstituted material (tracing of a photomicrograph), where the three constituents interface: electrode, particle of varistor material, and plastic binder. Here, the particle of varistor material is formed of several grains of ZnO with the varistor boundary layer between the individual ZnO grains. If the particles of varistor material were smaller than the ZnO grains, there would not be any barrier layer left and the varistor action would be lost. Thus, there is a need to optimize the grain size, the particle size, and the proportion of binder, depending on the specific application.

Thick Film

The reconstituted material just described might also be useful for producing varistor layers with thick-film technology. Currently, the reconstituted material is obtained by pressing the particles of varistor material in a cavity. Producing a paint-like film would require some development in order to maintain the particle/binder proportion leading to
Figure 13. Two-pin connector with reconstituted varistor protection.

Figure 14. Plug-type insert of reconstituted varistor.
Figure 15. V-I characteristic of two-pin connector.

Figure 16. V-I characteristic of plug-type varistor.
the behavior as a varistor. Should the particles become isolated from one another by an excess of binder, the varistor action would be lost.

With a suitable paint-like compound, it would be possible to silk screen varistor layers directly between conductor runs. The clamping voltage would be controlled by the distance between the runs, while the current density would be controlled by the combined thickness of the layer and length along the run.

As an alternate to painting layers of reconstituted varistor material between runs, a more advanced concept would include the sintering of varistor layers on a suitable substrate, on which runs and other components could be added in subsequent processing. This concept has not been reduced to practice, but illustrates the wide range of options that may be available beyond the commercial packaging of discs.

**Substrates**

Another approach to the integration of varistor protection to circuits would be the use of substrate slabs of sintered varistor material, on which the circuit runs could be metallized. Two possible concepts are proposed. In Figure 17, the substrate acts only along a thin layer between two runs located on the same side of the substrate. Because of the highly nonlinear property of the material, there is little fringing of current into the depth of the substrate.

![Figure 17. Varistor substrate with surface conduction between runs.](image)

In Figure 18, the substrate acts as an insulating barrier between the runs on one side and the ground plane on the other side. Under normal operating voltages, the substrate layer offers a very large resistance and, thus, is nearly an insulator. Upon development of a high voltage between any part of the circuit and ground, the surge is immediately diverted to the ground plane without any additional lead length.

![Figure 18. Varistor substrate with bulk conduction to ground plane.](image)

In the latter approach, development of optimum compositions would allow producing a substrate with sufficient mechanical strength and a voltage which is still low enough. Here we are dealing with the suppression of transients of moderate energy at the point where they may be induced (presumably as the circuit would be illuminated by an EM field).

**Varistor Materials Technology**

Many of the unconventional varistor applications listed above required tailoring the varistor materials properties for optimum implementation of the system requirements under consideration. Using technology available or under development, we can modify the varistor material in a variety of ways. For example, we may:

1. Increase the varistor breakdown field, thereby decreasing the space necessary for a given breakdown voltage.
2. Decrease (or increase) the varistor dielectric constant.
3. Fabricate screen-printable thin film varistors.
4. Fabricate large (up to about \(10^3\) g weight and 30 cm diameter) devices for extremely large energy or current carrying capacities.
5. Directly press varistors in special shapes (e.g., rings or slabs) to fit specific connector systems.
6. Fabricate low-voltage varistors with breakdown voltages as low as 15 to 20 volts.
7. Vary the degree of varistor nonlinearity.

In general, optimization of a particular varistor material parameter will involve a trade-off with other material properties. The proper component design will involve the synergistic combination of materials technology and circuit design expertise.

**CONCLUSIONS**

Varistor materials have been shown suitable for applications other than the conventional lead-mounted device used in electronic circuit boards or the large blocks used in power system surge arresters. The feasibility of some approaches has been shown, with details of implementation to be developed. For other approaches, the development is still at the conceptual stage.
ACKNOWLEDGMENTS

The concepts and technologies discussed in this report represent the contributions of many others, associated with metal oxide varistor technologies, some since their early introduction into the United States by General Electric. The authors listed in the references contributed to the concepts discussed in the present report and should be acknowledged as implicit co-authors of this report.

REFERENCES


